Air Force Institute of Technology

AFIT Scholar

3-2003

Developing a Leadership Commitments Inventory

Abigail L. White

Follow this and additional works at: https://scholar.afit.edu/etd

⊙ Part of the Leadership Studies Commons, and the Performance Management Commons

DEVELOPING A LEADERSHIP COMMITMENTS INVENTORY

THESIS

Abigail L. White, Captain, USAF

AFIT/GLM/ENV/03-03

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

AFIT/GLM/ENV/03-03

DEVELOPING A LEADERSHIP COMMITMENTS INVENTORY

THESIS

Presented to the Faculty

Department of Operational Sciences

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Logistics Management

Abigail L. White, B.S.

Captain, USAF

March 2003

DEVELOPING A LEADERSHIP COMMITMENTS INVENTORY

Abigail L. White, B.S.
Captain, USAF

Approved:

_____// signed // _____   05 March 2003
Paul W. Thurston, Ph.D. (Chairman)   date

_____// signed // _____   05 March 2003
William A. Cunningham, III, Ph.D. (Member)   date

_____// signed // _____   05 March 2003
Daniel T. Holt, Ph.D. (Member)   date

Acknowledgments

I would like to express my sincere appreciation to my faculty advisor, Maj Paul Thurston, for always taking the time and interest necessary to ensure the completion of my thesis was a true learning experience and not simply an academic exercise. I would also like to thank my thesis committee members, Dr. William Cunningham, III and Maj Daniel Holt for their patience as well as their guidance throughout the course of this thesis effort.

I am also indebted to my thesis team, Capts Theresa Wardak and Matt Douglas, without whose support I never could have accomplished this thesis.

Table of Contents

List of Figures

List of Tables

AFIT/GLM/ENV/03-03

Abstract

Organizations have been modifying performance appraisal systems to collect data from multiple sources to guide supervisor development. Multiple rater programs provide leaders with more than one source of feedback, and give them a more complete, and possibly more accurate picture of their performance, which can facilitate meaningful behavioral change. The purpose of this thesis was to develop and validate a Leadership Commitments Inventory (LCI) from a 360-degree feedback approach that reflected a six-factor leadership taxonomy composed of twelve distinct indices. Using successive content adequacy tests, 48 items were validated and the web-based LCI was administered to 278 participants. Internal consistency estimates for the twelve commitment scales differed for supervisors and observers. The internal consistency estimates (indexed by Chronbach's alpha) for supervisors ranged from .27 to .79 (8 of 12 were .68 or greater). Internal consistency estimates for the observer sample were much better, with all twelve scales greater than .70. A nested confirmatory factor analysis using LISREL demonstrated that the hypothesized six-factor model clearly provided the best explanation of the LCI's underlying factor structure, with 11 of the 12 commitments loading on their hypothesized practice. The instrument is now available on the Internet for supervisor's who wish to develop into better leaders.

Developing a Leadership Commitments Inventory

I. Introduction

Background

In the last decade, many organizations have modified their performance appraisal

systems to collect data from multiple sources to guide the development of supervisors

(London & Smither, 1995). As the term suggests, multiple source programs provide

leaders with more than one source of feedback, such as subordinates, peers, or customers.

By giving the supervisors a more complete, and possibly more accurate picture of their

performance, supervisors can make meaningful behavioral changes. Multiple rater

feedback ultimately focuses on providing supervisors with developmental feedback.

Developmental feedback programs provide an opportunity to improve leadership

practices by giving supervisors feedback about their performance from a number of

sources that include themselves, supervisors, peers, subordinates, and team members.

This thesis expanded on Patton's (2002) upward feedback instrument (UFI) pilot study

that developed and validated an instrument that measured leadership behaviors at the

commitment level for supervisors at the Aeronautical Systems Center (ASC) and the Air

Force Security Assistance Center (AFSAC). Specifically, this thesis intended to develop

an updated feedback instrument that contained refined leadership statements to better

measure supervisor's leadership behavior at work.

The Leadership Practices Inventory (LPI), a commercial instrument developed by

Posner and Kouzes (1988), measured five leadership constructs at the practice level. The

UFI, developed by Patton (2002), measured six leadership constructs at the commitment

level. Five constructs, called practices, were modeled after those presented by Posner

and Kouzes and were named *Challenge the Process, Inspire a Shared Vision, Model the Way, Enable Others to Act,* and *Encourage the Heart*. Each practice had two commitments that defined actions the leader should take into consideration to develop the leadership behaviors described by that construct. A leader could *Challenge the Process* through the two commitments *seek out challenges to innovate and improve* and *try ideas, take risks and learn from mistakes*. *Inspire a Shared Vision* was defined by the two commitments *create a vision* and *attract others to a common purpose*. The commitments included in the *Model the Way* construct were *set the example* and *motivate and build commitment through small victories*. *Enable Others to Act* entailed *sharing information and power* and *encouraging trust and cooperation*. And finally, the leader *recognized and rewarded individual performance* and *celebrated team accomplishments* to *Encourage the Heart*.

The UFI's sixth construct, developed at the request of AFSAC and ASC, was termed *Have Fun* and reflected the extent to which supervisors engaged in behaviors that were designed to relieve stress and tension in the workplace. The two commitments for the *Have Fun* construct were *allow humor to reduce stress and boredom* and *promote fun activities to relax and unwind*.

The Leadership Commitments Inventory (LCI) was modeled and developed after the research conducted by Kouzes and Posner (1997, 2002) and Patton (2002). It maintained the original five practices from Kouzes and Posner's LPI, and the sixth practice from Patton's UFI now called *Enjoy the Workplace*. Three significant factors led to the development of the new instrument.

First, based on the content validity and factor analysis from Patton's (2002) pilot UFI study, the six constructs and 12 commitments needed further refinement. Specifically, the 12 commitments were not viewed as distinctly different constructs and the leadership concepts overlapped based on the pilot UFI items. Due to the high inter-correlations among the items on the UFI, each practice and commitment was newly defined, which led to over half of the pilot questions being modified or deleted.

Secondly, in September 2002, Kouzes and Posner released an updated version of *The Leadership Challenge*. The most significant change in their updated version was the revision of the *Model the Way* practice. Previously, *Model the Way* was captured by the two commitments *sets the example* and *achieve small wins*. They revised the practice so that *achieve small wins* was incorporated into *Challenge the Process*, and instead they replaced it with *find your voice*. In addition to modifying the *Model the Way* practice, they also re-ordered the practices in the book to better reflect the leadership development process. The new order of the practices is: (a) *Model the Way*, (b) *Inspire a Shared Vision*, (c) *Challenge the Process*, (d) *Enable Others to Act*, and (e) *Encourage the Heart* (Kouzes and Posner, 2002). Based on the research by Kouzes and Posner, *find your voice* was revised for the LCI to read *shares personal values*.

The third factor that led to the new LCI was a shift from a purely upward feedback instrument to a 360-degree approach. While the UFI focused on feedback from subordinates only, the LCI now measures feedback from subordinates, peers, and team members in an attempt to replicate an entire 360-degree approach.

Research Focus

As part of the follow-on research to Patton's (2002) development of the UFI, this thesis developed and validated a refined feedback instrument that generates reliable and valid data regarding leader behavior at both the practice and more specific commitment level. The instrument measures leadership principles valued by ASC and AFSAC and provides specific statements to measure leader development. The information provided by the leadership commitments inventory identify gaps between leaders' and subordinates', team members', and peers' beliefs about that leader's behaviors.

Overview of Paper

The remainder of this paper is divided into four chapters. Chapter II begins by reviewing existing research literature on upward feedback and multiple rater feedback. The literature review first focuses on the growing popularity of multiple rater feedback, its value, usefulness, and the benefits of accurate multiple rater feedback for developing better leaders. The chapter then presents alternative leadership taxonomies posited by the Ohio State University, the University of Michigan, and Posner and Kouzes (1988). Posner and Kouzes' developed the Leadership Practices Inventory, which posited a 5-factor leadership taxonomy; its psychometric properties and supporting research are discussed. The next section discusses the upward feedback instrument developed by Patton (2002), its six-factor leadership taxonomy, and includes an in-depth look at the definitions and items he proposed for each of his 12 commitments. The last section of the chapter presents the newly refined practices and commitments for the leadership commitments inventory.

Chapter III outlines the development of the final 48 items that were included in the leadership commitments inventory based on the new definitions of the practices and commitments. Two different content analyses were performed on the items. The first section reports the results of the first version of content analysis, while the second section reports the results of the second content analysis which ultimately led to the final items. The third section reports the descriptive statistics of the self and observer scales, and concludes with a discussion on nested confirmatory factor analysis. Chapter IV presents and analyzes the results of the LCI nested model confirmatory factor analysis. The thesis concludes with a discussion of the research findings, the studies limitations, implications, and recommendations for future research.

II. Literature Review

Introduction

The following literature review contains four sections. The first section addresses the growing popularity and benefits of multiple-rater feedback programs. This section also addresses the effects of multiple rater feedback on supervisor's leadership development. The second section discusses the research of Posner and Kouzes (1988) which led to the development of the Leadership Practices Inventory (LPI). This section presents the psychometric properties of the LPI and discusses other studies that have employed the instrument for leadership development. The third section presents the Upward Feedback Instrument (UFI) developed by Patton (2002). The UFI was modeled after Kouzes and Posner's (1997) leadership taxonomy, and forms the basis for the Leadership Commitments Inventory (LCI) developed in this study. The final section of this chapter discusses the changes from the UFI to the LCI, and the specific definitions of the LCI practices and commitments.

The Value of Multiple-rater Feedback Programs

Organizations have been augmenting their performance appraisal systems with programs designed to collect data from multiple sources for a decade. London and Smither (1995) report that use of multiple rater feedback "is nearly universal among Fortune 500 firms" (p. 804). Sources of feedback can include supervisors, subordinates, peers, team members, and customers. Many feedback programs also provide an opportunity for supervisors to rate themselves. Self-assessments facilitate comparisons between how leaders view themselves and how their constituents view them. By giving leaders a more complete, and possibly more accurate picture of their performance,

advocates of these programs hope that leaders can make meaningful behavioral changes that will benefit the organization. Benefits come largely from the connection that employee productivity and morale is a result of effective leadership (Pinsonnault, 1992). Employees should have effective management, and managers should possess the skills to be effective. Multiple rater feedback is a tool to ensure that competence is obtained and maintained by organizational leaders (Cockfield, 1996). As multiple rater feedback programs are becoming more popular, organizations hope that these programs will help develop better leaders.

Multiple rater feedback is different from traditional downward directed performance appraisals. Downward feedback is often used in performance appraisals and is usually administratively focused rather than developmentally focused. While formal appraisal systems are frequently deemed critical to promotion and pay they are less effective for improving performance. Feedback provided through formal and permanently documented appraisal systems are often ineffective for employee development because supervisors are hesitant to provide critical feedback when associated with potentially punitive outcomes. Developmental feedback (a critical component of multiple rater feedback) is typically not part of the performance appraisal process. Organizations use multiple rater feedback programs to measure leadership behavior and then offer feedback to the supervisor specifically for developmental purposes. Consequently, multiple rater feedback programs provide the opportunity for leaders to receive honest and critical feedback without fear of administrative punishment.

Multiple rater programs provide comprehensive feedback from a variety of different perspectives. Subordinates, peers, and team members interact with a supervisor

in distinctly different contexts, each group observing different behaviors. Multiple rater feedback programs capture a larger proportion of a leader's behaviors and thus leads to more accurate and complete feedback.

The developmental focus combined with multiple perspectives provides a more accurate assessment of the leader's behaviors. Leaders who perceive feedback as more accurate are also more likely to act on the information. Ilgen, Fisher, and Taylor (1979) posited that the feedback process can be broken into four stages: perception of feedback, acceptance of feedback, desire to respond to feedback, and the intended response. Perception of feedback is concerned with how accurately the recipient perceives the feedback from any given source. Acceptance of feedback refers to recipients' beliefs that the feedback is an accurate portrayal of their behavior or performance. Acceptance is the most critical aspect of a recipient acting on feedback. Brett and Atwater (2001) applied the feedback model and learned that if recipients accepted their feedback as accurate, they were more likely to actually change their behavior. Feedback is most likely accepted as accurate if the source was reliable and credible and if the feedback was positive.

Feedback for Developmental Leadership

Multiple-rater feedback programs can be used successfully for leadership development purposes. According to McEnvoy and Buller (1987), subordinate and peer ratings are considered to be extremely useful when they are used for developmental purposes. London and Smither (1995) provided two assumptions for developmental feedback. First, this type of feedback helps leaders understand how others view them, which allows them to develop an accurate sense of goal accomplishment and self-

competence. Second, developmental feedback provides information on possible areas for skill development and performance improvement. Given these assumptions, leaders should take measures to improve the targeted behaviors. Therefore, developmental feedback moves the leader through the stages of the feedback process.

For the purpose of this study, upward feedback describes a specific multiple rater feedback program in which subordinates rate their supervisors' work performance (London & Wohlers, 1991) for developmental purposes. Upward feedback does not rely on the established downward feedback processes and forms that are traditionally used within the organization's formal appraisal systems. Instead, upward feedback programs usually involve subordinates using a survey to rate their manager's leadership behavior in the specific work situations.

In a study of upward feedback done by London and Wohlers (1991), 65% of the raters said their feedback should be used only for their manager's leadership development. Additionally, the raters were more willing to be accurate and honest when they believed that their feedback was only to be used for developmental rather than formal appraisal purposes. Subordinates and supervisors both thought that feedback for developmental purposes would produce fewer negative outcomes than if the feedback had been used for formal appraisals and administrative decision making (Bettenhausen & Fedor, 1997).

A developmental feedback program should focus on providing positive change for the leader and the organization through specific feedback (Bracken, 1994). Specific feedback is very useful for leadership development because it can highlight precise areas for improvement (Atwater & Roush, 1995).   A good program will underscore

development areas for the leader, encourage the leader to set goals based on desired improvement in those areas, and ultimately lead to behavior change (London & Smither, 1995). To implement a successful program, top management should introduce the program to let leaders know that developmental feedback is important to the organization. Leaders must also be given the discretion to seek feedback, develop action plans, and change behavior based on the feedback (Alimo-Metcalfe, 1998).

According to London and Smither (1995), multiple rater programs assume that feedback will help managers better understand how others view them and highlight areas for performance improvement and skill development. As Vinson (1996) stated, "the objective is to identify behavioral areas for improvement" (p. 11).

Usefulness of Upward Feedback

According to Adsit, London, Crom, and Jones (1997), upward feedback enhances supervisors' knowledge of their strengths and weaknesses, and is a key dimension in supervisor-subordinate relationships. Waldman and Atwater (2001) stated that most leadership behavior is directed towards subordinates, and they therefore "constitute a logical input source for feedback" for leadership development (p.189). Therefore, upward feedback is an excellent tool in the organization to provide feedback to supervisors on their work performance.

A number of program characteristics influence the extent to which upward feedback is useful, namely, the way the data are collected, the way the data are presented, and the instrument used to collect the data. First, several researchers (e.g., London & Wohlers, 1991; Baron, 1996; Atwater & Roush, 1995) agree that upward feedback is most effective if done anonymously. Anonymity allows observers to accurately and

honestly provide feedback to the manager without fear of reprisal. At least two subordinates rate on one supervisor and their scores are averaged, with the mean response reported back to that supervisor (London & Wohlers, 1991). Second, the way the data are collected and presented influences the leader's reaction. During the upward feedback process, managers rate themselves and subordinates rate the managers. The differences between self and subordinate ratings are compared. Items that have relatively large discrepancies between self and subordinate ratings indicate areas for assessment and improvement. Additionally, leaders tend to prefer individualized feedback (specific feedback from their raters) compared to normalized feedback (feedback on how the average supervisor was rated; Smither & Wohlers, 1995). Finally, because leaders tend to act on accurate information, the feedback instrument chosen for the developmental feedback program should be reliable and valid (Bracken, 1994; Fletcher, Baldry, & Cunningham-Snell, 1998; Morical, 1999; Vinson, 1996) and should have been developed based on statistical methods (Vinson, 1996).

A possible shortcoming of upward feedback is rater leniency.   Subordinates sometimes have a difficult time honestly appraising their supervisor, and are often reluctant to give negative feedback because they fear reprisal, or they think they will negatively impact their boss. According to London and Wohlers (1991), the best way to overcome rater leniency is to instill trust in the upward feedback process over time through training and communication. Furthermore, supervisors should show appreciation for their feedback and develop action plans based on the results of their self and subordinate ratings (London & Wohlers, 1991). Upward feedback is also limited in that it only offers one perspective to the supervisor.

Comprehensive Multiple Rater Feedback Leads to Increased Accuracy

Compared to multiple rater feedback programs, upward feedback systems may offer an incomplete picture due to the limited interaction between a supervisor and their subordinate. The multiple rater approach allows different perspectives by collecting data from those who should have more frequent, recent, and salient interactions.

The value of multiple rater feedback is that it goes beyond the formal appraisal system and provides managers with a more comprehensive picture of their strengths and weaknesses (London and Smither, 1995). A 360-degree feedback program gathers multiple sources of information, thereby providing a more diverse perspective that offers managers broad feedback on relevant and measurable behaviors (Waldman & Atwater, 1998; London & Wohlers, 1991). Theoretically, as supervisors receive more feedback for developmental purposes they will perceive that feedback as accurate and change their behavior to better improve their weak areas as they become more self-aware, which is an integral part of managerial competence (Atwater & Roush, 1995; McCarthy & Garavan, 1999). Because of this, 360-degree feedback provides the manager information that would not be available through more traditional downward feedback or appraisal systems. Ideally, a successful 360-degree program will lead to better leadership development, increased levels of trust and communication in the organization, and improved customer satisfaction (Waldman & Atwater, 1998).

Hoojiberg and Choi (2000) argue that 360-degree feedback is more effective than upward feedback because it provides more comprehensive feedback for the leader. Frequently, leaders face demands from more people at work than just their subordinates. However, leaders do not typically understand the demands and expectations of their

customers and constituents as well as they understand their subordinate's demands and expectations (Hoojiberg & Choi, 2000). By collecting feedback from direct reports, indirect reports, peers, and team members, it is assumed that different rating sources offer performance information on different leadership dimensions that add "incremental validity to the performance evaluation" (Brutus, London, & Martineau, 1999, p.678). The more feedback perspectives the manager receives, the more likely he or she is to accept that feedback as accurate and agree with it (Waldman & Atwater, 2001). Therefore, the 360-degree feedback will more accurately reflect the manager's "all-around" leadership behavior.

Multiple Rater Feedback Program Summary

Whether using upward feedback or 360-degree feedback, the overarching purpose of both programs is to provide accurate advice that supervisors will accept so they can improve their leadership behaviors. Both feedback processes provide managers with valuable criticism that highlights their strengths and weaknesses, based on the input from multiple raters. As reported by Reilly, Smither, and Vasilopoulos (1996), sometimes the most valuable aspect of multiple rater feedback is the exposure to desired organizational behaviors as noted in the type of feedback collected. When subordinates, peers, and team members note what organizational values they are asked about in regards to their leader's behavior, they better understand what is important to the organization. Consequently, the items in the instrument should focus on behaviors (not traits) that are important to the organization and highlight the organization's values (Bracken, 1994; London & Smither, 1995; Morical, 1999).

The Leadership Practices Inventory

Emphasis on feedback as a tool to improve leadership performance suggests that what leaders do, their observable behaviors, is an important determinant of organizational effectiveness. Many researchers have attempted to develop leadership behavior taxonomies through successive leadership behavior studies in an attempt to classify observable leadership behaviors. In the 1950's, both the Ohio State University and the University of Michigan attempted to identify relevant leadership categories which adequately described different leadership styles. The Ohio State study classified 150 different leadership behaviors into two broad categories which they called *Consideration* and *Initiating Structure* (Yukl, 2002). The Ohio State studies led to the development of several surveys to measure leadership behavior (e.g., the Leadership Behavior Description Questionnaire, Supervisory Behavior Description Questionnaire, Leadership Opinion Questionnaire, and the Leadership Behavior Description Questionnaire III). Recent studies have used the two-factor taxonomy proposed by the Ohio State study and have renamed the categories *Task* and *Relations* (Yukl, 2002).

At the same time as the Ohio State studies, the University of Michigan studied the relationship between leader behavior, group processes, and measures of the group's performance. They found three distinct leadership categories emerged to describe leadership behaviors: *Task-oriented Behavior*, *Relations-oriented Behavior*, and *Participative Leadership* (Yukl, 2002). The Michigan studies were the first to examine leadership from both a supervisor and peer perspective.

However, both the two-factor and three-factor leadership taxonomies have been criticized by follow-on studies because they have produced contradictory results and have

been called "weak and inconsistent" in determining the most effective leadership style

(Yukl, 2002, p. 52). Furthermore, recent leadership experts argue that the Ohio State

University and University of Michigan studies are over-simplified and do not adequately

capture the complexities of leadership behaviors. Both studies sparked a great interest in

finding the most effective taxonomy for classifying leadership behaviors.

Posner and Kouzes' (1988, 1993, 2000) research suggest five categories of

leadership behaviors contribute to positive organizational outcomes. The authors asked

1,100 participants to describe "their personal best experience as a leader" (1988, p. 484).

Additionally, they conducted 38 in-depth interviews with managers about their personal

best experiences. The researchers then analyzed the content of the surveys and

interviews and found that five distinct leadership practices emerged from the data. These

five practices "were then validated by two separate outside raters" (1988, p.484). Posner

and Kouzes found that more than 80% of the leadership behaviors described in the

surveys and interviews emerged in one of five leadership practices: (a) *Challenge the*

Process, (b) *Inspire a Shared Vision*, (c) *Enable Others to Act*, (d) *Model the Way*, and

(e) *Encourage the Heart*. Kouzes and Posner (1997, 2002) further separated each

practice into two commitments. For instance, *Challenge the Process* contained the

commitments *search for opportunities* and *experiment and take risks*.

