The Structure of the Leadership Construct: A Test of Factorial Invariance Using Structural Equation Modeling

Liliana Rodriguez-Campos
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations
Part of the Educational Leadership Commons

Recommended Citation
https://scholarworks.wmich.edu/dissertations/1329

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
THE STRUCTURE OF THE LEADERSHIP CONSTRUCT: A TEST OF FACTORIAL INVARIANCE USING STRUCTURAL EQUATION MODELING

by

Liliana Rodríguez-Campos

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Educational Studies

Western Michigan University
Kalamazoo, Michigan
August 2002

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
THE STRUCTURE OF THE LEADERSHIP CONSTRUCT:
A TEST OF FACTORIAL INVARIANCE USING
STRUCTURAL EQUATION MODELING

Liliana Rodríguez-Campos, Ph.D.
Western Michigan University, 2002

Structural equation modeling (SEM) was used in this dissertation to investigate the factorial invariance (i.e., equivalence) of the leadership construct as perceived by different groups of people in education. Specifically, the author focuses on three groups (364 teachers, 419 principals, and 369 superintendents) from a statewide survey, where the central concern is whether components of the measurement model of leadership are invariant across those particular groups.

In seeking evidence of multigroup invariance, the author was interested in finding the answer to the following research questions: (1) to what extent do some leadership factorial structures fit a group better than others? and (2) to what extent is the factorial structure of leadership invariant across the three groups of teachers, principals, and superintendents? Respondents’ attitudes toward leadership are related to how they perceive and interpret mental models for this construct. Hence, evidence to support the invariance of factorial structure across groups serves to strengthen the research on educational leadership.

The author drew on leadership theory to develop the basic structure of two measurement models, then LISREL 8 was used to estimate the parameters describing the relationships between the observed variables and the factors proposed in these
models. As to the best fitting model, teachers' conception on leadership is more consistent with the four-factor model, while principals' and superintendents' conceptions are more consistent with the two-factor model. Furthermore, configural and weak invariance across groups for the two measurement models are supported by the data. Therefore, limited inferences were made (from a more robust measurement position) about the relationship between respondents' administrative level and the emergence of differences in the mental models held by them for the leadership construct. Implications of the findings are discussed in the dissertation.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.
ACKNOWLEDGMENTS

I wish to take this opportunity to acknowledge the contribution of many individuals who served as a source of guidance. I would particularly like to thank my committee chair, Dr. Jianping Shen for his continuous advice, support, and publication opportunities throughout my years at Western Michigan University. I deeply appreciate all the time he dedicated to giving me pertinent feedback on my efforts when they were needed. Dr. Shen, you have my heartfelt gratitude for all your guidance and assistance.

I want to thank the other members of my dissertation committee, Dr. James Sanders, and Dr. Subhash Sonnad. I am very grateful to Dr. Sanders because he introduced me formally to the evaluation field through his academic instruction. Dr. Sanders provided me with many professional opportunities at the Evaluation Center, and at the American Evaluation Association. Also, I want to acknowledge the encouragement that Dr. Sonnad has always given me throughout my graduate career, and the dissertation process. Dr. Sonnad has always been happy to be of assistance in anything he could, and all the time with a welcoming smile.

My warmest thanks are extended to Dr. Cynthia Phillips who was always willing to provide me with significant support and guidance throughout my dissertation. I also want to thank Dr. B'ooks Applegate, Dr. Gerhard Mels, and Dr. Karl Jöreskog, their expertise and generosity with their time made the analysis of the data a very enjoyable experience. Furthermore, I need to acknowledge the people who
participated in this study for their willingness to share with a stranger their leadership perceptions.

I would like to specially acknowledge the profound influence of my parents. Their hard work, support, encouragement, dedication, and unending love inspired me to pursue higher professional goals in my life. I will always be grateful to them for giving me such a loving, and marvelous background from which to develop into who I am today. Thank you so much for teaching me to follow my dreams, to always strive to be the best at whatever I attempt to do, and to always enjoy along the wonderful learning experience of my life journey. Also, I want to thank Alessandra, because she keeps continuously reminding me the importance of being a role model to her. Alessandra, I hope that this accomplishment will inspire you to never give up your dreams, no matter how long it might take you to achieve them.

I want to express my deepest gratitude and love to my husband, Rigo, for being with me in every step of my adult life, and for working together to make our dreams come true. Rigo, you have been my best friend, without asking you have instinctively known how to give me your support, and you have been always there whenever I needed you. It is a wonderful experience to share my life with you.

Finally, I wish to thank God for always being my source of strength, and for making possible to experience such a marvelous accomplishment.

Liliana Rodríguez-Campos
# TABLE OF CONTENTS

**ACKNOWLEDGMENTS** ............................................................................................ ii

**LIST OF TABLES** ......................................................................................................... viii

**LIST OF FIGURES** ....................................................................................................... ix

**CHAPTER**

I. **INTRODUCTION** ............................................................................................. 1

   Leadership and Structural Equation Modeling ............................................ 2

   Statement of the Problem ............................................................................... 4

   Purpose of this Study ...................................................................................... 5

   Research Questions .......................................................................................... 5

   Research Justification ...................................................................................... 6

   Final Comments ............................................................................................... 8

II. **LITERATURE REVIEW** ..........................................................................................10

   Leadership ......................................................................................................... 10

      Two-Factor Theory ............................................................................. 11

      Four-Factor Theory ............................................................................ 16

      Leadership Patterns ........................................................................... 23

      Other Comments ................................................................................ 27

      Structural Equation Modeling (SEM)....................................................... 28

      Types of SEM Models ........................................................................ 29

      Scenarios for Testing SEM .................................................................. 36

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table of Contents—Continued

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factorial Invariance</td>
<td>37</td>
</tr>
<tr>
<td>Approaches to Factorial Invariance—Linear CFA Models</td>
<td>39</td>
</tr>
<tr>
<td>Testing Constraints for Factorial Invariance</td>
<td>41</td>
</tr>
<tr>
<td>Types of Factorial Invariance</td>
<td>44</td>
</tr>
<tr>
<td>Other Comments</td>
<td>51</td>
</tr>
<tr>
<td>Empirical Research Using SEM to Test Factorial Invariance</td>
<td>52</td>
</tr>
<tr>
<td>Gap in the Empirical Research on Leadership</td>
<td>54</td>
</tr>
<tr>
<td>Contribution of SEM to Research on Leadership</td>
<td>55</td>
</tr>
<tr>
<td>Final Comments</td>
<td>56</td>
</tr>
<tr>
<td>III. METHODOLOGY</td>
<td>58</td>
</tr>
<tr>
<td>Introduction</td>
<td>58</td>
</tr>
<tr>
<td>Part One of the Secondary Data Analysis—Groundwork</td>
<td>60</td>
</tr>
<tr>
<td>Step 1.1: Select Survey Instrument and Sampling Procedures</td>
<td>61</td>
</tr>
<tr>
<td>Step 1.2: Clean the Data and Impute Missing Values</td>
<td>70</td>
</tr>
<tr>
<td>Step 1.3: Generate Moment Matrices and Mean Vectors</td>
<td>72</td>
</tr>
<tr>
<td>Part Two of the Secondary Data Analysis—CFA</td>
<td>74</td>
</tr>
<tr>
<td>Step 2.1: Design the Measurement Models</td>
<td>74</td>
</tr>
<tr>
<td>Step 2.2: Identify, Estimate, and Assess the Best Fitting Model</td>
<td>78</td>
</tr>
<tr>
<td>Final Comments</td>
<td>87</td>
</tr>
<tr>
<td>Part Three of the Secondary Data Analysis—Invariance</td>
<td>88</td>
</tr>
<tr>
<td>Step 3.1: Configural Factorial Invariance</td>
<td>91</td>
</tr>
</tbody>
</table>
Table of Contents—Continued

Step 3.2: Weak Factorial Invariance ..................................................... 92

Step 3.3: Strong Factorial Invariance ................................................... 94

Step 3.4: Strict Factorial Invariance ................................................... 95

Final Comments ............................................................................................... 96

IV. RESULTS ................................................................................................................. 98

Best Fitting Model for Each Group ................................................................. 98

Factorial Invariance Across Groups ................................................................. 101

Configural Factorial Invariance ........................................................... 104

Weak Factorial Invariance .................................................................... 107

Strong Factorial Invariance .................................................................. 109

Strict Factorial Invariance ..................................................................... 111

Comparison of Parameter Estimates ....................................................... 113

Measured Variable Factor Loadings .................................................... 113

Measured Variable Intercepts ............................................................ 117

Measured Variable Error Variances .................................................. 118

V. DISCUSSION AND CONCLUSION ................................................................. 120

Findings .............................................................................................................. 120

Best Fitting Model .................................................................................. 121

Factorial Invariance ............................................................................ 123

Application of Findings ...................................................................... 127

Limitations of the Study and Suggestions for Further Research ................. 129
Table of Contents—Continued

Comparative Study on Different Respondents ........................................... 132
Longitudinal Study on the Same Respondents ........................................... 132
Alternative Measurement Models ............................................................... 133
Partial Metric Invariance .......................................................................... 135
Ordinal vs. Continuous Data ..................................................................... 136

Conclusion ..................................................................................................... 137

APPENDIX

A. Concept of Leadership According to Several Authors ................................ 140
B. General Information on Leadership ......................................................... 148
C. Leadership Patterns Found in the Literature .......................................... 154
D. Instrument Duplication Permission ......................................................... 159
E. Human Subjects Institutional Review Board Permission ............................. 161
F. The Educational Leadership Survey ......................................................... 163
G. Desirable Knowledge/Skill Base for Educational Leaders Survey ............. 170

BIBLIOGRAPHY ............................................................................................ 174
LIST OF TABLES

1. Fit Indexes for the Two-Factor and Four-Factor Models of Leadership
   Tested ....................................................................................................................... 100
2. Fit Indexes for Alternative Measurement Models Across Groups ..................... 103
3. Hypothesized Pattern Matrix for Configural Invariance of the Two-Factor
   Model ........................................................................................................................... 105
4. Hypothesized Pattern Matrix for Configural Invariance of the Four-Factor
   Model ........................................................................................................................... 106
5. Differences in Fit of Alternative Measurement Models Across Groups ............. 110
6. WLS Estimates of Intercept, Factor Loading, Error Variance, and Squared
   Multiple Correlation (SMC) for the Two-Factor Model ................................. 114
7. WLS Estimates of Intercept, Factor Loading, Error Variance, and Squared
   Multiple Correlation (SMC) for the Four-Factor Model ................................. 115
8. Comparison of the Four Types of Factorial Invariance ................................. 125
LIST OF FIGURES

1. Framework of Leadership Patterns ......................................................... 24
2. General Example of a Measurement Model of Correlated Independent
   Latent Variables ...................................................................................... 32
3. General Example of the Measurement and Structural Components of
   a Model .................................................................................................. 35
4. Flowchart for the SEM Analysis ............................................................. 62
5. Hypothesized Two-Factor Model of Leadership ..................................... 76
6. Hypothesized Four-Factor Model of Leadership ..................................... 77
CHAPTER 1

INTRODUCTION

Structural equation modeling (SEM) was used in this dissertation to investigate the factorial invariance (i.e., equivalence) of the leadership construct as perceived by teachers, principals, and superintendents. Furthermore, the importance of determining and reporting the extent to which comparison groups share the same mental model for leadership was affirmed in this study.

The assessment of factorial invariance has implications for improving the valid and reliable use of leadership scores, and introduces a way with which to address many of the measurement challenges inherent in studying educational leadership. Indeed, the implicit assumption that the relations among measured items are invariant lies at the heart of all research involving comparisons among multi-item concepts (Taris, Bok, & Meijer, 1998).

This dissertation contains five chapters. An overview of this research is provided in the present chapter, which consists of (a) leadership and structural equation modeling, (b) statement of the problem, (c) purpose of this study, (d) research questions, (e) research justification, and (f) final comments. The second chapter consists of a literature review that provides background on the topic of SEM and its application to the study of leadership. The SEM methodology employed to answer the research questions posed in this study is described in the third chapter.
The results obtained are presented in the fourth chapter, and the discussion and conclusions drawn from this research are presented in the fifth chapter.

Leadership and Structural Equation Modeling

Leadership has been a topic for many theoretical investigators and professionals for a long time in history. Numerous definitions and theories of leadership have been provided (see Appendix A). For example, according to Block (1993), "The strength in the concept of leadership is that it connotes initiative and responsibility" (p. 13). Also, Burns (1978) stated that leadership "... is exercised when persons with certain purposes mobilize, in competition or in conflict with others, institutional, political, psychological and other resources so as to arouse and satisfy the motives of followers" (p. 18).

It was revealed during the review of leadership literature that although the leadership construct is very complex, its elements can be analyzed from two specific perspectives: (a) the two-factor theory and (b) the four-factor theory. The two-factor theory consists of two particular elements that describe the leadership construct: (a) people orientation and (b) task orientation. From a broader point of view, the four-factor theory is an expansion of the two-factor theory, with two more elements of the leadership construct. Those four elements are: (a) leader, (b) follower (similar to people orientation), (c) organization, and (d) task (similar to task orientation). Therefore, this study is to test the validity of the two-factor and four-factor theories of leadership.
Shen (2001) stated, “History reminds us of the importance of comparing theory to practice” (p. 124). By moving to the field of practice and using SEM to investigate the structure of the leadership construct, it is possible to gain valuable knowledge on this topic. Moreover, it is important to notice that much of SEM’s attractiveness is due to the fact that it enables the researcher to specify structural relationships among the latent variables, as it is needed with the leadership construct in this study (Bollen & Long, 1993).

SEM techniques enable us to specify the number of factors in a model, and the items expected to associate with each factor (Kalliath, Bluedorn & Gillespie, 1999). Each factor structure reflects the “mental model” expressed by a group of respondents for a given construct (Phillips, 2000). According to Senge (1990, p. 8), “Mental models are deeply ingrained assumptions, generalizations, or even pictures or images that influence how we understand the world and how we take action.” Therefore, the use of SEM methodology will help to capture and interpret the leadership construct.

The power of SEM derives from its ability to assess the fit between theoretically derived predictions and the data (Kelloway, 1998). Leadership is recognized as being of substantial importance but is extremely difficult to be defined and measured. As Schumacker and Marcoulides (1998) stated, SEM provides a setting in which a variety of complex relations can be appropriately investigated. Furthermore, SEM helps to establish the relationship between latent variables or constructs given a theoretical perspective (Schumacker & Lomax, 1996). For this reason, it is clear that SEM can be used to investigate the leadership construct.
Statement of the Problem

The central concern of this dissertation is whether components of a measurement model of leadership are invariant (i.e., equivalent) across particular groups. Shen (2001) commented that "...despite the research, we know little about how teachers' and principals' leadership evolved in practice in the past few years... We know even less about whether there is a congruence between principals and teachers in terms of their perceived leadership" (p. 124). Therefore, in an effort to address this situation, in this study the researcher inquired into whether superintendents, principals, and teachers hold a similar conception for leadership.

There are three reasons for this study. First, according to Senge (1990), the structure, clarity, and stability of the mental picture that survey respondents use to describe a given construct (the leadership construct in this study) is an important concern. Second, previous studies (e.g., Shen, 1998) found that teachers, principals, and superintendents seem to have different concepts of leadership; from teachers to principals and to superintendents, instructional matters are de-emphasized while political and managerial matters are emphasized. Therefore, it is important to understand the similarities and differences among those three groups' conceptions of leadership. Third, from the perspective of research and measurement, it is imperative to illustrate that in order to measure leadership, the issue of construct validity needs to be investigated. Consequently, if different groups hold various conceptions of leadership, it will be a challenge for researchers to conduct studies on leadership and propose general leadership theories.
Purpose of this Study

Mayurama (1997) stated that "one of the most common opportunities, yet one not used all that frequently thus far in the social science literature, is to compare structural models from different groups" (p. 257). Therefore, the purpose of this investigation was to describe an application of structural equation modeling (SEM) to investigate the factorial invariance of the leadership construct as perceived by different groups of people.

The extent to which factorial invariance can be demonstrated describes the degree to which respondents share the same perception for a given construct such that it is comparable, equivalent, and stable across groups, conditions, and/or time (Phillips, 2000). The researcher needed to find whether a factorial model of leadership construct, which is assumed to hold in the population, still holds across three groups of superintendents, principals, and teachers.

If factorial invariance cannot be demonstrated, then substantive evidence is provided of difference in the mental models held by groups of respondents for a given construct (Phillips, 2000). For this reason, evidence was provided through the assessment of factorial invariance that supports improving the valid and reliable use of leadership scores, and a new way to assess the leadership construct was provided.

Research Questions

A confirmatory factor analysis (CFA) approach, a particular application of SEM, was used in this study to determine the extent to which three groups—teachers, principals, and superintendents—perceive the leadership construct in the same way.
Therefore, in seeking evidence of multigroup invariance, there is interest in finding the answer to the following research questions: (1) to what extent do some leadership factorial structures fit a group better than others? and (2) to what extent is the factorial structure of leadership invariant across the three groups of teachers, principals, and superintendents? If measurement structures fail to demonstrate factorial invariance (i.e., equality across groups), there would seem to be some difference in how the leadership concept is perceived by the three groups.

The central concern of this dissertation is whether or not components of two measurement models drawn from two theories of leadership (i.e., the two-factor theory and the four-factor theory) are invariant across particular groups, and whether one of those two models fit a group better than the other. It is important to clarify that more than establishing factorial invariance, this investigation explores the value of SEM methodology to support the valid use of leadership scores. Also, if differences across groups are found, it is necessary to understand that SEM methodology does not provide information of why these differences occurred. However, as Phillips (2000) stated, these models “...can be used to isolate the ways in which groups differ on variables, providing a concise statistical representation of group differences and thus serve as a springboard for additional research designed to identify the sources of group differences” (p. 31).

Research Justification

Mental models are very important for human behaviors, because how people perceive things will dictate how they behave. According to Senge (1990), “Mental...
models focus on the openness needed to unearth shortcomings in our present ways of seeing the world” (p. 12). By inquiring into the invariance of the leadership factorial structure for different groups, the author intends to contribute to the knowledge on leadership. Also, the researcher explores in this dissertation the utility of SEM in leadership research. Therefore, this study has value in both content and methodology.

There are several reasons why SEM can contribute to our understanding of the leadership concept. Kelloway (1998) believed that the primary reason for adopting SEM “is the ability to frame and answer increasingly complex questions about our data” (p. 3). Also, Bollen and Long (1993) stated that much of SEM’s attractiveness is due to the fact that it enables the researcher to specify structural relationships among the latent variables (theoretical constructs that cannot be observed directly), as is needed with the leadership concept in this study.

Byrne (1998) stated that through SEM, the researcher can operationally define the latent variables of interest in terms of behavior believed to represent them. As such, the unobserved variable is linked through paths to one that is observable, thereby making its measurement possible. Also, through using SEM, simultaneous investigation of both direct and indirect effects (through other paths in the model) that would be overlooked within the typical multilevel regression study will be possible, and all of the hypothesized relations in the model can be specified (Heck & Thomas, 2000). Finally, Maruyama (1997) wrote with SEM, researchers are provided opportunities: (a) to disconfirm a hypothesized model and its alternative model, and (b) to distinguish among competing and nonequivalent theoretical models.
Final Comments

The importance of this dissertation on the leadership construct is that it links past and present educational research on leadership while it builds new understanding on this area. Furthermore, Rost (1991) stated, “... practitioners of leadership need to adopt postindustrial leadership models that help them make sense of what they do as leaders and followers in the postmodern world of the twenty-first century” (p. 36).

In regard to the methodology chosen for this study, Hoyle (1995) explained that CFA is a more comprehensive and flexible approach to research design and data analysis than any other single statistical model in standard use by social and behavioral scientists. Furthermore, Phillips (2000) explained that in addition to perhaps reducing the influence of treatment, design, and analysis constraints on the valid use of information generated from longitudinal self-report survey data, tests for factorial invariance add value above and beyond strengthening the psychometric properties of evaluative measures.

Finally, as Shen (1998) wrote, “The empirical findings on the similarities and differences in teachers’, principals’, and superintendents’ conceptions on leadership have implications for explaining the leadership phenomenon and educating future school leaders” (p. 7). Therefore, studying whether superintendents, principals and teachers hold similar conceptions of leadership will have implications for educating, selecting, and improving our educational leaders. Furthermore, understanding how teachers, principals and superintendents hold the concept of leadership will also help
improve the organizational culture and morale by reducing the conflict due to their different conceptions.
CHAPTER 2

LITERATURE REVIEW

This chapter comprises the following sections: (a) leadership, (b) structural equation modeling (SEM), (c) factorial invariance, (d) empirical research using SEM to test factorial invariance, (e) gap in the empirical research on leadership, (f) contribution of SEM to research on leadership, and (g) final comments. The primary purpose of this literature review is to provide information on leadership and the SEM methodology in order for the reader to understand the advantages that this statistical method brings to the understanding of the leadership construct.

Leadership

Leadership has been an interesting research topic for a long period of time, hence numerous definitions and theories have been provided (see Appendix A). According to Owens (1998), no one of those definitions on leadership will satisfy everyone; however those definitions generally agree on two things: (a) leadership is a group function, because it occurs only in the processes of two or more people interacting, and (b) leaders intentionally seek to influence the behavior of other people.

Several perspectives used to describe the elements that constitute the leadership construct are discussed in the literature (Appendix B shows a synthesis of
the main leadership theories described in the literature). According to Sergiovanni (1984), "Perspectives are images of reality and not truths in themselves... perspectives of practice are not truth seeking in the traditional sense but rather to enhance one's understanding and to illuminate one's view of the world" (p. 10).

Although the leadership construct is very complex, review of the literature reveals that its elements can be analyzed from two specific perspectives: (a) the two-factor theory, and (b) the four-factor theory (e.g., Blake & Mouton, 1985; Fiedler, 1967; Hersey, Blanchard, & Johnson, 1996; Likert, 1961; Schriesheim & Bird, 1979).

The two-factor theory consists of two particular elements that describe the leadership construct: (a) people orientation and (b) task orientation. From a broader point of view, the four-factor theory is an expansion of the two-factor theory, with two more elements added to the leadership construct. Those four elements are: (a) leader, (b) follower (similar to people orientation), (c) organization, and (d) task (similar to task orientation).

Two-Factor Theory

As a result of the literature review, it was found that several theories of leadership stipulate that two particular elements constitute the leadership construct: (a) people orientation, and (b) task orientation. The following is a review of the authors and/or studies on leadership that support the two-factor theory.

Ohio State Leadership Studies

Two researchers at Ohio State University, Stogdill and Shartle, contributed to
the expansion of the leadership ideas in the area of leadership research (Stogdill & Shartle, 1953). These researchers began to explore the notion that how a person acts determines that person’s leadership effectiveness. Hence, in the Ohio State Leadership Studies employees in various kinds of organizations were asked to describe their leaders in order to study the different dimensions of leader behavior, the impact on the performance, and satisfaction of followers (Schriesheim & Bird, 1979).

Stogdill and Shartle (1955) identified two independent dimensions of leadership as having a positive influence on followers’ behavior: (a) consideration, or the extent to which the leader is mindful of subordinates, supports open communication, and is oriented toward their subordinates’ welfare, and (b) initiating structure, or the extent to which the leader is task oriented, defining and structuring their own roles and the roles of subordinates toward the attainment of the organizations’ formal goals (Shartle, 1979).

**Michigan Leadership Studies**

The Michigan Studies reported by Likert (1961, 1967) studied the behavioral differences between effective and ineffective supervisors. According to several authors (Hersey, Blanchard, & Johnson, 1996; Katz & Kahn, 1952; Likert, 1979), the Michigan Leadership Studies defined employee-centered and job-centered leadership as opposite ends of a single leadership continuum.

Referring to these studies at the University of Michigan, Hersey, Blanchard, and Johnson (1996) explained that leaders with an employee orientation take interest
in their subordinates, accepting their individuality and personal needs, while leaders with a task orientation see employees as tools to accomplish organizational goals. Furthermore, Daft (1994) stated, “The most effective supervisors were those who focused on the subordinates’ human needs in order to build effective work groups with high performance goals” (p. 485).

**Group Dynamic Studies**

Lewin, Lippitt, and White (1939) explored different leadership styles or behaviors (autocratic, democratic, and laissez-faire) and their effect on group performance. Their findings show that the democratic style — involving group members — resulted in higher group performance. Also, Cartwright and Zander (1960) summarized studies about group dynamics at the University of Michigan. They believed that group objectives are divided in two categories: (a) the achievement of some specific group goal, and (b) the maintenance or strengthening of the group itself. Moreover, Hersey, Blanchard, and Johnson (1996) wrote that goal achievement seems to coincide with the task or production orientation, and group maintenance parallels the people or relationship concepts.

**Rensis Likert’s Management Systems**

Over three decades at the University of Michigan, Rensis Likert — trying to arrive at standards and leadership styles — studied four leadership factors or system styles of administration that contribute to the productivity of a group of workers and to the satisfaction of their members (Likert, 1967). This author described four
leadership or system styles of administration, used by managers of organizations (a) system 1, a task-oriented authoritarian management style, (b) system 2, where some control is delegated to lower levels, (c) system 3, where management does not have complete confidence in employees, and (d) system 4, a people- or relationship-oriented management style based on teamwork and confidence. Hence, this study defines the task-oriented and people-oriented management styles as opposite ends of a single leadership continuum. Hersey, Blanchard, and Johnson (1996) stated, “...Likert found that the closer the management style of an organization comes to system 4 (relationship-oriented), the more likely it will be to have a continuous record of high productivity” (p. 111).

**Tannenbaum and Schmidt's Leadership Continuum**

Tannenbaum and Schmidt (1973) developed a situational approach to leadership, where the leader selects one of seven possible leader behaviors depending on the forces among the leader, follower, and situation. Tannenbaum and Schmidt established a range of choices between (a) democratic, or relationship-oriented behavior, and (b) authoritarian, or task-oriented behavior. These dimensions are found also in the Michigan and Ohio State studies (Hersey, Blanchard, & Johnson, 1996).

**The Leadership Grid**

Blake and Mouton (1985) of the University of Texas proposed a leadership theory built on the work of the Ohio State and Michigan Studies. They developed a
managerial grid and process to enable leaders to determine the nature of their leadership style. The Leadership Grid has five different types of leadership, based on two dimensions similar to those identified in the Ohio State Leadership Studies: (a) concern for people, and (b) concern for production (Carlson, 1996). There is one significant difference between the Leadership Grid and the Ohio State Studies. The former tends to be an attitudinal model, whereas the latter attempts to include behavioral as well as attitudinal concepts (Hersey, Blanchard, & Johnson, 1996).

**Stinson-Johnson Model**

Stinson and Johnson found that the leadership behavior style (task or relationship orientation) the leader should use depends on the nature of the followers and the type of task the followers are performing (Hersey, Blanchard, & Johnson, 1996). Moreover, Stinson and Johnson proposed two options: (a) task structure, which can be highly or lowly structured, and (b) follower capacity, which refers to the degree of achievement motivation, need for independence, and task-relevant education and experience (Hersey, Blanchard, & Johnson, 1996).

**Burn’s Transactional and Transformational Theory**

Transformational and transactional theory represents a synthesis of James MacGregor Burns’ examination on political process and power across contemporary cultures in the world. Burns (1978) stated that his notion is that the processes of leadership must be seen as a part of the dynamics of conflict and of power, that
leadership is linked to collective purpose, and that its effectiveness must be judged by actual social change and by the satisfaction of human needs and expectations.

According to Burns (1978), transactional — or task orientation — leadership occurs when one person makes contact with others in order to exchange valued things. Furthermore, when referring to transformational — or people orientation — leadership, Burns (1978) stated, “The transforming leader looks for potential motives in followers, seeks to satisfy higher needs, and engages the full person of the follower” (p. 4). Finally, this author summarized, “That people can be lifted into their better selves is the secret of transforming leadership and the moral and practical theme of this work” (p. 462).

Four-Factor Theory

As a result of the literature review, it was also found that several theories of leadership built on the two-factor theory, adding two more elements into the leadership construct. Those four elements are: (a) leader, (b) follower (similar to people orientation), (c) organization, and (d) task (similar to task orientation).

Particularly, the main approach to leadership that supports the four-factor theory is Fiedler’s Contingency Theory, which is explained below.

Fiedler’s Contingency Theory

This contingency approach is an effort to combine the leadership style with the situation most favorable for the leader’s success (Fiedler, 1967). According to Hersey, Blanchard, and Johnson (1996), this theory has the following situational
variables: leadership-member relations, position power, and task structure. Leadership-member relations, which is closely related to the first two dimensions of the four-factor theory (leader and follower), describes the leader’s characteristics and relationship with his/her group members. Position power — associated with the organization dimension described in the four-factor theory — refers to the power and authority that the leader’s position provides within the organization. Task structure, which identifies the task dimension in the four-factor theory, refers to the degree of structure that the group has been assigned to perform.

According to Fiedler (1967) the most favorable situation for leaders is when they have good leader-member relations, when they have a powerful position within the organization, and when they are directing a highly structured job. On the other hand, the most adverse situation for leaders is when they are disliked, they have little organizational power, and they are directing unstructured tasks. Finally, Daft (1994) wrote, “An important contribution of Fiedler’s research is that it goes beyond the notion of leadership styles to show how styles fit the situation to improve organizational effectiveness” (p. 490).

Other Contributions to the Four-Factor Theory

Many leadership theorists have also moved beyond the two-factor theory. Although their theories are not as comprehensive as Fiedler’s approach, for the purposes of this dissertation, those theories are going to be synthesized and combined into the four-factor theory (see Appendix C).
Classical Leadership Theory. Before 1945, efforts to understand the concept of leadership focused on the leaders — first dimension of the four-factor theory — and their traits or personal characteristics of the leaders (Hersey, Blanchard, & Johnson, 1996). According to Daft (1994) the personal characteristics of the leaders can be divided into (a) physical characteristics, (b) social background, (c) intelligence and ability, (d) personality, (e) work-related characteristics, and (f) social characteristics. Carlson (1996) wrote, “Knowing the traits of potential leaders reveals only part of the story” (p. 124). Moreover, Yukl (1994) stated that the premise that some leader traits are absolutely necessary for effective leadership has not been substantiated in several decades of research.

Human Relations Approach to Leadership. According to Carlson (1996) researchers shifted their focus from the leaders’ characteristics to the followers’ characteristics, which refers to the second dimension of the four-factor theory (e.g., size, group cohesiveness, intimacy, autonomy, collaboration, morale, and informal group dynamics). One of the most notable research efforts during this period was the work of McGregor (1944), who developed Theory X and Theory Y. In the former, it is assumed that workers inherently dislike work, have little ambition, want safety above all, are not interested in assuming responsibility, and therefore they must be closely supervised. In the latter it is assumed that people — if properly motivated — can view work as natural, be self-directed and creative at work, and require less supervision (Carlson, 1996; Hersey, Blanchard, & Johnson, 1996; Owens, 1998).
Hersey and Blanchard’s Situational Theory. According to Hersey and Blanchard (1974) the situational theory is a contingency approach to leadership that links the leader’s style to the task maturity of subordinates. This theory is divided into two dimensions: (a) leader style, which is based on a combination of the two dimensions of the four-factor model, relationship behavior orientation and task behavior orientation, and (b) task maturity of followers, which varies depending on the subordinate’s ability, skills, or willingness to take responsibility for their own task behavior (Hersey & Blanchard, 1982). Furthermore, Daft (1994) wrote that this theory is easier to understand than Fiedler’s Theory, but it incorporates only the characteristics of followers, and not those of the situation.

Niehouse’s Leadership Approach. Niehouse (1998) stated that how one leads is his/her leadership style, which is a blend of two elements: (a) interactive behavior, where there is participation between leader and follower in the decision-making process, and (b) directive behavior, where the leader delegates responsibility for a task and monitors the results. Moreover, Niehouse (1998) described two factors regarding a subordinate’s readiness: (a) psychological readiness (e.g., willingness, self-confidence, and commitment to do whatever needs to be done), and (b) job readiness (e.g., job skills, competence, and expertise to fulfill a task). Therefore, Niehouse (1998) believed that making the leadership style decision is a matter of recognizing when a subordinate has reached a certain level of readiness.

Path Goal Theory. This is a contingency approach to leadership, where the leader’s responsibility is to increase the subordinate’s motivation, and to clarify the
paths or behaviors that will lead to attain desired goals (Georgopoulos, Mahoney & Jones, 1957). According to Evans (1970), this theory consists of three sets of contingencies: leader behavior style, situational, and use of rewards. Leader behavior style is classified as supportive (similar to follower or people orientation), directive (similar to task orientation), achievement-oriented, and participative styles. Situational contingencies are classified as personal characteristics of group members or followers (similar to Hersey and Blanchard's maturity level) and the organizational environment (including the degree of task structure, the nature of the formal authority system, and the organization itself). Use of rewards consists of clarifying the subordinate's path to the available rewards, or increasing the rewards that they may desire. According to House and Mitchell (1974) the major concern of Path-Goal is how the leader influences followers' perceptions of their work goals, personal goals, and paths to goal attainment.

Vroom-Yetten Contingency Model. In the contingency model they developed, Vroom and Yetten (1973) attempted to prescribe how much participation subordinates should be allowed in making decisions. This model consists of (a) situational variables such as followers and task demands, (b) personal attributes of the leader, such as experience or communication skills, (c) leader behavior, such as directive style, and (d) organizational effectiveness.

Hersey, Blanchard, and Johnson (1996) stated that this model is based on the assumption that "...situational variables interacting with personal attributes or characteristics of the leader result in leader behavior that can affect organizational
effectiveness” (p. 129). The Vroom-Yetten contingency approach is widely respected among researchers in leadership behavior, where the authors believe that people can be developed into more effective leaders. Vroom (1976) explained that this model is very important because leaders have the ability to vary their leadership styles to fit a particular situation.

**Bass and Avolio’s Leadership Approach.** Bass and Avolio (1994) stated how a transformational leader is expected to contribute to an organization’s efforts to improve its tasks and best use its human resources. The terms “transactional” and “transformational” leadership were introduced within the framework of a full-range model of leadership that goes from the highly ineffective laissez-faire (LF) leadership, to the highly active and effective influential leadership.

According to Bass and Avolio (1994), people can be assigned a variety of tasks that can influence how they should be structured and function, which is fundamental to long-range organizational improvement. Moreover, these authors explained that organizations depend on self-managed but fully led multifunctional teams to get tasks done effectively. Therefore, one goal of this approach is to supply strategies to enhance the leadership and quality-improvement efforts throughout the organization (Bass & Avolio, 1994).

**Substitutes for Leadership.** According to Kerr and Jermier (1978) this contingency approach states that situational variables can be so powerful that they actually substitute for or neutralize the need for leadership. Situational variables include (a) organizational variables (e.g., group cohesiveness, formalization,
positional power), (b) task characteristics (e.g., highly structured task, automatic feedback), and (c) group or follower characteristics (e.g., professionalism, experience). Finally, Daft (1994) stated, “Leaders should adopt a style with which to complement the organizational situation” (p. 497).

Leadership Influence. Mintzberg (1983) stated that organizational power is the potential ability to influence or affect the behavior and performance of followers. Also, Owens (1998) stated that there are five sources of organizational power: (a) legitimate (formal position in an organization), (b) reward (authority to reward others), (c) coercive (authority to punish), (d) expert (special knowledge or skill), and (e) referent (subordinates’ desire to emulate the leader). Finally, Daft (1994) believed that power represents the resource with which a leader changes the followers’ behavior, and leadership is the use of that power.

Leadership Competencies. According to Hersey, Blanchard, and Johnson (1996), the leader requires the following skills or competencies: (a) diagnosing, or cognitive competency, where the leader understands what the situation is now and what is expected in the future, (b) adapting, or behavioral competency, which involves altering the leader’s behavior and other resources available to meet the contingencies of the situation, and (c) communicating, or process competency, which involves interacting with others in a way that people understand and accept.

As a final comment, this section covered important ideas on the several leadership approaches and elements. (A summary of this section is shown in Appendix B). According to Hersey, Blanchard, and Johnson (1996), “… on the basis
of the definition of the leadership process as a function of the leader, the followers, and other situational variables, a single ideal type of leader seems unrealistic” (p. 114). Also, Vroom (1976) stated that he does not see a particular form of leadership as optimal for all situations, without considering the nature of the situation in which the leadership behavior is displayed.

Kerzner (1997) wrote, “Effective leaders are neither pure task or relationship behavioralists, but maintain a balance between them. However, in time of crisis, a leader may be required to demonstrate a pure behavioral style or a pure task style” (p. 257). Therefore, it is important to understand that empirical studies show that there is no best leadership style, and leaders adapt their behavior to meet the followers’ needs and their particular environment (Hersey, Blanchard, & Johnson, 1996).

Leadership Patterns

As it was explained before, it was shown in the review of the literature on leadership that although this construct is multifaceted, its components can be analyzed from two specific perspectives: (a) the two-factor theory, and (b) the four-factor theory. The two-factor theory consists of two particular components that constitute the leadership construct: (a) people orientation, and (b) task orientation. Furthermore, the four-factor theory is an expansion of the two-factor theory, with two more components added to the leadership construct. As shown in appendix C, those four components are: (a) leader, (b) follower (similar to people orientation), (c) organization, and (d) task (similar to task orientation). In this study, the researcher is using Fiedler’s Contingency Theory as the main conceptual framework to synthesize
those authors that have gone beyond the two-factor theory. Figure 1 provides the conceptual framework for viewing the four leadership components interactively.

![Diagram of Leadership Patterns](image)

Figure 1. Framework of Leadership Patterns.

**Leader**

Cooley and Shen (1999) wrote, "Leaders must demonstrate a willingness to re-engineer ineffective management and leadership practices" (p. 3). Also, Linver (1994) wrote, "Leaders have passion and are willing to show it... Whatever the source of this passion, it gives them the courage to say and do what they believe is right and take the consequences" (p. 147). Moreover, Kelley (1992) stated, "The best leaders are attuned to themselves and their relationships with others — they understand who they are as leaders, what their strengths and weaknesses are in the role, and how they affect followers" (p. 88).
As a result of the leadership literature reviewed in this study, it can be concluded that the component “leader” has three elements: (a) leader’s style, (b) leader’s competencies or skills, and (c) leader’s personal characteristics (Appendix C assists in the identification of this component).

**Follower**

According to Kelley (1992), “Followers determine not only if someone will be accepted as a leader but also if that leader will be effective. Effective followers are critical for a leader’s or an organization’s success” (p. 13). Moreover, Kelley (1992) wrote, “Without followers, leadership is meaningless and leaders don’t exist” (p. 46). Furthermore, Burns (1978) wrote that the essence of the leader-follower relation is the interaction of persons with different levels of motivations in pursuit of a common, or at least joint purpose.

Hersey, Blanchard, and Johnson (1996) stated, “It is also important to keep in mind that it is the followers’ perceptions of the messages they receive from you [the leader] that evoke behavior” (p. 341). Therefore, leaders need to relate to followers in ways that motivate them to unite with others in sharing and achieving the organization vision (Owens, 1998).

As a result of the leadership literature reviewed in this study, it can be concluded that the component “follower” has four elements: (a) follower’s style, (b) follower’s characteristics, (c) follower’s readiness or maturity, and (d) follower’s feedback (Appendix C assists in the identification of this component). Furthermore, Davis (1998) wrote, “Running schools is a people business. Few important
administrative tasks are either conducted or implemented in isolation from human interaction” (p. 8).

Organization

Carlson (1996) wrote, “... the concept of organization is from the Greek organon, meaning a tool or instrument” (p. 4). Also, Kerzner (1997) stated that organizations are groups of people who coordinate activities to meet organizational objectives. Moreover, Hersey, Blanchard, and Johnson (1996) wrote, “Organizations exist for various reasons and have different organizational goals. Organizational goals are targets toward which input, process, and output are directed” (p. 364).

Deal and Kennedy (1982) stated that leadership is the process of stimulating, developing, and working with people within an organization. As a result of the leadership literature reviewed in this study, it can be concluded that the component “organization” has four elements: (a) organizational subsystems, (b) sources of influence or power, (c) indicators of organizational health, and (d) external environment. (Appendix C assists in the identification of this component).

Task

According to Bass and Avolio (1994), leaders and followers can be assigned a variety of tasks, and the characteristics of the assigned task can influence how the team should be structured and how it will function. Some task behaviors include telling people what, when, where, and how to do something (Hersey, Blanchard, & Johnson, 1996). Also, Kerzner (1997) wrote that there are events or circumstances
that are likely to occur and that could have a very negative impact on the task to be performed. As a result of the leadership literature reviewed in this study, it can be concluded that the component “task” has four elements: (a) task structure or complexity, (b) core dimensions, (c) task demands, and (d) task obstacles. (Appendix C assists in the identification of this component).

Other Comments

The main purpose of this section was to review, collect, and summarize literature in order to describe the concept of leadership, analyze the relationship of its components, and provide patterns found — as a result of this study — from two specific perspectives: (a) the two-factor theory, and (b) the four-factor theory. Through using those perspectives on leadership we are allowed to examine a number of concepts, and ideas that have implications for how the nature of leadership is viewed.

New insights may be gained by using each of the leadership elements explored in this section individually; however, to accomplish a comprehensive diagnosis, the use of those components on a rotating and remixing basis is desirable. Carlson (1996) stated that leadership is a process in which the dynamics permit multiple levels of exchange and participation. Finally, “… it should be emphasized that within a systems approach there is a clear understanding that changes in one subsystem affect changes in other parts of the total system” (Hersey, Blanchard, & Johnson, 1996, p. 13).
Structural Equation Modeling

Byrne (1998) explained that researchers can use SEM to take a confirmatory approach to the multivariate analysis of a structural theory, as is the case with the leadership theory. Moreover, Kalliath, Bluedorn, and Gillespie (1999) wrote, "Structural equation techniques allow us to specify the number of dimensions (factors) in a model as well as the items expected to associate with each dimension" (p. 150). Furthermore, Cliff (1983) stated that SEM techniques are a statistical revolution, and not since the advent of analysis of variance has using a statistical technique so transformed social science research.

According to Kaplan (1955) the path analytic origins of SEM had its beginnings with the biometric work of Sewell Wright, who first applied path analysis to the problem of estimating size components for the measurement of bones. SEM was designed for use by researchers with substantive interests in understanding complex patterns of interrelationships among variables. Furthermore, with SEM the researcher is provided with estimates of the strength of all the hypothesized relationships between variables in a theoretical model (Maruyama, 1997).

By using SEM to test empirical data against leadership theories we can gain valuable knowledge on how practitioners perceive the concept of leadership and whether theory is consistent with practitioners' perceptions. For this reason, it is clear that SEM can be used to investigate the leadership construct, which is recognized as being of substantial importance in education but is extremely difficult to be defined.
and measured. For more information, please refer to Alreck & Settle, 1995; Cohen, 1998; Glass & Hopkins, 1996; Hinkle, Wiersma, & Jurs, 1994.

Types of SEM Models

Schumacker and Lomax (1996) wrote that SEM "involves developing measurement models to define latent variables and then establishing relationships or structural equations among the latent variables" (p. 63). Given the potential complication of SEM, a clear description of the model is critical. Therefore, sections on measurement models (confirmatory factor analysis) and structural models (latent variable relationships) are explained below.

Measurement Model

The measurement model is used to define relations between the observed and unobserved or latent variables (Byrne, 1998). In other words, it is used to provide the link between scores on a measuring instrument and the underlying constructs they are designed to measure.

Jöreskog and Sörbom (1988) wrote that this type of model specifies the pattern by which each measure loads on a particular unobserved variable, and describes the measurement properties (e.g., reliability, validity) of the observed variables. Also, Schumacker and Lomax (1996) wrote, "The measurement model involves specifying which observed variables define a construct (i.e., it is associated with confirmatory factor analysis) and reflects the extent to which the observed variables are assessing the latent variables in terms of reliability and validity" (p. 64).
Therefore, the researcher specifies this model to allow for certain relationships between the unobserved and the observed variables.

According to Schumacker and Lomax (1996), confirmatory factor analysis (CFA) methods reflect measurement models in which observed variables define constructs or factors. Therefore, CFA is used to test hypotheses that a pre-specified factor is defined by a specified subset of variables (Gorsuch, 1983). Furthermore, Kelloway (1998) stated, "...applications of confirmatory factor analysis are particularly appropriate when there is a debate about the dimensionality of factor structure of a scale or measure" (p. 54).

Confirmatory factor analysis, an application of SEM, is both more rigorous and more parsimonious that the "more traditional" techniques of exploratory factor analysis (Kelloway, 1998). According to Kelloway (1998), "...structural equation modeling casts factor analysis in the tradition of hypothesis testing, with explicit tests of both the overall quality of the factor solution and the specific parameters (e.g., factor loading) composing the model" (p. 2).

Byrne (1998) stated that because the CFA model is concerned only with the way in which observed measurements are mapped to particular factors, and not with causal relations among factors, it is termed a measurement model. Furthermore, Schumacker and Lomax (1996) explained that measurement models can contain the following elements:

1. Independent latent variable, in which an independent factor (or latent variable) is defined by observed variables. Arrows pointing toward the observed variables
indicate each variable's measurement error, and the loading of the observed variable on the latent variable.

2. Dependent latent variable, in which a dependent factor is defined by observed variables. Again, arrows pointing toward the observed variables indicate each variable’s measurement error, and the loading of the observed variable on the latent variable.

3. Correlated independent latent variables, in which there are two or more independent factors. Each factor is defined by observed variables, and a curved arrow drawn from one factor to another indicates that they are correlated. As it was explained above, arrows pointing toward the observed variables indicate each variable's measurement error, and the loading of the observed variable on the latent variable. In this dissertation the measurement model of correlated independent latent variables is going to be used. (An example is provided in Figure 2).

4. Correlated error terms of independent latent variables, which is similar to (c) but it also has a curved arrow joining a pair of measurement error terms which indicates that they covary or are correlated.

Long (1983) wrote that in the confirmatory factor model, the researcher imposes constraints that determine (a) which pairs of common factors are correlated, (b) which observed variables are affected by which common factors, (c) which observed variables are affected by a unique factor, and (d) which pairs of unique factors are correlated.
Nunnally and Bernstein (1994) stated that CFA may be used to determine: (a) clusterings (groupings) of variables, (b) which variables belong to a particular group and how strongly they belonging, (c) how many dimensions are needed to explain the relations among the variables, (d) a frame of reference to describe the relations among the variables more conveniently, and (e) scores of individuals on such groupings.

Figure 2. General Example of a Measurement Model of Correlated Independent Latent Variables.

With CFA techniques, researchers are provided an unequivocal way to test the hypotheses related to factorial invariance and thus serve as a valuable tool to assess the leadership construct. When measurement structures fail to be equivalent, this in and of itself may provide evidence of group differences in the perception of
leadership. Therefore, with CFA techniques, researchers are provided a useful mechanism for assuring that any evaluative interpretation of leadership processes relies on the assertion of similarities and differences across groups.

**Structural Model**

Byrne (1998) wrote, “The structural model defines relations among the unobserved variables” (p. 10). Accordingly, in structural modeling, which factor(s) or unobserved variable(s) directly or indirectly causes change in the values of other factors in the model are specified (Byrne, 1998). Furthermore, with structural equation models the relationship among latent variables is established, and are also referred as latent-variable analysis or linear structural relationships (Loehlin, 1992).

As a contrast between measurement models (CFA) and structural models, Schumacker and Lomax (1996) stated, “The confirmatory factor-analytic techniques assess how well the observed variables define the latent variables of interest. In structural equation models, both the independent and dependent latent-variable measurement models are used” (p. 69). Moreover, with the structural model, the researcher is provided an assessment of predictive validity, and with the measurement model, the researcher is provided an assessment of convergent and discriminant validity (Anderson & Gerbin, 1998).

Schumacker and Lomax (1996) wrote that the researcher specifies the structural model to allow for certain relationships among the factors depicted by directed lines or arrows. Moreover, these authors explained that structural models can be divided in:
1. One predictor, in which a single independent factor can predict a single dependent factor.

2. Reciprocal prediction, in which both dependent factors may have a reciprocal relationship.

3. Two predictors that covary, in which two independent factors can correlate in the prediction of a single dependent factor.

4. Mediated prediction, in which one independent factor can predict another factor, which in turn predicts a third factor.

These types of models have become frequently used in the social and behavioral sciences, given the importance of establishing relationships among theoretical constructs (Bollen & Ting, 1991; Fassinger, 1987).

**Two-Step Modeling Approach**

A two-step modeling approach was proposed by James, Mulaik, and Brett (1982), and it describes the analysis of the two conceptually distinct factor models explained above (structural and measurement models). The authors stated that this approach expands the idea of assessing the fit of the observed variables to the factors (measurement model) independently of assessing the fit of the structural equation model among factors (structural model).

Schumacker and Lomax (1996) stated, "Confirmatory factor analysis yields a measurement model for defining and assessing the latent variables, and structural equations are specified between the independent latent variables and the dependent
latent variables to indicate the structural model” (p. 73). Figure 3 shows a general example of the measurement and structural components of a model.

An important point to emphasize in this two-step approach is that the relationships among the latent variables are subject to substantive theory (Schumacker, Lomax, 1996). Also, the use of this approach is impacted by other requirements such as sample size, missing data, outliers, parameter identification, multivariate normality, and the research hypothesis being tested (Fornell, 1983).