After defining each practice and its corresponding commitments, Posner and

Kouzes (1988) wrote "behavior-based statements" (p. 485) to reflect each practice. They

asked 120 MBA students to complete the pilot survey, with a 5-point Likert type scale to

rate how often they engaged in the leadership statements. Through several content and

construct validity tests, they validated the survey. The updated instrument, the Leadership

Practices Inventory (LPI), was ready to be administered in full. The remainder of this section examines the psychometric properties of the LPI and then looks at the utility of the instrument in providing developmental feedback to leaders.

Psychometric Properties of the LPI

Posner and Kouzes (1988, 1993, 2000) conducted three large studies that explored the fundamental properties of the LPI. The initial launch of the LPI included over 2,800 participants. Approximately 700 completed the LPI-self, which was the supervisor specific survey that measured their own leadership behaviors. About 2,100 participants completed the LPI-other (now called the LPI-observer), which was the survey for "others" to rate the leader's behaviors. The researchers analyzed the internal consistency of each practice-scale in the LPI. For the LPI-self, the reliability (coefficient alpha) was .73 for the *Challenge* practice, .83 for the *Inspire* practice, .70 for the *Enable* practice, .72 for the *Model* practice, and .90 for the *Encourage* practice. On the LPI-other, the reliability was .79 for *Challenge*, .89 for *Inspire*, .86 for *Enable*, .81 for *Model*, and .91 for *Encourage* (Posner & Kouzes, 1988). All internal consistency estimates proved to be sufficient for the practice-scales, based on the commonly accepted level of .70 (Nunnally and Bernstein, 1974).

Posner and Kouzes (1988) also checked for scale stability with a test-retest reliability format using 57 MBA students. The test-retest reliability coefficient after 10 days, averaged across the five practices was .94, which demonstrated excellent temporal stability (Posner & Kouzes, 1988). They further tested for contamination with extraneous response determinants and checked for social desirability response bias. Posner and Kouzes used the Marlowe-Crowne Personal Reaction Inventory to test whether

participants were answering the questions in an effort to create a desirable picture of him or herself. They found that none of the correlations were statistically significant ($p < .01$); therefore no contamination was present.

In 1993, Posner and Kouzes published an updated analysis of their LPI model that included a considerably larger, more diverse sample ($N = 36,226$ managers and their subordinates who had never taken the LPI previously). They used principle factor analysis with varimax rotation to analyze the factor structure of the larger sample size. All items loaded on the five factors as expected and the results mirrored the original LPI factor structure (Posner & Kouzes, 1993). The internal consistency (coefficient alpha) for the LPI-self ranged from .70 to .85, and the LPI-observer reliabilities ranged from .81 to .92 (Posner & Kouzes, 1993). The estimates of test-retest reliability (time unknown) were determined using 157 MBA students, and were reported at above .93 (Posner & Kouzes, 1993).

Posner and Kouzes (1993) specifically studied differences between self and observer scores and differences among gender, cultural, and ethnic backgrounds. They determined that male and female managers differed significantly with *Model the Way* and *Encourage the Heart*. However, there were no significant differences in LPI scores across ethnic or cultural backgrounds. Based on the results of Posner and Kouzes' studies, the opportunities to develop better leaders probably will not be compromised by gender, ethnic background, or organizational differences.

In 2000, Posner and Kouzes again reviewed the psychometric properties of their LPI. Based on weaknesses in discriminating between leadership practices, which was found in studies such as Bowles and Bowles (2000), they changed the Likert scale from a

5-point scale to a 10-point scale so that participants could better differentiate their answers. They also refined several questions and re-administered the LPI to a sample of over 17,000 managers and their direct reports, co-workers, and others (Posner & Kouzes, 2000). The estimates of internal consistency ranged from a low of .75 to a high of .87 (coefficient alpha) on the LPI-self, and from a low of .88 to a high of .92 (coefficient alpha) for the LPI-observer; these results were consistent with the 1988 and 1993 samples (Posner & Kouzes, 2000). Table 1 summarizes the reliability coefficients, test-retest reliabilities, and social desirability results from Posner and Kouzes' (1988, 1993, 2000) studies.

Construct validity was tested using exploratory factor analysis, which yielded the same five factors as previous studies. Table 2 summarizes the factor loadings found by Kouzes and Posner ($N = 43,899$), as of 1997. It does not include data from their 2000 study. While the factor structures emerged as hypothesized, Posner and Kouzes later explained that, "some statements loaded on more than one factor" (2000, p. 2) which was also noted by Carless (2001) and Patton (2002).

Several studies have tested the five-factor taxonomy. Fields and Herold (1997) conducted the first known confirmatory factor analysis of the LPI. They tested the five-factor model that was posited by Kouzes and Posner and found it had "an acceptable fit to the data" (1997, p. 578). Fields and Herold's sample included over 2,300 managers and their subordinates who anonymously completed the LPI. The researchers conducted a first order confirmatory factor analysis on the data and validated the psychometric properties found by Posner and Kouzes (1993). The coefficient alphas of the LPI scales ranged from a low of .82 to a high of .92, and all of the paths linking the LPI items to the

Table 1

Reliability Results from 3 Studies Evaluating the LPI

		Coefficient alpha	Test-Retest Reliability	Social Desirability
1988 Study	Challenging	.77	.93[a]	.13
N = 2,876	Inspiring	.88	.94	.04
	Enabling	.84	.94	.24
	Modeling	.80	.95	.29
	Encouraging	.90	.93	.27
1993 Study	Challenging	.80	.93[b]	
N = 36, 226	Inspiring	.87	.93	
	Enabling	.85	.94	
	Modeling	.81	.95	
	Encouraging	.91	.93	
2000 Study	Challenging	.84		
N = 17, 908	Inspiring	.90		
	Enabling	.82	.90 (average)	
	Modeling	.82		
	Encouraging	.90		

Note: Reliability coefficients above .60 were considered adequate (Posner and Kouzes, 2002)
[a] Test-Retest estimates based on a 10-day interval
[b] Time between administrations on test-retest assessment not reported.

Table 2

Factor Loadings for the Leadership Practices Inventory (N=43,899)

Practice	Challenge	Inspire	Enable	Model	Encourage
Challenge	**.664**	.235	.173	.046	.185
Challenge	**.641**	.285	.188	.22	.153
Challenge	**.577**	.250	.147	.157	.156
Challenge	**.577**	.220	.023	.234	.094
Challenge	**.388**	.152	.246	.259	.158
Inspire	.239	**.697**	.164	.109	.236
Inspire	.262	**.662**	.162	.128	.183
Inspire	.281	**.594**	.187	.232	.235
Inspire	.375	**.505**	.267	.254	.117
Inspire	.421	**.480**	.220	.037	.288
Inspire	.300	**.439**	.317	.141	.223
Enable	.032	.074	**.717**	.096	.238
Enable	.188	.194	**.701**	.246	.231
Enable	.115	.153	**.689**	.189	.234
Enable	.118	.124	**.577**	.018	.144
Enable	.224	.252	**.506**	.215	.239
Enable	.119	.251	**.469**	.248	.233

(*table continues*)

Table 2 (*continued*)

Practice	Challenge	Inspire	Enable	Model	Encourage
Model	.221	.221	.220	**.588**	.195
Model	.156	.076	.327	**.527**	.190
Model	.220	.309	.186	**.468**	.200
Model	.220	.128	.365	**.408**	.163
Model	.238	.342	.110	**.378**	.138
Model	.230	.311	.251	**.369**	.173
Encourage	.183	.209	.153	.109	**.755**
Encourage	.121	.225	.140	.119	**.726**
Encourage	.119	.141	.370	.128	**.711**
Encourage	.146	.181	.391	.168	**.708**
Encourage	.164	.109	.327	.198	**.695**
Encourage	.233	.231	.203	.201	**.577**

Note. Modified from *The Leadership Challenge* (p. 343), by J.M. Kouzes and B.Z. Posner, 1997, San Francisco: Jossey-Bass. Copyright 1997 by Jossey-Bass.

five factors were significant (Fields & Herold, 1997).

Carless (2001) used confirmatory factor analysis to assess the construct validity of Posner and Kouzes' (1993) LPI model. She sampled over 1,400 subordinates and used confirmatory factor analysis to test three different models of leadership (Carless, 2001). Contrary to Fields and Herold's (1997) study, Carless' models showed that distinct transformational leadership behaviors could not be distinguished in the LPI due to the high correlations among the items. Since her factor analyses were not able to identify

five distinct leadership behaviors, Carless concluded that the LPI had weak discriminant validity and suggested that "feedback may be misleading and detrimental" (2001, p. 237).

Similar to the weaknesses discussed in Carless's (2001) study, Patton (2002) did not find five distinct leadership constructs in his confirmatory factor analysis. Patton noticed significant cross loading between the *Challenge* and *Inspire* behaviors and also observed extremely high correlations between *Inspire Others to Act* and *Model the Way*. The results indicated that perhaps subordinates were not able to truly distinguish between the five different leadership practices posited by Posner and Kouzes (1988; Patton, 2002).

In summary, Posner and Kouzes (2000) have amassed over 350,000 data points over two decades to validate the LPI. When Posner and Kouzes (2000) studied variation due to demographics, they found that, "LPI scores have been found, in general, to be unrelated with various demographic characteristics … or organizational features" (p. 8). Despite the successive internal reliabilities and consistent factor analysis conducted by Posner and Kouzes, researchers such as Carless (2001) and Patton (2002) were not able to replicate their results.

LPI Validity Found in Other Studies

Several researchers have studied the effectiveness of the LPI and whether it is able to influence intentions to change or actual behavioral changes in leaders. Studies that successfully used the LPI to differentiate between leader groups indicate the instrument is able to discriminate leadership effectiveness, thereby giving some evidence that the LPI may measure distinct leadership constructs. Wunderly, Reddy, and Dember (1998) compared the results of the LPI with those of the Kirton Adaption-Innovation Inventory and a measure of optimism and pessimism. The three measures were used in

concert to assess changes in manager's optimism and pessimism over time. The sample of 196 respondents completed all three instruments and the means of the LPI-observer were consistent with the results reported by Posner and Kouzes (1993). Wunderly et al. found that *Inspire a Shared Vision* and *Encourage the Heart* had a significant, positive correlation with optimism. This study supports the theory that the LPI does measure meaningful constructs and that by positively relating two LPI practices with optimism, the LPI can help distinguish two groups: optimists and pessimist. This was the first study to relate optimism and pessimism scores with LPI-observer results and offers further support for the convergent validity of the LPI model.

Bowles and Bowles (2000) used the LPI to discriminate between two different nursing groups in England. The Nursing Development Units were developed as centers of nursing excellence, innovation, and leadership development, while non- Nursing Development Unit nurses worked in conventional clinical settings. The LPI results of two equal samples of 70 nurses found there were significant inter-group differences between the two groups. The Nursing Development Unit nurses had higher LPI scores, which "suggests that Nursing Development Unit leaders have enhanced leadership potential" (Bowles & Bowles, 2000, p.71). Their findings provide evidence of discriminant validity for the LPI model because it was able to discriminate between individuals that were expected to display different levels of leadership behavior. As Bowles and Bowles said, "this finding tends to support the claim that the LPI does measure different aspects of leadership" (p. 74).

Shoemaker (1999) conducted a third study that used the LPI to measure the leadership effectiveness of sales managers. She regressed the five LPI practices against

self-efficacy, role clarity, and job satisfaction. She found a significant relationship between all five practices and job satisfaction, and four of the five practices were significant when regressed against role clarity. Based on the results of her study, she found that the use of the five practices "enable the sales managers to positively impact his or her individual salespeople" (Shoemaker, 1999, p.10). Therefore, she found that the five distinct leadership practices exist and contributed to the accomplishment of her sales managers.

Kusy and Essex (1995) used the LPI in conjunction with a demographic survey followed up by one-on-one interviews with physician leaders in the United States. They found that doctors in their study used the five leadership behaviors more frequently than the population from Posner and Kouzes (1988) studies (i.e., the physicians had higher mean scores for each practice). They further found that the older the physician, the more frequently he or she used the behaviors. Therefore, their results indicate that the five leadership constructs can be learned and practiced. With their study, Kusy and Essex appeared to link Posner and Kouzes' leadership practices to successful physicians who were known to be effective: "The leaders in our study appeared to report even stronger than average utilization of behaviors associated with effectiveness by the Leadership Practices Inventory" (Kusy & Essex, 1995, p.16). Kusy and Essex's study lend further evidence to the validity of the LPI.

Summary

Based on the multitude of studies using the LPI, the instrument may generalize across different populations, times, and settings, all indicative of excellent external validity. The only conflicting validation arose with the results of the confirmatory factor

analyses done to replicate the five leadership factors. Consequently, construct validity

may not be as sound for the LPI and there may be problems classifying the different

leadership constructs on future leadership instruments. The validity studies seemed to

indicate that the LPI does successfully discriminate between different leader groups, and

it is an effective tool at measuring distinct leadership behaviors. Therefore, the measures

that Posner and Kouzes (1988) developed should adequately assess the leadership

behaviors of interest.

Although Kouzes and Posner (2002) describe their behavior taxonomy at the

commitment level, their index measures leadership behaviors at the practice level. This

drawback has led to the development of a similar instrument titled the Upward Feedback

Instrument, which measures leadership behaviors at the more specific commitment level.

The Upward Feedback Instrument

The Upward Feedback Instrument (UFI) developed by Patton (2002) was

designed to measure six different leadership constructs. Five were modeled after those

presented in Posner and Kouzes' (1988, 1993, & 2000) studies. The sixth construct was

originally termed *Have Fun* and reflected the extent to which supervisors engaged in

behaviors that were designed to relieve stress and tension in the workplace. As with the

five practices presented by Posner and Kouzes, the *Have Fun* practice was separated into

two commitments, namely *allow humor to reduce stress and boredom*, and *promote fun*

activities to relax and unwind.

Patton (2002) deviated from Kouzes and Posner by focusing his feedback

instrument at the commitment level. Patton hoped that commitment level feedback

would increase the utility of the assessment, prove to be more specific, more direct, and

need less interpretation (Patton, 2002).   Therefore, Patton developed an instrument that tested 12 different leadership commitments, with each commitment measured by five separate items.

Definitions of 12 UFI Commitments

The first two commitments Patton (2002) defined measured the *Challenge* practice. The first commitment (C1) was called *seek out challenges to innovate and improve*. He said leaders should "seek out challenges and encourage their subordinates to seek out challenges as well … leaders need to devote sufficient time to consider new ideas that may results in an improved product or service" (Patton, 2002, p. 23). He developed five items to measure the commitment such as "I encourage my people to look outside our work group to find better ways of doing things."

The second *Challenge* commitment (C2) was called *try ideas, take risks, and learn from mistakes.* Patton (2002) defined the second commitment as leaders who "promote new ideas that might increase workgroup effectiveness … volunteer their workgroup for tough assignments that are important to the organization" and turn a setback into "a valuable learning experience" (p. 24).   Five items were developed to measure the commitment, such as "I promote new ways of doing things that make us more effective."

The first commitment to measure the second practice, *Inspire*, was called c*reate a vision* (I1)*.* Leaders who *create a vision* "portray his or her unit as having a unique contribution critical to the success of the organization … underscores both the uniqueness and value of the unit's efforts" and also clearly explain their vision to subordinates (Patton, 2002, p. 25). An example of one of the five items written to

measure the first *Inspire* commitment was "I clearly explain my vision of the teams future."

The second *Inspire* commitment (I2) was titled *attract others to a common purpose.* According to Patton (2002), leaders *attract others to a common purpose* by "promoting common causes that can be supported by all members of the work group" and ensuring that "the unit goals should appeal to his or her subordinate's intrinsic desire to contribute to the success of the organization" (p. 26). One of the five items written to measure the second *Inspire* commitment read, "I explain how personal goals can be met by attaining the group's goals."

The first commitment designed for the *Enable* practice (E1) was called e*ncourage trust and cooperation.* Patton (2002) defined enabling leaders as those who "show respect for a subordinate's ideas and apply them whenever possible [and] show trust in subordinates' judgment" (p. 27). An example item written to measure the *Enable* commitment said, "I encourage my people to work with people outside of the unit."

The second *Enable* commitment (E2) was titled s*hare information and power.* Patton (2002) stated that a leader "shares information and power when he or she ensures that subordinates have the information necessary to make good judgments on their own and includes subordinates in decisions" (p. 28). Of the five items written to measure the second *Enable* commitment, one read, "I include my people when making important decisions."

The first commitment Patton (2002) defined to measure the *Model* practice (M1) was called s*ets the example.* Leaders who set the example "set priorities that are consistent with the unit's values and then act in a manner that is consistent with these

priorities" (Patton, 2002, p. 29). Five items were written to measure these leadership behaviors, and one of them said, "I set priorities that are consistent with my unit's values."

The second commitment in the *Model* practice (M2) was called *motivate and build commitment through small victories.* According to Patton (2002), an important component of the *Model* practice is to divide "large, complex tasks into smaller pieces that are more easily understood, accepted, and accomplished" and "successful leaders keep their subordinates focused on the long-term while reminding them to take things a step at a time" (p. 31). An example of one of the five statements written to measure this commitment read, "I divide large tasks into smaller pieces that are more easily understood and accepted."

In the fifth practice, *Encourage*, the first commitment (H1) was titled r*ecognize and reward individual performance.* According to Patton (2002), leaders should "be capable of tailoring his or her rewards to those things each individual subordinate value[s] … [and] show a genuine concern for their subordinates well-being and a true appreciation for their work performance" (p. 32). One of the five items written to measure the first *Encourage* commitment read, "I publicly reward individual members when they have done a good job."

The second commitment of the *Encourage* practice (H2) was called c*elebrate team accomplishments.* Patton (2002) stated that "leaders should cheer team actions that are consistent with achieving unit goals as well as take time to publicly recognize the unit's accomplishments … [and] ensure that the organization's senior leaders learn of the

group's successes" (p. 33). Five items were written to measure this commitment, and one of the items said, "I cheer actions that are consistent with achieving our unit's goals."

Patton (2002) developed the sixth practice, *Have Fun*, at the request of his sponsoring organization. Based on the model of Southwest Airlines, the *Have Fun* commitment was rooted in the theory that, "humor can help people thrive during change, remain creative under pressure, work more effectively, play more enthusiastically, and stay healthier in the process" (Frieberg & Frieberg, 1996, p.64). The first commitment in the *Have Fun* practice (F1) was called a*llow humor to reduce stress and boredom.* Leaders who display this commitment "show a willingness to laugh at himself or herself … encourage non-offensive humor and have fun with others" (Patton, 2002, p. 45). An example item written to reflect this commitment read, "I am willing to laugh and have fun with others."

The last commitment to measure the new *Have Fun* practice (F2) was called *promote fun activities to relax and unwind*. According to Patton (2002), leaders "should partake in fun activities … take advantage of any lull in the schedule and encourage simple, quick and fun activities [and] be willing to do something fun as a unit" (p. 46). One of the five items written to measure this second commitment said, "I encourage simple, quick and fun activities that lift spirits at work."

Scale Statistics and Reliability of the UFI

Table 3 shows the UFI scale means, standard deviations, and internal reliabilities from the pilot survey. The UFI used a 7-point Likert-type scale, with frequencies ranging from *not observed* to *almost always*. As can be seen from the mean scores, the majority

Table 3

Scale Means, Standard Deviations, and Reliability Indexes for the Upward Feedback Instrument

Scale	*M*	*SD*	α	*skew*	*Kurt*
Seek out challenges to innovate and improve (C1)	4.95	1.65	.91	-1.04	.42
Try ideas, take risks, learn from mistakes (C2)	4.79	1.68	.88	-.91	.05
Create a vision (I1)	5.00	1.63	.89	-.95	.17
Attract others to a common purpose (I2)	4.36	1.76	.89	-.66	-.42
Encourage trust and cooperation (E1)	5.34	1.49	.89	-1.31	1.38
Shares information and power (E2)	5.54	1.34	.89	-1.51	2.27
Set the example (M1)	4.94	1.60	.89	-1.06	.69
Motivate, build commitment with small wins (M2)	4.73	1.60	.89	-.84	.11
Recognize & reward individual performance (H1)	4.96	1.65	.87	-.85	-.06
Celebrate team accomplishments (H2)	4.70	1.79	.88	-.78	-.30
Allow humor to reduce stress & boredom (F1)	5.48	1.53	.90	-1.39	1.46
Promote fun activities to relax and unwind (F2)	4.54	1.86	.90	-.58	-.69

Note. Modified from *Developing an Upward Feedback Instrument For Supervisor Development* (p. 61), by D. Patton, 2002, AFIT: Wright-Patterson AFB.

of the UFI participants rated themselves or their leaders on the higher end of the frequency scale. Each UFI commitment scale's internal reliability was calculated using Cronbach's alpha (α), with a standard limit of .70 (Nunnally and Bernstein, 1974). The internal scale reliabilities ranged from .87 to .91, indicating the scales appear to have a high degree of internal consistency.

The test-retest reliability of the UFI was calculated using Pearson's correlation coefficients, and ranged from a low of .54 to a high of .80 (Patton, 2002). These results do not offer good evidence of temporal stability of the UFI over time. The low test-retest reliability could be due to sampling error in the small sample size (n = 28), or the instrument may not be reliable.

Confirmatory Factor Analysis of the UFI

Using confirmatory factor analysis, Patton (2002) determined that a six-factor structure best modeled the 12 leadership commitments compared to a five-factor structure. However, the constructs correlated very highly with each other, ranging from a low of .77 to a high of .98. Because the correlations among the commitments were very high, the results "cast doubt as to the true distinctiveness of the constructs as measured by the 12 UFI commitment scales" (Patton, 2002, p. 76). In conclusion, the UFI measured leadership behaviors at the more specific commitment level. However, the results of his pilot instrument indicate that further refinements should be made to ensure the instrument is truly measuring 12 different leadership constructs.