Figure 3. General Example of the Measurement and Structural Components of a Model.
According to Schumacker and Lomax (1996), "...the researcher must first specify a measurement model to indicate that the latent variables are measured well (are valid and reliable) by selected observed variables. Afterwards, we can specify a structural model" (p. 83). Furthermore, Jöreskog and Sörbom (1996) wrote that the measurement model should be tested before the structural relationships are tested. It may be useful to do it for each construct, then for two constructs at a time, and then for all constructs. Finally, Schumacker and Lomax, (1996) stated that once factors are adequately defined or measured, then the next step is to examine their factor relationships using a structural model.

Scenarios for Testing SEM

Several authors explained three types of scenarios for testing SEM models (e.g., Bollen & Long, 1993; Jöreskog, 1993; Schumacker & Lomax, 1996).

1. Strictly confirmatory, in which the researcher postulates a single model based on theory, collects the appropriate data, and then tests the fit of the hypothesized model to the sample data, then rejects or fails to reject the model without further modifications.

2. Alternative model, in which the researcher proposes several competing models grounded in theory, and after analyzing a single set of empirical data, the researcher selects one model as most appropriate in representing the sample data.
3. Model generating, in which after postulating and rejecting a theoretically derived model on the basis of its poor fit to the sample data, the researcher proceeds in an exploratory way to modify and reestimate the model.

In testing structural equation models in this dissertation, the researcher is going to use the alternative model approach, where two competing models grounded in theory are going to be proposed, as was explained at the beginning of this chapter.

Factorial Invariance

Application of CFA techniques for testing factorial invariance originated in the early 1970s (Jöreskog, 1971; McGaw & Jöreskog, 1971; Reise, Widaman, & Pugh, 1993). Factorial invariance includes the “study of similarities and differences in the covariation patterns of item-factor relations” (Windele, Iwawaki, & Learner, 1988, p. 551). According to Widaman and Reise (1997) the following question falls under the rubric of “invariance” testing: How can researchers establish that a test measures the same trait dimension, in the same way, when administered to two or more qualitatively distinct groups? For test scores to be comparable across distinct examinee populations, the observed test items, or indicators must have identical (invariant) quantitative relationships with the latent variable for each population of interest (Meredith, 1993; Widaman & Reise, 1997). In other words, one must assume that the test has “measurement invariance” across groups, or that the numerical values under consideration are on the same measurement scale (Drasgow, 1987).
Mayurama (1997) wrote that one of the most common opportunities, yet one not used all that frequently thus far in the social science literature, is to compare measurement models from different groups. Illustrations of multiple-sample comparisons can be found in manuals for most SEM programs including EQS (e.g., Dunn, Everitt, & Picles, 1993), AMOS (e.g., Arbuckle, 1997; Byrne, 1998), and LISREL (e.g., Aish & Jöreskog, 1990; Jöreskog & Sörbom, 1988; Jöreskog et al., 2000). Clearly, the assumption that the relations among a set of measured items and a given construct are invariant is central to this dissertation.

Kaplan (1955) wrote, “The problem of factorial invariance concerns the extent to which a factor model that is assumed to hold in a parent population also holds in subpopulations formed by means of some selection criterion” (p. 74). Taking the leadership concept as an example, the problem of factorial invariance concerns whether a factor model of perceptions of leadership, which is assumed to hold in the population, still holds within teachers, principals, and superintendents, where the types of groups are not necessarily formed by random assignment.

In the field of psychology, factorial invariance has been frequently used to investigate the issues of structure (validity) and stability (reliability) associated with longitudinal and cross-cultural data (e.g., Byrne, 1998; Drascow; Drascow & Kanfer, 1985; Frederiksen, 1987; Linn & Harnisch, 1981; Phillips, 2000). In factorial invariance, for a linear composite score or trait to be comparable across groups, the observed items must have the same relationship with the latent variables for each
group of interest, so that the units of measure, or the scale and the scale’s interpretation are assured to be the same (Meredith, 1993).

**Approaches to Factorial Invariance — Linear CFA Models**

In the review of the literature it was shown that the most common approach to factorial invariance testing relies on the use of CFA techniques to test the equivalence of covariance structures (e.g., Jöreskog, et al. 2000; Phillips, 2000). Jöreskog and Sörbom (1996) wrote, “Covariance structure analysis may be used to study differences in test performance when the tests have been constructed by assigning items or subtests according to objective features of content or format to subclasses of a factorial or hierarchical classification” (p. 208). According to Widaman and Reise (1997), the general equation for estimating parameters and assessing factorial invariance across groups using the CFA approach of covariance structure model — that are meant to be applied only to covariance matrices — is described in equation (1) below:

\[
S_g = \hat{\Lambda}_g \Phi \hat{\Lambda}_g + \hat{\Theta}_{rg} = \hat{S}_g
\]  

(1)

In the above formula, \( S_g \) is a \((p \times p)\) observed sample covariance matrix among measured variables. Also, \( \hat{\Lambda}_g \) is the \((p \times m)\) matrix of the loadings of \( p \) measured variables on \( m \) latent variables, and \( \hat{\Lambda}_g^\prime \) is the matrix transposed. Moreover, \( \hat{\Theta}_{rg} \) is the \((p \times p)\) matrix of covariances among the measurement residuals, and \( \Phi \) is
the \((m \times m)\) matrix of covariances among the factor scores. Furthermore, \(\hat{\Sigma}_g\) is the \((p \times p)\) estimated matrix of covariances among the population of \(p\) measured variables.

Meredith (1993) argued that in addition to the matrices used in the traditional CFA approach of testing the equivalence of covariance structure models in multiple group analysis, it is necessary to include the \(\tilde{\tau}_g\) (measured variable intercepts), and the \(\tilde{\kappa}_g\) (factor mean) matrices in order to test stronger forms of factorial invariance. Therefore, stronger forms of invariance require the consideration of means on the measured and latent variables. For example, without the inclusion of a test for the equality of the \(\tilde{\tau}_g\) matrices, the researcher is unable to ascertain from among a number of possible linear combinations whether that identified for each group is equivalent (Phillips, 2000). Because inclusion of these matrices requires that moment structure models be employed, Meredith (1993) used LISREL to analyze measurement models based on moment matrices.

The CFA approach of moment structure model is done in the same manner as covariance structure models, except that each score is left in its raw score form, and not as a deviation from its mean (Widaman & Reise, 1997). The general equation for estimating parameters and assessing factorial invariance across groups using the CFA approach of moment structure model proposed by Widaman and Reise (1997) is described in equation (2) below:

\[
M_g \equiv \tilde{\tau}_g - \tilde{\kappa}_g - \Phi_g - \tilde{\Phi}_g \hat{\Lambda}_g - \Phi_g \hat{\Lambda}_g \hat{\Theta}_g \hat{\Theta}_g = \tilde{M}_g
\]  \(2\)
In the above formula, $M_g$ is a $(p \times p)$ observed moment matrix (or matrix of raw sums of squares and cross-products of the measured variables), and $\tilde{M}_g$ is the estimated population moment matrix assuming the model is correctly specified. Also, $\hat{\Lambda}_g$ is the $(p \times m)$ matrix of the loadings of $p$ measured variables on $m$ latent variables, and $\hat{\Lambda}'_g$ is the matrix transposed. Moreover, $\hat{r}_g$ is the $(p \times p)$ matrix of the measured variable intercepts, and $\hat{r}'_g$ in the matrix transposed. Furthermore, $\hat{\Theta}_g$ is the $(p \times p)$ matrix of covariances among the measurement residuals, and $\hat{\Phi}_g$ is the $(m \times m)$ matrix of covariances among the factor scores. $\hat{\kappa}_g$ is the $(m \times m)$ matrix for the factor means, and $\hat{\kappa}'_g$ is the matrix transposed. It is important to notice that the $g$ subscript indicates that the matrices described were derived from the $g$th group or sample. The rest of this dissertation is going to present this terminology without the $g$ subscript (unless it is strictly necessary to mention a specific group).

Testing Constraints for Factorial Invariance

In order to compare variables and their relationships across samples, a researcher may decide that the relationships between measures and the underlying latent variables or factors they assess need to be identical across the samples (Maruyama, 1997). Therefore, the most basic way in which to compare solutions across samples is to fit the exact same model with data from different samples and to compare the goodness of fit and the model parameter estimates (Jöreskog, 1993).
This allows an overall fit test of separately estimated samples fitted to a single theoretical model. Then, the solution can force estimates of various parameters to be the same across samples, so a single estimate is generated for each one of any number of specified parameters. That estimate maximizes fit (or minimizes discrepancies) across all the samples simultaneously.

As an illustration, look back at Figure 2, imagining that three different samples are available (teachers, principals, and superintendents). Equality constraints could be imposed on any part of the model, including the relationships of latent variables to observed measures, the residuals for the observed measures, and the relationships among any of the latent variables. According to Maruyama (1997), “then the fit of the solution with constraints could be compared with the fit of a solution that allowed the parameters to be estimated separately for each group” (p. 260).

Finally, the critical question is whether or not the fit of the model to the data gets worse as the constraints and degrees of freedom are added. If the overall chi-square increased substantially, the estimates from the constrained model would not fit as well as those from the unconstrained model. Therefore, the samples must differ in terms of the parameters being constrained (Jöreskog, 1993).

**Imposing Equality Constraints**

Maruyama (1997) stated that by imposing different types of constraints, various assumptions about the relationships can be tested and comparability of latent variables in different samples can be increased. This author explained that “...the
central issue is how to ensure that the variables are defined the same, either across
time or across samples” (p. 263). For example, how can we be sure that what is being
called leader in Sample 1 (teachers) is the same as what is being called leader in
Sample 2 (principals) and in Sample 3 (superintendents)?

It is very difficult to compare causal processes if the factors being compared
differ from one another in different samples. Therefore, if the factors are different,
then the processes being compared cannot be the same. For this reason, Maruyama
(1997) stated, “Most important, the decision on constraining needs to be driven by
theory, not methodological elegance” (p. 264).

The software application LISREL 8 was introduced by Jöreskog (1993). This
software enables the investigation of similarities and differences among factor
structures across groups. The researcher relies on theory to develop the basic structure
of the model, then LISREL is used to estimate the parameters describing the
relationships between observed indicators and the latent constructs proposed in the
model.

The researcher can assign arbitrary values or constrain parameters to be
invariant across particular groups — such as factor loadings, factor intercepts, or error
variances — and estimate the equivalence of the relationships among variables and
factors proposed by a simple model and their fit to the data. What is freely estimated
and what is specified as fixed, is subjective and related to the parameters of greatest
interest to the study at hand (Maruyama, 1998).
Types of Factorial Invariance

There are several types of factorial invariance that show the extent to which groups share a mental model for the construct(s) being investigated. According to several authors (e.g., Meredith, 1993; Phillips, 2000; Taris, Bok, & Meijer, 1998; Widaman & Reise, 1997) the types of factorial invariance are (a) configural factorial invariance, (b) weak factorial invariance, (c) strong factorial invariance, and (d) strict factorial invariance.

Phillips (2000) stated that each type of factorial invariance has an increasing number of equality constraints on the parameter estimates derived across groups. Therefore, when conditions of the first type of factorial invariance are met across groups, then the researcher proceeds to examine the next type of factorial invariance, progressively placing more stringent restrictions on the parameter estimates that are tested using CFA techniques. This increases the strength of the comparative statements that can be made across groups of teachers, principals, and superintendents.

In this dissertation, four types of factorial invariance are analyzed for determining the degree to which respondents share the same perception on leadership across groups: (a) configural (simple structure), (b) weak (factor pattern), (c) strong (factor pattern and intercept), and (d) strict (factor pattern, intercept, and error variance). These can be investigated using LISREL 8 and the CFA models (Golembiewski, et al., 1976; Horn, McArdle & Mason, 1983; Meredith, 1963; Millsap & Hartog, 1988; Widaman & Reise, 1997).
Configural Factorial Invariance

Under the condition of configural factorial invariance the central requirement is that the same pattern of fixed and free loadings in the \( \hat{\Lambda} \) matrix is constrained to be equal for each group (Widaman & Reise, 1997). The researcher can only be sure that the items load on the same factors, but it is uncertain whether the factor loadings (\( \hat{\Lambda} \)), intercepts (\( \hat{\tau} \)), and error variance (\( \hat{\Theta}_e \)) matrices are the same across groups.

When the same pattern of zero and nonzero loadings, that comprise the \( \hat{\Lambda} \) matrices, holds across teachers, principals, and superintendents (for the two and four-factor models), then the same configuration of loadings on factors is observed. Therefore, if this type of invariance holds, then “the same simple structure exists in the subpopulations” (Meredith, 1993, p. 540).

This means that only the factors of the two- and four-factor models studied, and the items related to their particular factors, are hypothesized to be the same across groups. The rest of the parameters are free to vary across groups (e.g., Horn, McArdle & Mason, 1983; Widaman & Reise, 1997). Therefore, lack of reasonable fit shows that the number of factors or the compositions of those factors in the proposed two- and four-factor models differ across groups.

If configural factorial invariance test fails, then the interpretation of group differences to any extent is severely impaired because there is evidence that the leadership construct being measured is not perceived the same across groups (Taris, et al., 1998; Widaman & Reise, 1997). For example, a construct such as leadership
may evolve to mean different things to teachers, principals, and superintendents. This is especially true if the years of experience in those administrative positions increased their understanding of the leadership construct.

According to Widaman and Reise (1997), "Configural invariance has some utility within a set of models reflecting invariance across groups, but the interpretation of group differences is severely compromised" (p. 292). Therefore, the utility of the configural invariance model stems from its role as a baseline model. As a result, different ways of fixing the scale of measurement of the latent variables may yield different parameter estimates and, more importantly, different interpretations of the results.

Because interpretations of the results may vary under the different rescaling, Widaman and Reise (1997) referred to this situation as not ARF invariance (invariance under appropriate rescaling factors). Therefore, parameter estimates under configural invariance are not ARF invariant. Horn and McArdle (1992) stated that unfortunately, configural invariance leaves ambiguous the question of whether particular components of a measurement operation do, or do not, measure in the same manner under different conditions.

**Weak Factorial Invariance**

In this type of factorial invariance the elements that comprise the factor-loading ($\hat{A}$) matrices are constrained to equality across groups, whereas the elements of the intercept ($\hat{\tau}$), and error variance ($\hat{\Theta}_\varepsilon$) matrices are free to vary across groups.
(Meredith, 1993; Widaman & Reise, 1997). If the factor loading ($A$) matrices are equivalent across the three groups studied in this dissertation, then the groups weight the items the same. However, if the test for this form of factorial invariance fail, the interpretation of group differences are limited in that changes in the magnitude of factor loadings across groups indicate a variation or recalibration of scale across the conceptual domain (Phillips, 2000; Van de Vliert, et al., 1985).

People may have different perceptions of reality, given different estimates of what is happening, or they may highlight different aspects of this reality (Taris, et al., 1998). Therefore, a response-shift bias frequently occurs when respondents are more able to accurately assess their real level of functioning on a given construct (Howard & Dailey, 1978; Schaubroeck & Green, 1987; Schmitt, 1982).

According to Widaman and Reise (1997) if the model fits the data adequately, then the group differences in latent-variable variances and covariances become identified in an ARF-invariant, or appropriate rescaling factors, fashion. That is, any substantive interpretation of group differences in variances or covariances among latent variables will remain invariant over rescalings of the latent variables. In other words, any method of identifying a model will provide invariant interpretations of across-group differences in factor variances and covariances. Therefore, if there is not a significant difference of the present model in comparison with the less constrained model studied before, then the hypothesis of weak factorial invariance across groups is accepted, and the researcher can proceed to the next level of factorial invariance.
**Strong Factorial Invariance**

According to Meredith (1993), strong factorial invariance is tested when the loading elements that comprise the $\hat{A}$ matrices and the measured variable intercepts (item means) that comprise the $\hat{f}$ matrices are constrained to equality across groups. In the meanwhile, the elements of the error variance, variance/covariance, and factor mean matrices are free to vary across the three groups of teachers, principals, and superintendents.

Strong factorial invariance supports the hypothesis that the entire linear model that shows the relationship among factors to a given observed variable, in terms of both the raw-score regression weight (loading) and the intercept term is invariant across groups (Meredith, 1993). If this hypothesis holds, then group differences in means and variances on the factors are reflected in group differences in means and variances on the observed variables. Therefore, group differences in the mean level on the factors become identified in the ARF-invariant fashion (Meredith, 1993; Widaman & Reise, 1997; Phillips, 2000).

Widaman and Reise (1997) explained that under strong factorial invariance, both the means on the latent variables as well as the variances and covariances among the factors are ARF invariant. This means that interpretations of differences across teachers, principals and superintendents with regard to mean level, variance-covariance, or both on the latent variables are invariant across rescalings.
Furthermore, any method of identifying a model will provide substantive invariant interpretations of group differences in factor means and variances.

Constraints on the $\Lambda$ and the $\vec{r}$ matrices are considered crucial because this condition establishes that the same latent variables or factors are identified for each group under comparison (Meredith, 1993; Widaman & Reise, 1997). According to Phillips (2000), strong factorial invariance is the minimum factorial invariance condition required for substantive interpretation of group mean differences.

If this strong factorial invariance holds, group differences in both means and variances on the factors are reflected in group differences in means and variances on the measured variables (Widaman & Reise, 1997). However, if this invariance test fails, then the interpretation of group differences, other than those with respect to variance/covariances, are limited.

**Strict Factorial Invariance**

Strict factorial invariance is tested when the $\hat{\Theta}_e$ matrices are constrained to equality across groups, in addition to the previous constraints prescribed by the strong factorial invariance condition. When this condition holds, invariance of the diagonal elements of the $\hat{\Theta}_e$ matrices determines the extent to which measurement error is equivalent across groups. According to Widaman and Reise (1997) an important advantage is that if group differences in the $\hat{\Lambda}$, $\vec{r}$ and $\hat{\Theta}_e$ matrices are negligible, then group differences in means and variances on the measured variables are a function only of group differences in means and variances on the common factors. Therefore,
all group differences on the measured variables are attributable to group differences on the common factors.

Strict factorial invariance test provides evidence that scales show different error that may be dependent on situation (Taris, et al., 1998; Widaman & Reise, 1997). For example, respondents may not be equally able to understand and provide answers to the leadership items across the three groups of teachers, principals, and superintendents.

It is important to notice that this hypothesis is not often met with most data sets, and it is not required for substantive interpretation of group differences. For this reason, it is realistic to expect that the $\Theta_x$ matrices will vary across groups under sampling from a population (Meredith, 1993; Widaman & Reise, 1997). Therefore failing to meet this type of factorial invariance does not present serious interpretation problems in this dissertation.

In addition to the configural, weak, strong, and strict typology, Phillips (2000) explained that other two types of factorial invariance can be investigated across groups. These are (a) variance/covariance, in which the factor loadings, intercepts, and variance/covariance are constrained to be equal across groups and, (b) factor mean level, which constrains factor loadings, intercepts, and factor means to be equal across groups.
Other Comments

This section explained how SEM approaches can be used to compare structural relationships among data collected across three groups of teachers, principals, and superintendents. Therefore, multiple-sample comparisons offer a valuable extension of basic latent variable SEM approaches. As Maruyama (1997) wrote, “Yet with the extension comes greater complexity; for multiple-sample comparisons, there are important issues that will have to be resolved a priori about the most likely way in which to ensure comparability of processes across samples” (p. 265).

With the CFA approach to the investigation of factorial invariance, the researcher is provided with a powerful tool to identify the different quantitative and qualitative aspects of the leadership process, and the explanatory power of the findings are increased (Bollen & Long, 1993). Widaman and Reise (1997) argued that, of the various forms of factorial invariance, constraints on the \( \Lambda \) and \( \Phi \) matrices — strong factorial invariance — are the most important when framing substantive questions. Furthermore, these authors explained that additional constraints on the \( \Gamma \) matrices — strict factorial invariance — are nice but not necessary.

It is important to understand that CFA models are not ends in themselves. Even if one detects differences between groups in crucial CFA model parameters, the CFA models do not indicate why these differences occur (Phillips, 2000). However, the two-factor and four-factor models can be used to isolate the ways in which the groups of teachers, principals, and superintendents differ on variables. These offer a
concise statistical representation of these groups' differences and serve as a basis for further research designed to identify the sources of group differences on the latent and observed variables.

Empirical Research Using SEM to Test Factorial Invariance

Because establishing factorial invariance of a test across groups is critical to progress in many domains, in the review of the literature it was shown that tests of factorial invariance have been applied to many fields, such as education (Bae, Bachman, & Lyle, 1998; Bandalos & Benson, 1990; Huang & Michael, 2000; Maller, et al., 1998; Shen, 1997), sociology (Bledsoe & Baber, 1979; Fergusson, Norwood, & Lynskey, 1994; Marcus, 1999; Motl, et al., 2000; Motl & Conroy, 2001; Oetting, et al., 1998; Rettig & Pasamanick, 1962; Rhodes & Fisher, 1993; Schaie, Maitland & Willis, 1998; Shaw, et al., 1992; Tang, 1998), psychology (e.g., Byrne & Shavelson, 1987; Drasgow & Kanfer, 1985; Frederiksen, 1987; Widaman & Reise, 1997), public affairs and law (Baldwin, 1986; Bove, 1981; Chiou, 1995; Dauphinee, 1993; Elliott, 1970; Whiteside-Mansell, 1995), life sciences (Brakel & Visser, 1996; Gasior, 1999; Giggs & Mather, 1975; Logsdon, Dimenez, & Allmaras, 1996; Nesselroade, 1983), and medicine (Chen, Wang, & Ju, 1993; George, et al., 1995; Rhoades, et al., 1993; Reynolds, et al., 2000; Schaie, et al., 1998; Solomon, et al., 1996).

As was mentioned above, the preceding studies have used SEM to test the factorial invariance of a construct. For example, in the field of social sciences, Shaw, et al., (1992) used factor analysis in samples of high school and college students to investigate the factorial invariance of reckless behavior, where a two-factor model
(driving beyond the speed and drug use) best fit the data for both samples. Also, Marcus (1999) tested the factorial invariance on antisocial behavior, showing that this construct is multidimensional, and has different factor structures for males and females. Furthermore, several authors (e.g., Rhodes & Fisher, 1993; Fergusson, Norwood, & Lynskey, 1994) tested the factorial invariance of antisocial behavior for females and males, and found that they differ in form, variability and frequency.

Several authors have also applied SEM when there is presence of ordinal data, as is the case in this dissertation (Aish & Jöreskog, 1990; Jöreskog, 1990; Jöreskog, 1994; Muthen, 1984; Lee, Poon & Bentler, 1990; Poon & Lee, 1987). The use of ordinal data in structural equation models requires techniques other than those that are used for continuous data. Jöreskog (1994) explained that typically in these types of studies, the estimation of the measurement model is performed through two steps: (a) estimation of polychoric and other correlations for the observed variables, and (b) estimation of the parameters of the model by weighted least squares, which must be a consistent estimate of the asymptotic covariance matrix of the correlations estimated in the first step. Furthermore, in the case of SEM applied to ordinal data, Jöreskog (1994) explained, “The asymptotic covariance matrix of the estimated polychoric correlations is derived for the case when the thresholds are estimated from the univariate marginals and the polychoric correlations are estimated from the bivariate marginals for given thresholds” (p. 381).
Gap in the Empirical Research on Leadership

A desire to identify the concept of leadership has dominated much of organizational research, and the topic has been addressed in many studies and continues to be addressed today (e.g., Portin, Shen & William, 1998; Shen, 1998; Shen, 2001; Shen & Hsieh, 1999; Cooley & Shen, 1999; Weiss & Cambone, 1994).

Particularly, Shen (1998) analyzed qualitative data collected from teachers, principals and superintendents respectively, and found that there are both similarities and differences among the three groups, and their perceptions of leadership are contingent upon their positions in the school system. Moreover, Shen (2001) explained “To make the teachers’ and principals’ perceptions congruent is a daunting task facing us in this new era of school leadership” (p. 128).

Although the attempt to solve the leadership puzzle has generated creative information in this area, in the review of the literature it was found that there has been a lack of effort to use SEM to test the factorial validity of this construct (Hess & Wagner, 1999; Schriesheim & Scandura, 1992). There is also a lack of information as to whether components of the measurement model of leadership are invariant (i.e., equivalent) across groups, and whether a particular leadership factorial structure fits a group better than others. For this reason, in this study the researcher used an SEM approach to assess the structure of the leadership construct, based on two theoretical models (two-factor theory and four-factor theory), and to test its invariance across three groups of teachers, principals, and superintendents.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Contribution of SEM to Research on Leadership

Byrne (1998) stated that researchers are often interested in studying theoretical constructs that cannot be directly observed. These abstract phenomena are termed latent variables, or factors. In this study the latent variables are (a) leader, (b) follower, (c) organization, and (d) task (see Appendix C). Because latent variables are not observed directly, it follows that they cannot be measured directly. Through SEM, the researcher can operationally define the latent variables of interest in terms of behavior believed to represent them. As such, the unobserved variable is linked to one that is observable, thereby making its measurement possible.