The Leadership Commitments Inventory

The LCI was modeled and developed after the research conducted by Kouzes and Posner (1997, 2002) and Patton (2002). It maintained the original five practices from Kouzes and Posner's LPI, and the sixth practice from Patton's UFI. Three significant factors led to the development of the new instrument.

First, based on the factor analysis from Patton's (2002) pilot UFI study, the six constructs and 12 commitments needed further refinement. Due to the high inter-correlations among the items on the UFI, the 12 commitments were not viewed as

distinctly different constructs and the leadership concepts overlapped. Therefore, each practice and commitment was newly defined, which led to over half of the pilot questions being modified or deleted.

Secondly, in September 2002, Kouzes and Posner released an updated version of *The Leadership Challenge*. The most significant change in their updated version was the revision of the *Model the Way* practice. Previously, *Model the Way* was captured by the two commitments *sets the example* and *achieve small wins*. They revised the practice so that *achieve small wins* was incorporated into *Challenge the Process*, and instead they replaced it with *find your voice*. Based on the research by Kouzes and Posner, *find your voice* was revised for the LCI to read *shares personal values.*   In addition to modifying the *Model* practice, they also re-ordered the practices in the book. The new order of the practices is: (a) *Model the Way*, (b) *Inspire a Shared Vision*, (c) *Challenge the Process*, (d) *Enable Others to Act*, and (e) *Encourage the Heart* (Kouzes and Posner, 2002). According to Kouzes and Posner, the new order of the practices reflects the leadership development process evolving from a personal inward inspection out to an external examination of leadership behaviors.

The third factor that led to the new LCI was a shift from a purely upward feedback instrument to a 360-degree approach. Possibly, the UFI may not have measured all of the leadership behaviors it was attempting to. While the UFI focused on feedback from subordinates only, the LCI now focuses on feedback from direct reports, indirect reports, peers, team members, and bosses in an attempt to replicate an entire 360-degree approach. The resulting items written to measure each leadership behavior therefore focused on the team approach rather than the subordinate viewpoint.

LCI Practice and Commitment Definitions

Before new and meaningful items could be written, the six leadership practices and 12 commitments had to be defined, which would then allow for the systematic assessment of each item's content adequacy. In an effort to ensure that the newly developed items "are measuring their intended content domain" (Schriesheim, Powers, Scandura, Gardiner, & Lankau; 1993, p. 397), a definition for each practice and commitment identified in Patton's (2002) pilot instrument was developed. While defining each practice and commitment, the goal was to capture the essence of Kouzes and Posner's (2002) definitions while including concepts from other leadership experts. Additionally, the practice definitions were written to be compatible with existing leadership theory that has been discussed in the literature. The development of the operational definitions and their constructs is presented in the following discussion, in an order mirroring the new organization determined by Kouzes and Posner. The practices and their commitment definitions are summarized in Table 4.

Model the Way

Based on research by Kouzes and Posner (2002), Yukl (2002), Crocker (2000), Maxwell (1999), Axelrod (1999), and Maister (2001), leaders who *Model the Way* know what they stand for, communicate their values, and act in ways that are consistent with those values. *Modeling* is about character, as much as it is about process and effects. Leaders who demonstrate high standards and act in ways consistent with their stated values set an example for others within their organization. *Model the Way* is reflected in two commitments, '*shares personal values*' and '*sets the example*.'

Table 4

Leadership Practices and Commitments of the Updated LCI

Leadership Practice	Commitment	Definition
Model the Way	1. Set the example	Demonstrate and teach the values and standards espoused by the leader.
	2. Share personal values	Communicate the core beliefs that fundamentally guide the way the leader thinks and acts.
Inspire a Share Vision	3. Create a vision	Convey a vivid image of the organization's future.
	4. Attract others to a common purpose	Show and communicate how aspirations are mutually beneficial to work group members and the organization.
Challenge the Process	5. Seek Innovation	Search for and encourage others to search for opportunities to improve the efficiency and effectiveness of the organization.
	6. Take risks and learn from mistakes	Create opportunities for team members to experiment with new ideas.

(*table continues*)

Table 4 (*continued*)

Leadership Practice	Commitment	Definition
Enable Others to Act	7. Promote cooperation	Encourage the open exchange of information and ideas among work group members and promote good working relationships with outside organizations
	8. Empower	Provide members with the necessary resources, support, and skills to take control of their jobs, and make significant inputs to the organization.
Encourage the Heart	9. Recognize individual accomplishments	Reward individual progress and contributions that meet high standards of performance.
	10. Celebrate team accomplishments	Personally highlight and recognize the work group's attainment of key objectives and goals.

Table 4 (*continued*)

Leadership Practice	Commitment	Definition
Enjoy the workplace	11. Allow humor to reduce stress and boredom	Encourage humor to break tension and create an enjoyable workplace.
	12. Promote fun activities to relax and unwind	Encourage creative and fun activities to increase morale and job satisfaction.

Shares personal values. Personal values are the core beliefs that fundamentally guide the way leaders think and act. Leaders must first learn what makes them tick. Leaders learn about their values by exploring their passions, frustrations, fears, and joys. They learn by contemplating the troubles that keep them awake at night. Knowing oneself is just the first step; leaders must then clarify and communicate their personal values so that others may know what they stand for as well.

Sets the example. Refers to daily behaviors that demonstrate and teach the values and standards espoused by the leader. Leaders who do what they say, spend time on what they believe is important, and lead by example, appear more credible to others.

Inspire a Shared Vision

According to research by Kouzes and Posner (1997), Kotter (2002), Blanchard (2002), Hiam (2002), Yukl (2002), and Maxwell (2001), leaders who *Inspire a Shared Vision* convey a vivid image of the future to develop a general understanding of that vision among organization's members. A clear vision helps guide people towards a common purpose. Shared visions are broad strategic roadmaps that provide a clear image of the organization's mission and ways in which workgroups and individual members can contribute to accomplishing that mission. *Inspire a Shared Vision* is reflected in two commitments, '*create a vision*' and '*attract others to a common purpose.*'

> *Create a vision.* Refers to leadership behaviors that convey a vivid image of the organization's future.

> *Attract others to a common purpose.* Refers to leadership behaviors that show and communicate how aspirations are mutually beneficial to work group members and the organization.

Challenge the Process

Based on research by Kouzes and Posner (2002), Hiam (2002), Maxwell (1999), Yukl (2002), and Blanchard (1998), leaders who *Challenge the Process* encourage their people to search for opportunities to challenge the status quo, take risks, and then learn from their mistakes. Their behaviors go beyond the mundane reactive activities associated with near and long-term demands placed on the manager by others. Actions that *Challenge the Process* move beyond task completion, to continued improvement of the effectiveness and efficiency of the organization. Experimentation requires controlled change and assessment. Leaders must take note of the change, the environment in which

the change was implemented, the results of the change, and how the change could be improved. *Challenge the Process* is reflected in two commitments, '*seek innovation*' and '*take risks and learn from mistakes.*'

> *Seek innovation.* Refers to leadership behaviors that search for and encourage others to search for opportunities to improve the efficiency and the effectiveness of the organization.
>
> *Take risks and learn from mistakes.* Refers to leadership behaviors that create opportunities for team members to experiment with new ideas in order to gain critical knowledge about the best ways to add value to the customer.

Enable Others to Act

Research done by Kouzes and Posner (1997), Yukl (2002), Crocker (2000), Maxwell (1999), Axelrod (1999), and Blanchard (1998) suggest that leaders who *Enable Others to Act* foster teamwork among their organization's members and create an environment of mutual respect and trust. Enabling behaviors provide team members with the necessary support to perform at their best. *Enable Others to Act* is reflected in two commitments, '*promote cooperation*' and '*empower.*'

> *Promote cooperation.* Refers to leadership behaviors that encourage the open exchange of information and ideas among work group members, promote good working relationships with outside organizations, and build teams both within and outside of the organization.
>
> *Empower.* Refers to leadership behaviors that provide members with the necessary resources, support, and skills to take control of their jobs and make significant inputs to the organization.

Encourage the Heart

According to research by Kouzes and Posner (1997), leaders who *Encourage the Heart* recognize individuals for their progress and contributions, provide rewards for exceptional performance, and celebrate the accomplishments of the work group. *Encourage the Heart* is reflected in two commitments, '*recognize individual contributions*' and '*celebrate team accomplishments*.'

Recognize individual contributions. Refers to leadership behaviors that reward individual progress and contributions that meet high standards of performance.

Celebrate team accomplishments. Refers to leadership behaviors that personally highlight and recognize the work group's attainment of key objectives and goals.

Enjoy the Workplace

In addition to the five practices outlined by Kouzes and Posner (1997), ASC leadership identified a sixth leadership practice, called *Enjoying the Workplace.* Creating fun in the workplace is a subject drawing increasing attention in the commercial sector and has been a popular topic in industry periodicals for several years. Consequently, Patton (2002) included this additional practice in his pilot UFI study.

In order to define and measure the sixth practice, Patton (2002) surveyed a sample of ASC and AFSAC employees using critical incident questionnaires. Employees were asked to describe their perceptions of what *Enjoy the Workplace* might encompass and then describe an experience where a leader exhibited behavior that best represented their concept of *Enjoy the Workplace*. Content analysis of the responses by both researchers and the ASC Human Resources leadership team led to a better understanding of the dimensions of *Enjoy the Workplace* (Patton, 2002, p. 45). Furthermore, the results of

Patton's research lends evidence that *Enjoy the Workplace* is a distinct sixth practice with two well-defined commitments which are valued by both supervisors and subordinates. Based on Patton's conclusions, the LCI will include *Enjoy the Workplace*, and also attempt to validate it as a separate leadership construct.

Based on studies and work experiences of Clark (2000), Freiberg and Freiberg (1996), London, Paul, and Christensen (2000), McManus (2000), Overby (2001), Patton (2002), Paulson (2001), and Zbar (1999), leaders who *Enjoy the Workplace* create a playfully productive atmosphere at work by encouraging humor and promoting fun activities. *Enjoy the Workplace* is reflected in two commitments, '*allow humor to reduce stress and boredom*' and '*promote fun activities to relax and unwind.*'

Allow humor to reduce stress and boredom. Refers to leadership behaviors that encourage humor to break tension and create an enjoyable work place.

Promote fun activities to relax and unwind. Refers to leadership behaviors that encourage creative and fun activities to increase morale and job satisfaction.

This study attempted to develop reliable and content valid measures of the 12 leadership commitments. The study also tested the factor structure underlying the 12 measures of the Leadership Commitments Inventory.

Hypothesis 1: The Leadership Commitments Inventory will yield 12 reliable and content valid indices that measure supervisor work behavior from a 360-degree feedback process.

Hypothesis 2: Variation in the measures of two commitments not included in Kouzes and Posner's (2002) five-factor taxonomy will be best explained by a sixth leadership practice called Enjoy the Workplace.

Hypothesis 3: The six-dimension taxonomy from the Leadership Commitments Inventory will better explain leadership behavior than alternative theoretical frameworks with one-, two-, and three-dimension taxonomies.

Chapter Summary

Multiple rater feedback programs are becoming increasingly more popular in today's organizations. As such, certain Air Force organizations are attempting to develop flexible, inexpensive feedback instruments that provide their leaders with useful, specific feedback that can lead to improved leadership development. Based on the research done by Kouzes and Posner (1997), Patton (2002) developed the UFI for the Aeronautical Systems Center to develop better leaders. At their request, the UFI included the five leadership practice from Kouzes and Posner's LPI, and also a sixth practice called *Have Fun*. Based on the results of the pilot UFI study, the LCI was created to more distinctively measure leadership behaviors in the workplace at the commitment level. Over half of the leadership behavior statements were revised or deleted as a result of updating the practices and commitments. In an effort to mirror the updated research by Kouzes and Posner (2002), the practices were re-ordered and *Model the Way* was revised with a new commitment called *shares personal values.*

III. Methodology

Introduction

The following methodology chapter contains four sections. The first section outlines the content analysis of the initial version of Leadership Commitment Inventory items, which will be called version 1 for the remainder of this chapter. It describes the sample, procedure, and results of the first content analysis conducted on the Leadership Commitments Inventory. The second section of this chapter describes the second content analysis questionnaire, which will be called version 2 for the remainder of this chapter. It also includes the sample, procedure, and results of version 2. The third section describes the administration and descriptive statistics of the LCI and details the internal consistency of the commitment scales. The fourth section introduces structural equation modeling and nested confirmatory factor analysis.

Schriesheim, Powers, Scandura, Gardiner, and Lankau (1993) documented the importance of content adequacy when developing new or modified scales. To ensure content adequacy, Schriesheim et al. said, "the sample of items contained on that measure [should] be a representative sample of the content universe of the underlying theoretical construct … all the major subdomains of a construct must be covered" (p. 388). According to Schriesheim et al., demonstrating content adequacy is "necessary for judging a measure as having reasonable construct validity" (p. 386). The first objective was to develop a pool of items that reflected the updated definitions of the leadership commitments and practices, and then test whether those new items adequately represented the leadership constructs as intended.

According to Hinkin (1998), the purpose of content validity (adequacy) is to identify which items should be kept, deleted, or refined. Following Hinkin's recommendations, the new items were assessed with two iterations of content analysis (e.g., version 1 and version 2). The first version had three different samples, with each new sample examining a new set of items that had been refined based on the data from the previous sample. The second version had one larger sample and analyzed the final results from version 1.

LCI Content Analysis, Version 1

A pool of 60 items that were designed to reflect the 12 commitments from Patton's (2002) pilot LCI was analyzed. Based on previously listed leadership experts and the formally defined practices and commitments, 33 of the original 60 items were revised or discarded. The remaining 27 original items accurately reflected the updated definitions. Thirty-nine additional items were then created that were consistent with the new definitions, for an initial pool of 66 items (5-7 items per commitment). All items were designed to be consistent with recommendations by Hinkin (1998): short, simple, consistent perspective, specific to a single construct, socially neutral, and described an observable behavior. Each item was analyzed using the *Flesch-Kincaid* reading statistic to ensure it did not exceed the 8th-grade reading level (Microsoft Office, 2000). After the items had been generated and checked to ensure they adhered to the specifications set forth, version 1 of the content adequacy tests was conducted.

According to Schriesheim et al. (1993), the content reviewers did not have to be expert judges. Rather, they needed to be intelligent enough to read and understand the

instructions, items, definitions, and be able to match the behavior-based statements to the theoretically defined categories.

Version 1 Sample

Three different groups of judges completed version 1 of the content adequacy assessment. The first sample consisted of six individuals, four military officers who attended a graduate program in engineering management, and two recent college graduates. The sample was mostly male ($n = 5$) and the average age was 25 years ($SD = 1.9$). The second sample consisted of six military officers, all male graduate students at a small engineering management school in the Midwest. The average age was 29.8 years ($SD = 4.7$). The third sample consisted of five military officers, all graduate students at a small engineering management school in the Midwest. The majority of the sample was male ($n = 4$) and the average age was 32.4 years ($SD = 4.8$).

Version 1 Procedure

Each of the 66 items was typed and affixed to an index card. The 12 commitment definitions were typed and affixed to an envelope (one commitment per envelope). Each participant received a packet that included the 66 index cards randomly sorted and the 12 commitment definition envelopes. The participants were told to physically sort the index cards and match each leadership behavior with the appropriate leadership commitment. The final instructions encouraged the participants to categorize every item and let the researcher know which items they had trouble classifying. The second and third versions of the sorting exercise were the same procedurally; however the participants received fewer items (62 and 48, respectively) on the successive versions.

Analysis. The items were evaluated by determining what proportion of participants categorized the item as intended. Anderson and Gerbing (1991) proposed an index used to assess the extent to which the items reflected its intended content, called the proportions of substantive agreement (PSA). The PSA was calculated for each participant and each item, which indicated the percentage of participants that correctly assigned an item to its intended construct. Items with PSA values less than .66 were discarded or revised; this standard was less stringent than recommended by Greenberg (1986), who recommended using .75 as the limit. However, because content adequacy was tested a second time, a less stringent standard was applied during version 1.

The second index proposed by Anderson and Gerbing (1991) was called the substantive-validity coefficient (SVC), and it reflected the extent to which participants assigned an item to its intended leadership commitment more than they assigned it to another particular commitment. The substantive-validity coefficient ranged from -1.0 to 1.0, where extreme positive values indicated the item was assigned to its intended commitment more than it was assigned incorrectly. Extremely small and negative values for the coefficient indicated that the item was assigned to a commitment it did not intend to reflect (Anderson & Gerbing, 1991). For this test, the critical value used was .60.

Refinements. As each group completed the sorting exercises, the items were refined based on the results. After the first sorting exercise, 25 items were discarded or revised because they did not measure their intended leadership commitment; that is, their PSA values were less than .66. For instance, the item "Cheers actions that are consistent with achieving unit's goals" was intended to reflect *celebrating team accomplishments*, but only one of the participants was able to correctly classify the item. Participants

instead classified the item as *recognized individual contributions*, and *attract others to a common purpose*. After the 25 items were discarded, 20 new items were written. An example of a new item included "Takes part in celebrating team accomplishments," which was written to measure the leadership commitment *celebrate team accomplishments*. The original 66 items and their status, deleted or retained, are listed in Appendix A.

Based on the results of the next two samples of the sorting exercise, the items were further refined. The second sample included 62 items; 18 were discarded and four were re-written. Of the 18 discarded, six were thrown out because their PSA value was below the critical level. Additionally, the number of items was standardized to four per commitment. Therefore, 12 items were thrown out to standardize the commitments and provide a concise questionnaire which would not lend itself to lengthy items and careless responses, thereby reducing the total number of items from 62 to 48. An example of an item thrown out because it was below the allowable PSA limit included "Sets the tone for a friendly and supportive workplace." The item was written to reflect *allow humor to reduce stress and boredom*, however it was misclassified as *sets the example* and *promote cooperation*. The 62 items and their status, deleted or retained, are listed in Appendix B. Of the 48 items on sample 3, one item was misclassified and re-written to better measure the proper leadership commitment. These items are depicted in Appendix C.

Version 1 Results

First, the PSA for each participant was collated and the percentage of behaviors they correctly classified was computed. Of the 16 participants in version 1, the PSA values for the participants themselves correctly classifying the items ranged from a low

of 72% from sample 1 correctly classifying the items up to a high of 92% in the third

sample. The average PSA and SVC values per item across all three samples of version 1

are displayed in Table 5. Generally, the PSA values are satisfactory, and range from a

low of .60 to a high of 1.0 (a perfect categorization). Of the final 48 items analyzed in

the initial version, over half (25 items) had perfect PSA values. Additionally, the SVC

values were also satisfactory, ranging from a low of .20 to a high of 1.0. The same 25

items with perfect PSA scores also had perfect SVC scores. An unsatisfactory SVC score

of .20 indicated that not only were the participants unable to correctly classify the item as

intended, they in fact misclassified that item to another construct.

As mentioned, one item had an unsatisfactory PSA value of .60 and an

unsatisfactorily SVC value of .20, which was "Challenges team members to experiment

with new ideas." Consequently, a new item was written to replace the old item. The

items with n/a (not applicable) in the PSA and SVC columns are newly written items to

be analyzed in version 2 of the content adequacy tests. In addition to the new item

previously mentioned, five new items were written to replace the old *motivate with small

wins* commitment and with the new *shares personal voice* commitment.

<p style="text-align:center">*LCI Content Analysis, Version 2*</p>

In September 2002, Kouzes and Posner released an updated version of *The

Leadership Challenge*. The most significant change in their updated version was the

revision of the *Model the Way* practice. In accordance with previous research, five new

items were written to measure the new commitment. The items were written to

Table 5

Average PSA and SVC Values for Version 1

Statement	Commitment	PSA	SVC	Note
Challenge				
Devotes time to discuss innovation and change.	C1	0.94	0.88	
Encourages others to seek out better ways of doing things.	C1	0.90	0.82	
Looks for ways that challenge the status quo.	C1	0.62	0.36	
Looks for ways to improve the unit's effectiveness.	C1	0.89	0.76	
Challenges team members to experiment with new ideas.	C2	0.60	0.20	Deleted
Lets team members experiment with new ideas.	C2	n/a	n/a	New
Encourages people to take risks.	C2	0.94	0.88	
Finds ways to turn setbacks into learning events.	C2	0.93	0.88	
Takes measured risks based on the team's capabilities.	C2	1.00	1.00	
Inspire				
Clearly explains a vision of the team's future.	I1	1.00	1.00	
Creates vivid images that help convey our mission.	I1	1.00	1.00	
Portrays the unit as having an important impact on the future.	I1	0.92	0.82	
Appeals to each member's desire to contribute to the success of the organization.	I2	0.83	0.76	
Helps focus team on a common purpose.	I2	0.62	0.45	
Helps team members relate their own aspirations with the unit's mission.	I2	1.00	1.00	
Promotes common causes that can be supported by all members of the work group.	I2	0.82	0.76	

(*table continues*)

Table 5 (*continued*)

Statement	Commitment	PSA	SVC	Note
Enable				
Assigns tasks that require team members to cooperate with each other.	E1	1.00	1.00	
Builds long-term relationships with others outside of the unit.	E1	0.94	0.88	
Encourages the open exchange of information and ideas.	E1	1.00	1.00	
Persuades team to cooperate with others in order to build strong partnerships.	E1	0.93	0.88	
Allows us to decide the best way to get our jobs done.	E2	0.94	0.88	
Gives us important work to do on critical tasks.	E2	0.93	0.88	
Grants us the appropriate authority to do our work.	E2	1.00	1.00	
Makes sure that we have the ability to make good judgments on our own.	E2	1.00	1.00	
Model				
Acts in ways that are consistent with stated values.	M1	0.94	0.88	
Leads by example.	M1	1.00	1.00	
Makes decisions that are consistent with the unit's values.	M1	1.00	1.00	
Sets a personal example of what is expected of unit members.	M1	1.00	1.00	
Communicates what he or she is passionate about.	M2	n/a	n/a	new
Shares personal values with team members.	M2	n/a	n/a	new
Lets us know what causes him or her frustration.	M2	n/a	n/a	new
Clarifies to others what leads him or her to become impatient.	M2	n/a	n/a	new

(*table continues*)

Table 5 (*continued*)

Statement	Commitment	PSA	SVC	Note
Builds team's confidence by showing that small steps can make a big difference.	M2	1.00	1.00	Deleted
Divides large tasks into smaller pieces to gather momentum on a project.	M2	1.00	1.00	Deleted
Reminds us to take things a step at a time.	M2	1.00	1.00	Deleted
Sets short term goals that lead to visible signs of success.	M2	1.00	1.00	Deleted
Encourage				
Publicly rewards individual members when they do a good job.	H1	0.89	0.82	
Rewards only those who meet or exceed challenging standards.	H1	1.00	1.00	
Tailors rewards to things we each individually value.	H1	0.82	0.76	
Takes note of high performers.	H1	1.00	1.00	
Makes sure leaders know about the unit's successes.	H2	0.94	0.88	
Shows appreciation for the team's hard work.	H2	1.00	1.00	
Takes part in celebrating team accomplishments	H2	0.82	0.73	
Takes time out to publicly recognize the unit's endeavors.	H2	1.00	1.00	
Enjoy the Workplace				
Allows humor to break through during tense moments.	W1	1.00	1.00	
Encourages non-offensive humor as a way to make the workplace more fun.	W1	1.00	1.00	
Not afraid to laugh at himself/herself.	W1	0.88	0.82	
Willing to laugh and have fun with others.	W1	0.88	0.82	

(*table continues*)

Table 5 (*continued*)

Statement	Commitment	PSA	SVC	Note
Encourages simple, quick, and fun activities that lift spirits at work.	W2	1.00	1.00	
Finds ways to offset hardships caused by work with some fun outcome or activity.	W2	0.94	0.88	
Takes advantage of lulls in the schedule for relaxing and fun activities.	W2	1.00	1.00	
Willing to take a break during busy periods to do something fun as a unit.	W2	1.00	1.00	

Note: PSA is the proportion of substantive agreement. *SVC* is the substantive-validity coefficient. *C1* is the first challenge commitment and *C2* is the second challenge commitment. *I1* is the first inspire commitment, while *I2* is the second inspire commitment. *E1* is the first enable commitment and *E2* is the second enable commitment. *M1* is the first model commitment, which changed with the updated Kouzes and Posner (2002) *Leadership Challenge*. *M2* is the second model commitment. *H1* is the first encourage commitment and *H2* is the second encourage commitment. *W1* is the first workplace commitment and *W2* is the second workplace commitment.

measure the ability of a leader to communicate his or her values, and to measure the subordinates' understanding of that leader's personal values. In reaction to the new commitment, version 2 of the content adequacy analyzed the five new items and also attempted to validate the rewritten item "Lets team members experiment with new ideas" from version 1.

Version 2 Sample

The sample consisted of 19 students at a small graduate school in the Midwest. As noted earlier, Schrieshiem et. al. (1993) recommended the participants have sufficient

cognitive ability to judge a series of statements with respect to a set of theoretically defined categories. Therefore, the sample of graduate students was deemed sufficient to complete the task.

Version 2 Questionnaire

Unlike the previous categorization task, the participants completed a paper and pencil questionnaire during their content adequacy exercise. The questionnaire included the definitions of the 12 leadership commitments and 49 items that were written to measure each of the 12 commitments (4-5 items per commitment). The 12 leadership commitments and their definitions were labeled **A** through **L**. Of the 49 items on the questionnaire, 43 items had been previously screened by the initial samples and five items were new, designed to reflect Kouzes and Posner's (2002) revised commitment. Additionally, one item was re-written because it had been classified incorrectly in version 1. Therefore, six items were being content analyzed for the first time, while 43 of the items were repeats. A copy of the questionnaire used in this final assessment is presented in Appendix D.

In order to capture the essence of each leadership commitment, shorter modified definitions were used on the first page of the categorization exercise. For instance, the shorter version of *share personal values* read, "refers to leaders communicating the core beliefs that fundamentally guide the way they think and act. Leaders must clarify and communicate their personal values so that others may know what they stand for."

The 49 leadership behaviors were presented in quasi-random order. The instructions asked participants to carefully read each statement and indicate what leadership commitment (A through L) they felt the behavior-based statement described.

They were asked to place the letter (A through L) in the left hand column that corresponded to the ONE leadership behavior that they felt BEST described the statement.

The final instructions requested that participants assign a leadership behavior to each statement, and not to omit any. Participants were also told that eleven of twelve commitments were represented by four behaviors, while one commitment was represented by five behaviors.

Version 2 Procedure

The categorization questionnaire was administered to the students at the end of their normal class period and they were given approximately 30 minutes to complete it. Prior to the administration, the purpose of the research was briefly explained and students were encouraged to ask any questions that would clear up any confusion. During this time, the instructions were read out loud and participants were directed to assign a commitment to every behavior on the questionnaire. As the instructions concluded, the instructor asked that participants not include their name anywhere on the questionnaire.

Analysis. The data from the questionnaire were analyzed using the PSA and SVC methods described earlier. First, the PSA for each participant was collated and the percentage of behaviors they correctly classified was computed. Participants who classified less than 80% (40 out of 49) correctly were discarded. Consequently, the original sample of 19 participants was decreased to 12 when seven participants were thrown out because they were not able to correctly classify enough items. Of the seven participants thrown out, three did not complete the second page of the questionnaire. Of

the remaining 12 participants, their PSA values ranged from .82 to .98 correctly classified items.

The second PSA and SVC computation determined the percentage of total points for each item in each of the 12 leadership commitments, and whether participants categorized items to their intended construct, or to an incorrect commitment. According to Schriesheim et al. (1993) and Greenberg (1986), an acceptable cut-off limit for PSA values is .75. Accordingly, items with PSA and SVC values of .75 and higher were considered adequate and retained. Items were revised if fewer than nine of the participants correctly assigned them to their intended commitment. A more stringent standard of .75 was used for version 2 of the content analysis in order to ensure the content was adequately validated.

Version 2 Results

The intent of this step was three fold. The questionnaire was designed to assess the content adequacy of the five new items written to measure *shares personal values*, ensure the repeat 43 items were still adequate, and to validate the revised item from version 1. As can be seen from Table 6, all five behavior-based statements written to measure the new commitment proved excellent measures of *shares personal values*. Two items had the same PSA and SVC values of .92 and .83 respectively, while three items had perfect PSA and SVC values of 1.0. Therefore, the new items all appeared to measure the new correct commitment. In order to standardize the commitment and ensure it only had four items, the behavior "Lets us know what causes him or her frustration" was discarded.

Table 6

PSA and SVC Values for Version 2

Statement	Commitment	PSA	SVC
Challenge			
Devotes time to discuss innovation and change.	C1	1.00	1.00
Encourages others to seek out better ways of doing things.	C1	0.83	0.67
Looks for ways that challenge the status quo.	C1	0.75	0.50
Looks for ways to improve the unit's effectiveness.	C1	0.83	0.67
Encourages people to take risks.	C2	1.00	1.00
Finds ways to turn setbacks into learning events.	C2	0.75	0.67
Lets team members experiment with new ideas.	C2	0.58	0.33
Takes measured risks based on the team's capabilities.	C2	1.00	1.00
Inspire			
Clearly explains a vision of the team's future.	I1	0.92	0.83
Creates vivid images that help convey our mission.	I1	0.92	0.83
Portrays the unit as having an important impact on the future.	I1	0.83	0.75
Provides a vision that helps the team stay energized, focused, and confident.	I1	1.00	1.00

(*table continues*)

Table 6 (*continued*)

Statement	Commitment	PSA	SVC
Appeals to each member's desire to contribute to the success of the organization.	I2	0.83	0.67
Helps focus team on a common purpose.	I2	0.92	0.83
Helps team members relate their own aspirations with the unit's mission.	I2	0.83	0.75
Promotes common causes that can be supported by all members of the work group.	I2	0.67	0.33
Enable			
Assigns tasks that require team members to cooperate with each other.	E1	0.92	0.83
Builds long-term relationships with others outside of the unit.	E1	0.92	0.83
Encourages the open exchange of information and ideas.	E1	0.83	0.75
Persuades team to cooperate with others in order to build strong partnerships.	E1	0.92	0.83
Allows us to decide the best way to get our jobs done.	E2	1.00	1.00
Gives us important work to do on critical tasks.	E2	0.75	0.67
Grants us the appropriate authority to do our work.	E2	1.00	1.00
Makes sure that we have the ability to make good judgments on our own.	E2	0.75	0.67

(*table continues*)

Table 6 (*continued*)

Statement	Commitment	PSA	SVC
Model			
Acts in ways that are consistent with stated values.	M1	0.83	0.75
Leads by example.	M1	1.00	1.00
Makes decisions that are consistent with the unit's values.	M1	0.50	0.25
Sets a personal example of what is expected of unit members.	M1	0.92	0.83
Communicates what he or she is passionate about.	M2	1.00	1.00
Shares personal values with team members.	M2	0.92	0.83
Lets us know what causes him or her frustration.	M2	0.92	0.83
Clarifies to others what leads him or her to become impatient.	M2	1.00	1.00
Shares with us what keeps him or her awake at night.	M2	1.00	1.00
Encourage			
Publicly rewards individual members when they do a good job.	H1	1.00	1.00
Rewards only those who meet or exceed challenging standards.	H1	1.00	1.00
Tailors rewards to things we each individually value.	H1	0.83	0.75
Takes note of high performers.	H1	1.00	1.00

(*table continues*)

Table 6 (*continued*)

Statement	Commitment	PSA	SVC
Makes sure leaders know about the unit's successes.	H2	0.92	0.83
Shows appreciation for the team's hard work.	H2	1.00	1.00
Takes part in celebrating team accomplishments	H2	1.00	1.00
Takes time out to publicly recognize the unit's endeavors.	H2	1.00	1.00
Enjoy the Workplace			
Allows humor to break through during tense moments.	W1	1.00	1.00
Encourages non-offensive humor as a way to make the workplace more fun.	W1	0.92	0.83
Not afraid to laugh at himself/herself.	W1	1.00	1.00
Willing to laugh and have fun with others.	W1	0.92	0.83
Encourages simple, quick, and fun activities that lift spirits at work.	W2	1.00	1.00
Finds ways to offset hardships caused by work with some fun outcome or activity.	W2	0.92	0.83
Takes advantage of lulls in the schedule for relaxing and fun activities.	W2	1.00	1.00
Willing to take a break during busy periods to do something fun as a unit.	W2	1.00	1.00

Note: PSA is the proportion of substantive agreement. *SVC* is the substantive-validity coefficient.

Of the 43 items that had been assessed in the previous content adequacy test, two did not meet the PSA or SVC standards of .75. In version 1, neither of the items had difficulty, but they were revised based on the low percentage of participants who were able to correctly classify them on version 2, and the number of times they were misclassified as the wrong construct. The item "Makes decisions that are consistent with the unit's values" had a PSA value of .50 and a SVC value of 0.25, and was rewritten to read "Makes decisions that are consistent with his or her stated values." This behavior is intended to measure *sets the example*, one of the *Model the Way* commitments. Therefore, the updated commitments and definitions could have led this item to be poorly classified. The second item poorly classified for the first time was "Promotes common causes that can be supported by all members of the work group." It was written to measure the leadership commitment *attract others to a common purpose*, but it had a low PSA value of only .67 and low SVC value of 0.33 and was confused with *promote cooperation* by four participants. Therefore, it was re-worded to say, "Directs our attention to common goals that can be supported by all team members."

The third intent of the version 2 categorization exercise was to validate the item "Lets team members experiment with new ideas." The behavior was re-written after it was misclassified in version 1, and was intended to measure the commitment *take risks and learn from mistakes*. However, it was not adequately classified as such, and had a low PSA value of .58 and a SVC value of 0.33. Consequently, the item was revised once again, this time to say: "Willing to experiment with new ideas."

Consequently, as a result of version 2 of the content analysis three items emerged with low PSA values. One extra item that measured *shares personal values* was

discarded, and the remaining 45 items were further validated. A complete copy of the final 48 items is presented in Appendix E, in accordance with the new order presented by Kouzes and Posner (2002).

Administration of the Leadership Commitments Inventory

The Leadership Commitments Inventory was tested and then administered over a web-based survey. Thirty seven graduate students and faculty members tested the pilot LCI. Participants were given an overview of the system, its function, specific reporting instructions, and the link to the system via e-mail. All participants (students and faculty) were instructed to identify any bugs in the system and provided feedback on its ease of use. Participants were encouraged to provide any comments about the system to the research team via e-mail or in person if necessary.

After the initial pilot instrument was "debugged," it was administered in full to potential participants. The LCI was originally fielded to students and faculty at the graduate school. The LCI was available through a web based developmental feedback tool designed and implemented by another member of the LCI research team (Douglas, 2003). The tool was also advertised to part-time graduate students at another school in the area. Potential participants were given an overview of the feedback instrument, its function, and the link to the LCI via e-mail. The leaders were then presented with the home page and could navigate the website as desired. The LCI website was self-sufficient and required no researcher actions to operate. A complete description of the capabilities of the tool can be found in Douglas's (2003) research. Researchers retrieved data from the system via the school's contractor. The contractor extracted the data from the database and provided an Excel spreadsheet to the researcher.

The data used in analyzing the LCI included 68 supervisors and 210 observers (e.g., boss, peer, indirect report, direct report, or team member). The leaders completed the LCI survey on themselves and then asked observers to complete an assessment of their leadership behaviors. Of the 68 supervisors, six were faculty at a graduate school in the Midwest, twenty were full time graduate students, forty were part time graduate students, and two were administrators. Fifteen of the full time graduate students and all of the part time graduate students performed their assessments based on their current or most recent full time work positions. Demographic data on the 210 observers was not available because their responses were collected anonymously.

Reliability Estimates of LCI Commitment Scales

The 48 observable leadership items on the leadership commitments inventory were measured on a 7-point Likert scale ranging from 1 = *Almost never*, to 7 = *Almost always*. The scale reliability for each commitment was estimated by calculating the internal consistency of each 4-item scale as indexed by Cronbach's coefficient alpha (α). Table 7 shows the means, standard deviations, reliability, skewness, and kurtosis for the LCI-self commitment scales. The means did not vary much across the 12 commitments, and ranged from a low of 4.66 (*shares personal values*) to a high of 5.98 (*sets the example*). Interestingly, the highest and the lowest means are both from the *Model the Way* practice. According to leader self report data, the two most frequently used leadership behaviors in the work place are *sets the example* ($M = 5.98$, $SD = 0.68$) and *allow humor to reduce stress and boredom* ($M = 5.92$, $SD = 0.93$). The commitment scale alpha coefficients ranged from a low of .27 (*recognize individual contributions*) to a high of .79 (*attract others to a common purpose*). Half of the reliability estimates did not

Table 7

Means, Standard Deviations, and Reliability of the LCI Commitments, Self

Commitment, *n* = 68	*M*	*SD*	*α*	*skew*	*kurt*
Sets the example (M1)	5.98	0.68	0.76	-0.28	-0.58
Shares personal values (M2)	4.66	0.87	0.49	0.30	-0.32
Create a vision (I1)	4.91	0.93	0.72	-0.53	-0.20
Attract others to a common purpose (I2)	5.25	0.83	0.79	-0.52	-0.25
Seek innovation (C1)	5.53	0.67	0.56	-0.21	-0.33
Take risks and learn from mistakes (C2)	5.25	0.70	0.68	0.00	-0.41
Promote cooperation (E1)	5.52	0.68	0.46	-0.77	0.96
Empower (E2)	5.57	0.83	0.73	-0.63	-0.20
Recognize individual contributions (H1)	5.09	0.74	0.27	-0.07	-0.83
Celebrate team accomplishments (H2)	5.57	0.82	0.69	-0.88	1.54
Allow humor to reduce stress and boredom (W1)	5.92	0.93	0.77	-0.79	0.55
Promote fun activities to relax and unwind (W2)	5.03	1.07	0.74	-0.32	-0.36

Note: The scale mean (*M*) has been transformed back to the original metric by dividing by the number of items, (*SD*) is the standard deviation of the transformed scale values, (*skew*) and (*kurt*) are the skewness and kurtosis respectively, and measure the normality of the scale data.

meet the minimum acceptable limit of 0.70 (Nunnally & Bernstein, 1974). Specifically, M2 (*shares personal values*), C1 (*seeks innovation*), C2 (*take risks and learn from*

mistakes), E1 (*promote cooperation*), H1 (*recognize individual contributions*), and H2

(*celebrate team accomplishments*) had low reliability estimates. The low reliability

estimates could be due to the small sample size ($n = 68$), or the fact that supervisors are

more stringent when rating themselves. Therefore, perhaps supervisors do not

discriminate between behaviors at the commitment level when self reporting. The low

alpha coefficients for the scales indicate possible internal consistency problems within

the commitments. The skewness and kurtosis all fell within a small range from -0.88 to

1.54, indicating normality of the scale data (Neter, Kutner, Nachtscheim, & Wasserman,

1996).

The web based leadership development tool provided leaders the ability to invite

others (e.g., boss, peer, direct reports, indirect reports, team members) to complete

observer assessments using the LCI. Observers ($n = 210$) rated the supervisor using the

same 7-point Likert scale for observed leadership behaviors. The combined observer

data for each of the twelve commitments is presented in Table 8. The means ranged from

4.96 (*shares personal values*) to a high of 6.05 (*sets the example*). Similar to the LCI-self

data, the most commonly reported leadership behaviors were *sets the example* ($M = 6.05$,

$SD = 0.97$) and *allow humor to reduce and stress and boredom* ($M = 5.91$, $SD = 0.96$).

The reliability estimates for the commitments were much higher than the self reliability

estimates and were all above 0.70, ranging from a low of 0.71 (*shares personal values*) to

a high of .88 (*sets the example*). Similar to the self data, the two *Model the Way*

commitments represented the largest spread in data. According to Neter, Kutner,

Nachtscheim, and Wasserman (1996), skewness and kurtosis values should fall with + / -

Table 8

Means, Standard Deviations, and Reliability of the LCI Commitments, Observer

Commitment, *n* = 210	*M*	*SD*	*α*	*skew*	*kurt*
Sets the example (M1)	6.05	0.97	0.88	-1.40	2.39
Shares personal values (M2)	4.96	1.11	0.71	-0.24	-0.34
Create a vision (I1)	5.28	1.07	0.85	-0.47	-0.30
Attract others to a common purpose (I2)	5.51	1.00	0.86	-0.81	0.87
Seek innovation (C1)	5.42	0.93	0.82	-0.46	-0.02
Take risks and learn from mistakes (C2)	5.27	1.00	0.83	-0.59	0.39
Promote cooperation (E1)	5.64	0.90	0.74	-0.98	1.67
Empower (E2)	5.59	0.96	0.82	-1.10	1.76
Recognize individual contributions (H1)	4.98	1.09	0.80	-0.60	0.87
Celebrate team accomplishments (H2)	5.54	0.96	0.79	-0.57	0.32
Allow humor to reduce stress and boredom (W1)	5.91	0.96	0.84	-1.43	3.25
Promote fun activities to relax and unwind (W2)	5.21	1.15	0.85	-0.55	-0.13

Note: The scale mean (*M*) has been transformed back to the original metric by dividing by the number of items, (*SD*) is the standard deviation of the transformed scale values, (*skew*) and (*kurt*) are the skewness and kurtosis respectively, and measure the normality of the scale data.

3. Therefore, the kurtosis of W1 (*allow humor to reduce stress and boredo*m; *kurt* = 3.25) offers some question as to the normality for that scale data.

The scale statistics for the combined sample (self and observer) is displayed in Table 9 (*n* = 278). The mean scores exhibit the same characteristics as the previous tables; the commitment means ranged from a low of 4.89 (*shares personal values*) to a

Table 9

Means, Standard Deviations, and Reliability of the LCI Commitments

Scale, *n* = 278	*M*	*SD*	*α*	*skew*	*kurt*
Sets the example (M1)	6.03	0.91	0.86	-1.29	2.38
Shares personal values (M2)	4.89	1.06	0.67	-0.1	-0.37
Create a vision (I1)	5.19	1.04	0.83	-0.42	-0.28
Attract others to a common purpose (I2)	5.45	0.97	0.85	-0.63	0.59
Seek innovation (C1)	5.44	0.87	0.78	-0.47	0.14
Take risks and learn from mistakes (C2)	5.27	0.94	0.81	-0.54	0.53
Promote cooperation (E1)	5.61	0.85	0.70	-0.92	1.68
Empower (E2)	5.58	0.92	0.80	-1.02	1.5
Recognize individual contributions (H1)	5.01	1.01	0.73	-0.6	1.08
Celebrate team accomplishments (H2)	5.55	0.93	0.77	-0.62	0.53
Allow humor to reduce stress and boredom (W1)	5.91	0.96	0.82	-1.28	2.62
Promote fun activities to relax and unwind (W2)	5.17	1.13	0.82	-0.48	-0.22

Note: The scale mean (*M*) has been transformed back to the original metric by dividing by the number of items, (*SD*) is the standard deviation of the transformed scale values, (*skew*) and (*kurt*) are the skewness and kurtosis respectively, and measure the normality of the scale data.

high of 6.03 (*sets the example*). Eleven out of the twelve commitment scale reliabilities exceeded .70. *Shares personal values* (M2) reported a reliability of only .67.

Table 10 shows the zero-order correlation coefficients and covariance coefficients among the twelve leadership commitment scales. The variance / covariance matrix was analyzed using a structural equation modeling method described in the next section.

Confirmatory Factor Analysis

Nested confirmatory factor analysis was performed using the LISREL (Jöreskog and Sörbom, 1993) structural equation modeling program. The nested comparison of the proposed leadership behavior models provides a test of the hypothesis concerning the relationships of the twelve commitment variables to the underlying latent leadership dimensions. The hypothesized six-factor model based on Patton's (2002) UFI was compared to the 5-factor model proposed by Kouzes and Posner's (2002) leadership taxonomy and several plausible alternative models to determine the factor structure that best describes the covariance patterns in the LCI data.