There are several ways the use of SEM can contribute to the understanding of the leadership construct in the educational area. Kelloway (1998) believed that the primary reason for adopting SEM “is the ability to frame and answer increasingly complex questions about our data” (p. 3). Also, with SEM testing and specification of complex path models is possible, and SEM is more rigorous and more flexible than are the comparable techniques based on multiple regression (e.g., Cliff, 1983; Kelloway, 1998).

Bollen and Long (1993) stated that much of SEM’s attractiveness is due to the fact that the researcher is able to specify structural relationships among the latent variables, as is needed with the leadership construct in this study. Also, this approach has become a popular methodology for non-experimental research, where methods for testing theories are not well developed and ethical considerations make experimental design infeasible (Bentler, 1990).
In the case of leadership theory, SEM can be an important analytical tool for researchers when it is used to test models guided by strong substantive theory (Schumacker & Marcoulides, 1998). Also, Maruyama (1997) stated that the greatest strength of SEM is that it needs to be driven by theory not by statistical techniques. Moreover, Maruyama wrote that with SEM, the researcher is provided opportunities (a) to disconfirm a hypothesized model and its alternative models, and (b) to distinguish among competing and nonequivalent theoretical models. Furthermore, Kaplan (1955) stated, “Structural equation modeling represents an extremely important advancement in statistical modeling when the goal is accurate estimation and inference within complex systems” (p. 54).

The power of SEM is derived from the attempt to assess the fit of theoretically derived predictions and the data (Kelloway, 1998). Leadership is recognized as being of substantial importance but is extremely difficult to define operationally and measure. Therefore, Schumacker and Marcoulides (1998) explained that with SEM a setting is provided in which many complex relations can be appropriately investigated. For this reason, it is clear that SEM can be used to investigate a variety of leadership problems.

Final Comments

The main purpose of this chapter was to review, collect, and synthesize literature on leadership, SEM, and factorial invariance. New insights are gained every time research on leadership is investigated in particular settings or with particular methodologies, and the potential for achieving new insights through the application of
SEM is vast. The general support that can be obtained from SEM for the leadership construct describes an example of this potential.

With SEM, the researcher is provided an alternative and complementary methodology for examining plausibility of the hypothesized model of leadership. Phillips (2000) wrote, "In addition to perhaps reducing the influence of treatment, design, and analysis constraints on the valid use of information ... tests for factorial invariance add value above and beyond strengthening the psychometric properties of evaluative measures" (pp. 10-11). Although the appropriateness of a common-factor model and even the number of factors comprising it can never be assessed definitely (Kim & Mueller, 1978), the use of structural equation techniques can increase our confidence that the model is consistent with the true population parameters (Kallith, Bluedorn, & Gillespie, 1999).
CHAPTER 3

METHODOLOGY

The methodology for the CFA approach to the investigation of factorial invariance applied to the leadership construct is described in this chapter. The methodology used in this research including: (a) introduction, (b) part one of the secondary data analysis – groundwork, (c) part two of the secondary data analysis – CFA, (d) part three of the secondary data analysis – invariance, (e) limitations of this study, and (f) final comments are explained in the five sections of this chapter.

Introduction

Assuming that measurement models are invariant across groups is critical, and much discussion has been devoted to this topic (e.g., Byrne & Shavelson, 1987; Frederiksen, 1987; Hiu & Triandis, 1985; Widaman & Reise, 1997). Horn and McArdle (1992) argued that if there is no evidence indicating presence or absence of measurement invariance (i.e., factorial invariance), then findings of differences between individuals and groups cannot be unambiguously interpreted. Moreover, it is severely lacking the basis for drawing scientific inference.

One central principle, evident throughout the literature is that measurements are on the same scale (i.e., comparable) when the empirical relations between the observed variables (e.g., survey items) and the latent variables are invariant across
groups (Horn & McArdle, 1992; Reise, Widaman, & Pugh, 1993; Widaman & Reise, 1997). That is, for scores to be comparable across groups, the observed items, or indicators, must have identical, or invariant, quantitative relationships with the latent variable for each population of interest (Meredith, 1993; Widaman & Reise, 1997).

The premise of this chapter is that if leadership invariance does not occur among the groups of teachers, principals, and superintendents, it can be captured by evidence of alterations in the measurement structure determined by the SEM methodology. According to Phillips (2000), “…the measurement structure for a given set of respondents reflects the mental model they hold for the object or behavior under investigation, comparison of measurement structures across groups should be indicative of the extent to which groups ‘see’ the object or behavior in the same way” (pp. 34-35). Therefore, when perceptions about the leadership construct are different among groups, then the mental model of respondents is also different, and this may result in alterations in the model.

In this study, the researcher analyzed the extent to which leadership differences can be detected by SEM methods. Particularly, the researcher sought to determine the extent to which groups that belong to different administrative levels (teachers, principals, and superintendents) perceive the leadership construct in the same way. Therefore, in seeking evidence of multigroup invariance, there is interest in finding the answer to the following research questions: (1) to what extent do some leadership factorial structures fit a group better than others? and (2) to what extent is the factorial structure of leadership invariant across the three groups of superintendents, principals, and teachers? Therefore, if measurement structures fail to
demonstrate factorial invariance (i.e., specific model parameters constrained to equality across groups) it can be said that some difference in leadership perception has occurred.

The following is a description of the processes used to conduct this research. As shown in Figure 4, a flowchart is used to illustrate the three parts that constitute the SEM data analysis.

Part One of the Secondary Data Analysis – Groundwork

This part of the study was a preparation phase to conduct the basic procedures needed to perform the SEM analysis. In Step 1.1, the survey instrument was selected and respondent groups were categorized according to the type of administrative level reported (teachers, principals, and superintendents). In Step 1.2, the data from each category was cleaned and compared – for imputing missing values – in order to conduct the analyses. In Step 1.3, the moment matrices and mean vectors from each sample were generated as needed in order to conduct the second and third parts of this study (see Figure 4, Part 1).

Initially, the data set was recorded in SPSS 9.0\(^1\) for the categorization of groups according to their administrative levels. This procedure was done prior to exporting those files to PRELIS 2 for the identification of missing values and the generation of the covariance matrices and mean vectors needed. PRELIS 2 is a program that checks assumptions about the data on the observed variables, and it is a pre-processor for LISREL 8. Furthermore, LISREL 8 is a program that checks the

\(^1\) SPSS is the Statistical Package for the Social Sciences.
reasonableness of the hypothesized relationships in the models. Both programs can communicate with each other through files written by one program and read by another (Jöreskog, 1990).

Step 1.1: Select Survey Instrument and Sampling Procedures

Given the purpose of this study, it was important to identify a data set that captured evidence of the types of leadership discussed in the previous chapter. Obviously, the differences among groups could result in qualitatively different thinking about the leadership construct. Also, it was necessary to find a psychometrically sound instrument, representative of the population, and large enough to support the proposed SEM methodology. This is because treatment influence, research design, and analysis constraints, threaten construct validity and the credible application of survey methodology to this type of study (Cook & Campbell, 1979; Murnane, Singer, & Willett, 1988; Phillips, 2000; Porras & Berg, 1978).

Three important issues influenced the selection of the actual data set and the instrument used in this research. They are: (a) recent survey instrument, because it is important to have recent information to be analyzed, (b) sampling procedures used by the instrument, because a sample large enough to support SEM methodology was needed in order to demonstrate leadership perceptions, and (c) substantive content on leadership, because only those items that were most pertinent to the leadership patterns found in the theory could be included in the final data set analyzed.
PART 1 - GROUNDWORK.
Preparation for the SEM Data Analysis

**Groundwork**
Purpose: To conduct the basic procedures required in order to perform the SEM data analysis

Select Survey Instrument and Sampling Procedures

Group 1 (Teachers) (n=367)
Clean the Data and Impute Missing Values

Group 2 (Principals) (n=427)
Clean the Data and Impute Missing Values

Group 3 (Superintendents) (n=370)
Clean the Data and Impute Missing Values

Group 1 (Teachers) (n=364)

Group 2 (Principals) (n=419)

Group 3 (Superintendents) (n=369)
Generate Matrices and Mean Vectors

PART 2 - CFA.
Determination of the Best Measurement Model for Groups 1, 2, and 3.

Confirmatory Factor Analysis (CFA)
Purpose: To identify which of the two measurement models of leadership (two-factor and four-factor models) is the best fitting one for each of the groups in this data set.

Design the Measurement Models

Two-Factor Model of Leadership
Two elements constitute the leadership construct: (a) people orientation, and (b) task orientation.

Four-Factor Model of Leadership
Four particular elements constitute the leadership construct: (a) leader, (b) follower, (c) organization, and (d) task.

Identify, Estimate and Assess the Best Fitting Model

PART 3 - INVARIANCE.
Comparison of the Measurement Model Across Groups 1, 2, and 3.

Factorial Invariance
Purpose: To compare to what extent the factorial structure of the two measurement models of leadership (two-factor and four-factor models) is invariant across these three groups.

Impose Constraints

Configural Factorial Invariance
The patterns of factor loadings for each of the measured variables are constrained to be equal (items load on the same factors) across all three groups. It is uncertain whether the factor loadings, intercepts, error variance, and covariance/variance are the same across groups (they are allowed to vary).

Weak Factorial Invariance
The factor loadings for each of the measured variables are constrained to be equivalent across all three groups. It is uncertain whether the intercepts, error variance, and covariance/variance are the same across groups (they are allowed to vary).

Invariant? Yes

STOP HERE

No

Invariance? Yes

Invariance? No

Strong Factorial Invariance
The factor loadings and intercepts for each of the measured variables are constrained to be equivalent across all three groups. It is uncertain whether the error variance, and covariance/variance are the same across groups (they are allowed to vary).

3.1

3.2

3.3

3.4

Invariant?

Invariant?

Invariant?

Figure 4. Flowchart for the SEM Data Analysis.
The survey instrument selected as the best candidate for this secondary analysis covered the requirements established above, because (a) it was conducted from September 1998 until August 2001 in the state of Michigan, (b) it produced samples bigger than 300 subjects, and (c) it contains the specific items related to the patterns found in the theory on leadership. As it will be discussed later, in this study the researcher used data from a statewide survey of teachers, principals, and superintendents conducted by researchers at Western Michigan University.

**Description of the Instrument**

Worthen, Sanders, and Fitzpatrick (1997) stated that surveys have in common the fact that they are measures designed for respondents to report information. In a survey, the investigators select a group of respondents from a larger population through some type of probability sampling, collect information, and then analyze the information to answer the research questions (McMillan, 1996). In particular, self-report surveys are widely used to identify and examine differences between treatment and control or comparison groups (Braverman, 1996; Weisber, Krosnick, & Bowen, 1996).

Obviously, surveys are among the most important data collection tools available for the assessment of leadership. Therefore, the survey instrument selected as the best candidate (e.g., psychometrically sound and large enough to support SEM) for this present research is the statewide “Educational Leadership Survey,” which was used to monitor whether differences in leadership perceptions occur for particular groups of respondents in the state of Michigan.
The survey instrument has a six-point Likert-type frequency scale response format, ranging from 1 (least helpful) to 6 (most helpful) for all the survey items. Furthermore, this instrument has four different versions, because it addresses four types of respondents: (a) superintendents, (b) principals, (c) graduates, which includes among others, principals and teachers, and (d) current students, which also includes among others, principals and teachers (see Appendix F).

For the purposes of this secondary analysis, the four versions of the instrument and their data were provided by Dr. Jianping Shen (Professor at Western Michigan University) as an electronic attachment via email to the researcher (see Appendix D for the Instrument Duplication Permission). The information was stored as required by the Western Michigan University Human Subjects Institutional Review Board (HSIRB) for the duration of the study (a photocopy of the HSIRB permission form is provided in Appendix E).

In order to set apart the three groups examined in this study and save each of the groups (teachers, principals, and superintendents) in separate files, the following independent variables were chosen from the survey's demographic rubric: (a) current employment and (b) position obtained. Although each of the versions in this survey instrument has a different number of areas to be analyzed (and different number of total items per survey), all of them have the same 30 items that address the leadership perceptions of respondents. Therefore, the observed or measurement variables used in this research were obtained from those 30 items that specifically inquire into the leadership perceptions of respondents. Given the leadership theory, the sub-set of
items was identified as those items most closely related to the theories reviewed in Chapter 2 (a photocopy of the instrument is provided in Appendix F).

In this study, three items were selected for each factor (two-factor and four-factor models) of the measurement models explained in Part Two of this chapter, because (a) they produce better measurement than a single item, (b) they produce detailed measurement across a larger spectrum of a continuum, and (c) by diluting item-specific effects, they produce a better measure of commonality (Fowler, 1995). Furthermore, only those items that were most pertinent to the study, demonstrating a full range of variance in response, were included in the final data set analyzed in this study.

**Overview of the Psychometric Properties of the Instrument**

According to Litwin (1995) psychometrics is the branch of survey research that enables the researcher to determine how good the survey is. Therefore, in order to assure psychometric quality and to simplify the reporting of large amounts of survey data, statistics were used to determine the extent to which each composite was a robust measure for the leadership construct. A description of the instrument validation and scaling processes executed by Hsieh (1999) provides specific information of the psychometrics that underlies the survey instrument used in this dissertation.

To develop the leadership section on the survey, Dr. Shen and his team used the items from a survey designed by Goodlad (1994). Also, leadership literature was used to identify the areas required to adequately describe and bound the spectrum of leadership items employed by teachers, principals, and superintendents. Moreover,
Hsieh (1999) explained that psychometric studies, as were required by the leadership team to support the construction of the leadership items, were performed in an adequate manner. Furthermore, this author attests that more effort went into the determination of the instrument quality than is generally found in the educational field.

Jöreskog, et al., (2000) wrote that any variable with less than 16 distinct categories is considered to be ordinal. Therefore, the data of this instrument is assumed to be ordinal, because it represents responses to a set of six ordered categories (six-category Likert scale). Furthermore, it is assumed that a person who responds in a higher category has more of a characteristic than a person who responds in a lower category. Jöreskog and Sörbom (1996) argued, “Ordinal variables are not continuous variables and should not be treated as if they are…. Ordinal variables do not have origins or units of measurements” (p. 146).

Description of Data Collection and Sampling Procedures

Due to the limitation of resources and time constraints in this dissertation study, the author decided to use a reconstructed sample that included 364 teachers, 419 principals, and 369 superintendents. Certainly, it would be ideal to have representative samples across the nation or within the state of Michigan; but this would be very costly to do. McCall (1923) explained that as representativeness can be secured by the method of chance, also equivalence may be secured by chance, provided the number of subjects to be used is sufficiently numerous. Furthermore, Campbell and Stanley (1963) stated, “whereas the problems of internal validity are
solvable within the limits of the logic of probability statistics, the problems of external validity are not logically solvable in any neat, conclusive way” (p. 17).

Gage (1963) explained that generalization is attempted by guessing at laws and revising some of these generalizations in other equally specific but different conditions. Also, Campbell and Stanley (1963) wrote “the sources of external invalidity are thus guesses as to general laws in the science of a science: guesses as to what factors lawfully interact with our treatment variables, and, by implications, guesses as to what can be disregarded” (p. 17). Hence, the current sampling strategy could be viewed as a case study approach, in the sense that the author was interested in illustrating the possible differences among the three groups of teachers, principals, and superintendents. The empirical evidence of this study will provide solid background in order to continue this line of study with other samples.

Data collection procedures were developed to protect respondent confidentiality, while ensuring high quality in the data. Respondents were notified that information collected would only be reported as a group, that individual information would only be used for the purposes of administration of the surveys, and that under no circumstance would information identifying individual respondents be revealed.

For the purposes of this research, responses were considered anonymous in that no identification information was provided to the researcher conducting this analysis. There was no physical, psychological, or social discomfort or inconvenience to respondents. Also, there was no potential for disclosure of sensitive information. Moreover, all participation was voluntary, and consent was obtained from participants.
as they filled out the survey. Furthermore, participants could withdraw their consent at any time during the research process by notifying the researcher and her doctoral committee.

In the following section, how the researcher reconstructed the samples for the present analysis is explained.

**Teachers.** The data to be analyzed in this study was gathered from the combination of two surveys: (a) current students and (b) graduates. The survey of current students was conducted during regular class sessions in the Department of Educational Leadership at Western Michigan University. The survey instrument had a total of 155 items (30 items belonging to the leadership section). A 100% return rate was achieved without any follow-up procedure, and 178 subjects (specifically 129 teachers) returned usable data.

The survey of graduates was sent to alumni from 1992 to 1998 in the Department of Educational Leadership at Western Michigan University. The survey instrument had a total of 145 items (30 items belonging to the leadership section). Also, two follow-up postcards were sent out to all the subjects in order to improve response rate. Therefore, 198 subjects (specifically 88 teachers) returned usable data, with a return rate of 49%. Adding the responses from those two surveys, there is a total sample of 217 teachers.

Due to the fact that this study needs to use sample sizes of 300 and higher because of the ordinal characteristics of the items, the researcher developed another survey instrument: “Desirable Knowledge/Skill Base for Educational Leaders” (see...
Appendix G), which is based on the same 30 leadership items designed by Dr. Shen and Dr. Van Cooley.

Survey data was collected on teachers studying at Western Michigan University, specifically in the Educational Leadership Programs in the Teaching, Learning, and Leadership Department during summer of 2001 in regular class sessions. A sample size of 150 current students was chosen to take part in this survey, in order to fulfill the requirement of a sample size of teachers bigger than 300 (after imputation of the missing values) when programming in LISREL 8 with ordinal variables (Jöreskog & Sörbom, 1996). Therefore, initially there was a total sample of 367 teachers.

**Principal.** The data used in this study was gathered from the combination of three surveys: (a) principals, (b) current students, and (c) graduates. For the principals’ survey, a random sample was selected from the Directory of Michigan Educational Administration Association. The survey instrument had a total of 134 items (30 items belonging to the leadership section). Two follow-up postcards were sent out to all the subjects in order to improve response rate, and 307 respondents returned usable data, with a return rate of 67.2%.

The survey of current students was conducted during regular class sessions in the Department of Educational Leadership at Western Michigan University. The survey instrument had a total of 155 items (30 items belonging to the leadership section). A 100% return rate was achieved without any follow-up procedure, and 178 subjects (specifically 35 principals) returned usable data.
The survey of graduates was sent to alumni from 1992 to 1998 in the Department of Educational Leadership at Western Michigan University. The survey instrument had a total of 145 items (30 items belonging to the leadership section). Also, two follow-up postcards were sent out to all the subjects in order to improve response rate. As a result, 198 subjects (specifically 85 principals) returned usable data, with a return rate of 49%. Adding the responses from those three surveys, initially there was a total sample of 427 principals.

Superintendents. In the original study, 551 superintendents were selected from the Directory of Michigan Educational Administration Association. Questionnaires with 126 items (30 items belonging to the leadership section) were mailed, and one follow-up postcard for superintendents was sent to all the subjects in order to improve response rate. As a result, 370 respondents returned usable data, with a return rate of 67%. Therefore, initially there was a total sample of 370 superintendents for this study.

Step 1.2: Clean the Data and Impute Missing Values

Jöreskog (2001) explained that most raw data from surveys contain many variables, and before doing more elaborate analysis it is important to do a data screening to check for mistakes in the data. Using the program PRELIS 2, separate files were created for teachers, principals, and superintendents with the data set reduced to only the 12 leadership variables needed for the four-factor model, and to the six leadership variables needed for the two-factor model, respectively. Then, a data screening of the samples was performed to obtain the frequency distribution of
the observed variables (Aish & Jöreskog, 1990). Such a data screening was performed in order to reveal if there were specific patterns of missing values in the ordinal data.

As Schumacker and Lomax (1996) wrote, “A problem that researchers often confront is how to treat missing data, that is, the absence or unavailability of data on one or more measured (observed) variables for one or more cases” (p. 3). Jöreskog (Personal Communication, February 19, 2002) stated that there are few options for procedures to deal with missing values, and imputation by matching is a good procedure for the analysis of ordinal variables. With PRELIS 2 it is possible to impute (substitute) real values for the missing values. In other words, the researcher can impute missing values on a variable by matching on other variables. Basically, the value to be substituted is obtained from another case that has a similar response pattern over a set of matching variables (Jöreskog & Sörbom, 1996).

Due to the potential problem of missing data, or unavailability of data on one or more observed variables, the procedure of imputation was performed in PRELIS 2 based on matched variable similarity by using a vector of variables with incomplete data, and a vector of variables with complete data to impute the missing values (Schumacker & Lomax, 1996). According to several authors (Du Toit, Du Toit, & Hawkins, 2001; Jöreskog, 2001) this imputation procedure is based on the notion that if person $x$ has a missing value on variable $i$, and there are several persons with the same response patterns on the matching variables and with the same values on variable $i$, then there is a strong case for substituting this value for person $x$’s value on variable $i$. Obviously, missing values are only imputed if matching cases are found, hence there may still be missing values after imputation (Jöreskog, 2001).
Imputation of missing values was done with utmost care, since missing values were going to be replaced by other values that were going to be treated as real observed values (Jöreskog & Sorbom, 1996b). Therefore, after selecting the 12 and six leadership variables or items needed for the measurement models in this study, then matching variables were chosen from all the remaining leadership items (out of the total 30 leadership items) that were not going to be used in the LISREL 8 modeling. Furthermore, PRELIS 2 listed all the missing values per variable, both before and after imputation. As a result, a final sample was obtained to be used in this study of 364 teachers, 419 principals, and 369 superintendents from the state of Michigan.

Step 1.3: Generate Matrices and Mean Vectors

Jöreskog and Sörbom (1993) recommended the use of PRELIS 2 in this first stage to generate the matrices and mean vectors needed as input for Part 2 of this analysis. Given the ordinal nature of these data, Jöreskog (Personal Communication, June 23, 2001) recommended putting the data for all groups in one file – for the data set with six and 12 variables respectively – and running this through PRELIS 2 to compute the thresholds. Then, he recommended running each group separately in PRELIS 2 with fixed thresholds and computing the mean vector and the covariance matrix of the underlying variables for each group of teachers, principals, and superintendents.

Thresholds were computed in PRELIS 2 for the six and 12 items retained in the data set, respectively. They were used as a common scale for the CFA performed
on the three groups of teachers, principals, and superintendents (n=1152). It is important to note that when the same ordinal variable is measured one or more times, as in the case of this multigroup study, it is possible to estimate the means and variances of these variables (relative to a fixed origin and scale) by specifying the thresholds to be the same for the same variable over groups (Jöreskog, 1990). This makes it possible to put the continuous underlying variables on the same scale, at least artificially, where the average of the means is zero and the average of the variances is one (Aish & Jöreskog, 1990).

Whereas structural models are usually fit to covariance matrices, incorporating the $\tilde{\tau}_x$ (measured variable intercepts) matrices in this dissertation necessitates the fitting of structural models to moment matrices. Therefore, PRELIS 2 was used to allow the researcher to list the correlations among measured variables, along with the respective standard deviations and means of these variables, as input to the program. Then, the moment matrices were implicitly calculated and evaluated (Widaman & Reise, 1997).

PRELIS 2 estimates the asymptotic covariance matrix of the estimated sample variances and covariances under arbitrary non-normal distributions (Browne, 1984; Jöreskog, 1994). Therefore, this matrix can be used to correct for any violation of normal distribution in the samples analyzed (Jöreskog, 1994; Phillips, 2000). Furthermore, the information presented here was used in LISREL 8 – in the next part of this study – to compute a weight matrix for the weighted least squares (WLS) method. Then, LISREL 8 was used to estimate the two- and four-factor models with
the WLS method and to assess the fit of the models to the data (Chou & Bentler, 1995; Phillips, 2000; Jöreskog, 1994).

Part Two of the Secondary Data Analysis – CFA

The purpose of the second part of this study (see Figure 4) was to design the leadership measurement models as suggested in theory, and to identify the best fitting CFA model for the data of teachers, principals, and superintendents. Therefore, this section addresses the design of a confirmatory factor analysis study in order to test the leadership typologies constructed in Appendix C and find the answer to the first research question: to what extent do some leadership factorial structures fit a group better than others?

The following steps were used in this part of the study: (a) Step 2.1 consisted of the design of the measurement models to be used in this study, and (b) Step 2.2 consisted of the identification, estimation, and assessment of the best fitting model for each of the groups in the data set.

Step 2.1: Design the Measurement Models

According to Kelloway (1998) the most important requirement for any form of SEM is the specification of a model, which is an explanation of why two or more variables are related or not. Moreover, the measurement model "specifies how latent variable or hypothetical constructs depend upon or are indicated by the observed variables" (Jöreskog & Sörbom, 1993, p.1). Furthermore, the propositions composing the model are drawn from previous research or theory, taking into consideration that
informed judgment with dogmatic statements of belief should not be discarded (e.g., Bollen & Long, 1993; Kelloway, 1998).