Structural equation modeling methodology analyzes the observed covariance matrix of a set of variables in reference to a hypothesized structure. The analysis produces several fit indices that reflect the hypothesized model's ability to reproduce the original variance and covariance matrices given the constraints of proposed variable-construct relationships. The predicted variances and covariances are compared to the observed data's variances and covariances. Therefore, the fit indices test whether the model fits the sample data. If the hypothesized model represents the observed data, then the path coefficients and parameter estimates of the model are analyzed and the model is considered a viable explanation of the data (Jöreskog and Sörbom, 1993).

Table 10

Variance, Covariance, and Correlation Matrix of the LCI Commitments

	LCI Commitments											
	m1	*m2*	*i1*	*i2*	*c1*	*c2*	*e1*	*e2*	*h1*	*h2*	*w1*	*w2*
m1	**0.82**	0.41	0.62	0.70	0.59	0.60	0.68	0.70	0.47	0.55	0.49	0.39
m2	0.40	**1.12**	0.52	0.45	0.48	0.47	0.51	0.41	0.50	0.37	0.39	0.40
i1	0.58	0.57	**1.09**	0.82	0.71	0.69	0.62	0.66	0.58	0.66	0.48	0.45
i2	0.61	0.47	0.83	**0.94**	0.67	0.65	0.68	0.70	0.57	0.68	0.46	0.46
c1	0.47	0.44	0.65	0.56	**0.76**	0.76	0.56	0.57	0.54	0.55	0.35	0.35
c2	0.51	0.47	0.67	0.59	0.62	**0.87**	0.62	0.62	0.59	0.59	0.44	0.46
e1	0.53	0.46	0.55	0.56	0.42	0.49	**0.72**	0.75	0.50	0.58	0.54	0.48
e2	0.58	0.40	0.64	0.63	0.46	0.54	0.59	**0.85**	0.52	0.53	0.46	0.39
h1	0.43	0.53	0.61	0.56	0.47	0.56	0.43	0.48	**1.02**	0.64	0.36	0.46
h2	0.46	0.36	0.64	0.61	0.44	0.51	0.46	0.45	0.60	**0.85**	0.50	0.57
w1	0.42	0.40	0.48	0.43	0.29	0.39	0.44	0.41	0.34	0.44	**0.91**	0.61
w2	0.40	0.48	0.53	0.51	0.34	0.48	0.46	0.41	0.53	0.60	0.66	**1.28**

Note. N = 278. The covariances are show in the lower half of the triangle, the variances are show on the diagonal, and the correlations are show on the upper half of the triangle.

When determining if the model is consistent with the data, a common fit index is the Chi-square (X^2) index, which measures the discrepancy between the observed and predicted matrices and is directly proportional to the amount of discrepancy. The X^2 is reported with the number of degrees of freedom associated with the model. The degrees of freedom are a function of the number of covariances provided and the number of paths specified: $df = \frac{1}{2}(p + q)(p + q + 1) - t$ where p is the number of observed independent

variables, q is the number of observed dependent variables and t is the number of independent parameters estimated (Jöreskog and Sörbom, 1993). For a confirmatory factor analysis, all the observed variables are considered independent (p).

The properties of the X^2 allow nested models to be directly compared. A more specified model (fewer degrees of freedom) is nested in another less specified model if it contains all paths of the more parsimonious model. For each additional path proposed by the researcher and estimated by the structural equation modeling program, a degree of freedom is lost. In general, for a given model, the more parameters estimated, the more closely the structural equation modeling methodology can reproduce the observed covariance matrix (Jöreskog and Sörbom, 1993). The nested model with fewer degrees of freedom will have a lower X^2 (Jöreskog and Sörbom, 1993). If the reduction in X^2 due to the additional paths is sufficiently large given the loss of degrees of freedom, then the revised model provides a better fit. A statistically reliable reduction in the value of the model X^2 given the loss of the degrees of freedom implies that the alternative model provides a statistically reliable improvement over the comparison model.

The maximum likelihood estimation technique used in the LISREL (Jöreskog and Sörbom, 1993) assumes that the measured variables are continuous and have a multivariate normal distribution. However, LISREL is quite robust when dealing with data that only moderately violates the assumption of normality. The range of skewness and kurtosis found in the combined commitment variable distributions fall well within the LISREL program's level of robustness.

Five different plausible leadership behavior models will be compared in a nested fashion to determine the model with the best relative fit. The five models include the six-

factor model using the six leadership constructs originally proposed by Patton (2002) and tested in the LCI, the five-dimension model proposed by Kouzes and Posner (2002), the three-factor modeled proposed by the University of Michigan, the two-factor model posited by the Ohio State University, and a one-dimension leadership taxonomy. Figure 1 displays the six-factor structure tested to model the LCI. Each of the 12 commitment scales are used to explain the variance in the six latent leadership constructs. In the five-factor model, the *Enjoy* and *Encourage* commitments are combined to measure the *Encourage* practice. In the three factor model, the two *Challenge* commitments and the two *Inspire* commitments represent *task*, the *Enable* and *Model* commitments represent *participative*, and the *Encourage* and *Enjoy* commitments represent *relations*. In the two factor model, the *Challenge* and *Inspire* commitments continue to measure *task*, while all of the remaining commitments measure *relations*. In the last one-factor model structure, all twelve commitments together measure one latent leadership construct.

Chapter Summary

Based on the results of the three samples of version 1 of the content adequacy tests, the final pool of items was reduced from 66 to 48. Version 2 validated the new commitment and also highlighted three bad items that needed to be re-written. The resulting LCI will have four items per leadership commitment. Hinkin (1998) suggested that measures should have four to six items per construct; therefore, four items were deemed sufficient. However, four items represent the minimum sufficient number of items, which could pose a problem, as evidenced by the relatively low internal consistency estimates on several scales when measured as a self assessment.
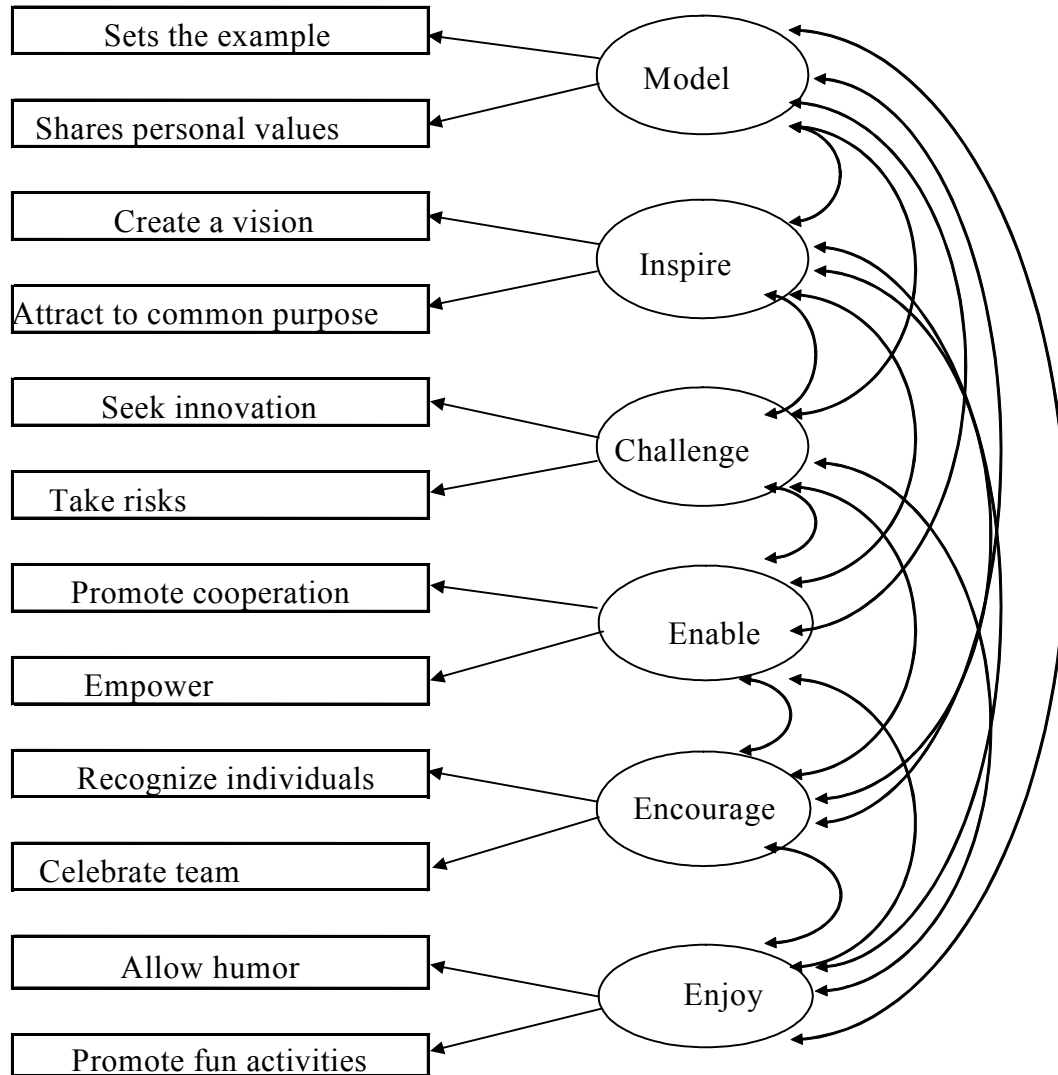
Figure 1. Six-dimension leadership taxonomy.

As a group, the items appear to adequately reflect the commitments they were intended to, as sorted by four independent panels of impartial judges. The improvement across each version for the proportions of substantive agreement and the substantive-validity coefficients provide some confidence in the content validity and quality of the items.

The descriptive statistics of the LCI indicate that eleven of the twelve commitments have adequate internal reliability. The M2 commitment, *shares personal values*, shows potential internal consistency problems.

Follow-on nested confirmatory factor analysis will empirically determine how many practices will emerge through factor analysis. Thus, although the analyses performed and reported in this chapter demonstrate content adequacy, the analyses of the following chapter are equally important for construct validity.

IV. Results and Analysis

Introduction

This chapter contains three sections, all of which address the reliability and validity of the leadership commitments inventory as a feedback instrument. The first section reports the results of the nested confirmatory factor analyses. The LCI's six-factor structure was compared to Kouzes and Posner's (2002) 5-factor leadership taxonomy, and three alternative leadership models in order to determine which model best represented the observed LCI data. The five goodness of fit indices were analyzed, and the standardized residuals and modification indices were investigated in order to assess the fit of each model. Although the hypothesized six-factor model proved to be a better fit than the comparison models, it displayed some serious indications of poor fit. Revised five and six factor structures were then built by moving one measured variable to another construct and then combining constructs. Comparison of fit indices indicated the revised six-factor model provided a reasonable fit to the data. The third section of this chapter analyzes the convergent validity between the LCI-self and LCI-observer data. According to the observed data, supervisors view leadership behaviors very differently from their observers, and the observer data had significant "halo" effect, while the self-data did not. The differences in the self and observer data are discussed.

Results of the Nested Confirmatory Factor Analysis

The nested comparison of the proposed leadership behavior models tested the relationships of the twelve commitment variables to the underlying latent leadership dimensions. The LCI's hypothesized six-factor model (model E) was compared to the 5-factor model (model D) and was also compared to several plausible alternative models

(models A - C) to determine the factor structure that best describes the covariance patterns in the LCI data. Three different classes of fit indices were used to assess the different models of leadership behavior. The three classes (e.g., absolute, parsimonious, and relative) should all be considered when evaluating the fit of a structural equation model. Absolute model fit compares the predicted and observed covariances and variances of the data. The most traditional measure of absolute model fit is the Chi-squared (X^2) fit index. The X^2 index formally tests the null hypothesis that the data has a perfect model fit (Jaccard & Wan, 1996). A X^2 statistic that is statistically significant (e.g., $p < 0.05$) indicates that there are differences between the observed and predicted covariance matrices. If the X^2 statistic is not significant, then the model provides a good fit to the observed data. The second absolute fit index used is the Goodness of Fit Index (GFI), which ranges from 0.0 to 1.0. The higher the number the better the fit, and according to Jaccard and Wan, values above 0.90 indicate a good fit. The third absolute fit index is the standardized root mean square residual (std RMR), which measures the average discrepancy between observed and predicted correlations. The smaller the standardized RMR value, the better the model fit, with values below .05 indicating a good model (Jaccard & Wan, 1996).

The second category of fit indices also looks at absolute fit, but includes a penalty for lack of parsimony. More complex models are penalized as more paths are specified, therefore ensuring the researcher does not arbitrarily add paths to perfectly identify their model (Jaccard & Wan, 1996). The root mean square error of approximation (RMSEA) takes parsimony and absolute fit into account when measuring the model fit. The smaller

the value the better the fit, with values equal to and smaller than 0.08 indicating a good model.

The third category of fit is relative, in that it compares absolute fit to an alternative model. The comparative fit index (CFI) measures relative fit and ranges from 0.00 to 1.00. Larger values indicate a better model fit, and it is recommended that models exceed 0.90 for a good fit (Jaccard & Wan, 1996).

Table 11 presents the goodness of fit indices for the five different models (i.e., models A-E), and the differences in performance of each model. Model E, with the six-factor taxonomy, demonstrates the best model fit across the indices compared to the other models. The six-factor leadership model has an acceptable standard RMR (0.036), an acceptable GFI (0.93), and a high CFI (0.98). However, it does not demonstrate perfect fit and has problems with the X^2 index ($p < .05$), indicating a significant model in which a perfect fit does not exist. It also does not meet the minimum RMSEA value (0.09), which is higher than the recommended 0.08.

When analyzing the differences in the performance of each factor structure, each more complex model is significantly ($p < 0.05$) better than the previous model. Therefore, the six-factor structure has a better model fit ($X^{2 \ (diff)} = 50.2$, $p < 0.05$) compared to the five-factor model and the other more parsimonious models. The standardized and unstandardized path coefficients for the twelve leadership commitment scales measuring the six leadership constructs and their correlations are depicted in Figure 2. The correlations among the six leadership practices range from .59 to 1.04. The correlation between the model and enable practices ($r = 1.04$) is especially disturbing. Correlation coefficients should not exceed 1.0 (a perfect correlation) if the model is

Table 11

Comparison of Nested Confirmatory Factor Analysis Models GOF Indices

Model	df	X^2	RMR	GFI	RMSEA	CFI
A. One-Factor	54	353.4	0.06	0.82	0.14	0.95
B. Two-Factor	53	311.9	0.06	0.84	0.13	0.95
C. Three-Factor	51	242.9	0.05	0.88	0.11	0.97
D. Five-Factor	44	184.7	0.05	0.90	0.11	0.98
E. Six-Factor	39	134.6	0.04	0.93	0.09	0.98

Nested model Comparisons	df	$X^{2(diff)}$	p
1. B to A	1	41.5	0.000
2. C to B	2	69.0	0.000
3. D to C	7	58.2	0.000
4. E to D	5	50.2	0.000

Note. $N = 278$; Standardized Root Mean Square Residual (*std RMR*), Goodness of Fit (*GFI*), Root Mean Square Error of Approximation (*RMSEA*), Comparative Fit Index (*CFI*).

accurate. Additionally, the error component of *shares personal values* (M2) is very large ($e = .70$), which further highlights concerns with that specific commitment and this six-factor model. Consequently, while this model performed better than the more parsimonious models, it does not fit the data accurately.
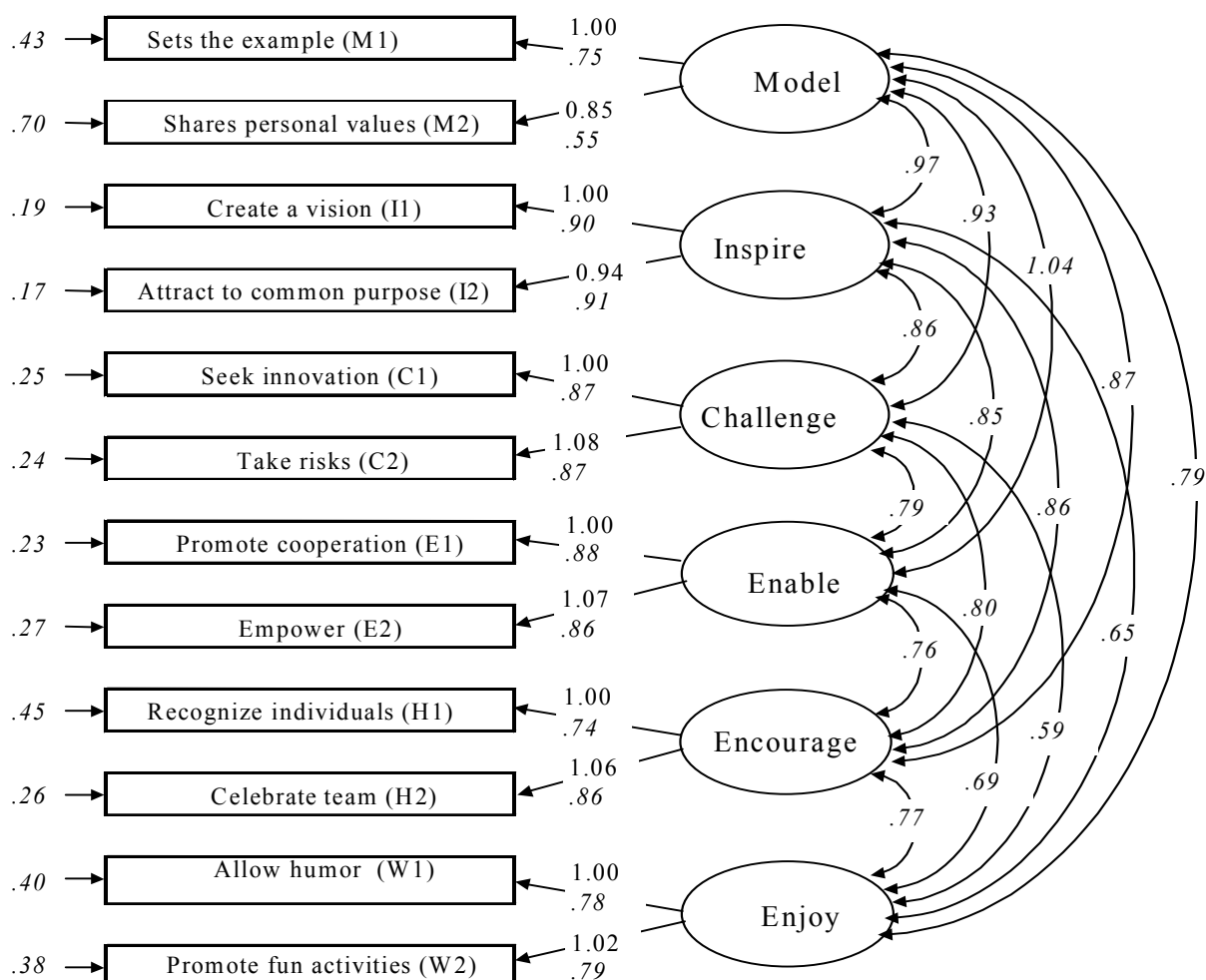
Figure 2. Resulting confirmatory factor structure of 12 leadership commitment scales corresponding to a six-factor leadership behavior taxonomy. All paths are statistically reliable at $p < .05$, except for the path between *Model* and *Enable*. All standardized paths appear in italics, $N = 278$, $X^2 (39) = 134.6$, *CFI* = .98.

Furthermore, an inspection of the fitted and standardized residuals of the six-factor structure revealed numerous areas of ill fit. Of the 66 residuals, 19 residuals were statistically significant (29%), which is far greater than the recommended five percent (Jaccard & Wan, 1996). Therefore, the different fit indices provide conflicting

conclusions about the model indicating that possibly a new model would better fit the data.

Confirmatory Factor Analysis of a Revised Model

Based on the impossible correlation between the *Model* and *Enable* practices and high error components in the six-factor structure, two revised structures were analyzed using the LISREL program. The first concern was the impossible correlation between the *Enable* and *Model* practices. Two observations led to the revised factor structure: M1, *sets the example*, correlated highly with both *Enable* commitments, E1 *promote cooperation* and E2 *empower*, and M2, *shares personal values* did not correlate well with the other *Model* commitment. Consequently, the first new structure maintained the original six practices, but reassigned the first *Model* commitment. The *Enable* practice was redefined with three commitments; it contained the two original *Enable* commitments and the M1 commitment. The *Model* practice contained only one commitment, M2. Creation of this version required an estimate for the error variance of the measured M2 construct. The LISREL program cannot create estimated errors for measured variables that are single indicators of constructs. The error variance for M2 was calculated by multiplying the measures estimated unreliability (1-Chronbach's alpha) and its variance.

The second revised model further assessed the impossibly high correlation between the *Enable* and *Model* practices. The structure had a five-factor taxonomy combining the *Model* and *Enable* practices into one *Model the Way* practice with four commitments (E1, E2, M1, and M2). The five practices tested with this model were: *Model, Inspire, Challenge, Encourage,* and *Enjoy*. The new five-factor structure differed

from Kouzes and Posner's (2002) taxonomy in that it included *Enjoy the Workplace*, but did not discriminate between the *Enable* and *Model* practices.

Table 12 presents the nested comparison and goodness of fit indices for the revised six- and five-factor models, compared to the previously presented six-factor model. The hypothesized and revised five- and six-factor models could not be directly compared because one was not nested within another. Both six-factor models, however, were nested in the revised five-factor model. As can be seen from the Table, both six-factor models demonstrated an improved fit over the revised five-factor model, which illustrates the need for six constructs to explain the observed data. The revised six-factor structure is significantly better than the revised five-factor structure ($X^{2(diff)} = 14.77$, $p <$ .005). Furthermore, the revised six-factor structure performs better across the indices, and is the first taxonomy to meet the acceptable RMSEA criteria (e.g., equal to or less than 0.08), in addition to meeting the goodness of fit indices requirements for the standard RMR, GFI, and CFI.

The revised six-factor model is depicted in Figure 3. The standardized and unstandardized path coefficients for the twelve leadership commitment scales measuring the six leadership constructs and their correlations are all displayed. The correlations among the six leadership practices range from 0.59 to 0.87. All path coefficients are reasonable. The error components are much smaller than the previous six-factor mode, with the highest error component being H1 at 0.45, which is far better than the original six-factor model that had M2 with an error component of 0.70. An inspection of the fitted and standardized residuals of the new six- factor structure still revealed numerous

Table 12

Comparison of Nested Confirmatory Factor Analysis Models GOF Indices, 5- versus 6-factor

Model	df	X^2	Std RMR	GFI	RMSEA	CFI
A. Revised 5-Factor	44	145.9	0.04	0.92	0.09	0.98
B. Revised 6-factor	40	131.1	0.03	0.93	0.08	0.98
C. Old 6-factor	39	134.6	0.04	0.93	0.09	0.98

Nested model Comparisons	df	$X^{2(diff)}$	p			
1. B to A	4	14.8	0.005			
2. C to A	5	11.3	.04			

Note. $N = 278$; Standardized Root Mean Square Residual (*std RMR*), Goodness of Fit (*GFI*), Root Mean Square Error of Approximation (*RMSEA*), Comparative Fit Index (*CFI*).

areas of ill fit. Of the 66 residuals, 15 residuals (23%) were statistically significant, which is far greater than the recommended five percent (Jaccard and Wan, 1996). Consequently, the model is incorrectly over or under estimating the difference in the covariance between the actual data and the predicted model. The worst points of ill fit are overestimated by .11 and underestimated by .20.
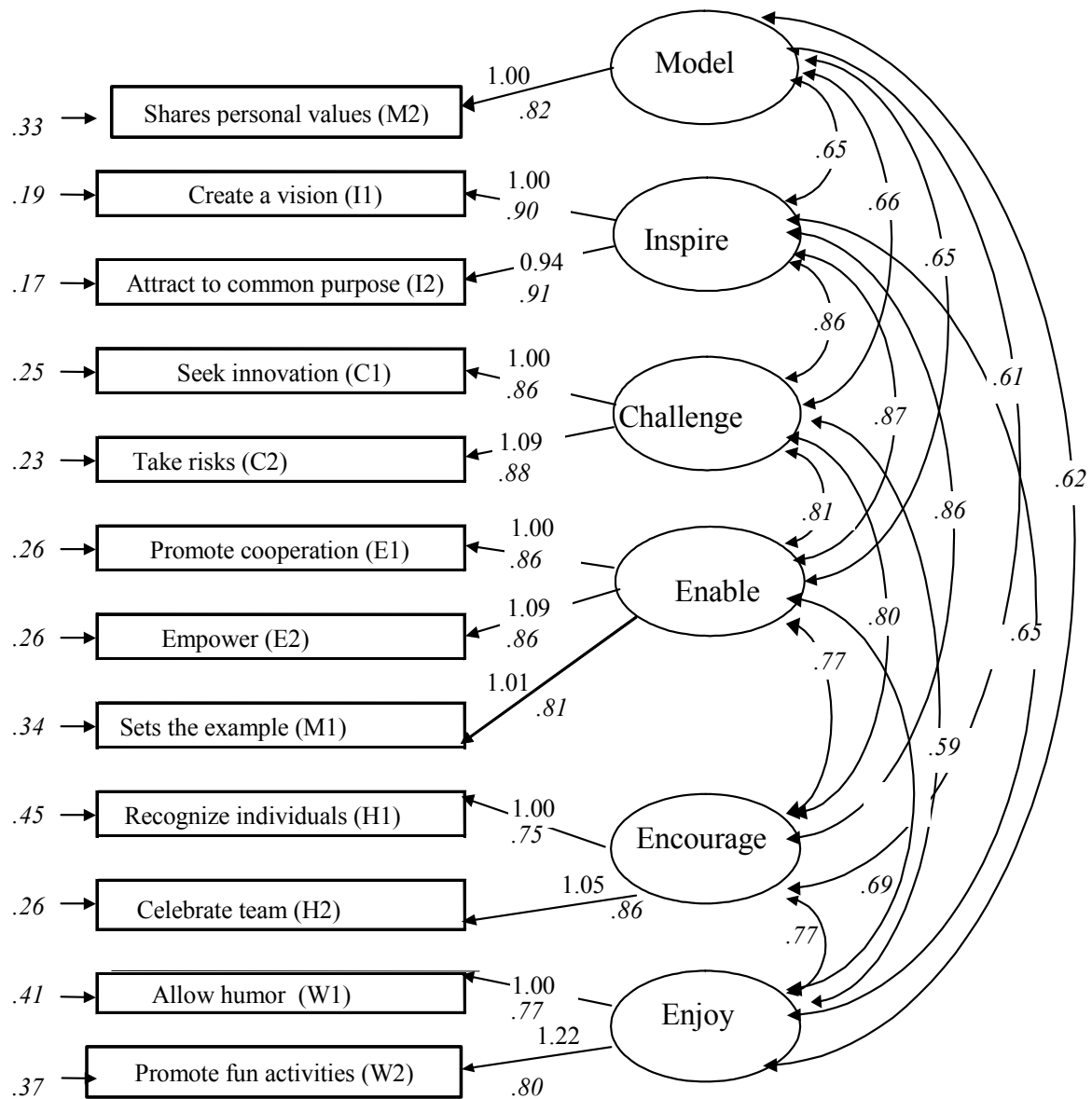
Figure 3. Revised confirmatory factor structure of 12 leadership commitment scales

corresponding to a six-factor leadership behavior taxonomy. All paths are statistically

reliable at $p < .05$. All standardized paths appear in italics, $N = 278$, $X^2 (40) = 131.1$, *CFI*

$= .98$.

Convergent Validity Results

The confirmatory factor analyses demonstrated that there are indeed six distinct

practices measured by the LCI. Four of the practices have two commitments (as

hypothesized), one practice has three commitments, and another has a single

commitment. Therefore, during the ensuing convergent validity discussion, the *Model*

the Way practice is represented with only M2, *shares personal values*. The M1

commitment, *sets the example*, loaded onto the *Enable* practice but the different

definitions do not compare. For instance, the *Enable* definition says: "leaders who enable

others to act foster teamwork among their organization's members and create an

environment of mutual respect and trust," which differs from the M1 commitment which

is defined as: "Demonstrate and teach the values and standards espoused by the leader."

Furthermore, specific items used to measure the M1 leadership behavior do not display

enabling characteristics, such as "Leads by example." Consequently, the M1

commitment, while reliable and valid, does not correlate well with the other practice, and

represents content not included in the *Enable* practice. Therefore, that commitment was

dropped and only eleven commitments are represented in the following anlayses.

The final analysis of the LCI data looked at the correlation coefficients among

and between self and observer data at the practice level. Table 13 shows the correlation

table between all of the practices from the supervisor's LCI-self data ($n = 68$). As can be

seen on the table, the only practices that are not significantly correlated are *Challenge*

and *Enjoy*. Five out of six of the reliabilities are above 0.60. The lack of correlation

between the practices is good, in that it indicates that supervisors are able to discriminate

between those two practices. It also appears that supervisors were able to distinguish

Table 13

LCI-Self Practice Scale Mean Pearson's Correlation Coefficients

	mean	*std dev*	Model	Inspire	Challenge	Enable	Encourage	Enjoy
Model	4.66	0.88	0.49					
Inspire	5.08	0.81	.47**	0.84				
Challenge	5.39	0.62	.48**	.69**	0.77			
Enable	5.54	0.69	.25*	.53**	.41**	0.77		
Encourage	5.33	0.65	.53**	.75**	.55**	.52**	0.64	
Enjoy	5.47	0.90	.22*	.33**	0.05	.34**	.29**	0.83

Note. n = 68, internal consistency (α) appears on the diagonal, * *p* < .05, ** *p* < 01.

between the *Model* and *Encourage*, and *Model* and *Enjoy* practices. Of note is that the

Model the Way practice had a low mean score and lower reliability than the others, which

is primarily due to the reduced number of items (4 rather than 8). The M2 commitment is

a brand new commitment that Kouzes and Posner (2002) presented in their updated *The*

Leadership Challenge.

Table 14 presents the practice correlations for the observer data, with M1 absent.

The "Halo Effect" exists when different leadership principles have very high correlations.

When present, the halo effect indicates that measures do not differentiate between

different principles. As seen in Table 14, every single correlation between every practice

is significant and above .48, which is indicative of the halo effect. Interestingly, the

Challenge and *Enjoy* practices that were not correlated on the self data are positively and

significantly correlated on the observer data. The six leadership constructs overlap,

Table 14

LCI-Observer Practice Scale Mean Pearson's Correlation Coefficients

	mean	*std dev*	Model	Inspire	Challenge	Enable	Encourage	Enjoy
Model	4.96	1.10	0.71					
Inspire	5.40	1.00	.51**	0.92				
Challenge	5.35	0.91	.52**	.79**	0.90			
Enable	5.61	0.87	.54**	.76**	.73**	0.87		
Encourage	5.26	0.94	.48**	.73**	.69**	.65**	0.87	
Enjoy	5.56	0.95	.49**	.59**	.57**	.61**	.65**	0.88

Note. n = 210, internal consistency (α) appears on the diagonal, ** *p* < 01.

and indicate a stronger halo effect in the observer data than in the supervisor data.

Table 15 compares the self data at the practice level matched against the observer data at the practice level (*n* = 36). The observer scores for each supervisor were averaged to form one composite observer score. Then, that observer score was matched with the appropriate supervisor by account identification, and the supervisor's practice scores were correlated against their averaged observer scores at the practice level. The only practice that significantly correlated between self and observer was the *Enable* practice (*p* < .05). All other corresponding practice correlations were low and not statistically significant.

The low correlations between practices indicate that there are gaps between what supervisors think they do and what their observers see them do. The means and standard deviations for the supervisor and observer pairs are presented in Table 16. The simple

Table 15

Matched LCI-Observer and LCI-Self Practice Scale Mean Pearson's Correlation

Coefficients

		Model	Inspire	Challenge	Enable	Encourage	Enjoy
				LCI Practices, Self			
LCI Practices, Observer	Model	0.13	0.19	0.19	0.06	0.25	0.13
	Inspire	-0.10	0.24	0.09	.35*	0.12	0.06
	Challenge	0.13	0.27	0.14	0.28	0.18	0.09
	Enable	-0.01	0.27	0.04	.33*	0.13	-0.01
	Encourage	-0.09	0.20	0.05	.31*	0.21	0.00
	Enjoy	0.17	.45**	0.21	.29*	.32*	0.12

Note. $n = 36$, * $p < .05$, ** $p < .01$.

differences appear small. These differences, however, do not reflect that some leaders

have positive discrepancies while others have negative discrepancies. Therefore, the

absolute differences better reflect the extent to which supervisors and observers do not

agree on the observed leadership behaviors within each practice. As can be seen from the

Table, the *Model* practice has the largest degree of difference between observers and

supervisors.

This also explains the low correlations exhibited between corresponding practices.

The correlations consider squared discrepancies for each supervisor, which does account

for the magnitude of the differences across each supervisor; large differences for each

supervisor account for the low correlations. These differences also indicate that there are

Table 16

Matched LCI-Observer and LCI-Self Practice Scale Means and Standard Deviations

	Supervisor		*Observers*		Difference	
	mean	*std dev*	*mean*	*std dev*	*simple*	*absolute*
Model	4.56	0.94	4.98	0.71	-0.42	0.96
Inspire	4.92	0.85	5.24	0.76	-0.32	0.74
Challenge	5.38	0.57	5.24	0.63	0.14	0.54
Enable	5.55	0.71	5.51	0.60	0.04	0.63
Encourage	5.14	0.61	5.15	0.81	-0.01	0.68
Enjoy	5.36	0.95	5.59	0.56	-0.23	0.79

Note. n = 36.

areas for improvement for the supervisor to focus on. As presented in the literature review, differences between supervisors self scores and their observer scores indicate areas for improvement. Therefore, this tool is highlighting those areas of concern where the supervisor's opinion of their own leadership behaviors is significantly different from their observers' view of their leadership behaviors. Better leaders are developed when supervisors are given feedback on the discrepancies between the self and observer data and they are then given the opportunity to change their behavior.

Correlations did exist between non-corresponding practices. Three of the six practices (i.e., *Inspire, Enable,* and *Encourage*) correlated with other practices. The most significant practice correlation was between self *Inspire* and observer *Enjoy* ($p < .01$). Apparently, supervisor's leadership behaviors that are intended to inspire their observers

are actually seen and interpreted as *Enjoying the Workplace*, or observers think that their

supervisors enjoy their work. Self *Encourage* and observer *Enjoy* also correlated with

each other ($p < .05$). However, observer *Encourage* showed zero relationship with self

Enjoy. Apparently, when supervisors think they are encouraging their observers, their

observers see these actions more as enjoying the workplace. On the other hand,

observers saw *Encouraging the Heart* and *Enjoying the Workplace* as completely

unrelated and non-overlapping leadership constructs.

Similarly, the self *Enable* practice correlated significantly ($p < .05$) with the

observer *Encourage, Inspire,* and *Enjoy* practices. As evident in the specific items used to

measure *Enable* (i.e. "Assigns tasks that require team members to cooperate with each

other" and "Gives us important work to do on critical tasks"), the leadership practice

lends itself to providing immediate and accurate feedback to the leader. Therefore, the

supervisor has a good idea of how well and how frequently they display those leadership

behaviors, and are able to accurately self report.

On the other hand, the self *Model*, self *Challenge*, and self *Enjoy* practices

correlated with no other practices. These practices do not avail themselves for immediate

feedback and therefore the supervisor is often not as accurate about their behavior in

these areas. As can be seen from the specific questions asked for these practices (i.e

"Shares personal values with team members" and "Willing to laugh and have fun with

others") the supervisor may not be given feedback in how they are doing in these areas,

and therefore they may have misconceptions about these three leadership practices.

Chapter Summary

This chapter attempted to show reliability and validity for the Leadership Commitments Inventory. Based on the nested confirmatory factor analysis for the five competing models, the six-factor structure performed the best across the five goodness of fit indices. However, the six-factor model had severe indications of ill fit, especially with the *Enable* and *Model* leadership practices and their associated commitments. Revised five- and six-factor models were developed to correct this problem. Further nested analyses between the revised and hypothesized models demonstrated that six practices were indeed required to fit the observed data. The revised six-factor structure showed that eleven of the twelve measured commitments could be explained by their hypothesized leadership practice. The *Enable* practice explained the variation in its hypothesized commitments as well as one of the commitments that had been hypothesized as part of the *Model* practice. Practice scales from both the self and observer perspectives were then created to analyze the convergent validity of these constructs.

The confirmatory factor analyses provide fairly strong evidence for the hypothesized structure (11 of 12 measured commitments correctly loaded on their hypothesized practice). Further, the resulting practice scales were fairly reliable for both self and observers. Self reports show less halo than observer reports, which is not surprising. This may be due to the lower reliability for self scales and possibly the supervisors are distinguishing more between specific types of behaviors.

The LCI-observer data showed moderately high correlations, and suffered from the halo affect. When comparing the LCI-self and LCI-observer practices very few

correlations were significant, which indicates there is a gap between what supervisors think they do and what observers see them do. Therefore, the two scales were not linearly related. The lack of correlations among the practices may indicate a disconnect between supervisors and their observers.

IV. Discussion

Overview

The intent of this research was to refine the upward feedback instrument (UFI) developed by Patton (2002) and develop a new reliable and valid leadership commitments inventory (LCI). The LCI differed from the UFI in three fundamental ways. First, the LCI modified over half of the original UFI items in order to adhere to the updated definitions of the practices and commitments. Second, the LCI reflected the updates that Kouzes and Posner (2002) made to *The Leadership Challenge*, which specifically included the new commitment *Model the Way* commitment, *find your voice*. Third, the LCI reflects a 360-degree feedback approach, opposed to the UFI's upward feedback only.

This chapter discusses the results of the confirmatory factor analyses and convergent validity performed in Chapter IV that tested the reliability and validity of the LCI. These results are discussed in reference to the three hypotheses posited in Chapter II, and also in light of the three reasons why the LCI was developed. Based on the results, conclusions regarding this research are drawn. Additionally, this chapter discusses the limitations of the research as well as theoretical and practical implications of the research results. The final section of this chapter suggests further research focusing on the Leadership Commitments Inventory and its web-based platform.

Hypothesis 1

Hypothesis 1 addressed the three reasons for the new LCI. Patton's (2002) UFI suffered from areas of ill fit, and he wrote "an evaluation and edit of scale content would reduce some areas of ill fit" (p. 84). Therefore, the practices and commitment definitions

were refined, and over half of the items were re-written to reflect the updated meanings. These new items were then content analyzed to ensure they met the stringent requirements for a well-worded leadership instrument. All of the 48 items performed well on the content adequacy tests, indicating that the newly refined LCI should lead to more stable commitment and practice scales.

Hypothesis 1 also posited that the LCI had an underlying factor structure comprised of 12 different indices measured with a 360-degree feedback approach. To test Hypothesis 1, the LCI was administered to 68 supervisors and 210 observers. The LCI practice and commitment scales proved to be highly reliable internally. Additionally, the confirmatory factor analyses demonstrated that the best model fit was indeed with six distinct leadership constructs. However, those six leadership constructs were actually composed of eleven indices, not the hypothesized twelve. The *Inspire, Challenge, Enable, Encourage,* and *Enjoy* practices all had two commitments as expected, but the *Model* practice only performed well with one commitment, M2 (*shares personal values*). The M1 commitment, *sets the example,* demonstrated very high internal reliability and content validity, but it did not measure its intended practice and was therefore eliminated.

The hypothesized six-factor model clearly provided the best explanation of the LCI's underlying factor structure. To a certain extent, this demonstrates the instrument's validity as the items were written and tested to measure the underlying latent structure of six factors with twelve corresponding commitment scales. However, as mentioned earlier, only eleven commitments were retained and the model did not have a perfect fit. The chi-squared fit index was significant (indicating a poor model fit), and there were

numerous areas of ill fit. Additionally, all six leadership constructs were highly correlated. Patton found 21 of 60 (31%) points of ill fit, and the revised model was only appreciably better, with 15 of 66 points of ill fit (23%).

Analyzing the cause of the six-factor model's ill fit is complex, with no definite answers. LCI participants were possibly unable to distinguish between the different commitments, which could be due to poorly worded items, a poorly selected sample, incorrect underlying constructs, or poorly defined constructs and commitments. Possibly, the six different leadership constructs do not actually exist, which is evident in the significant amount of cross loading between the different practices. However, the nested confirmatory factor analyses seem to indicate that there are six somewhat distinct leadership practices.

The second reason for refining the LCI was to better reflect Kouzes and Posner's (2002) updated *Leadership Challenge,* and especially their new *Model* commitment. Throughout the analyses of the LCI data, the *Model* commitments represented opposite ends of the spectrum. The M1 commitment, *sets the example,* continually had the highest mean scores and the highest reliabilities, while the M2 commitment, *shares personal values,* had the lowest mean scores and the lowest internal reliability. They never correlated with each other, and at no point did they display any relationship to each other. The M1 commitment most reflects the United States Air Force leadership instruction, which explains its high mean scores, internal reliability, and excellent content validity. Possibly, the LCI participants value setting and leading by example to such an extent, they think that it should be exhibited in all of the leadership practices. On the other hand, the M2 commitment is newly created and possibly hard to understand, which would

explain its low mean scores and low internal reliability. M2 did have good content validity, which indicates that participants could classify the items with their intended leadership behavior. This research supports the creation of the new M2 commitment because it is distinct from the other commitments.

The third LCI refinement was intended to reflect a 360-degree feedback approach opposed to the upward feedback used in the UFI. As mentioned in the literature review, a multiple rater approach would offer a broader perspective and therefore more accurate feedback that would most likely be accepted by the supervisor. In an effort to collect data from a 360-degree perspective, the 48 items measuring leadership behavior were written from a team member viewpoint, and the 210 observers were either a boss, direct report, indirect report, team member, or peer. Therefore, the observer data was not limited to only direct report feedback. This broader outlook provided a more complete feedback framework.

In general, the confirmatory factor analyses, descriptive statistics, and convergent validity tests offer partial support for Hypothesis 1. Six distinct leadership practices emerged as the underlying latent structure of the LCI data, but those practices were measured by only eleven (vice twelve) leadership indices. Furthermore, the six-factor model was plagued with numerous areas of ill fit and moderate correlations.

Hypothesis 2

Hypothesis 2 proposed that the LCI's six-factor taxonomy would better explain leadership behavior compared to Kouzes and Posner's (2002) five-factor structure. Nested confirmatory factor analysis was used to test whether the sixth practice, *Enjoy the Workplace*, is a distinct leadership practice that can be measured by its two commitments:

allow humor to reduce stress and boredom and *promote fun activities.* Based on the goodness of fit indices, the original six-factor model was significantly better than Kouzes and Posner's five-factor model. The more complex model outperformed the five-factor model on the standard RMR index, GFI, and RMSEA, while it was the same on the CFI. While neither model met the acceptable criteria for RMSEA, the six-factor structure performed better. When comparing the two nested models, both models had significant Chi-squared indices, indicating a poor model fit, but the six-factor taxonomy had a significantly improved Chi-squared value based on the number of degrees of freedom.

However, the original six-factor model had critical areas of ill fit and an impossible correlation between the *Enable* and *Model* practice, which led to the creation of a revised five- and six-factor model. The revised six-factor model (i.e., *Enable* had three commitments and *Model* had one commitment) was significantly better than the revised five-factor model, also lending support to Hypothesis 2. Additionally, the revised six-factor model had an acceptable fit for four out of the five goodness of fit indices, and had fewer areas of ill fit. Of particular note is that the revised six-factor structure was the only model to have an acceptable RMSEA, which tests for parsimony.