Kaplan (1998) explained that if the researcher is interested in comparing models, it is appropriate to develop rival models to contrast with the proposed one. Ideally, these rival models will stand in nested sequence with the model of interest to allow for the use of direct comparisons (Kelloway, 1998). The two proposed models of leadership designed in this dissertation are based on the typologies constructed as a result of the literature review (see Appendix C). As shown in Figures 5 and 6, note that the rectangles in those models represent observed variables, and the ellipses represent latent constructs (unobserved variables or factors) because they are not measured directly, but they can be indirectly measured or inferred through observable variables (Schumacker & Lomax, 1996).

In this study, leadership was conceived as a two-dimensional and four-dimensional construct. Figure 5 shows the leadership model as a two-dimensional construct, with one dimension corresponding to People Orientation, and the other dimension corresponding to Task Orientation. Consequently, they are the two latent variables (or factors) in this model. Furthermore, the four-factor theory is an expansion of the two-factor theory, with two more components added to the leadership construct. Figure 6 shows the leadership model as a four-dimensional construct, with the following dimensions: (a) Leader, (b) Follower, (c) Organization, and (d) Task. Then, those dimensions are the four factors in this model.

Although we distinguish the factors in each of the models as separate dimensions, it is clear that there are different aspects of a more general and abstract
concept of leadership. Therefore, the two-way arrows present among factors indicate that each of the factors is allowed to correlate with the other factors (i.e., the factors are oblique). In addition, Bagozzi and Heatherton (1994) reported that measurement models frequently demonstrate unsatisfactory fit when there are more than four or five items per factor and sample sizes are large.

![Figure 5. Hypothesized Two-Factor Model of Leadership.](image)

Other than the obvious reason that the model is inappropriate, the principal reason the procedures might fail to achieve invariance appears to be the utilization of an insufficient number of items per factors (Meredith, 1993). As a result, three items per factor were included in the two- and four-factor models of leadership.

Kelloway (1998) stated that structural relations that form a model are depicted in a path diagram in which variables are linked by: (a) unidirectional arrows, which
Figure 6. Hypothesized Four-Factor Model of Leadership.

represent causal relations, or (b) bidirectional curved arrows, which represent noncausal, or correlational, relationships. Therefore, the arrows shown in the diagram, from left to right, are as follows: (a) standardized error term, (b) validity...
(paths or loadings of the observed variables on the latent variables) coefficients of indicators, and (c) correlations between latent constructs (Crocker & Algina, 1986).

According to Byrne (1994) each of the hypothesized models can then be tested statistically in a simultaneous analysis of the entire system of variables to determine the extent to which they are consistent with the data. If goodness of fit is adequate, the model argues for the plausibility of postulated relations among variables; if it is inadequate, the tenability of such relations is rejected. Finally, it is important to understand that although the hypotheses underlying model development may be causal in nature, assessing the fit of a model does not provide a basis for causal inference (Brannick, 1995; Kelloway, 1998; Williams, 1995).

Step 2.2: Identify, Estimate, and Assess the Best Fitting Model

Though justification for estimating the parameters is contingent upon the identification of the model, identification and estimation are distinct issues (Wonnacott & Wonnacott, 1979). Hoyle (1995) explained that identification concerns the correspondence between the information to be estimated – the free parameters – and the information from which it is to be estimated – the observed variances and covariances. Therefore, identification is concerned with whether parameters of the model are uniquely determined (Long, 1983). Once a model has been specified, the next task, estimation, involves using a set of observed data to make estimates of the free parameters. Hence, estimation assumes that the model is identified; otherwise, it can result in arbitrary estimates of the parameters and meaningless interpretations (Long, 1983).
According to Long (1983), in a CFA specific hypotheses are tested. Schumacker and Lomax (1996) explained that through the assessment of model fit, the degree to which the structural equation model fits the sample data can be determined. Bollen and Long (1993) argued that fit indices should not be regarded as measures of usefulness of a model, because they only contain some information about the lack of fit of a model, but none about plausibility. Moreover, these authors believed that fit indices “should not be used in a mechanical decision process for selecting a model. Model selection has to be a subjective process involving the use of judgment” (Bollen & Long, 1993, p. 157).

Identification

Issues of identification deal with whether a unique solution for the model (or its components’ parameters) can be obtained (Bollen, 1989). Moreover, this author wrote that confirmatory factor analyses models are identified if: (a) there are at least three indicators (observed variables) for each latent variable (factor), or (b) there are at least two indicators for each factor and the factors are allowed to correlate (i.e., an oblique solution). In both the two-indicator and three-indicator rules, the researcher must assume that the unique factor loadings (i.e., error terms) are uncorrelated. In this study, there are three indicators for each factor, and the factors are allowed to correlate (see Figures 5 and 6). Therefore, the basic stipulation that three indicators are required to identify a latent variable well was followed (Jöreskog & Sörbom, 1989).
Anderson and Gerbin (1984) found in their Monte Carlo study, that with small samples (e.g., n < 100), there were both convergence failures and improper solutions for CFA when using only two indicators for each latent variable. Therefore, in this dissertation a sample size per group above 300 and three indicators for each latent variable were used, so both convergence failures and the occurrence of improper solutions could be almost eliminated (Kelloway, 1998).

**Conditions for identification.** Long (1983) proposed three easily verifiable conditions that determine unambiguously whether a model is identified: (a) necessary conditions, which if not satisfied indicate that a model is not identified, but if satisfied do not necessarily mean that the model is identified, (b) sufficient conditions, which if met imply that the model is identified, but if not met do not imply that the model is unidentified (although it may be that it is unidentified), (c) necessary and sufficient conditions, which if satisfied imply that the model is identified, and if not satisfied imply that the model is not identified.

Several authors (e.g., Duncan, 1975; Long, 1983; Schumacker & Lomax, 1996) explained that the most effective way to demonstrate that a model is identified is to show – through algebraic manipulations of the model’s covariance equations – that each of the parameters can be solved in terms of the population variances and covariances of the observed variables. Obviously, this is a necessary and sufficient condition for identification and it was addressed in this dissertation.
Estimation

After identification has been established, the researcher can proceed with estimation. Long (1983) stated, “The general objective in estimating the factor model is to find estimates of the parameters that reproduce the sample matrix of variances and covariances of the observed variables as closely as possible in some well-defined sense” (p. 56).

If the model is overidentified, then, by definition, there are an infinite number of solutions, and software packages (e.g., LISREL 8, and AMOS) are designed to solve sets of structural equations, by using numerical methods to estimate parameters (e.g., Bentler & Bonet, 1980; Browne, 1974; Browne, 1984; Byrne, 1998). In particular, in this dissertation, the software LISREL 8 was used, because it solves for model parameters by a process of iterative estimation, which continues until some fitting criteria have been achieved (Kelloway, 1998). Three additional issues regarding model estimation should be noted: choice of estimators, choice of data, and sample size.

Choice of estimators. There are three very common fitting criteria: (a) ordinary least squares (OLS), (b) generalized least squares (GLS), and (c) maximum likelihood (ML). Kelloway (1998) writes that OLS is known as a partial information technique (each path value is estimated independently of the others), whereas both ML and GLS are full information techniques (all the parameters are estimated simultaneously). Under many circumstances, ML and OLS result in identical estimates, lending weight to the use of full information techniques (Williams, 1995).
Particularly, in SEM using ordinal variables, as in this leadership study, weighted least squares (WLS) has been chosen as the method of estimation because it is more efficient than ML estimation, and the chi-square and standard errors are correct (Aish & Jöreskog, 1990; Anderson & Gerbing, 1984; Kelloway, 1998). Moreover, the WLS procedure is also the least biased and most consistent estimator for models with ordinal variables (Bollen, 1989; Kalliath, Bluedorn, & Gillespie, 1999).

**Choice of data.** Some authors have suggested that the choice of a particular type of matrix for input is based on both theoretical concerns and the preferences of some disciplines (Cudeck, 1989; Kelloway, 1998). Specifically, Kelloway (1998) stated that the moment and covariance matrices are strongly recommended in virtually all instances, and their use is generally recommended.

With a comparative assessment of different exploratory and confirmatory procedures it was shown that the analysis of covariance and moment structures are the preferred techniques for investigating factor structure differences (Schmitt, Pulakos, & Lieblein, 1984; Widaman & Reise, 1997). As it was explained in Chapter 2, in order to test stronger forms of factorial invariance (e.g., $\bar{\tau}_s$ or measured variable intercepts matrix), analyses of measurement models were based on moment matrices (Meredith, 1993; Phillips, 2000; Widaman & Reise, 1997).

**Sample size.** SEM is a large-sample technique, where several authors have presented guidelines on the definition of "large" (e.g., Anderson & Gerbing, 1984; Bentler & Chou, 1987). As addressed in this leadership study, a sample size of at least
300 observations would be an appropriate minimum for models of moderate complexity (e.g., Boomsma, 1983; Kelloway, 1998; Tabachnick & Fidell, 1996).

In the present dissertation, the sample sizes chosen for each group of teachers, principals, and superintendents are 364, 419, and 369 subjects, respectively. Therefore, these three samples fulfill the requirement of having sizes bigger than 300 for programming in LISREL 8 when using ordinal or categorical variables (Jöreskog & Sörbom, 1996).

Assessment of Fit

Byrne (1994) stated, "The focal point in analyzing structural equation models is the extent to which the hypothesized model fits, or, in other words, adequately describes the sample data" (p. 53). Given the models described in Figures 5 and 6, assessing the fit of the two-factor and four-factor models is based on (a) whether one model fits better than the rival specification, and (b) whether the model provides a good absolute fit to the data (Kelloway, 1998). Furthermore, Byrne (1998) wrote, "A model never can be confirmed. It can be disconfirmed (it does not fit the observed data), or it can fail to be disconfirmed (it fits)" (p. 139).

According to Bentler and Bonett (1980), "There is a danger in current practice of overemphasizing goodness-of-fit tests while ignoring or minimizing the practical importance associated with various model comparisons" (p. 603). Also, an overemphasis on probability values is particularly dangerous with large-sample data (Reise, Widaman & Pugh, 1993), as in the case of this study. For this reason, several research reports have been inappropriately rejected in instances in which the proposed
models clearly represented a superior understanding of a phenomenon, compared with competing or inadequately specified theories (Bentler & Bonett, 1980; Reise, Widaman & Pugh, 1993). Therefore, Guttman (1977) observed, “A test of statistical significance is not a test of scientific importance” (p. 92).

Reise, Widaman, and Pugh (1993) explained that there are two typical ways of judging the adequacy of an estimated CFA model. First, as it was performed in this study, with certain methods of estimation (i.e., weighted least squares) the researcher is provided a likelihood ratio chi-square statistic to test whether the moment matrix $\hat{M}_g$ reproduced from the estimated parameters, differs significantly from the observed sample moment matrix $M_g$. The statistical acceptability of each model depends on how close the estimated matrices $\hat{M}_g$ are to the observed moment matrices $M_g$.

The most common index is the chi-square ($\chi^2$) test of model fit (Widaman & Reise, 1997). Therefore, in this study, if the chi-square value is statistically significant, then there is a statistical basis for rejecting the tested model in favor of the alternative model. However, if the chi-square value is statistically non-significant, the model studied is an adequate representation of the data. In addition, this study used the $\chi^2/df$ ratio only in a descriptive manner to assess relative fit. This is because several authors (e.g., Marsh, Balla & McDonald, 1988; Widaman & Reise, 1997) argued that the $\chi^2/df$ ratio – as the $\chi^2$ index – is influenced by sample size, and there is no general agreement about its optimal magnitude (e.g., the $\chi^2/df$ ratio has to be below 3.0 or the model should be rejected).
According to several authors (Bentler & Bonett, 1980; Marsh, Balla & McDonald, 1988; Tucker & Lewis, 1973), if the sample size is large – as in this current study) – the likelihood ratio chi-square statistic appears to be overly sensitive to trivial discrepancies between $\hat{M}_g$ and $M_g$. For this reason, Bentler and Bonett (1980) explained that a problem in assessing goodness-of-fit is that the probability of accepting a model increases as the sample size decreases. Thus, one's favorite model will stand the best chance of being retained when tested against the data of small samples. Therefore, as a second way to evaluate CFA models, it is necessary to use "practical" indices of fit.

**Assessing the practical fit of a model.** Because the $\chi^2$ index is highly dependent on sample size, most researchers recommend computing at least two practical indices of model fit; if these are in acceptable ranges, then the model is assumed to be an adequate representation of the data (Widaman & Reise, 1997). Bentler and Bonett (1980) explained that there is a clear distinction between statistical significance and practical significance. Practical indices of fit can provide information about practical significance, in which a statistically significant effect can be evaluated for its practical usefulness in explaining the data. Such an index of information gained in the comparison of competing models is independent of sample size.

In using practical indices of fit, it is important to follow the following principles: (a) no CFA model should be retained or rejected on statistical ground alone (theory, judgment, and persuasive argument should play a key role in defending
the adequacy of a CFA model), and (b) it is useful to calculate at least two indices of practical fit when evaluating a model to ensure that similar characterizations of model fit are obtained (Widaman & Reise, 1997). Additionally, Reise, Widaman, and Pugh (1993) stated, “In brief, researchers applying CFA procedures to data typically stress practical indices of fit at least as strongly as the likelihood ratio chi-square statistical index of fit” (p. 564).

To assess fit of the model for each group, the researcher used the likelihood ratio chi-square statistical index and the following practical fit indices: (a) the root mean square error of approximation (RMSEA), (b) the non-normed fit index (NNFI), and (c) the comparative fit index (CFI). Several authors (Browne, 1990; Browne & Cudeck, 1993; Reise, Widaman, and Pugh, 1993) demonstrated that the RMSEA is an absolute fit measure that performs well as an index of practical fit. The lower bound of the RMSEA is 0 if a model fits a set of data perfectly. Furthermore, RMSEA values of about 0.05 indicate a close fit of a model to data, and values of about 0.08 reflect reasonable fit of a model.

The two practical fit statistics NNFI and CFI are relative fit indices that indicate roughly the proportion of covariation among indicators explained by the model relative to a null model of independence in the indicators (Bentler, 1990; Hoyle & Panter, 1995). Several authors (Marsh, Balla, & McDonald, 1988; Reise, Widaman, and Pugh, 1993) argue that the NNFI and CFI are among the best of the available indices of practical fit, were values greater than .90 are usually considered satisfactory.
Assessing the fit of nested structural models. Statistical and practical fit indices are also used in the comparison of two alternative nested models (e.g., types of factorial invariance). Therefore, in addition to evaluating the chi-square value for each model, in this research the difference in chi-square value for nested models was also evaluated. According to several authors (Aish & Jöreskog, 1990; Bentler & Bonnett, 1980; Reise, Widaman, & Pugh, 1993) the difference in chi-square ($\Delta \chi^2$) values for two nested models is distributed as a chi-square value with degrees of freedom equal to the difference in degrees of freedom ($\Delta df$) for the two models.

If the $\Delta \chi^2$ value is statistically significant, the less restrictive model provides a significantly better fit to the data. In addition, if the more restricted nested model results in a nonsignificant increase in chi-square over the less restricted model, then that particular hypothesis (i.e., the more restricted model) of factorial invariance is tenable (Widaman & Reise, 1997). Finally, as with the chi-square statistic, the researcher also computed the difference in practical-fit indices between nested models (e.g., RMSEA, NNFI, CFI). If practical-fit indices change greatly when restrictions are placed on a model, the restrictions must be fully justified theoretically or they should be relaxed (Reise, Widaman & Pugh, 1993).

Final Comments

Kelloway (1998) wrote that a CFA should be concluded with the presentation of a sample results section. Also, several authors provided a useful guide to reporting the results of SEM (e.g., Bentler, 1990; Jöreskog & Sörbom, 1992; Kelloway, 1998; Raykov, Tomer, & Nesselroade, 1991). Therefore, this dissertation includes: (a) a
graphic presentation of the structural equation models, (b) parameters for the structural equation run, (c) an assessment of model fit, (d) an examination of the obtained solution, and (e) nested model comparisons.

As a summary for this section, it is important to notice that model specification refers to the two- and four-factor models formulated on the basis of leadership theory and past research in the area. Identification determines whether it is possible to find unique values for the parameters of the specified models. Selection of estimation techniques is determined by the distributional properties of the variables being analyzed. Then, after the estimates are obtained, the researcher can test whether the model is consistent with the data (e.g., Bollen & Long, 1993; Browne, 1984; Cliff, 1983; Cudeck & Browne, 1983; MacCallum, Roznowski, & Necowitz, 1992).

Part Three of the Secondary Data Analysis – Invariance

This part of the study refers to the factorial invariance phase (see Figure 4). According to Tisak and Meredith (1991), factorial invariance addresses the extent to which the same constructs are being measured for each group, and under each condition. Therefore, the central concern in this section is whether or not components of the measurement model of leadership are invariant (i.e., equivalent) across three particular groups studied in this dissertation. In other words, to what extent teachers, principals, and superintendents shared the same mental model, or measurement structure on leadership.

In testing for invariance, equality constraints are imposed on particular parameters and, thus, the data for all groups must be analyzed simultaneously to
obtain efficient estimates (Bentler, 1995; Jöreskog & Sörbom, 1996). SEM programs (e.g., LISREL 8) have options that allow simultaneously estimating a single solution across a number of samples (Maruyama, 1997). By comparing fits of different solutions, researchers are able to draw additional inferences about overall model comparability.

According to Jöreskog (1993) tests of hypotheses related to group invariance begin with scrutiny of the measurement model. Once it is known which measures are group-invariant, those parameters are constrained equal while subsequent tests of parameters are conducted. As each new set of parameters is tested, those known to be group-invariant are constrained equal. Therefore, the process of determining nonequivalence of measurement parameters across the three groups of teachers, principals, and superintendents involves the testing of a series of increasingly restrictive hypotheses (Byrne, 1994).

In this study, the three groups of teachers, principals, and superintendents were simultaneously compared, using the multiple sample feature of LISREL 8, against the increasingly restrictive factorial invariance conditions shown at the bottom of Figure 4. Specifically, a set of four increasingly more restrictive CFA models was analyzed to assess invariance for factor pattern, factor loadings, intercept, and error variance. It is important to notice that the terms “constrained” and “restrictive” are used somewhat interchangeably in this dissertation to refer to the extent to which parameters are set or fixed – not freely estimated – to an established value or equality (Phillips, 2000; Widaman & Reise, 1997).
As shown in Figure 4, the analysis proceeded according to a predetermined framework, which flowed sequentially from testing the least restricted configural invariance (only the factor pattern constrained to equality across groups, all others freely estimated) to the most highly restricted strict invariance (factor loadings, intercepts, and error variance constrained to equality across groups). The investigation of factorial invariance was planned to stop when a hypothesis detailing a specific degree of parameter equality across groups failed to be supported by these data (Meredith, 1993; Widaman & Reise, 1997).

In practice, most data sets are found to meet the least constrained forms of factorial invariance (factor pattern), some meet the minimum accepted requirement for comparison of factor means across groups (factor loadings), and very few meet any of the more constrained forms of factorial invariance (LaBouvie & Ruetsch, 1995). In the present research, rejection of $H_0$ is going to argue for the nonequivalence of the groups of teachers, principals, and superintendents, and for the subsequent identification of the source of noninvariance. On the other hand, if $H_0$ cannot be rejected, the groups of superintendents, principals, and teachers are going to be considered to be equivalent, tests for invariance are going to be unjustified, and all subsequent investigative work is going to be based on single-group analyses. (Pitts, West, & Tein, 1996).

According to Maruyama (1997), "...the central issue is how to ensure that the variables are defined the same, either across time or across samples" (p. 263). For example, how can we be sure that what is being called leader in Group 1 (teachers) is
the same as what is being called leader in Group 2 (principals) and in Group 3 (superintendents)? Therefore, by imposing the constraints to be explained below – configural invariance, weak factorial invariance, strong factorial invariance, and strict factorial invariance – various assumptions about the relationships can be tested, and comparability of latent variables in different samples can be increased (Maruyama, 1997).

Step 3.1: Configural Factorial Invariance

Configural factorial invariance is performed as the first activity of this analysis, where only the factor pattern is constrained to equality across groups of teachers, principals, and superintendents (see Figure 4). Therefore, the researcher must assess whether the same number of factors is present and whether the pattern of factor loadings is the same across groups (Taris, et al., 1998). If this assumption does not apply, then further analysis in this dissertation is meaningless.

Configural invariance is consistent with the presumption that similar, but not identical, latent variables in the two- and four-factor models are present across the three groups of teachers, principals, and superintendents. Therefore, configural factorial invariance means that items load on the same factors across the groups being compared (Widaman & Reise, 1997). In other words, only the number of factors and salient items are hypothesized to be the same across groups (Phillips, 2001).

A shift in the pattern of factor loadings is indicative of a different frame of reference, perception, or a redefinition/reconstitution of the conceptual domain (e.g., Golembiewski et al., 1976; Phillips, 2000; Schaubroeck & Green, 1989; Taris et al.,
Furthermore, a lack of reasonable fit shows that the dimensionality (either composition of factors or number of factors) of the factor model proposed differs across groups (Taris et al., 1998).

If the hypothesis of configural invariance across groups is rejected, then the analysis should proceed to identify the extent and nature of the structural differences among groups and describe the evidence of differences in leadership perceptions exhibited as a function of administrative experience. On the other hand, when reasonable indices of fit are obtained for this step, the hypothesis that the conditions of configural invariance are met across groups of teachers, principals, and superintendents should be retained. Therefore, acceptable fit values are obtained for a particular model (two- and/or four-factor model) specified in step 3.1. Then, the analysis should proceed to the next level of invariance – weak factorial invariance – as a series of increasingly restricted nested invariance models.

**Step 3.2: Weak Factorial Invariance**

According to Widaman and Reise (1997), weak factorial invariance holds if the invariance constraints invoked involve the relations of the factors to their indicators (p. 293). The researcher would expect that the item loadings are the same across situations. If this assumption is not met, then a comparison of the factor scores across situations is not warranted, because the magnitude of the factor loadings is not the same across situations (Taris et al., 1998).

In order to test this hypothesis, the model (two-factor and/or four-factor model) specified in step 3.1 is adapted by adding the constraint that the factor-loading
matrices ($\tilde{\Lambda}$) be invariant across groups (see Figure 4). Then, the $\chi^2$ that results from this step 3.2 model is compared with the value obtained for the configural invariance model tested in step 3.1 (Bentler & Bonnett, 1980).

If the restricted model (two-factor and/or four-factor model) results in a non-significant increase in $\chi^2 (\Delta \chi^2)$ over the less constrained model then, the hypothesis of weak factorial invariance across teachers, principals, and superintendents is retained. Therefore, any substantive interpretation of group differences in variances or covariances among latent variables will remain invariant over rescalings of the latent variables (Widaman & Reise, 1997). If the hypothesis of weak factorial invariance is retained, the researcher will continue to place additional invariance constraints on the model(s). If the conditions of weak factorial invariance are not met, then the researcher needs to determine the nature and extent of the structural differences among groups (Taris et al., 1998).

Several authors (Alwin & Jackson, 1981; MacCallum & Tucker, 1991; Reise, Widaman, & Pugh, 1993; Sörbom, 1974) explained that measurement invariance in CFA models is established if factor loadings are invariant across groups. In other words, those authors argue that in order to meet the requirement of "full measurement invariance", it is only necessary to retain the hypothesis of weak factorial invariance ($H_0: \hat{\Lambda}_1=\tilde{\Lambda}_2=\tilde{\Lambda}_3$) assuming that the two-factor and/or four-factor models hold exactly in the population.

It is important to notice that comparison of groups is possible even if the hypothesis of weak factorial invariance is rejected, because the only requirement to
compare groups on a latent variable is that partial measurement invariance be established (Byrne, Shavelson & Muthen, 1989). According to Reise, Widaman and Pugh (1993), "To ensure the non-arbitrariness of the across-group comparisons, a majority of the items on a given latent variable should have loadings that are invariant across groups" (p. 556).

**Step 3.3: Strong Factorial Invariance**

Strong factorial invariance involves one set of additional constraints (the intercepts of the measured variables) over and above those defining the weak factorial invariance model (Meredith, 1993). Therefore, if acceptable fit values are obtained for the model specified in step 3.2, then this model will be modified by adding the constraint that the intercept matrices (\( \bar{\alpha} \)) be invariant across groups (see Figure 4).

According to Widaman & Reise (1997), "Once the strong factorial invariance constraints are imposed on the \( \hat{\Lambda} \) and \( \bar{\alpha} \) matrices, any method of identifying a model will provide substantively invariant interpretations of across-group differences in factor means and variances" (p. 295). For this reason, strong factorial invariance allows for ARF invariant interpretation of estimated parameters. Therefore, similar to step 3.2 above, if the more highly constrained model results in a non-significant increase in \( \chi^2 (\Delta \chi^2) \) over the less constrained model, then the hypothesis of strong factorial invariance (or equal intercepts across groups) is retained, and the researcher will proceed to test the hypothesis of strict factorial invariance.
Step 3.4: Strict Factorial Invariance

The additional constraints that define the strict factorial invariance model involve the measurement residuals contained in the $\hat{\Theta}_e$ matrices (Meredith, 1993). Therefore, if acceptable fit values are obtained for the model specified in step 3.3, then this model is adapted by adding the constraint that the diagonal elements of the error matrices ($\hat{\Theta}_e$) be invariant across groups.

The test of equal error variances indicates whether the reliability of the measurement is also equal across situations (Taris et al., 1998). If this hypothesis of equal variance is retained, then this test of equality across error variances indicates that the measurement error is invariant across groups. Furthermore, several authors explained that if the constrained model holds up, the respondents are equally well able to understand and provide answers to the items, regardless of the situations they are in. However, the reliability may be found to be dependent on the situation (Taris et al., 1998; Widaman & Reise, 1997).

If this constrained model results in a non-significant increase in $\chi^2 (\Delta \chi^2)$ over the less constrained model, then the hypothesis of equal error variance across groups is retained. This implies equivalence across respondents in their ability to understand and provide answers to the items, regardless of group membership; however, if the hypothesis is rejected, the error variance of the items may depend on the situation (Taris et al., 1998; Phillips, 2000).

It is important to notice that strict factorial invariance rarely occurs, and this occurs for a variety of reasons. Widaman and Reise (1997) believed that one of the
most noticeable reasons is the potential increase in residual variance of a observed variable as its variance caused by the factor, or latent variable, increases. However, several authors believed that this is not a major problem because group differences in means and variances on the latent variables are still identifiable in an ARF-invariant fashion if strong factorial invariance holds (Meredith, 1993; Widaman & Reise, 1997).

Final Comments

In this chapter two measurement models based on knowledge of the related theory on leadership have been postulated. Using those hypothesized models the researcher tested their plausibility based on sample data with all the observed variables in the models, as described in the next chapter. Through this model-testing procedure the goodness of fit between each of the hypothesized models and the sample data was determined. Then, results of the statistical testing and parameter estimates were reported. On the basis of this information, the researcher decided whether each of the models seems like a good fit to the data (Byrne, 1998).

In order to compare results across groups with confidence and rigor, it is necessary to first establish that an invariant relationship exists for each construct across the conditions pertinent to the investigation conducted (Pitts, West, & Tein, 1996). However, because measuring instruments are often group-specific in the way they operate (Byrne, 1998), baseline models, as the two- and four-factor models, are not going to be expected to be identical across the three groups of teachers, principals, and superintendents.
Kaplan (1955) wrote, “The practical implications of factorial invariance, as it applies to multiple group modeling, concerns the potential for explicitly testing whether selection mechanisms are present that militate against arguing for the causal effect of treatments” (p. 75). As it has been explained in this section, with the invariance approach, the researcher can constraint parameters or assign arbitrary values to be invariant across particular groups – such as, factor loadings, intercepts, or error variances – and estimate the equivalence of the relationships among variables and factors proposed by the model and their fit to the data (Phillips, 2000). Successfully addressing issues of comparability across samples is critical for SEM approaches (Dunn, Everitt, & Picles, 1993).

As a final comment for this section, Maruyama (1997) stated that multiple-sample comparisons offer a valuable extension of basic latent variable SEM approaches, but with the extension comes greater complexity. Therefore, “…there are important issues that will have to be resolved a priori about the most likely way in which to ensure comparability of processes across samples” (p. 265).
CHAPTER 4

RESULTS

Throughout this study it has been explained that measurements are on the same scale (i.e., comparable) when the empirical relations between the trait indicators (e.g., survey items) and the trait of interest are invariant across groups (Meredith, 1993). Hence, in this chapter information of the results found in this present study are presented, including: (a) best fitting model for each group, (b) factorial invariance across groups, and (c) comparison of parameter estimates.

Best Fitting Model for Each Group

The overall ($\chi^2$) and practical fit statistics ($\chi^2/df$, RMSEA, NNFI, and CFI) for the two-factor and four-factor models of leadership tested are shown in Table 1. Chi-square ($\chi^2$) values for each of the models were significant at the $p < 0.001$ level, with the exception of the principals’ two-factor model ($p > 0.01$) and the superintendents’ two-factor model ($p > 0.05$). Therefore, the influence of large sample size on the chi-square statistic used to assess absolute model fit made it necessary to rely on the comparative measures of “practical fit” rather than the statistical significance of the chi-square alone for decisions as to retain or reject each of the models (Widaman & Reise, 1997).
In the case of the teachers’ data, as shown in Table 1, the four-factor model is a slightly better fit of the data in comparison with the two-factor model. Although results for the teachers’ two-factor model showed NNFI and CFI values greater than 0.98, they also showed a significant $\chi^2 (p < 0.01)$, a $\chi^2/df$ ratio greater than 3, and a RMSEA greater than 0.05. As a contrast, although the four-factor model for teachers had a significant $\chi^2 (p < 0.01)$, results also showed a $\chi^2/df$ ratio of less than 2, a RMSEA smaller than 0.05, and NNFI and CFI values of 0.99 and greater. Therefore, those results support that the proposed four-factor model is the best representation of the teachers’ data from a “practical fit” point of view (Widaman & Reise, 1997).

The two-factor model is a better fit for the principals’ data than the four-factor model from a “practical fit” point of view, as indicated in Table 1 (Widaman & Reise, 1997). Although the two-factor model for principals had a significant $\chi^2 (p < 0.05)$, this value was better in comparison with the four-factor model, as well as the values reported for the $\chi^2/df$ ratio, and RMSEA. Furthermore, it was shown that principals’ data for both models had NNFI and CFI values greater than 0.90, but those values were higher for the two-factor model than for the four-factor model.

As reported in Table 1, all the fit indices support accepting that the proposed two-factor model is the best representation of the superintendents’ data from a “statistical and practical fit” point of view (Widaman & Reise, 1997). It is important to notice than an adequate representation of the data is accepted on the basis of a non-significant $\chi^2 (p > 0.05)$, $\chi^2/df$ ratio of less than 2, RMSEA of 0.051, as well as NNFI and CFI values of 0.96 and greater. Although results for the four-factor model
showed a NNFI and CFI values greater than 0.90, results also showed a significant $\chi^2$ ($p < 0.01$), $\chi^2/df$ ratio greater than 2, and RMSEA greater than 0.05.

Table 1

Fit Indexes for the Two-Factor and Four-Factor Models of Leadership Tested

<table>
<thead>
<tr>
<th>Model Comparison</th>
<th># Items</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$\chi^2/df$</th>
<th>RMSEA</th>
<th>NNFI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group 1: Teachers</strong>&lt;br&gt;Sample (n=364)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two-Factor Model</td>
<td>6</td>
<td>25.73**</td>
<td>8</td>
<td>3.22</td>
<td>0.078</td>
<td>0.981</td>
<td>0.990</td>
</tr>
<tr>
<td>Four-Factor Model</td>
<td>12</td>
<td>90.54**</td>
<td>48</td>
<td>1.89</td>
<td>0.049</td>
<td>0.990</td>
<td>0.992</td>
</tr>
<tr>
<td><strong>Group 2: Principals</strong>&lt;br&gt;Sample (n=419)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two-Factor Model</td>
<td>6</td>
<td>19.49*</td>
<td>8</td>
<td>2.44</td>
<td>0.059</td>
<td>0.979</td>
<td>0.989</td>
</tr>
<tr>
<td>Four-Factor Model</td>
<td>12</td>
<td>127.63**</td>
<td>48</td>
<td>2.66</td>
<td>0.064</td>
<td>0.971</td>
<td>0.979</td>
</tr>
<tr>
<td><strong>Group 3: Superintendents</strong>&lt;br&gt;Sample (n=369)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two-Factor Model</td>
<td>6</td>
<td>15.84</td>
<td>8</td>
<td>1.98</td>
<td>0.051</td>
<td>0.967</td>
<td>0.983</td>
</tr>
<tr>
<td>Four-Factor Model</td>
<td>12</td>
<td>126.96**</td>
<td>48</td>
<td>2.65</td>
<td>0.068</td>
<td>0.940</td>
<td>0.956</td>
</tr>
</tbody>
</table>

Note. RMSEA = root-mean square error of approximation; NFI = non-normed fit index; CFI = comparative fit index.
* $p < 0.05$. ** $p < 0.01$.

Browne and Cudeck (1993) explained that a value of the RMSEA of about 0.05 or less would indicate a close fit of the model in relation to the degrees of
freedom. Also, these authors explained that a value of about 0.08 or less for the RMSEA would indicate a reasonable error of approximation, but the researcher would not want to employ a model with a RMSEA greater than 0.1. Given the results reported in this study, Mels (Personal Communication, February 19, 2002) argued that since the RMSEA point estimates for both models indicate that both models provide a reasonable approximation to the data for all the groups, this is a motivating argument for testing for the invariance of both models across groups. Furthermore, Jöreskog (Personal Communication, February 19, 2002) also agreed that results indicate that the factorial invariance of the two-factor and four-factor models can be tested in this study across teachers, principals, and superintendents.

Factorial Invariance Across Groups

According to Meredith (1993) for scores to be comparable across groups, the observed variables must have identical, or invariant, relationships with the latent variable for each population of interest. Therefore, the researcher’s goal for this factorial invariance study was to determine whether the two-factor and four-factor models under investigation demonstrate sufficient similarity across the three groups of teachers, principals, and superintendents in order to support valid comparison of their scores.

The fit indices for the two-factor and four-factor models tested with the invariance hypotheses (configural, weak, strong, and strict) are reported in Table 2. All the chi-square values for each of the increasingly restricted two-factor and four-factor models were significant at the $p < 0.01$ level. As it was mentioned in the last
section, it is necessary to rely on what Widaman and Reise (1997) referred to as comparative measures of "practical fit" rather than the statistical significance of the chi-square alone for decisions as to retain or reject each of the increasingly restricted nested models and their related invariance hypothesis.

The Bentler and Bonett (1980) method of chi-square difference was used as an alternative to test the invariance hypotheses for the set of nested models. With this method the researcher was able to establish the extent to which factorial invariance exists using the strength of statistical testing to retain or reject the set of nested invariance hypotheses.

When restrictions are placed on one model to create another more constrained model, such as holding a parameter invariant, the more constrained model is said to be nested within the less constrained model (e.g., Bentler & Bonnett, 1980; Phillips, 2000). Reise, Widaman, and Pugh (1993) explained that under these conditions, the difference in chi-square values ($\Delta \chi^2$) for the nested model pair is distributed as a chi-square variate with degrees of freedom equal to the difference in degrees of freedom between the two models ($\Delta df$). Then, the $\Delta \chi^2$ value is used to test the statistical significance of the difference in fit between the nested models.

For each of the increasingly constrained models in this study (see Table 2), the extent to which factorial invariance holds was determined by testing whether constraining a given matrix (such as $\hat{\Lambda}$, $\hat{\tau}$, or $\hat{\Theta}$) to invariance across groups resulted in significant deterioration in model fit compared to that for the less constrained model from which it was constructed (Widaman & Reise, 1997).
Table 2
Fit Indexes for Alternative Measurement Models Across Groups

<table>
<thead>
<tr>
<th>Model Comparison</th>
<th>( \chi^2 )</th>
<th>( df )</th>
<th>( \chi^2/df )</th>
<th>RMSEA</th>
<th>NNFI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Configural Invariance</strong> (Patterns Invariant)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two-Factor Model</td>
<td>61.06</td>
<td>24</td>
<td>2.54</td>
<td>0.064</td>
<td>0.978</td>
<td>0.988</td>
</tr>
<tr>
<td>Four-Factor Model</td>
<td>345.13</td>
<td>144</td>
<td>2.40</td>
<td>0.061</td>
<td>0.975</td>
<td>0.982</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Weak Invariance</strong> (( \hat{\lambda} )s Invariant)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-Factor Model</td>
<td>76.19</td>
<td>32</td>
<td>2.38</td>
<td>0.060</td>
<td>0.981</td>
<td>0.986</td>
</tr>
<tr>
<td>Four-Factor Model</td>
<td>375.55</td>
<td>160</td>
<td>2.35</td>
<td>0.060</td>
<td>0.976</td>
<td>0.981</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Strong Invariance</strong> (( \hat{\lambda} )s and ( \hat{\tau} )s Invariant)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-Factor Model</td>
<td>497.60</td>
<td>44</td>
<td>11.31</td>
<td>0.165</td>
<td>0.856</td>
<td>0.859</td>
</tr>
<tr>
<td>Four-Factor Model</td>
<td>2753.40</td>
<td>184</td>
<td>14.96</td>
<td>0.193</td>
<td>0.754</td>
<td>0.772</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Strict Invariance</strong> (( \lambda )s, ( \hat{\tau} )s, and ( \hat{\Theta}_e )s Inv.)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-Factor Model</td>
<td>532.635</td>
<td>56</td>
<td>9.51</td>
<td>0.150</td>
<td>0.881</td>
<td>0.852</td>
</tr>
<tr>
<td>Four-Factor Model</td>
<td>2894.864</td>
<td>208</td>
<td>13.92</td>
<td>0.186</td>
<td>0.773</td>
<td>0.761</td>
</tr>
</tbody>
</table>

Note. All chi-square values were significant at the \( p < 0.01 \) level. RMSEA = root-mean square error of approximation; NFI = non-normed fit index; CFI = comparative fit index.
According to Phillips (2000), when the nested invariance restriction ($\Delta \chi^2$) did not result in degradation of model fit ($p > 0.05$), that invariance hypothesis was retained and that level of invariance, as defined by the parameter constraint tested, was said to hold. Furthermore, when $\Delta \chi^2$ did result in significant degradation of model fit ($p < 0.05$), the invariance hypothesis was rejected and the less constrained model was said to provide a significantly better fit to the data. The results that follow present the extent to which the nested set of factorial invariance hypotheses described in the previous chapter were upheld.

**Configural Factorial Invariance**

Configural invariance requires that the factor pattern matrices are equivalent across the three groups of teachers, principals, and superintendents. This means that for each group, the measured or observed variables (items) relate to the latent variables (factors) in the same general way. Specifically, the pattern of zero (an item does not load on a given factor) and non-zero (an item does load on a given factor) loadings should be the same across those three groups (see Tables 3 and 4).

In this study, those minimal conditions were tested when all matrices for the two-factor and four-factor models were freely estimated for each of the three groups studied, with the exclusion of those constraints imposed on the loading patterns across groups (configural invariance). These two-factor and four-factor baseline models then served as the starting point against which the fit of more restricted forms of invariance were compared.
### Table 3

**Hypothesized Pattern Matrix for Configural Invariance of the Two-Factor Model**

<table>
<thead>
<tr>
<th>Measured Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>People Orientation</td>
<td>Task Orientation</td>
</tr>
<tr>
<td>Development of Skills In Group Dynamics and Group Process</td>
<td>$\lambda_{1,1}$</td>
<td>0</td>
</tr>
<tr>
<td>Collaborative Group Decision Making</td>
<td>$\lambda_{2,1}$</td>
<td>0</td>
</tr>
<tr>
<td>Staff, Faculty, Student, and Community Input</td>
<td>$\lambda_{3,1}$</td>
<td>0</td>
</tr>
<tr>
<td>Implementation of District Policies and Procedures</td>
<td>0</td>
<td>$\lambda_{4,2}$</td>
</tr>
<tr>
<td>Use of School Law Knowledge to Plan and Intervene in Situations</td>
<td>0</td>
<td>$\lambda_{5,2}$</td>
</tr>
<tr>
<td>Planning and Management of School Budget</td>
<td>0</td>
<td>$\lambda_{6,2}$</td>
</tr>
</tbody>
</table>

As shown in Table 2, the data provide evidence that supports that the two-factor and four-factor models have an acceptable level of fit from a practical point of view (Widaman & Reise, 1997). Specifically, both models have a $\chi^2/df$ ratio smaller than 3, and both NNFI and CFI relative fit indices are greater than 0.90. Thus, although $\chi^2$ was statistically significant at $p < 0.01$, and the RMSEA was slightly greater than 0.05, both models can be said to fit the data reasonably well considering the large sample size for all three groups relative to the small number of variables included in the models. Therefore, these results support retaining the hypothesis of
<table>
<thead>
<tr>
<th>Measured Variable</th>
<th>Factor 1 Leader</th>
<th>Factor 2 Follower</th>
<th>Factor 3 Organization</th>
<th>Factor 4 Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Needs Assessment for Improvement and Intervention</td>
<td>$\lambda_{1,1}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mission Articulation to Staff and Community</td>
<td>$\lambda_{2,1}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Use of Different Leadership Styles as Needed</td>
<td>$\lambda_{3,1}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Development of Skills in Group Dynamics and Group Process</td>
<td>0</td>
<td>$\lambda_{4,2}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Collaborative Group Decision Making</td>
<td>0</td>
<td>$\lambda_{5,2}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Staff, Faculty, Student, and Community Input</td>
<td>0</td>
<td>$\lambda_{6,2}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Establishment of Effective School/Community Relationships</td>
<td>0</td>
<td>0</td>
<td>$\lambda_{7,3}$</td>
<td>0</td>
</tr>
<tr>
<td>Climate That Promotes Growth, Learning, and Excellence</td>
<td>0</td>
<td>0</td>
<td>$\lambda_{8,3}$</td>
<td>0</td>
</tr>
<tr>
<td>Use of Organizational Theories for Development and Change</td>
<td>0</td>
<td>0</td>
<td>$\lambda_{9,3}$</td>
<td>0</td>
</tr>
<tr>
<td>Implementation of District Policies and Procedures</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\lambda_{10,4}$</td>
</tr>
<tr>
<td>Use of School Law Knowledge to Plan and Intervene in Situations</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\lambda_{11,4}$</td>
</tr>
<tr>
<td>Planning and Management of School Budget</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\lambda_{12,4}$</td>
</tr>
</tbody>
</table>
configural invariance for the two-factor and four-factor models. Under these
conditions, the researcher can assert that similar, but not identical, latent variables
(factors) are present across the three groups of teachers, principals, and
superintendents.

**Weak Factorial Invariance**

Upon accepting the baseline hypothesis of configural invariance, the
hypothesis of weak factorial invariance was analyzed \( (H_0: \hat{\Lambda}_1 = \hat{\Lambda}_2 = \hat{\Lambda}_3) \). Under the
conditions of weak factorial invariance, the loading, or regression coefficient, for
each of the measured variables on their respective factor (latent variable) was
hypothesized to be equivalent across groups. To test this hypothesis, the baseline two-
factor and four-factor models were modified by imposing the constraint that the
\( \hat{\Lambda} \) matrix, or factor loading, be held invariant across groups.

As shown in Table 2 the chi-square for both modified models were
statistically significant \( (p < 0.01) \), which under the stringent conditions of absolute
measures of fit would lead to rejection of these models. However, both models have a
\( \chi^2/df \) ratio smaller than 3, RMSEA of 0.06, and both NNFI and CFI relative fit indices
are greater than 0.90. These results provide evidence that supports retaining the
hypothesis of weak invariance from a practical point of view (Widaman & Reise,
1997).

The differences in fit indices for alternative measurement models is shown in
Table 5. Using the \( \Delta \chi^2 \) nested model method to compare absolute and relative fit
between weak vs. configural models, the additional constraint of holding the \( \hat{\Lambda} \)
matrix to invariance across groups led to a statistically nonsignificant decrease in model fit.

As shown in Table 5, moving from configural to weak factorial invariance there was no evidence of fit deterioration for both modified two-factor and four-factor models, because \( \Delta \chi^2 \) were not statistically significant \((p > 0.05)\), and the \( \Delta \chi^2/df \) ratios were 1.89 and 1.90 (for the two-factor and four-factor models), which were even smaller than the \( \chi^2/df \) ratios of 2.54 and 2.40 respectively, for the two baseline models. This suggested that the additional constraints invoked on the more constrained models led to an even smaller lack of fit per degree of freedom as held for the baseline models, hence the more restricted models would be preferred (Widaman & Reise, 1997). In addition, RMSEA and NNFI exhibited some degree of improvement for the more restricted two-factor and four-factor models, while the CFI was only marginally worse. Consequently, the modified models appeared to be clearly preferable, because they represent a better fitting alternative than the baseline models. Thus, the hypothesis of weak factorial invariance across the three groups of teachers, principals, and superintendents was retained.

Widaman and Reise (1997) explained that by retaining this hypothesis, group differences in factor variances and covariances become identified in an ARF-invariant fashion. Therefore, any substantive interpretation of group differences in variances or covariances among factors will remain invariant over rescalings of the factors. Furthermore, according to several authors (Alwin & Jackson, 1981; MacCallum & Tucker, 1991; Reise, Widaman, & Pugh, 1993; Sörbom, 1974) the requirement for
meeting "full measurement invariance" was established, hence this hypothesis \( H_0: \)

\[
\hat{\Lambda}_1 = \hat{\Lambda}_2 = \hat{\Lambda}_3
\]

for two-factor and four-factor models holds in the population.

**Strong Factorial Invariance**

In addition to the constraint specified before (holding the \( \hat{\Lambda} \) matrix invariant),
the \( \hat{\tau} \) matrix, or item intercepts, were held invariant for the two-factor and four-factor
models and across the three groups of teachers, principals, and superintendents (\( \hat{\tau}_1 = \hat{\tau}_2 = \hat{\tau}_3 \)). As shown in Table 2 the chi-square values for both models were
statistically significant \( (p < 0.01) \), which under the stringent conditions of absolute
measures of fit would lead to rejection of these models. Moreover, for both models
the \( \chi^2/df \) ratios are greater than 3, RMSEA were immensely greater than 0.05, and
both NNFI and CFI relative fit indices were smaller than 0.90, providing evidence
that support rejection of the hypothesis of strong factorial invariance from a "practical
fit" point of view (Widaman & Reise, 1997).

As shown in Table 5, moving from weak to strong factorial invariance there
was evidence of significant fit deterioration, because \( \Delta \chi^2 \) was statistically significant
\( (p < 0.01) \). Therefore, constraining the \( \hat{\tau} \) matrix to invariance did result in a
statistically significant decrease in model fit from that obtained for constraining the
\( \hat{\Lambda} \) matrix alone.

There was a large decrease in the fit per degree of freedom difference \( \Delta \chi^2/df \)
for the more constrained models compared to the less constrained two-factor and
four-factor models. In addition, the RMSEA, NNFI, and CFI exhibited a very large
amount of deterioration. Therefore, the hypothesis of strong factorial invariance across the groups of teachers, principals, and superintendents was rejected (the less constrained model provides a significantly better fit to the data).

Table 5
Differences in Fit of Alternative Measurement Models Across Groups

<table>
<thead>
<tr>
<th>Model Comparison</th>
<th>$\Delta \chi^2$</th>
<th>$\Delta df$</th>
<th>$\Delta \chi^2/\Delta df$</th>
<th>RMSEA</th>
<th>NNFI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Configural vs. Weak</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two-Factor Model</td>
<td>15.12*</td>
<td>8</td>
<td>1.89</td>
<td>-0.003</td>
<td>0.003</td>
<td>-0.002</td>
</tr>
<tr>
<td>Four-Factor Model</td>
<td>30.42*</td>
<td>16</td>
<td>1.90</td>
<td>-0.001</td>
<td>0.001</td>
<td>-0.001</td>
</tr>
<tr>
<td><strong>Weak vs. Strong</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two-Factor Model</td>
<td>421.42**</td>
<td>12</td>
<td>35.12</td>
<td>0.105</td>
<td>-0.125</td>
<td>-0.127</td>
</tr>
<tr>
<td>Four-Factor Model</td>
<td>2377.86**</td>
<td>24</td>
<td>99.08</td>
<td>0.133</td>
<td>-0.222</td>
<td>-0.209</td>
</tr>
<tr>
<td><strong>Strong vs. Strict</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two-Factor Model</td>
<td>35.04**</td>
<td>12</td>
<td>2.92</td>
<td>-0.015</td>
<td>0.024</td>
<td>-0.007</td>
</tr>
<tr>
<td>Four-Factor Model</td>
<td>141.46**</td>
<td>24</td>
<td>5.89</td>
<td>-0.007</td>
<td>0.019</td>
<td>-0.011</td>
</tr>
</tbody>
</table>

Note. For all model comparisons, the second-listed model is more restricted than, and is nested within, the first-level model. For example, for the Configural vs. Weak comparison, Weak is nested within Configural. Given the way in which indices of practical fit were computed and interpreted, negative difference values of the RMSEA indicate better fit for more restricted model, whereas negative difference values of the NNFI and CFI indicate worse fit for the more restricted model. RMSEA = root-mean-square error of approximation; NNFI = non-normed fit index; CFI = comparative fit index.