The success of the six-model suggests that *Enjoy the Workplace* may be a unique leadership behavior construct. Therefore, this research supports Hypothesis 2, and the six-factor taxonomy better explains the leadership behaviors measured in the LCI.

Hypothesis 3

Hypothesis 3 suggested that the six-dimension taxonomy would perform better than alternative one-, two-, and three-factor leadership models. Similar to Hypothesis 2, this hypothesis was tested using nested confirmatory factor analysis. The one-factor

model insinuated that only one underlying leadership construct would emerge, while the two-factor suggested that *task* and *relations* practices would become apparent, and the three-factor taxonomy posited that *task*, *participative*, and *relations* leadership practices would emerge. However, the more complex six-factor model was significantly better than all of the lesser complex models.

The one- and two-factor models only met the CFI index requirements (CFI = .95), while the three-factor model met the standard RMR (.05) and the CFI (.97) requirements. The six-factor model met the standard RMR (.04), the GFI (.93), and the CFI (.98) requirements for a good model. The revised six-factor model met the standard RMR (.04), the GFI (.93), RMSEA (.08) and the CFI (.98) requirements for a good model.

Based on the results of the nested confirmatory factor analysis, supervisors appear to value more specific feedback offered by the six-factor model because they are able to better accept that feedback as accurate and improve their leadership behaviors. Furthermore, leadership is apparently more complicated than researchers at the Ohio State University and the University of Michigan originally thought, because both the observers and supervisors were able to distinguish between six distinct leadership constructs.

Therefore, the results of this research offer good support for Hypothesis 3, and the six-factor taxonomy outperformed the more parsimonious alternative theoretical frameworks.

Limitations

There were several limitations to this research. First, the web-based system used to administer the LCI and collect feedback was extremely complex to develop and field.

Therefore, it took over 6 months to develop, test, and implement the LCI web based

system, as reported in Douglas' (2003) research. The eventual successful implementation

of the LCI web site occurred in January 2003, which did not allow much time to actually

collect data. Therefore the sample size is smaller than expected ($N = 278$), and reflects

only 3 weeks worth of data collection. According to Jaccard and Wan (1996), an

adequate sample size should be a 10:1 ratio between items and participants, which would

indicate that the supervisor sample size should have been closer to 480, not including

their accompanying observers. While the data does represent a cross-section of mid-level

managers, it does not capture data from senior managers. Therefore, the high-level

manager perspective is not represented in this LCI data. Additionally, the timing of the

web-based system did not allow for a test-retest analysis. Consequently, there is no

indication if the practice and commitment scales are reliable over time.

A second limitation is also related to the small sample size. In an effort to collect

data in a timely manner, not all of the supervisors were able to get feedback from

observers. As a result, the matched LCI-self and LCI-observer data sample was 36,

indicating that only half of the supervisors were able to ask for and receive feedback from

observers. Ideally, all of the supervisors would have had an opportunity to receive

feedback from each of the various observer categories (i.e., boss, indirect report, direct

report, team member, and peer), and then be given feedback on their averaged responses

in each category. Additionally, if there had been a larger more diverse sample size, then

concurrent validity could have been analyzed by comparing the responses of the different

observer categories for each supervisor to determine if the observer ratings correlated to

each other for each specific supervisor. However, there were not enough supervisors

who had sufficient observer data to adequately conduct the concurrent validity analysis (sample size would have been five).

Another limitation with the fielding of the LCI web-based system is that several functions were not yet implemented. Specifically, the stakeholder survey, automatic feedback reports, and action plans functions were not working when the data was collected. Therefore, the entire web-based system could not be used, nor could the benefits of using the system (i.e. feedback reports automatically generated for the supervisor) be realized. Additionally, stakeholder data, such as job satisfaction, could not be used for further concurrent validity analysis. Ideally, observer ratings should be correlated with their level of job satisfaction.

The LCI participants noted the last major limitation of this research. Many of the observers were not comfortable with the 7-point Likert scale that asked them to rate the frequency with which they noted their leader exhibiting the leadership behaviors. Many thought that an extra category labeled "not observed" or "not appropriate" should have been available. However, in an effort to collect useful, informative data, those categories were not included for fear they would be used whenever an observer was reluctant to honestly answer a survey item. As mentioned in the literature review, observers often have a hard time candidly assessing supervisors if they think the assessment is negative. However, the LCI data will only be used for developmental purposes, and answering on the low end of the Likert scale (i.e. "almost never") is not a negative response. Providing a "not observed" category is no different from the current low end of the frequency scale that participants were asked to use. However, this may have affected the responses of the LCI participants and provided less variability in the

answers because the observers and supervisors did not want to record a possibly "negative" response.

<center>*Theoretical Implications*</center>

The success of the revised six-factor modeled suggests that six distinct leadership constructs exist. The sixth construct, *Enjoy the Workplace*, is a relatively new practice that warrants further study in the field of leadership development. Hopefully this research will add to the theoretical study of *Enjoy the Workplace* and lead to further investigation into the sixth latent leadership construct.

Furthermore, this study may be the first to use nested confirmatory factor analysis to compare different leadership models to Kouzes and Posner's (2002) updated *Leadership Challenge.* Only five original models were analyzed in this research, but there are many more competing leadership theories for leaders to use. Leaders must balance the ease of simple models against the complexity of more sophisticated models. This research has provided support for the six-factor leadership taxonomy that does not sacrifice sophistication for simplicity.

Lastly, this research highlighted some theoretical concerns about the *Model the Way* practice and its two commitments. Further research needs to study the new M2 commitment, determine the proper relationship between M1 and M2, and how they both fit into the six-factor structure. This research has shown that the M2 commitment acts independently of the M1 commitment. Additionally, the M1 commitment may be so important to supervisors and their observers that it cannot be classified into one specific practice.

Practical Implications

The Leadership Commitments Inventory is now a development tool available to individuals and organizations.   It is a free, accessible alternative to expensive copyrighted leadership developmental instruments such as Kouzes and Posner's (2002) Leadership Practices Inventory. The LCI measures leadership behavior from an all-encompassing 360-degree perspective so that supervisors can gain a more accurate assessment of their leadership behaviors. Furthermore, it measures leadership behaviors at the more specific commitment level and therefore provides more precise feedback for supervisors to act on. Practically, the more feedback the supervisor receives, the more accurate it will be perceived, and the more likely the supervisors will be to act on that feedback. Therefore, the LCI is a tool to highlight areas for a supervisor to improve.

Additionally, the LCI is now a web-based system. Therefore, it is easy to access for unlimited use. As long as organizations or individuals have access to the Internet, they can use the LCI system and get feedback from others. The system was designed to provide an effective and inexpensive platform for organizations to conduct developmental feedback programs. It could also provide a template for organizations that desire to design and build their own programs.

Lastly, the web-based system will eventually be able to automatically generate feedback reports and recommend action plans. This will enhance the LCI because not only will feedback be given, but also an automatic report will provide timely and accurate results to the supervisor. Previously, supervisors had to wait for the reports to be manually generated, which could cause delays in the supervisor receiving their feedback. The new web-based LCI system will facilitate timely reports and suggested action plans

for the supervisor to follow based on discrepancies between the supervisor and observer's scores.

Areas For Future Research

There are several areas for future research. First and foremost, the LCI is in its pilot testing stages and needs to be administered to a larger, more diverse group of managers to further test its reliability and validity. While the current research was able to reach some conclusions about reliability and validity, lower level, mid-level, and especially senior level supervisors and their observers should complete the LCI. By administering the developmental tool to a more diverse group of participants, external validity could be evaluated. Furthermore, test-retest statistics could be analyzed along with concurrent validity.

The *Model the Way* practice seems to have some classification problems that should be studied further with a larger sample. The commitments may be misclassified, poorly defined, poorly measured (i.e. bad items), or may not exist.

The full LCI web-based system should be implemented, tested, and then administered in full. Currently, the feedback reports section and stakeholder survey are working, and soon the action plans should be up and running. Future research should integrate all of the functions on the LCI web-based system developed by Douglas (2003) and test its usability and utility for leadership development. Ideally, supervisors should get feedback, an automatic feedback report should be generated, and action plans should highlight the greatest areas for the supervisor to improve. Additionally, the stakeholder survey should provide further information on the relationship between leadership development and job satisfaction. It would be beneficial to determine if these

relationships exist and if the integrated LCI web-based system is an effective platform for

leadership development.

References

Adsit, D.J., London M., Crom, S., & Jones, D. (1997). Cross-cultural differences in upward ratings in a multinational company. *The International Journal of Human Resource Management, 8,* 385-401.

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

Anderson, J.C., & Gerbing, D.W. (1991). Predicting the performance measures in a confirmatory factor analysis with a pre-test assessment of their substantive validities. *Journal of Applied Psychology, 76,* 732-739.

Atwater, L., & Roush, P. (1995). The influence of upward feedback on self – and follower ratings of leadership. *Personnel Psychology, 48,* 35-60.

Axelrod, A. (1999). *Patton on leadership: Strategic Lessons for Corporate Warfare.* Paramus NJ: Prentice Hall.

Baron, R.A. (1996). 'LaVie En Rose' revisited. *Management Communication Quarterly, 9,* 338- 349.

Bennis, W., & Goldsmith, J. (1997). *Learning to Lead.* Perseus Publishing: New York.

Bettenhausen, K.L., & Fedor, D.B. (1997). Peer and upward appraisals. *Group and Organizational Management, 22,* 236-264.

Blanchard, K. (1998). *Gung ho!* New York: William and Morrow Co.

Blanchard, K. (2002). *Zap the Gaps.* New York: Harper Collins Publisher.

Bowles, A., & Bowles, N.B. (2000). A Comparative Study of Transformational Leadership in Nursing Development Units and Conventional Clinical Settings. *Journal of Nursing Management, 8*, 69-76.

Bracken, D.W. (1994). Straight talk about multirater feedback. *Training & Development, 48,* 44-51.

Brett, J.F. & Atwater, L.E. (2001). 360 degree feedback: Accuracy, reactions, and perceptions of usefulness. *Journal of Applied Psychology, 86,* 930-942.

Brutus, S., London M., & Martineau, J. (1999). The impact of 360-degree feedback on planning for career development. *Journal of Management Development, 18,* 676-693.

Carless, S.A. (2001). Short Research Note: Assessing the Discriminant Validity of the Leadership Practices Inventory. *Journal of Occupational and Organizational Psychology, 74*, 233-239.

Clark, R.D. (2000). Giggle while you work: Make it a point to have fun at the office. *Black Enterprise, 31*, 214-216.

Cockfield, R. (1996). Upward feedback. *Teaching in Higher Education, 1,* 387-391.

Crocker, H.W. (2000). *Robert E. Lee on leadership: Executive lessons in character, courage, and vision.* Roseville, CA: Prima.

Fields, D.L., & Herold, D.M. (1997). Using the Leadership Practices Inventory to Measure Transformational and Transactional Leadership. *Educational and Psychological Measurement, 57,* 569-579.

Fletcher, C., Baldry C., & Cunningham-Snell, N. (1998). The psychometric properties   of 360-degree feedback: An empirical study and a cautionary tale. *International Journal of Selection and Assessment, 6,* 19-34.

Freiberg, K.L., & Freiberg, J.A. (1996). *Nuts! Southwest airlines' crazy recipe for business and personal success.* Austin TX: Bard Press, Inc.

Greenberg, J. (1986). Determinants of perceived fairness of performance evaluations. *Journal of Applied Psychology, 71,* 340-342.

Hair, J.F. (1995). *Multivariate Data Analysis.* NJ: Prentice-Hall.

Hiam, A. (2002). *Making Horses Drink.* Canada: Entrepreneur Press.

Hinkin, T.R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1,* 104-121.

Hoojiberg, R., & Choi, J. (2000). Which leadership roles matter to whom? An examination of rater effects on perceptions of effectiveness. *Leadership Quarterly, 11,* 341-364.

Huck, S. W., & Cormier, W. H. (1996). *Chapter 4. Reliability and validity. Reading statistics and research* (2nd ed., pp. 75-88). New York: Harper Collins.

Hughes, R.L., Ginnett, R.C. & Curphy, G.J. (1999). *Leadership.* Boston: Irwin McGraw-Hill.

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

Jaccard, J., & Wan, C. K. (1996). *LISREL approaches to interaction effects in multiple regression.* Newbury Park California: Sage Publications, Inc.

Jöreskog, K., and Sörbom, D. (1993). *LISREL 8.* Chicago: Scientific Software International.

Kotter, J.P. (2002). *The Heart of Change.* Boston: Harvard Business School Press.

Kouzes, J. M., & Posner, B. Z. (1997). *The leadership challenge: How to keep getting extraordinary things done in organizations (2nd ed.).* San Francisco CA: Jossey-Bass.

Kouzes, J. M., & Posner, B. Z. (2002). *The leadership challenge: How to keep getting extraordinary things done in organizations* (3rd ed). San Francisco CA: Jossey-Bass.

Kusy, M., & Essex, L. (1995). No longer a solo practice: How physician leaders lead. *Physician Executive, 21,* 11-16.

Locke, E.A. & Latham, G.P. (1990). *A Theory of Goal Setting and Task Performance.* Englewood Cliffs, NJ: Prentice-Hall.

London, M., & Smither J.W. (1995). Can multi-source feedback change perceptions of goal accomplishments, self-evaluations, and performance related outcomes? *Personnel Psychology, 48,* 803-840.

London, M., & Wohlers, A. (1991). Agreement between subordinate and self-ratings in upward feedback. *Personnel Psychology, 44,* 375-391.

London, S.C., Paul H., & Christensen J. (2000). *Fish!* New York: Hyperion.

Maister, D.H. (2001). *Practice what you Preach: What managers must do to create a High Achievement Culture.* New York: Free press.

Maxwell, J.C. (1999). *The 21 Indispensable Qualities of a leader.* Nashville: Thomas Nelson Publishers.

Maxwell, J.C. (2001). *The 17 Indisputable laws of teamwork.* Nashville: Thomas Nelson Publishers.

McCarthy, A.M., & Garavan, T.N. (1999). Developing self-awareness in the managerial career development process: The value of 360-degree feedback and the MBTI. *Journal of European Industrial Training, 23,* 437-445.

McEnvoy, G.M., & Buller, P.F. (1987). User acceptance of peer appraisals in an industrial setting. *Personnel Psychology, 40,* 785-797.

McManus, K. (2000). What are you in for? *IIE Solutions, 32,* 18.

Microsoft Office [software manual] (2000). Seattle: Microsoft Corporation.

Morical, K.E. (1999). A product review: 360 assessments. *Training & Development, 53,* 43-47.

Neter, J., Kutner, M., Nachtscheim, C., & Wasserman, W. (1996). *Applied linear statistical models* (4th ed.). Chicago: Irwin.

Nunnally, J.C., & Bernstein, I. H. (1974). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Overby, S. (2001). Creating a fun workplace, [on-line]. Available: http://www.inc.com/leadershipandstrategy/guides/20684-print.html.

Patton, D. (2002). *Developing an Upward Feedback Instrument for supervisor development.* Unpublished Masters thesis, AFIT/GLM/ENS02-14. School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH, March 2002.

Paulson, A. (2001). Fun at the firm: The role of play at work. *Christian Science Monitor, 94*, 16.

Pinsonnault, J. (1992). Grading the Boss's performance. *CMA Magazine, 66,* 10-13.

Posner, B.Z., & Kouzes, J.M. (1988). Development and Validation of the Leadership Practices Inventory. *Educational and Psychological Measurement, 48*, 483-496.

Posner, B.Z., & Kouzes, J.M. (1993). Psychometric properties of the Leadership Practices Inventory—Updated. *Educational and Psychological Measurement, 53*, 191-199.

Posner, B.Z., & Kouzes, J.M. (2000). Leadership Practices Inventory Psychometric Properties, [on-line]. Available: *http://www.josseybass.com,* 1-3.

Reilly, R.R., Smither, J.W., & Vasilopoulos, N.L. (1996). A Longitudinal study of upward feedback. *Personnel Psychology, 49,* 599-612.

Schriesheim, C., Powers K. J., Scandura T.A., Gardiner C.C., & Lankau M.J. (1993). Improving construct measurement in management research: Comments and a quantitative approach for assessing the theoretical content adequacy of paper-and-pencil survey-type instruments. *Journal of Management*, 19: 385-417.

Shoemaker, M. (1999). Leadership practices in sales managers associated with the self-efficacy, role clarity, and job satisfaction of the individual industrial salespeople. *Journal of Personal Selling & Sales Management, 19*, 1-19.

Smither, J.W., & Wohlers, A. J. (1995). A Filed Study of reactions to normative versus individualize upward feedback. *Group and Organizational Management, 20,* 61-90.

SPSS$^{©}$ base 11.0 applications guide [software manual] (1999). Chicago, IL: SPSS, Inc.

Tseng, A. (2000). *Leadership Practices in Higher Education in Mongolia.* Unpublished Doctoral Dissertation, Virginia Polytechnic Institute and State University, April 2000: p.31-32.

Vinson, M. (1996). The pros and cons of 36-degree feedback: Making it work. *Training and development, 50,* 11-13.

Waldman, D.A., & Atwater, L.E. (1998). Has 360-degree feedback gone amok? *Academy of Management Executive, 12,* 86-95.

Waldman, D.A., & Atwater, L.E. (2001). Attitudinal and behavioral outcomes of an upward feedback process. *Group and Organizational Management, 26,* 189-206.

Walker, A.G. & Smither, J.W. (1999). A Five-Year Study of Upward Feedback: What Managers Do With Their Results Matters. *Personnel Psychology, 13*, 393-410.

Wunderly, L.J., Reddy, W.B., & Dember, W.N. (1998). Optimism and Pessimism in Business Leaders. *Journal of Applied Social Psychology, 28*, 751-760.

Yukl, G. (2002). *Leadership in Organizations.* New Jersey: Prentice Hall.

Zbar, J. (1999). Are we having fun yet? *Computer world, 33,* 70.

Appendix A: Original 66 Items

Item	Practice	Commitment	Statement	Status
1	Challenge	Seek innovation	Challenges people to improve their performance.	Deleted
2	Challenge	Seek innovation	Challenges the status quo.	Deleted
3	Challenge	Seek innovation	Devotes time to discuss innovation and change.	
4	Challenge	Seek innovation	Encourages others to find better ways of doing things.	Deleted
5	Challenge	Seek innovation	Looks for ways to improve the unit's effectiveness.	
6	Challenge	Seek innovation	Open to suggestions and new ideas.	Deleted
7	Challenge	Take risks	Encourages people to take risks.	
8	Challenge	Take risks	Finds ways to turn setbacks into learning events.	
9	Challenge	Take risks	Lets others take risks and fail without negative consequences.	
10	Challenge	Take risks	Supports others to experiment with new ideas.	Deleted
11	Challenge	Take risks	Takes measured risks based on the team's capabilities.	
12	Inspire	Create a vision	Clearly explains a vision of the team's future.	
13	Inspire	Create a vision	Helps the team create "site pictures" – vivid images that help convey our mission.	Deleted
14	Inspire	Create a vision	Portrays the unit as having a positive impact on the future.	Deleted
15	Inspire	Create a vision	Provides a vision helps the team stay energized, focused, and confident.	
16	Inspire	Create a vision	Tells the team how they can contribute to accomplishing the mission.	Deleted
17	Inspire	Common purpose	Appeals to each member's desire to contribute to the success of the organization.	
18	Inspire	Common purpose	Communicates a vision that draws on the shared values and interests of the team.	Deleted
19	Inspire	Common purpose	Promotes common causes that can be supported by all members of the work group.	
20	Inspire	Common purpose	Shows how the unit's vision is consistent with the teams' own beliefs and values.	Deleted

21	Inspire	Common purpose	Tries to help team members become passionate about the unit's vision.	Deleted
22	Enable	Promote cooperation	Breaks down barriers by encouraging people to work with others outside of their group.	
23	Enable	Promote cooperation	Builds long-term relationships with others outside of the unit.	
24	Enable	Promote cooperation	Persuades team to cooperate with others in order to build strong partnerships.	
25	Enable	Promote cooperation	Respects our ideas and applies them when possible.	Deleted
26	Enable	Promote cooperation	Shows others a willingness to trust their judgment.	Deleted
27	Enable	Empower	Allows us to decide the best way to get our jobs done.	
28	Enable	Empower	Gives us important work to do on critical tasks.	
29	Enable	Empower	Grants us the appropriate authority to do our work.	
30	Enable	Empower	Includes us when making important decisions.	Deleted
31	Enable	Empower	Keeps us up to date on critical issues facing the unit.	Deleted
32	Enable	Empower	Makes sure that we have the information needed to make good judgments on our own.	
33	Enable	Empower	Makes sure we get the chance to develop the skills needed to make good decisions.	
34	Model	Set the example	Acts in ways that are consistent with stated values.	
35	Model	Set the example	Acts in ways that generates trust.	
36	Model	Set the example	Sets a personal example of what is expected of unit members.	
37	Model	Set the example	Sets priorities that are consistent with the unit's values.	Deleted
38	Model	Set the example	Takes time to teach and emphasize the unit's values.	Deleted
39	Model	Small wins	Builds an environment that suggests success is just around the corner.	Deleted
40	Model	Small wins	Builds team's confidence by showing that small steps can make a big difference.	
41	Model	Small wins	Divides large tasks into smaller pieces to gather momentum on a project.	
42	Model	Small wins	Motivates us to tackle big problems with small steps.	
43	Model	Small wins	Reminds us to take things a step at a time.	

44	Model	Small wins	Sets achievable goals.	Deleted
45	Model	Small wins	Sets short term goals that lead to visible signs of success.	
46	Encourage	Recognize Individuals	Gets out from behind the desk and catches people doing things right.	Deleted
47	Encourage	Recognize Individuals	Publicly rewards individual members when they do a good job.	
48	Encourage	Recognize Individuals	Recognizes effort, not just results.	
49	Encourage	Recognize Individuals	Sets high expectations.	Deleted
50	Encourage	Recognize Individuals	Tailors rewards to things we each individually value.	
51	Encourage	Celebrate team	Celebrates events that are important to the unit's members.	
52	Encourage	Celebrate team	Celebrates milestones as a way to acknowledge progress toward group goals.	Deleted
53	Encourage	Celebrate team	Cheers actions that are consistent with achieving the unit's goals.	Deleted
54	Encourage	Celebrate team	Makes sure leaders know about the unit's successes.	
55	Encourage	Celebrate team	Takes part in the unit's celebrations.	Deleted
56	Encourage	Celebrate team	Takes time out to publicly recognize the unit's endeavors.	
57	Enjoy	Allow humor	Allows humor to break through during tense moments.	
58	Enjoy	Allow humor	Encourages non-offensive humor as a way to make the workplace more fun.	
59	Enjoy	Allow humor	Not afraid to laugh at himself/herself.	
60	Enjoy	Allow humor	Sets the tone for a friendly, supportive, and fun workplace.	Deleted
61	Enjoy	Allow humor	Willing to laugh and have fun with others.	
62	Enjoy	Promote fun	Encourages simple, quick, and fun activities that lift spirits at work.	
63	Enjoy	Promote fun	Finds ways to offset hardships caused by work with some fun outcome or activity.	
64	Enjoy	Promote fun	Takes advantage of lulls in the schedule for relaxing and fun activities.	
65	Enjoy	Promote fun	Takes part in social activities organized by the unit.	
66	Enjoy	Promote fun	Willing to take a break during busy periods to do something fun as a unit.	

Appendix B: 62 Items from Version 1

Item	Practice	Commitment	Statement	Status	Reason
1	Challenge	Seek innovation	Challenges processes -- asks "why do we do it this way?"	Deleted	Extra
2	Challenge	Seek innovation	Devotes time to discuss innovation and change.		
3	Challenge	Seek innovation	Encourages others to seek out better ways of doing things.		
4	Challenge	Seek innovation	Looks for ways that challenge the status quo.		
5	Challenge	Seek innovation	Looks for ways to improve the unit's effectiveness.		
6	Challenge	Take risks	Encourages people to take risks.		
7	Challenge	Take risks	Finds ways to turn setbacks into learning events.		
8	Challenge	Take risks	Lets others take risks and fail without negative consequences.	Deleted	Bad
9	Challenge	Take risks	Makes sure the team has opportunities to experiment with new ideas.	Changed	Bad
10	Challenge	Take risks	Takes measured risks based on the team's capabilities.		
11	Inspire	Create a vision	Clearly explains a vision of the team's future.		
12	Inspire	Create a vision	Communicates a vision that draws on the shared values and interests of the team.	Deleted	Bad
13	Inspire	Create a vision	Creates vivid images that help convey our mission.		
14	Inspire	Create a vision	Portrays the unit as having an important impact on the future.		
15	Inspire	Create a vision	Provides a vision helps the team stay energized, focused, and confident.		
16	Inspire	Common purpose	Appeals to each member's desire to contribute to the success of the organization.		
17	Inspire	Common purpose	Gains vision acceptance by showing how it is consistent with the teams' beliefs and values.	Deleted	Bad
18	Inspire	Common purpose	Helps focus team on a common purpose.		
19	Inspire	Common purpose	Helps team members relate their own aspirations with the unit's mission.		

20	Inspire	Common purpose	Promotes common causes that can be supported by all members of the work group.		
21	Inspire	Common purpose	Shares with team members, the importance of their efforts to accomplish the vision.	Deleted	Extra
22	Enable	Promote cooperation	Assigns tasks that require team members to cooperate with each other.		
23	Enable	Promote cooperation	Breaks down barriers by encouraging people to work with others outside of their group.	Deleted	Extra
24	Enable	Promote cooperation	Builds long-term relationships with others outside of the unit.		
25	Enable	Promote cooperation	Encourages the open exchange of information and ideas.		
26	Enable	Promote cooperation	Persuades team to cooperate with others in order to build strong partnerships.		
27	Enable	Empower	Allows us to decide the best way to get our jobs done.		
28	Enable	Empower	Gives us important work to do on critical tasks.		
29	Enable	Empower	Grants us the appropriate authority to do our work.		
30	Enable	Empower	Makes sure that we have the information needed to make good judgments on our own.	Changed	
31	Enable	Empower	Makes sure we get the chance to develop the skills needed to make good decisions.	Deleted	Extra
32	Model	Set the example	Acts in ways that are consistent with stated values.		
33	Model	Set the example	Acts in ways that generates trust.	Deleted	Extra
34	Model	Set the example	Leads by example.		
35	Model	Set the example	Makes decisions that are consistent with the unit's values.		
36	Model	Set the example	Sets a personal example of what is expected of unit members.		
37	Model	Small wins	Builds team's confidence by showing that small steps can make a big difference.		
38	Model	Small wins	Divides large tasks into smaller pieces to gather momentum on a project.		

39	Model	Small wins	Motivates us to tackle big problems with small steps.	Deleted	Extra
40	Model	Small wins	Reminds us to take things a step at a time.		
41	Model	Small wins	Sets short term goals that lead to visible signs of success.		
42	Encourage	Recognize Individuals	Catches people doing things right.	Deleted	Extra
43	Encourage	Recognize Individuals	Publicly rewards individual members when they do a good job.		
	Encourage	Recognize Individuals	Recognizes effort, not just results	Deleted	Extra
45	Encourage	Recognize Individuals	Rewards only those who meet or exceed challenging standards.		
46	Encourage	Recognize Individuals	Tailors rewards to things we each individually value.		
47	Encourage	Recognize Individuals	Takes note of high performers.		
48	Encourage	Celebrate team	Celebrates events that are important to the unit's members.	Deleted	Bad
49	Encourage	Celebrate team	Makes sure leaders know about the unit's successes.		
50	Encourage	Celebrate team	Schedule events to show appreciation for the team's hard work.	Changed	
51	Encourage	Celebrate team	Takes part in celebrating team accomplishments		
52	Encourage	Celebrate team	Takes time out to publicly recognize the unit's endeavors.		
53	Enjoy	Allow humor	Allows humor to break through during tense moments.		
54	Enjoy	Allow humor	Encourages non-offensive humor as a way to make the workplace more fun.		
55	Enjoy	Allow humor	Not afraid to laugh at himself/herself.		
56	Enjoy	Allow humor	Sets the tone for a friendly and supportive workplace.	Deleted	Bad
57	Enjoy	Allow humor	Willing to laugh and have fun with others.		
58	Enjoy	Promote fun	Encourages simple, quick, and fun activities that lift spirits at work.		
59	Enjoy	Promote fun	Finds ways to offset hardships caused by work with some fun outcome or activity.		
60	Enjoy	Promote fun	Takes advantage of lulls in the schedule for relaxing and fun activities.		
61	Enjoy	Promote fun	Takes part in social activities organized by the unit.	Deleted	Extra
62	Enjoy	Promote fun	Willing to take a break during busy periods to do something fun as a unit.		

Appendix C: Version 1 Final 48 Items

Item	Practice	Commitment	Statement	Status
1	Challenge	Seek innovation	Devotes time to discuss innovation and change.	
2	Challenge	Seek innovation	Encourages others to seek out better ways of doing things.	
3	Challenge	Seek innovation	Looks for ways that challenge the status quo.	
4	Challenge	Seek innovation	Looks for ways to improve the unit's effectiveness.	
5	Challenge	Take risks	Encourages people to take risks.	
6	Challenge	Take risks	Finds ways to turn setbacks into learning events.	
7	Challenge	Take risks	Challenges team members to experiment with new ideas.	Changed
8	Challenge	Take risks	Takes measured risks based on the team's capabilities.	
9	Inspire	Create a vision	Clearly explains a vision of the team's future.	
10	Inspire	Create a vision	Creates vivid images that help convey our mission.	
11	Inspire	Create a vision	Portrays the unit as having an important impact on the future.	
12	Inspire	Create a vision	Provides a vision that helps the team stay energized, focused, and confident.	
13	Inspire	Common purpose	Appeals to each member's desire to contribute to the success of the organization.	
14	Inspire	Common purpose	Helps focus team on a common purpose.	
15	Inspire	Common purpose	Helps team members relate their own aspirations with the unit's mission.	
16	Inspire	Common purpose	Promotes common causes that can be supported by all members of the work group.	
17	Enable	Promote cooperation	Assigns tasks that require team members to cooperate with each other.	
18	Enable	Promote cooperation	Builds long-term relationships with others outside of the unit.	
19	Enable	Promote cooperation	Encourages the open exchange of information and ideas.	
20	Enable	Promote cooperation	Persuades team to cooperate with others in order to build strong partnerships.	
21	Enable	Empower	Allows us to decide the best way to get our jobs done.	

22	Enable	Empower	Gives us important work to do on critical tasks.
23	Enable	Empower	Grants us the appropriate authority to do our work.
24	Enable	Empower	Makes sure that we have the ability to make good judgments on our own.
25	Model	Set the example	Acts in ways that are consistent with stated values.
26	Model	Set the example	Leads by example.
27	Model	Set the example	Makes decisions that are consistent with the unit's values.
28	Model	Set the example	Sets a personal example of what is expected of unit members.
29	Model	Small wins	Builds team's confidence by showing that small steps can make a big difference.
30	Model	Small wins	Divides large tasks into smaller pieces to gather momentum on a project.
31	Model	Small wins	Reminds us to take things a step at a time.
32	Model	Small wins	Sets short term goals that lead to visible signs of success.
33	Encourage	Recognize Individuals	Publicly rewards individual members when they do a good job.
34	Encourage	Recognize Individuals	Rewards only those who meet or exceed challenging standards.
35	Encourage	Recognize Individuals	Tailors rewards to things we each individually value.
36	Encourage	Recognize Individuals	Takes note of high performers.
37	Encourage	Celebrate team	Makes sure leaders know about the unit's successes.
38	Encourage	Celebrate team	Shows appreciation for the team's hard work.
39	Encourage	Celebrate team	Takes part in celebrating team accomplishments
40	Encourage	Celebrate team	Takes time out to publicly recognize the unit's endeavors.
41	Enjoy	Allow humor	Allows humor to break through during tense moments.
42	Enjoy	Allow humor	Encourages non-offensive humor as a way to make the workplace more fun.
43	Enjoy	Allow humor	Not afraid to laugh at himself/herself.
44	Enjoy	Allow humor	Willing to laugh and have fun with others.
45	Enjoy	Promote fun	Encourages simple, quick, and fun activities that lift spirits at work.

46	Enjoy	Promote fun	Finds ways to offset hardships caused by work with some fun outcome or activity.
47	Enjoy	Promote fun	Takes advantage of lulls in the schedule for relaxing and fun activities.
48	Enjoy	Promote fun	Willing to take a break during busy periods to do something fun as a unit.

ITEM CATEGORIZATION EXERCISE

The purpose of this exercise is to determine what leadership commitments are described by various statements. Beginning on the next page, a list of behavior-based leadership statements is provided. Each statement describes a leadership behavior that a supervisor may display in a work environment. Each statement describes one of the following 12 leadership commitments:

A. *"Seek innovation"* refers to leadership behaviors that search for and encourage others to search for opportunities to improve the efficiency and effectiveness of the organization.

B. *"Take risks and learn from mistakes"* refers to leadership behaviors that create opportunities for team members to experiment with new ideas.

C. *"Create a vision"* refers to leadership behaviors that convey a vivid image of the organization's future.

D. *"Attract others to a common purpose"* refers to leadership behaviors that show and communicate how aspirations are mutually beneficial to work group members and the organization.

E. *"Support Cooperation"* refers to leadership behaviors that encourage the open exchange of information and ideas among work group members and promotes good working relationships with outside organizations.

F. *"Empower"* refers to leadership behaviors that provide members with the necessary resources, support, and skills to take control of their jobs and make significant inputs to the organization.

G. *"Set the example"* refers to daily behaviors that demonstrate and teach the values and standards espoused by the leader.

H. *"Share personal values"* refers to leaders communicating the core beliefs that fundamentally guide the way they think and act. Leaders must clarify and communicate their personal values so that others may know what they stand for.

I. *"Recognize individual contributions"* refers to leadership behaviors that reward individual progress and contributions that meet high standards of performance.

J. *"Celebrate team accomplishments"* refers to leadership behaviors that personally highlight and recognize the work group's attainment of key objectives and goals.

K. *"Allow humor to reduce stress and boredom"* refers to leadership behaviors that encourage humor to break tension and create an enjoyable workplace.

L. *"Promote fun activities to relax and unwind"* refers to leadership behaviors that encourage creative and fun activities to increase morale and job satisfaction.

CATEGORIZATION TASK

INSTRUCTIONS

Carefully read each statement. Then, think about what leadership commitment (category A through L from the first page) you feel that the particular statement best describes.

In the left most column, place the letter (**A** through **L**) that corresponds to the **ONE** leadership behavior that you feel **BEST** describes the statement.

Please be sure to assign a leadership behavior to each statement, and do not omit any. Eleven of twelve commitments are represented by four behaviors. One commitment is represented by five behaviors.

We recommend you separate page one from the categorization exercise so that you may easily reference the 12 leadership commitments when assigning them to each statement.

Category	Statement
	1. Devotes time to discuss innovation and change.
	2. Encourages people to take risks.
	3. Clearly explains a vision of the team's future.
	4. Appeals to each member's desire to contribute to the success of the organization.
	5. Assigns tasks that require team members to cooperate with each other.
	6. Allows us to decide the best way to get our jobs done.
	7. Acts in ways that are consistent with stated values.
	8. Communicates what he or she is passionate about.
	9. Publicly rewards individual members when they do a good job.
	10. Makes sure leaders know about the unit's successes.
	11. Allows humor to break through during tense moments.
	12. Encourages simple, quick, and fun activities that lift spirits at work.
	13. Finds ways to offset hardships caused by work with some fun outcome or activity.
	14. Encourages non-offensive humor as a way to make the workplace more fun.
	15. Shows appreciation for the team's hard work.
	16. Rewards only those who meet or exceed challenging standards.
	17. Shares personal values with team members.
	18. Leads by example.
	19. Builds long-term relationships with others outside of the unit.

_____ 20. Helps focus team on a common purpose.

_____ 21. Creates vivid images that help convey our mission.

_____ 22. Gives us important work to do on critical tasks.

_____ 23. Finds ways to turn setbacks into learning events.

_____ 24. Encourages others to seek out better ways of doing things.

_____ 25. Helps team members relate their own aspirations with the unit's mission.

_____ 26. Encourages the open exchange of information and ideas.

_____ 27. Grants us the appropriate authority to do our work.

_____ 28. Lets us know what causes him or her frustration.

_____ 29. Tailors rewards to things we each individually value.

_____ 30. Takes part in celebrating team accomplishments.

_____ 31. Not afraid to laugh at himself/herself.

_____ 32. Takes advantage of lulls in the schedule for relaxing and fun activities.

_____ 33. Makes decisions that are consistent with the unit's values.

_____ 34. Looks for ways that challenge the status quo.

_____ 35. Lets team members experiment with new ideas.

_____ 36. Portrays the unit as having an important impact on the future.

_____ 37. Sets a personal example of what is expected of unit members.

_____ 38. Clarifies to others what leads him or her to become impatient.

_____ 39. Promotes common causes that can be supported by all members of the work group.

_____ 40. Makes sure that we have the ability to make good judgments on our own.

_____ 41. Takes note of high performers.

_____ 42. Takes time out to publicly recognize the unit's endeavors.

_____ 43. Persuades team to cooperate with others in order to build strong partnerships.

_____ 44. Willing to take a break during busy periods to do something fun as a unit.

_____ 45. Looks for ways to improve the unit's effectiveness.

_____ 46. Willing to laugh and have fun with others.

_____ 47. Takes measured risks based on the team's capabilities.

_____ 48. Provides a vision that helps the team stay energized, focused, and confident.

_____ 49. Shares with us what keeps him or her awake at night.

Appendix E: Final 48 Items on LCI

Item	Practice	Commitment	Statement
1	Model	Shares personal values	Communicates what he / she is passionate about.
2	Model	Shares personal values	Shares personal values with team members.
3	Model	Shares personal values	Clarifies to others what leads him / her to be impatient.
4	Model	Shares personal values	Shares with us what keeps him / her awake at night.
5	Model	Sets the example	Acts in ways that are consistent with stated values.
6	Model	Sets the example	Leads by example.
7	Model	Sets the example	Makes decisions that are consistent with his or her stated values.
8	Model	Sets the example	Sets a personal example of what is expected of unit members.
9	Inspire	Create a vision	Clearly explains a vision of the team's future
10	Inspire	Create a vision	Creates vivid images that help convey our mission.
11	Inspire	Create a vision	Portrays the unit as having an important impact on the future.
12	Inspire	Create a vision	Provides a vision that helps the team stay energized, focused, and confident.
13	Inspire	Common purpose	Appeals to each member's desire to contribute to the success of the organization.
14	Inspire	Common purpose	Helps focus team on a common purpose.
15	Inspire	Common purpose	Helps team members relate their own aspirations with the unit's mission.
16	Inspire	Common purpose	Directs our attention to common goals that can be supported by all team members.
17	Challenge	Seek innovation	Devotes time to discuss innovation and change.
18	Challenge	Seek innovation	Encourages others to seek out better ways of doing things.
19	Challenge	Seek innovation	Looks for ways that challenge the status quo.
20	Challenge	Seek innovation	Looks for ways to improve the unit's effectiveness.
21	Challenge	Take risks	Willing to experiment with new ideas.
22	Challenge	Take risks	Encourages people to take risks.
23	Challenge	Take risks	Finds ways to turn setbacks into learning events.
24	Challenge	Take risks	Takes measured risks based on the team's capabilities.
25	Enable	Promote cooperation	Assigns tasks that require team members to cooperate with each other.

26	Enable	Promote cooperation	Builds long-term relationships with others outside of the unit.
27	Enable	Promote cooperation	Encourages the open exchange of information and ideas.
28	Enable	Promote cooperation	Persuades team to cooperate with others in order to build strong partnerships.
29	Enable	Empower	Allows us to decide the best way to get our jobs done.
30	Enable	Empower	Gives us important work to do on critical tasks.
31	Enable	Empower	Grants us the appropriate authority to do our work.
32	Enable	Empower	Makes sure that we have the ability to make good judgments on our own.
33	Encourage	Recognize individuals	Publicly rewards individual members when they do a good job.
34	Encourage	Recognize individuals	Rewards only those who meet or exceed challenging standards.
35	Encourage	Recognize individuals	Tailors rewards to things we each individually value.
36	Encourage	Recognize individuals	Takes note of high performers.
37	Encourage	Celebrate team	Makes sure leaders know about the unit's successes.
38	Encourage	Celebrate team	Shows appreciation for the team's hard work.
39	Encourage	Celebrate team	Takes part in celebrating team accomplishments
40	Encourage	Celebrate team	Takes time out to publicly recognize the unit's endeavors.
41	Enjoy	Allow humor	Allows humor to break through during tense moments.
42	Enjoy	Allow humor	Encourages non-offensive humor as a way to make the workplace more fun.
43	Enjoy	Allow humor	Not afraid to laugh at himself/herself.
44	Enjoy	Allow humor	Willing to laugh and have fun with others.
45	Enjoy	Promote fun	Encourages simple, quick, and fun activities that lift spirits at work.
46	Enjoy	Promote fun	Finds ways to offset hardships caused by work with some fun outcome or activity.
47	Enjoy	Promote fun	Takes advantage of lulls in the schedule for relaxing and fun activities.
48	Enjoy	Promote fun	Willing to take a break during busy periods to do something fun as a unit.

REPORT DOCUMENTATION PAGE		*Form Approved* *OMB No. 074-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to an penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)* 25-03-2003	2. REPORT TYPE **Master's Thesis**	3. DATES COVERED *(From – To)* Mar 2002 – Mar 2003
4. TITLE AND SUBTITLE DEVELOPING A LEADERSHIP COMMITMENTS INVENTORY		**5a. CONTRACT NUMBER**
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) White, Abigail L., Captain, USAF		**5d. PROJECT NUMBER**
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 P Street, Building 640 WPAFB OH 45433-7765		**8. PERFORMING ORGANIZATION REPORT NUMBER** AFIT/GLM/ENV/03-03
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) ASC/HR Attn: Ms. Pat Houston 1865 4th Street DSN: 785-4477 WPAFB OH 45433-7765 e-mail:patricia.houston@wpafb.af.mil		**10. SPONSOR/MONITOR'S ACRONYM(S)**
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

    APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Organizations have been modifying performance appraisal systems to collect data from multiple sources to guide supervisor development. Multiple rater programs provide leaders with more than one source of feedback, and give them a more complete, and possibly more accurate picture of their performance, which can facilitate meaningful behavioral change. The purpose of this thesis was to develop and validate a Leadership Commitments Inventory (LCI) from a 360-degree feedback approach that reflected a six-factor leadership taxonomy composed of twelve distinct indices. Using successive content adequacy tests, 48 items were validated and the web-based LCI was administered to 278 participants. Through confirmatory factor analyses, this research demonstrated that the hypothesized six-factor model clearly provided the best explanation of the LCI's underlying factor structure. To a certain extent, this demonstrates the instrument's validity as the items were written and tested to measure the underlying latent structure of six factors with twelve corresponding commitment scales. All three hypotheses were fully or partially supported in this research, and the instrument is now available on the Internet for supervisor's who wish to develop into better leaders.

15. SUBJECT TERMS

    Multiple rater feedback, leadership development, upward feedback, leadership practices, leadership behavior taxonomy, leadership behavior psychometric properties, leadership commitments inventory, leadership practices inventory

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Paul W. Thurston, Maj, USAF (ENV)
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	130	(937) 785-7777, ext 3276; e-mail: paul.thurston@afit.edu

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18