*p > 0.05. **p < 0.001
These measures indicate that the strong factorial invariance model does not fit the data better than the less restricted weak factorial invariance model against which fit was compared. Hence, the data provide evidence that does not support the assertion that item means are invariant across the three groups of teachers, principals, and superintendents. Furthermore, it is important to notice that imposing across-groups constraints on the $\hat{f}$ matrices resulted in too great a cost in terms of model fit (Widaman & Reise, 1997). Therefore, the weak factorial invariance model provides the optimal CFA two-factor and four-factor models for our data.

**Strict Factorial Invariance**

Strong factorial invariance serves as a prerequisite to strict factorial invariance, which specifies that the measurement residuals (error variance) be equivalent across groups. Therefore, in addition to the constraints specified before, holding the $\hat{A}$ and $\hat{f}$ matrices invariant, the constraint was imposed that the $\hat{\Theta}_{e}$ matrices be held invariant across the three groups of teachers, superintendents, and principals ($\hat{\Theta}_{e1} = \hat{\Theta}_{e2} = \hat{\Theta}_{e3}$).

As shown in Table 2, for the two-factor and four-factor models the test of strict factorial invariance resulted in $\chi^2$ statistically significant ($p < 0.01$), which under the stringent conditions of absolute measures of fit leads to rejection of these models. Furthermore, both models had $\chi^2/df$ ratios greater than 3, RMSEA was immensely greater than 0.05, and both NNFI and CFI relative fit indices were smaller than 0.90, providing also evidence that supports rejection of the hypothesis of strict factorial invariance from a practical point (Widaman & Reise, 1997).
As shown in Table 5, the absolute and relative fit between strong and strict factor models was compared to determine the effect on model fit of constraining the $\Theta_e$ matrix to invariance across groups. Moving from strong to strict factorial invariance, there was evidence of significant fit deterioration, because the difference in statistical fit ($\Delta \chi^2$) between models was large ($p < 0.01$). Therefore, constraining the $\Theta_e$ matrix to invariance did result in a statistically significant decrease in model fit from that obtained for constraining the $\hat{\Lambda}$ and $f$ matrices alone.

The $\Delta \chi^2/df$ ratios for the change in chi-square were 2.92 and 5.89 (for the two-factor and four-factor models), which were larger than the $\chi^2/df$ ratios of 2.54 and 2.40 respectively, for the two baseline models. Although the RMSEA and CFI exhibited a slight degree of improvement, the NNFI showed deterioration in its measures. Thus, the hypothesis of strict factorial invariance across the four groups of teachers, principals and superintendents was rejected. The strict factorial invariance models do not fit the data better than the less restricted (strong factorial invariance) models against which fit was compared.

As in the case of this study, when strict factorial invariance does not hold, group differences on the measured variables are not entirely attributable to group differences on the latent variables. In addition, these data provided evidence to support the assertion that there are differences in the reliability of measures across groups (Phillips, 2000). The nature and extent of these differences will be presented in the following section on parameter estimates.
Comparison of Parameter Estimates

By observing the differences in absolute and relative fit for the two-factor and four-factor models tested across groups, it was found that the weak factorial invariance model best represented the data among the alternatives tested. In this type of factorial invariance the elements that comprise the factor loading (A) matrices were constrained to equality across groups, whereas the elements of the intercept (f), and error variance (Ω) matrices were free to vary across groups (Meredith, 1993; Widaman & Reise, 1997).

Invariance of the A matrix has established that the same latent variables, or factors, are identified in each of the three groups of teachers, principals, and superintendents. Alternatively, elements in the f and Ω matrices were found not to be invariant across those groups. Therefore, parameter estimates from the weak factorial invariance model were examined to detect the extent to which similarities and differences were evident across groups.

Measured Variable Factor Loadings

The WLS estimates completely standardized to a common metric for factor loadings across groups are shown in Tables 6 and 7. Given that the A matrix was found to be invariant (H₀: A₁=A₂=A₃) for the two-factor and four-factor models respectively, then, the values are equivalent across the three groups of teachers (GR1), principals (GR2), and superintendents (GR3).
<table>
<thead>
<tr>
<th>Measured Variable</th>
<th>Intercept ((\hat{\alpha}))</th>
<th>Loading ((\hat{\Lambda}))</th>
<th>Error Variance ((\hat{\Theta}_e))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GR1</td>
<td>GR2</td>
<td>GR3</td>
</tr>
<tr>
<td>Development of Skills In Group Dynamics and Group Process</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>Collaborative Group Decision Making</td>
<td>0.07</td>
<td>0.08</td>
<td>0.06</td>
</tr>
<tr>
<td>Staff, Faculty, Student, and Community Input</td>
<td>0.06</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>Implementation of District Policies and Procedures</td>
<td>0.06</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>Use of School Law Knowledge to Plan and Intervene in Situations</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>Planning and Management of School Budget</td>
<td>0.04</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Measured Variable</td>
<td>Intercept ($\hat{\alpha}$)</td>
<td>Loading ($\hat{\lambda}$) (Communality Estimate)</td>
<td>Error Variance ($\hat{\Theta}_e$) (Squared Multiple Correlation)</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>----------------------------</td>
<td>-------------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>GR1</td>
<td>GR2</td>
<td>GR 3</td>
</tr>
<tr>
<td>Needs Assessment for Improvement and Intervention</td>
<td>0.07</td>
<td>0.07</td>
<td>0.08</td>
</tr>
<tr>
<td>Mission Articulation to Staff and Community</td>
<td>0.06</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>Use of Different Leadership Styles as Needed</td>
<td>0.07</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>Development of Skills In Group Dynamics and Group Process</td>
<td>0.06</td>
<td>0.07</td>
<td>0.05</td>
</tr>
<tr>
<td>Collaborative Group Decision Making</td>
<td>0.08</td>
<td>0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>Staff, Faculty, Student, and Community Input</td>
<td>0.07</td>
<td>0.08</td>
<td>0.08</td>
</tr>
</tbody>
</table>
Table 7—Continued

WLS Estimates of Intercept, Factor Loading, Error Variance, and Squared Multiple Correlation for the Four-Factor Model

<table>
<thead>
<tr>
<th>Measured Variable</th>
<th>Intercept (( \hat{\beta} ))</th>
<th>Loading (( \hat{\lambda} )) (Communality Estimate)</th>
<th>Error Variance (( \hat{\Theta}_e )) (Squared Multiple Correlation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GR1</td>
<td>GR2</td>
<td>GR 3</td>
</tr>
<tr>
<td>Establishment of Effective School/Community Relationships</td>
<td>0.08</td>
<td>0.07</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Climate Which Promotes Growth, Learning, and Excellence</td>
<td>0.08</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of Organizational Theories for Development and Change</td>
<td>0.05</td>
<td>0.06</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementation of District Policies and Procedures</td>
<td>0.07</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of School Law Knowledge to Plan and Intervene in Situations</td>
<td>0.06</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planning and Management of School Budget</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Factor loadings for the $\hat{A}$ matrix are provided in Table 6 for the two-factor model. Following the criteria established by several authors (Comrey & Lee, 1992; Widaman & Reise, 1997), 67% of the measured variables in this study were shown to be excellent markers for their respective factors (People Orientation and Task Orientation), with loadings of 0.76 and greater (communality estimates of 57% and greater). Furthermore, 33% of the measured variables were found to be very good markers with loadings between 0.63 and 0.71 (communality estimates from 39% to 49%).

Factor loadings for the $\hat{A}$ matrix are provided in Table 7 for the four-factor model. Again, using the criteria established by several authors (Comrey & Lee, 1992; Widaman & Reise, 1997), 92% of the measured variables in this study were shown to be excellent markers for their respective factors (Leader, Follower, Organization, and Task), with loadings of 0.72 and greater (communality estimates of 50% and greater). Furthermore, 8% were found to be very good markers with loadings of 0.68 (communality estimates of 46%).

**Measured Variable Intercepts**

Given that the $\hat{c}$ matrix was not found to be invariant for the two-factor and four-factor models, the values shown in Tables 6 and 7, respectively, are not equivalent across the three groups of teachers, principals, and superintendents. In the two-factor model, common metric completely standardized estimates of the elements in the $\hat{c}$ matrix varied from 0.04 to 0.08 (see Table 6). The intercept estimates for the People Orientation factor ranged from 0.06 to 0.07 for the group of teachers, from
0.06 to 0.08 for the group of principals, and from 0.05 to 0.07 for the group of superintendents. Furthermore, the intercept estimates for the Task Orientation factor ranged from 0.04 to 0.06 for the group of teachers, from 0.05 to 0.06 for the group of principals, and from 0.06 to 0.07 for the group of superintendents.

In the four-factor model, elements in the $\hat{\psi}$ matrix varied from 0.05 to 0.09 (see Table 7). The intercept estimates for the Leader factor ranged from 0.06 to 0.07 for the group of teachers, from 0.06 to 0.07 for the group of principals, and from 0.07 to 0.08 for the group of superintendents. Also, the intercept estimates for the Follower factor ranged from 0.06 to 0.08 for the group of teachers, from 0.07 to 0.08 for the group of principals, and from 0.05 to 0.08 for the group of superintendents. Moreover, for the Organization factor, the intercept estimates ranged from 0.05 to 0.08 for the group of teachers, from 0.06 to 0.09 for the group of principals, and from 0.08 to 0.09 for the group of superintendents. Furthermore, the intercept estimates for the Task factor ranged from 0.05 to 0.07 for the group of teachers, from 0.05 to 0.06 for the group of principals, and from 0.06 to 0.07 for the group of superintendents.

**Measured Variable Error Variances**

As shown in Tables 6 and 7, for the two-factor and four-factor models, neither the error variance nor the squared multiple correlation (SMC) of measured variables were invariant across the three groups of teachers (GR1), principals (GR2), and superintendents (GR3). According to Phillips (2000) the SMC in CFA studies is similar to the communality estimate in traditional factor analytic approaches, and
serves to assess both the reliability and the proportion of variance of a measured variable accounted for by a given factor.

As shown in Table 6, in the two-factor model for the most part, the reciprocal relationship predicted by classical test theory — as error variance decreases, reliability increases and the converse — was not observed (Phillips, 2000; Widaman & Reise, 1997). Only for the first two measurement variables loading on the People Orientation factor (33%) the reliability decreased as the error variance increased moving from the group of teachers, to principals, and to superintendents. For the rest of the measurement variables loading on their respective factors (67%), the reliability decreased as the error variance decreased moving throughout these three groups studied.

In the four-factor model for the most part, the reciprocal relationship predicted by classical test theory — as error variance decreases, reliability increases and the converse — was not observed (see Table 7). Only for the first measurement variable of the Follower and Task factors (17%) the reliability decreased as the error variance increased moving throughout the three groups studied in this dissertation. For the rest of the measurement variables loading on their respective factors (83%), the reliability decreased as the error variance decreased moving from the group of teachers, to principals, and to superintendents.
CHAPTER 5

DISCUSSION AND CONCLUSION

A discussion of the findings with respect to the leadership construct studied in this dissertation is provided in this chapter. It is divided in the following sections: (a) findings, (b) limitations of the study and suggestions for further research, and (c) conclusion.

Findings

This dissertation was designed to study the extent to which SEM could be used to address some of the leadership challenges faced by educational researchers. In particular, regardless of the conditions under which the research was conducted or the groups surveyed, the researcher needs evidence that speaks to the substantive coherence and measurement equivalence of pertinent constructs (Phillips, 2000).

Two measurement models, the two-factor and four-factor models, were used in this study as a way to examine the measurement equivalence of the multi-dimensional construct of leadership across three specific groups. The methods presented here demonstrate the extent to which factorial invariance holds across the groups of teachers, principals, and superintendents. In addition, with these methods the researcher was able to make limited inferences (from a more robust measurement position) about the relationship between respondents’ administrative level (teacher,
principal, and superintendent) and the emergence of differences in the mental models held by them for the leadership construct.

**Best Fitting Model**

Several authors (Clinchy, 1995; Sarason, 1990; Shen, 1997) explained that in educational practice there is a tendency to accentuate certain aspects of educational leadership at the expense of others. Also, Shen (1997) explained that this practice of emphasizing certain goals at the expense of others might be justified in a context in which certain educational goals fail to receive enough emphasis. Obviously, this could lead to the emergence of different leadership perceptions held by groups depending on their administrative level (teachers, principals, and administrators).

Therefore, in this dissertation the researcher contrasted the two-factor and four-factor models of leadership across the groups of teachers, principals, and superintendents.

In this study, results show that the two-factor model best fit the data of principals and superintendents. Hence, the two-factor model is a better conceptual framework for developing instruments, accumulating data, communicating, and reporting findings pertaining to these two groups. The relative harmony of the two-factor model may have resulted because principals and superintendents have more years of experience and an administration-related understanding of the concept of educational leadership. Therefore, principals’ and superintendents’ conception on leadership are more pragmatic (due to their experience) than just theoretical, which help them arrive at the more parsimonious two-factor model of leadership.
In the best fitting model for teachers, results illustrate that the four-factor model was slightly better than the two-factor model. This suggests that the four elements distinguished in the teachers' mental model are: leader, follower, organization, and task. Hence, teachers hold mental models that place more emphasis on the four-factor model, in comparison with the principals' and superintendents' emphasis on the two-factor model. Teachers appear to have a sophisticated, yet idealistic understanding of the concept of leadership. Obviously, how to integrate the different leadership perceptions held by teachers, principals, and superintendents is a challenge in the field of education.

The differences of the leadership perceptions between teachers, on one hand, and principals and superintendents, on the other, could lead to problems and difficulties as far as school leadership is concerned. However, the mental model of leadership held by teachers is thought to be subject to change as they become familiar and more comfortable with new administrative practices (e.g., becoming a principal and then a superintendent). According to several authors (Knapp, 1997; Phillips, 2000; Spillane & Zeuli, 1999), as new practices are accepted and included, the ways in which teachers perceive the various aspects of their profession (e.g., leadership) may result in differences in the strength and direction of the relationships among those perceptions.

In general, findings in this dissertation suggest that differences in the administrative level (among other factors) are related to the conceptual clarity with which teachers, principals, and superintendents perceive the two-factor and four-factor models of leadership. Smithson and Porter (1994) explained the relationship
between training (i.e., professional development) and the accuracy of self-report behavioral data. Therefore, it was not surprising that the clarity as respondents perceive the leadership construct varies with the increase in their administrative level. Teachers — not yet exposed to higher administrative levels — were found to have a more sophisticated, yet idealistic mental model of leadership (four-factor model). On the other hand, principals and superintendents — exposed to higher administrative levels — appear to have come to express a more tightly focused and pragmatic mental model of leadership (two-factor model).

**Factorial Invariance**

Several authors (Alwin & Jackson, 1981; MacCallum & Tucker, 1991; Reise, Widaman, & Pugh, 1993; Sörbom, 1974) explained that measurement invariance in CFA models is established if factor loadings are invariant across groups. In other words, those authors argued that it is only necessary to retain the hypothesis of weak factorial invariance ($H_0: \lambda_1 = \lambda_2 = \lambda_3$) in order to meet the requirement of full measurement invariance, and assumed that the two-factor and four-factor models hold exactly in the population. Hence, weak factorial invariance is a sensible and sufficiently demanding test of psychometrically sound measurement invariance (Horn & McArdle, 1992; Reise, Widaman & Pugh, 1993).

Given the results reported in this study, Mels (Personal Communication, November 25, 2001) argued that configural and weak invariance across groups are supported by the data of teachers, principals, and superintendents, while strong and strict invariance across groups are not supported by these data. These findings are true.
for both the two-factor and four-factor models, and it seems that the major difference across groups is the difference between the intercepts (item means) across teachers, principals, and superintendents. Therefore, results provide support for the invariance of factor pattern and for the invariance of factor loadings across those three groups (see Table 8).

First, acceptance of the configural invariance hypothesis established that an equivalent factor pattern described the leadership construct across groups irrespective of their administrative level. In other words, the same configuration of loadings of items on factors is observed across groups for the two-factor and four-factor models. Thus, the way in which teachers, principals, and superintendents perceive leadership items (observed variables) relate to leadership factors was found to be equivalent across groups.

Second, acceptance of the weak factorial invariance hypothesis established that factor loadings were equivalent across teachers, principals, and superintendents (irrespective of their administrative level). Therefore, group differences in factor variances and covariances became identified in an ARF-invariant or appropriate rescaling factors fashion (Widaman & Reise, 1997). Although tests of weak factorial invariance were accepted, this really only meant that the differences in factor loading (\( \lambda \)) matrices present across groups were not sufficiently large to detect statistically significant differences, and not that there was no difference.

Third, rejection of the strong factorial invariance hypotheses established that the intercept (\( \tau \)) matrices were not equivalent across teachers, principals, and superintendents. Therefore, there was evidence to suggest that the entire linear model
Table 8
Comparison of the Four Types of Factorial Invariance

<table>
<thead>
<tr>
<th>Factorial Invariance</th>
<th>Constraints in the CFA Model</th>
<th>Interpretation of the Analysis</th>
<th>Appropriate Rescaling Factors</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configural Factorial Invariance</td>
<td>The same pattern of fixed and free loadings in the $\Lambda$ matrix is constrained to equality across groups.</td>
<td>The items load on the same factors across groups (not identical loadings). Only the number of factors and salient items are hypothesized to be the same across groups. The rest of the parameters are free to vary across groups.</td>
<td>Group differences in factor variances and covariances become identified in an ARF-invariant fashion (remain invariant over rescalings of the factors). Same rescaling of parameter estimates for samples.</td>
<td>(e.g., Horn, McArdle &amp; Mason, 1983; Taris, et al., 1998; Widaman &amp; Reise, 1997).</td>
</tr>
<tr>
<td>Weak Factorial Invariance</td>
<td>The elements that comprise the factor loading ($\lambda$) matrices are constrained to equality across groups.</td>
<td>The requirement for meeting &quot;full measurement invariance&quot; is established. The loading for each of the observed variables on their respective factor is hypothesized to be equivalent across groups (the groups weight the items the same).</td>
<td>Group differences in factor variances and covariances become identified in an ARF-invariant fashion (remain invariant over rescalings of the factors). Same rescaling of parameter estimates for samples.</td>
<td>(e.g., Alwin &amp; Jackson, 1981; MacCallum &amp; Tucker, 1991; Meredith, 1993; Widaman &amp; Reise, 1997).</td>
</tr>
<tr>
<td>Strong Factorial Invariance</td>
<td>The elements that comprise the $\Lambda$ matrices, and the observed variable intercepts (item means) that comprise the $\tilde{r}$ matrices are constrained to equality across groups.</td>
<td>Minimum condition required for substantive interpretation of group mean differences on the factors. Group differences in both means and variance on the factors are reflected in group differences in means and variances on the observed variables.</td>
<td>Group differences in means and variances on the factors are identified in an ARF-invariant fashion, as are the covariances (and correlations) among the factors.</td>
<td>(e.g., Alwin &amp; Jackson, 1981; MacCallum &amp; Tucker, 1991; Meredith, 1993; Widaman &amp; Reise, 1997).</td>
</tr>
<tr>
<td>Strict Factorial Invariance</td>
<td>The elements that comprise the $\tilde{\Theta}$ matrices (measurement error or error variance for items) are constrained to equality across groups, in addition to the previous constraints.</td>
<td>Group differences in means and variances on the observed variables are a function only of group differences in means and variances on the &quot;common factors&quot;. Furthermore, the reliability of the measures (items) is equivalent across groups.</td>
<td>Group differences in error variances are identified in an ARF-invariant fashion, in addition to the previous statement of strong factorial invariance.</td>
<td>(e.g., Meredith, 1993; Phillips, 2000; Taris, et al., 1998; Widaman &amp; Reise, 1997).</td>
</tr>
</tbody>
</table>

Note. Shaded row denotes the maximum factorial invariance condition met in this study.
that describes the relationship among factors to a given observed variable, in terms of both the regression weight (loading) and the intercept term is not invariant across conditions compared (Meredith, 1993; Widaman & Reise, 1997). Furthermore, the interpretation of group differences, other than those with respect to variances/covariances, is limited (Widaman & Reise, 1997).

Fourth, rejection of the strict factorial invariance hypotheses established that the error variance for items, or the diagonal elements of the error (\( \Theta \)) matrices, was not equivalent across groups. Therefore, because this hypothesis was rejected, there was evidence to suggest that respondents may not be equally well able to understand and respond to the survey items across groups. Moreover, results show that error variance and item reliability, as measured by squared multiple correlation may depend, to some extent, on the group situation (Taris et al., 1998). It is important to notice that this hypothesis is not often met with most data sets, and it is not required for substantive interpretation of group differences (Widaman & Reise, 1997).

Kalliath (1999) explained that it should be obvious that configural factorial invariance, weak factorial invariance, strong factorial invariance, and strict factorial invariance are idealizations. Also, their validity and existence in the real world of psychological measurement and research can never be finally established in practice; however, they are very useful idealizations in their application to educational leadership theory. Furthermore, Horn, McArdle and Mason (1983) argued that the more complex the phenomena we must create to make sense — the less we know — the more reasonable it is to expect that the research will disclose only very imperfect indications of invariance. Therefore, with methods that indicate only subtle, "weak
evidence" of invariance, researchers may be provided with more interesting information (indicative of things not previously understood), and results that more accurately represent the true complexity of the leadership construct.

Given that the data were found to support the condition of weak factorial invariance across groups, it is reasonable to assert that exposure to administrative levels (teacher, principal, superintendent) in and of itself was not associated with alterations in factor loadings for the mental model of leadership. However, these data do not provide evidence of invariance in the intercepts and error variances. Furthermore, interpretation of these results suggest that reporting of scores on leadership without addressing issues of factorial invariance in effect can bury significant information about the undercurrents of the leadership construct and limit the valid use of those leadership scores.

**Application of Findings**

With the results of this dissertation, the researcher showed an example of how simple and direct ways of testing crucial hypotheses related to factorial invariance can be provided with SEM (Widaman & Reise, 1997). SEM is used to focus on the processes that give rise to the theoretical phenomena — leadership construct — while an unique analysis is provided that simultaneously assesses the quality of measurement and examines predictive relationships among constructs (e.g., Kaplan & Elliot, 1997; Kelloway, 1998; Muthen, 1994; Willet & Sayer, 1996). Hence, with the two- and four-factor models, an adaptable and powerful set of tools for investigating similarities and differences on leadership perceptions across groups is provided.
Hsieh and Shen (1998) stated that “it appears that conceptions of leadership change when teachers are promoted to principalship and principals to superintendents; instructional matters are de-emphasized and managerial and political affairs are emphasized as a person is promoted to a higher position in the school system” (p. 118). Evidently, this may explain why results in this dissertation show that teachers’ perception of leadership is more sophisticated, yet idealistic, whereas principals’ and superintendents’ perception of leadership is more parsimonious and pragmatic.

Phillips (2000) wrote, “An abstract construct may evolve to mean different things to respondents over time, especially if the intervention being evaluated included sessions intended to increase the respondent’s understanding of the concept” (p. 25). Also, as people assume new responsibilities they tend to add new practices to their existing repertoire of traditional methods (Knapp, 1997; Spillane, 1994). Moreover, Schmitt (1982) indicated that experience in a work environment could systematically shift or transform response patterns in ways that alter the meaning of work-related concepts (as the leadership construct) over time. Therefore, Hsieh and Shen (1998) wrote that “if we understand the similarity and differences among teachers’, principals’ and superintendents’ conceptions of leadership we can explain the difficulties and, therefore, allow the reform agenda to proceed smoothly” (p. 108).

Systemic reform requires teachers, principals, and superintendents to revise their leadership roles and responsibilities in order to acquire the necessary practices to perform successfully in their profession. Therefore, there are important implications of the findings in this dissertation for teachers, principals, and superintendents.
According to Rahim and Magner (1996) opportunities should be given to people to acquire more education, training, and job experience as this can enhance the knowledge and perceptions of teachers, principals, and superintendents of the leadership construct.

The current national and state systemic reform policy agenda requires the advance of the educational environment in order to have an optimal impact on student achievement (Darling-Hammond, 1997). Also, Phillips (2000) explained that the impact of current policies on student achievement is limited by the extent to which teachers, and the institutions where they work, have the capacity to contribute toward the evolution of classroom practice. Hence, results in this dissertation contribute to the knowledge on leadership and offer information in support to understanding the leadership trends sought by the school reform movement. Furthermore, Hsieh and Shen (1998) explained that “to associate leadership position in a system with perspective on leadership might be one of the approaches to integrating perspectives on leadership” (p. 107).

Limitations of the Study and Suggestions for Further Research

According to Widaman and Reise (1997), even if one detects differences between groups in crucial CFA model parameters, the CFA models do not indicate why these differences occur. Many psychological, cultural, or economic factors may contribute to the group differences observed. Therefore, the CFA models presented in this dissertation are not ends in themselves. Instead, the two- and four-factor models may be used to isolate the ways in which groups differ on their leadership
perceptions, providing a concise statistical representation of group differences. Furthermore, this can serve as a facilitator for additional research designed to identify the sources of group differences on the latent and measured variables on leadership.

An important limitation of SEM and its application to leadership is that it simply identifies clusters of variables. Attempts to determine why these clusters of variables exist are speculative and should be viewed as hypotheses subject to additional verification (Crocker & Algina, 1986). Also, the results of the research presented in this dissertation with the particular groups of teachers, principals, and superintendents may differ from the results based on different types of samples. Furthermore, Kaplan (2000) explained, “A problem with estimation methods that explicitly address non-normality is the reliance on very large sample sizes or unrealistic small models” (p. 85). Therefore, another limitation of this dissertation was the need to use sample sizes of 300 and higher, due to the fact that this study uses ordered categories (ordinal variables), specifically Likert-type scaled (Jöreskog & Sörbom, 1996).

According to Jöreskog and Sörbom (1989) there are many sources of error other than just sampling error, that may have major effects on inferences drawn from the application of SEM, and could affect the results of its analysis. Therefore, researchers have to be aware of several types of error that can produce different results than those expected by the researcher. Some of those many types of errors are: (a) nonresponse error, when there is failure to collect data on all persons in the sample, (b) measurement error, errors in recorded responses due to several factors (e.g., respondents’ inability to answer questions, wording of survey questionnaires),
(c) specification errors in the model, when there are incorrect distributional assumptions, and (d) technical errors, when there are incorrect procedures in the analysis of data (Groves, 1989).

Fit indices reported in this dissertation should not be regarded as measures of usefulness of a model (Bollen & Long, 1993). This is because fit indices contain some information about the lack of fit of a model, but none about plausibility. Consequently, they should not be used in a mechanical decision process for selecting a model, because this has to be a subjective process involving the use of judgment. Furthermore, it is evident that the model that ultimately best fit the respondents' data from among those tested was only one of the many models that could have been specified for teachers, principals, and superintendents. Therefore, Jöreskog (1993) cautioned users of CFA to avoid the situation of failing to recognize that they have not tested all the models, just a select few. In noting the subjective nature of this type of study, Phillips (2000) stated that it is important "to alert the reader to the possibility that other configurations and analyses may have provided alternative scenarios and explanations of the phenomenon at hand" (pp. 92-93).

Several methodological limitations exist for both the execution and interpretation of the secondary analysis presented in this study. It is important to observe that these methodological limitations can help further research in this area to be pursued. For this reason, some of the suggestions for further research can be summarized as: (a) comparative study on different respondents, (b) longitudinal study on the same respondents, (c) alternative measurement models, (d) partial metric invariance, and (e) ordinal vs. continuous data.
Comparative Study on Different Respondents

It would be an excellent additional line of investigation to use CFA in order to determine the invariance of the factorial structure of leadership across educational institutions and gender (Hox, 1994; Muthen, 1991). In addition, variability in outcomes across institutions due to the influence of leadership perspectives (e.g., formal, political, collegial, cultural) has been never examined nor utilized to assess the plausibility of the leadership assumptions (Bush, 1995). Therefore, this may serve as another possibility for future research on leadership.

The groups studied in this dissertation, because of their professional background, might have an understanding of leadership different from that of members of groups such as the general public, parents, or students at various stages. It would be interesting to explore the factorial validity of the leadership construct by using data collected from those other groups. Results of this type of study would further the understanding of the factorial validity and provide information useful to explain phenomena related to the leadership construct.

Longitudinal Study on the Same Respondents

McMillan (1996) explained that in longitudinal studies the same group of subjects is studied over a specified length of time. Moreover, data are collected at different times, usually over several years. The advantage of this type of study is that the limitations of cross-sectional designs are avoided. Collins (1991) argued that the costs of survey designs are often very high when they include repeated measures. However, this may be the best approach for the study of leadership over time.
If it is chosen not to follow the same respondents over time, then it may be difficult to study the on-going leadership perceptions of teachers, principals, and superintendents. The interpretation of any results from a non-experimental approach such as this must take the often-cited threats to validity into account (Mehrens & Lehmann, 1991; Trochim, 1998). In particular, the researcher must assume that the groups of teachers, principals, and superintendents were comparable prior to the study of their leadership perceptions, and that social threats such as diffusion of treatment influence were minimal (Phillips, 2000). Obviously, this also presents an interesting and useful line of further research.

**Alternative Measurement Models**

In the present study, two measurement models of leadership were tested, but to obtain further evidence of construct validity, the subscales should be examined in the context of structural models involving variables that have been associated with the bases of leadership in prior theoretical and empirical work. Therefore, data collected in previous studies may be reanalyzed with SEM to investigate the construct validity on leadership. Also, different items or indicators from the present instrument studies can be selected, and this study repeated in order to compare results. Kelloway (1998) explained that “although the three-indicator rule is perhaps the most commonly cited, the empirical evidence supports the use of two indicators for each latent variable when the sample size is large” (p. 63).

There are many alternative models that could have been used in order to fit the data equally well or better than the two-factor and four-factor models assessed in this
study. It is important to be aware that another model, with a decidedly different form, could fit the data even better than did any of the models considered in this study (Widaman & Reise, 1997). Reise, Widaman, and Pugh (1993) stated that many experts on CFA modeling believe that modification indices are dangerous, because they allow mere post hoc modification of models without theoretical justification. However, given that model modifications need to be replicated across samples to be supported, those modification indices provide an efficient way for respecifying models and remolding theories. This is an interesting idea for future research, but it is important to notice that when a model is modified empirically rather than theoretically, cross-validation for assuring that the statistical theory is not violated becomes essential (Bentler & Bonett, 1980). Furthermore, Meredith (1993) explained that the results of simultaneous model fitting are clearly informative no matter which case holds. Hence, the models are meaningful and acceptable if the results fit into a substantive theoretical framework.

Reise, Widaman, and Pugh (1993) stated that CFA and item response theory (IRT) models provide interesting ways of representing data in the social and behavioral sciences. According to Mehrens and Lehmann (1996) IRT is the process of examining the responses to each survey item in order to judge the quality of the item. Therefore, future investigators may further explore the relations between CFA and IRT leadership models and their differential utility for representing data and testing theoretical hypothesis. Whereas CFA models account for the covariance between test items, IRT models account for examinee item responses (Lord, 1980).
Hence, the outcome of this work would be a more rational framework for the use of current methods of psychometric analysis on the leadership construct.

**Partial Metric Invariance**

Future studies may test other types of restrictive forms of factorial invariance, by using the partial metric invariance methods proposed by Byrne, Shavelson, and Muthen (1989). Under partial metric factorial invariance, only some of the parameter estimates in the $\hat{\Lambda}$, $\hat{\tau}$, and $\hat{\Theta}_e$ matrices may be constrained to invariance across groups, and the remaining estimates may vary freely across groups. Therefore, when invariance constraints are imposed across groups on a particular matrix, not necessarily all parameters in the given matrix need to be constrained to invariance across groups (Byrne, Shavelson, & Muthen, 1989; Reise et al., 1993).

In this study, only one of the matrices ($\hat{\Lambda}$) was invariant across the three samples of teachers, principals, and superintendents. Evidently, results supported retaining the hypothesis of weak factorial invariance for the two-factor and four-factor models. However, the other matrices ($\hat{\tau}$, and $\hat{\Theta}_e$) had values that varied across samples. Therefore, the use of partial matrix invariance methods would lead to a model with partial metric invariance of the $\hat{\tau}$, and $\hat{\Theta}_e$ matrices, along with full metric invariance of the $\hat{\Lambda}$ matrix.

It is important to note that the notion of partial metric factorial invariance is open to some disagreement in the literature. Although partial metric factorial invariance models may be specified, and group differences in means and variances on
the latent variables may be identified in an invariant fashion, many experts on SEM (Byrne et al., 1989; Reise et al., 1993; Widaman & Reise, 1997) disputed that partial metric factorial invariance models are not viable models for demonstrating that true ARF-invariant latent variables are identified. Therefore, this clearly is a topic that will require further attention, and it may present new opportunities to strengthen the knowledge of the leadership construct.

**Ordinal vs. Continuous Data**

As it was explained in previous chapters, the data used in this study were obtained from a six-point Likert scale. Therefore, the data are considered to be ordinal (Jöreskog et al., 2000). CFA and the related SEM techniques are based on stringent assumptions of linearity, and Jöreskog (1993) asserted that rating scale data may provide only weak support for such assumptions. This could account for some degree of model misfit in the two- and four-factor models. Therefore, stronger conclusions may be obtained in future analyses if the data, on which the analyses are based, have stronger (i.e., continuous measurement scales) measurement properties (Widaman & Reise, 1997).

Many authors made convincing arguments as to why the practice of assuming and underlying interval scale is acceptable without the laborious modifications suggested by Jöreskog (1993) and implemented in this dissertation. For example, ordinal variables with underlying continuous latent-variable attributes have been used with tetrachoric or polychoric correlation in several studies (e.g., Muthen, 1984; Muthen & Kaplan, 1985; Jöreskog, 1994). Therefore, a future study would be to
determine the impact of the methodological decision of analyzing ordinal data by relying on the assumption of underlying continuous variables (or unconverted ordinal scale data), rather than to use the fixed threshold method, as suggested by Jöreskog (1993), to treat ordinal data. Clearly, it would be very appealing to compare the conclusions for converted vs. unconverted ordinal data, and to find if different results can be drawn from these analyses.

Conclusion

In this study the importance of determining and reporting the extent to which teachers, principals, and superintendents share the same perceptions for the leadership construct under investigation was affirmed. The researcher drew on leadership theory to develop the basic structure of the two-factor and four-factor models, then LISREL 8 was used to estimate the parameters describing the relationships between the observed variables and the factors proposed in these models. Therefore, this study offers an empirically tested guide to study the extent to which survey respondents relate observed measures to factors on leadership in the same way across the groups of teachers, principals, and superintendents. Furthermore, this study presents a helpful point of view for addressing many of the measurement challenges inherent in the research on educational leadership.

According to Shen (1998) “The empirical findings on the similarities and differences in teachers’, principals’ and superintendents’ conceptions of leadership have implications for explaining the leadership phenomenon and educating future school leaders” (p. 7). To study whether teachers, principals, and superintendents
hold similar perceptions of leadership can help improve the progress of our school leaders. Also, it can help improve the organizational culture and morale by reducing their conflict due to different leadership perceptions on how schools should operate (Carlson, 1996; Trice, 1991). Furthermore, these results can have important implications for the selection of future educational leaders who share a congruent mental model of leadership with colleagues in the school system.

The CFA techniques used in this dissertation help to better appreciate the complex nature of the leadership process. Respondents’ attitudes toward leadership are related to how they perceive and interpret mental models for this construct. Therefore, evidence to support the equivalence of measurement structures across teachers, principals, and superintendents serves to strengthen the research on educational leadership. It is important to note that if the groups of interest do not share the same mental model for the construct under investigation, then the use of composite scores to compare the level and/or relationship among constructs across groups should not be supported (Meredith, 1993).

Research on the leadership construct provides new insights every time it is investigated in particular settings or with particular methodologies, and the potential for achieving new insights through the application of SEM is vast (Kalliath, 1999). The general support obtained in this dissertation for the leadership construct provides one example of this potential. Tests of factorial invariance were justified by both strong theoretical argument and large sample sizes. Moreover, as suggested by Widaman and Reise (1997), this study justified on theoretical grounds why the respondents were divided into groups in a particular way, and how those groups
differed with regard to the measurement models. Obviously, no CFA model should be accepted or rejected on statistical grounds alone; theory and judgment should play a key role in defending the adequacy of any estimated CFA model (Reise, Widaman, & Pugh, 1993).

Hopefully, this dissertation will provide a valuable learning opportunity for researchers to gain new knowledge on the leadership mental model held by teachers, principals, and superintendents. According to Widaman and Reise (1997) measurement models provide stronger tests of our measurement theories. Clearly, researchers are encouraged to use the CFA models discussed in this study to evaluate the factorial invariance of their measures more stringently.

With this dissertation, the researcher has clarified issues in testing leadership factorial invariance in ways that illustrate the interesting questions to be asked of data as well as many possibilities that need to be addressed by educational researchers. Hopefully, this dissertation will be a useful source for future inquiry into testing the invariance of relations among leadership measures. Obviously, the appropriateness of a common-factor model of leadership, and even the number of factors comprising it can never be assessed definitely (Kim & Mueller, 1978). However, the use of structural equation techniques can increase our confidence that the model is consistent with the true population parameters (Kalliat, Bluedorn, & Gillespie, 1999).
APPENDIX A

Concept of Leadership According to Several Authors
Concept of Leadership According to Several Authors

History indicates that by a long period of time, leadership has been an interest matter for many theoretical, investigators and professionals, producing several definitions, or concepts. By effect of this review of literature it will be considered convenient to mention some of them, where it can be seen the theoretical direction taken by these authors.

Leadership according to Block

“The strength in the concept of leadership is that it connotes initiative and responsibility.” (Block, 1993, p. 13). Moreover, Block (1993) writes “The attraction of the idea of leadership is that it includes a vision of the future, some transforming quality that we yearn for... Leaders bring spirit, even integrity, into play.” (p. 14).

Leadership according to Boles and Davenport

Leadership is a process of social interaction or exchange, in which each person gives certain things for something in return (Boles & Davenport, 1983).

Leadership according to Burns

Burns (1978) defines leadership as the process of “Inducing followers to act for certain goals that represent the values and the motivations—the wants and needs, the aspirations and expectations—of both leaders and followers” (p. 19). Moreover, Burns (1978) states that leadership “… is exercised when persons with certain purposes mobilize, in competition or in conflict with others, institutional, political,
psychological and other resources so as to arouse and satisfy the motives of followers.” (p. 18).

**Leadership according to Carlson**

“Leadership is very dynamic, highly interactive, and, depending on the circumstances, may be shared by different people at different times; it may be viewed as a total of experiences that work to move ideas and people.” (Carlson, 1996, p. 127).

**Leadership according to Daft**

“Leadership occurs between people, involves the use of influence, and is used to attain goals.” (Daft, 1994, p. 478).

**Leadership according to Davis**

Leadership is the ability to make things happen by encouraging and channeling the contribution of others, taking a stand on addressing important issues, and acting as a catalyst for change and continuous improvement (Davis, 1996).

**Leadership according to Deal and Kennedy**

Leadership is the process of stimulating, developing, and working with people within an organization (Deal & Kennedy, 1982).

**Leadership according to DePree**

Leadership is much more an art, a belief, a condition of the heart, than a set of things to do. The leader is the servant of the followers in that he removes the obstacles that prevent them from doing their jobs. The true leader enables his or her followers to realize their full potential (DePree, 1989).
Leadership according to Etzioni

"Ability, based on the personal qualities of the leader, to achieve on the part of the followers the voluntary fulfillment of a wide variety of matters." (Etzioni, 1965, p. 688).

Leadership according to Gibson, Ivancevich and Donnelly

Leadership is an interactional phenomenon arising when group formation takes place (Gibson, Ivancevich & Donnelly, 1997).

Leadership according to Hemphill and Coons

Leadership is the behavior of an individual when he is directing the activities of a group toward a shared goal (Hemphill & Coons, 1950).

Leadership according to Hersey, Blanchard, and Johnson

Leadership is any attempt to influence the behavior of another individual or group (Hersey, 1996).

Leadership according to Hollander

A leadership process involves a two-way influence relationship aimed primarily at attaining mutual goals (Hollander, 2000).

Leadership according to Jessamon

Leadership is an interaction between persons in which one presents information of a sort and in such a manner that the other becomes convinced that his outcomes will be improved if he behaves in the manner suggested of desired (Jessamon, 1973).
Leadership according to Kahn

Leadership is the process or act of influencing. A leader has the authority to decide what should happen and who should do it, the responsibility to make it happen and the accountability for what does actually happen (Kahn, 1979).

Leadership according to Kerzner

“Leadership can be defined as a style of behavior designed to integrate both the organizational requirements and one’s personal interests into the pursuit of some objective.” (Kerzner, 1997, p. 253).

Leadership according to Koontz & O’Donnell

“...leadership is defined generally as a influence, that is to say, the art or process of influencing on the persons so that attempt with good will and enthusiasm the achievement of the goals of the organization.” (Koontz, & O’Donnell, 1973, p.665).

Leadership according to Melinkoff

Leadership is the art of getting others to want to do something you are convinced should be done (Melinkoff, 1986).

Leadership according to Niehouse

Leadership is a strategic skill that can be defined as the process of attempting to influence the behavior of one or more persons toward reaching a goal or accomplishing a task (Niehouse, 1988).
Leadership according to Owens

"Leadership is not something that one does to people, nor is it a manner of behaving toward people: it is working with and through other people to achieve organizational goals." (Owens, 1998, p. 206).

Leadership according to Owens and Steinhoff

Leadership may be viewed as a process in which others are influenced to act to achieve goals in a specific situation (Owens & Steinhoff, 1976).

Leadership according to Pareles

"Capacity to direct and guide the attitudes and actions of individual groups in function of some objectives, they will be formal institutional, or informal institutional, positive or negative." (Pareles, 1985, p. 43).

Leadership according to Penland

Leadership is the process of influencing others to achieve mutually agreed upon purposes for the organization (Penland, 1974).

Leadership according to Peters and Waterman

Leadership is the capacity to influence and organize meaning for the members of the organization (Peters & Waterman, 1982).

Leadership according to Powell

Leadership is the process of influencing the activities of an organized group toward goal attainment (Powell, 1988).
Leadership according to Robbins

“Leadership is the capacity of influencing a group for the fulfillment of the goals.” (Robbins, 1987, p.244).

Leadership according to Rosenberg

Leadership is the ability to influence, and shape values, beliefs, and behaviors consistent with increased follower commitment to the mission of the organization (Rosenberg, 1989).

Leadership according to Tannenbaum

“Leadership is the interpersonal influence exercised in situation and guided through communication process toward the achievement of a goal or specific goals.” (Tannenbaum, 1971, p.29).

Leadership according to Tannenbaum, Weschler, and Massarik

Leadership is an interpersonal influence exercised in a situation and directed toward the attainment of a specialized goal or goals (Tannenbaum, Weschler, & Massarik, 1959).

Leadership according to Terrence and Peterson

Leadership is the activity of influencing people to strive willingly for group objectives. (Terrence & Peterson, 1994).

Leadership according to Yukl

Leadership is the exercise of influence resulting in enthusiastic commitment by followers (Yukl, 1994).
As a final comment on leadership, Burns (1978) summarizes, "To elevate the goals of humankind, to achieve high moral purpose, to realize major intended change, leaders must thrust themselves into the most intractable processes and structures of history and ultimately master them" (p. 421).
APPENDIX B

General Information on Leadership
General Information on Leadership

Leadership Behavior
(Ohio State Studies)
(Michigan Studies)
(Leadership Grid of Blake & Mouton)
(Tannenbaum-Schmidt Continuum)
(Group Dynamics Studies)

Consideration = Leader is sensitive to subordinates, respects their ideas, and feelings, mutual trust, friendly, open communication, teamwork, oriented to subordinate's welfare.

Employee-centered leaders = Leader establishes high performance goals and displays supportive behavior toward subordinates.

Concern for People = The highest of this scale (Country Club Management), gives thoughtful attention to the needs of people because satisfying relationships leads to comfortable, friendly organization atmosphere.

Democratic = Similar to the relationship oriented definitions explained above.

Group Maintenance = The leader maintains and strengthens the group itself.

Initiating structure = Leader is task oriented, directs work activities toward goal attainment. Gives Instructions, spends time planning, deadlines, schedules.

Job-centered leaders = Leader in favor of meeting schedules, keeping costs low, and achieving production efficiency.

Concern for Production = The highest of this scale (Authority-Compliance), says that efficiency in operations results from arranging conditions of work in such a way that human elements interfere to a minimum degree.

Authoritarian = Similar to the task oriented behaviors explained above.

Goal achievement = The leader clarifies the task and develops a procedural plan.

Contingency Approaches
Model that describes the relationship between leadership styles and specific organizational situations

Fiedler's theory
Hersey and Blanchard's situational theory
The path-goal model
Substitutes for Leadership

In general mean:

Leader's style
Subordinates nature
Situation characteristics

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Fiedler's Contingency Theory
Effort to combine leadership style and organizational situation into a theory of leadership (it matches leader's style with the situation most favorable for his/her success)

Leadership Style

Task-oriented leader = Motivated by task achievement, as in the initiating structure style above.

Organizational Situation

Leader-member relations = (Good or Poor) Group atmosphere, and members' attitude toward and acceptance of the leader. Trust, respect, confidence in the leader. Good, means favorable to the leader.

Task structure = (High or Low) Extent to which tasks performed by the group are defined, involve specific procedures, and have clear explicit goals. When task structure is high, the situation is considered favorable to the leader (ill-defined tasks, have a low degree of task structure).

Position power = (Strong or Weak) Extent to which the leader has formal authority over subordinates. It is high when leader has power to plan and direct work, and reward or punish others. Strong, means favorable.

Hersey and Blanchard's Situational Theory
(Four styles: telling, selling, participating, and delegating) A contingency approach to leadership that links the leader's two-dimensional style with the task maturity of followers. It incorporates only characteristics of follower, not those of the situation (leaders must carefully diagnose maturity level of followers)

Leader Style

Relationship Behavior = Already explained above.

Task Behavior = Already explained above.

Task Maturity of Followers

Followers vary from immature level to mature level. e.g., job related experience, ability, training, skills, confidence, and willingness to work and achieve. People is low in task maturity when they have little ability or training, or insecurity, and are unable to take responsibility for their own task behavior.

Niehouse's Leadership Approach

Leadership Style

Interactive Behavior = Participation between leader and subordinate in the decision-making process.

Directive Behavior = Leader delegates responsibility for a task, and monitors the results

Subordinate's Readiness

Psychological Readiness = Willingness, motivation, self-confidence, and commitment to do whatever needs to be done.

Job Readiness = Job skills, competence, expertise of a subordinate to fulfill an assigned task.
Path-goal theory
Leader's responsibility is to increase subordinate's motivation to attain goals, by clarifying the behaviors necessary for task accomplishment and rewards.

Leader Behavior (Not ingrained personality traits, rather every leader is able to adopt depending on situation)

Supportive = Leader behavior that shows concern for subordinates' well-being and personal needs. Behavior is open, friendly and approachable, treats subordinates as equals. Similar to consideration leadership.

Directive = Leader tells subordinates exactly what they are supposed to do. Includes planning, making schedules, setting performance goals, adherence to rules and regulations. Similar to initiating structure leadership.

Participative = Leader consults with his or her subordinates about decisions. Behavior includes asking for opinions and suggestions, encouraging participation in decision-making, and meeting with subordinates in their workplaces. Encourage group discussion and written suggestions

Achievement-oriented = Leader sets clear and challenging objectives for subordinates. Leader behavior stresses high-quality performance and improvement over current performance. Confidence in subordinates and assist them in learning how to achieve high goals.

Situational Contingencies

Personal characteristics of group members = Similar to Hersey and Blanchard's maturity level and include such factors as ability, skills, needs, and motivations (e.g., if employee has low level of ability, the leader may need to provide additional coaching to improve performance).

The work environment = Include the degree of task structure, the nature of the formal authority system, and the work group itself. Task structure is similar to the same concept described in Fiedler's Contingency Theory. Includes the extent to which tasks are defined and have explicit work procedures. The formal authority system includes the amount of legitimate power used by managers, and extent to which rules constrain employee's behavior. Work group characteristics are the educational level of subordinates and the quality of relationships among them.

Available Rewards = Leader clarifies the path to rewards already available for subordinates

Use of Rewards

New Rewards = Leader develops new rewards to meet the specific needs of a subordinate.
Stinson-Johnson Model
Leadership behavior style—task & relationship—depends on:

- **Task Structure**
  - **Low task structure** = Tasks are not clearly defined, and do not involve specific procedures.
  - **High task structure** = Tasks are clearly defined, and involve specific procedures.

- **Follower Capacity**
  - **High degree** = of achievement motivation, need for independence, and task relevant education and experience.
  - **Low degree** = of achievement motivation, need for independence, and task relevant education and experience.

Leadership Competencies or Skills (Hersey, Blanchard, and Johnson, 1996)

- **Diagnosing**
  - Cognitive or cerebral competency, where the leader understands what the situation is now and what is expected in the future

- **Adapting**
  - Behavioral competency that involves altering the leader's behavior and other resources available to meet the contingencies of the situation

- **Communicating**
  - Process competency that involves interacting with others in a way that people understand and accept

Substitutes for Leadership
It suggests that situational variables can be so powerful that they actually substitute for or neutralize the need for leadership. This approach outlines those organizational settings in which a leadership style is unimportant or unnecessary. (e.g., highly professional subordinates who know how to do their tasks do not need a leader who tells them what to do, because they do not need much direction or consideration)

- **Organizational variables** → Group cohesiveness, formalization, inflexibility, low positional power, physical separation.
- **Task Characteristics** → Highly structured task, automatic feedback, intrinsic satisfaction
- **Group Characteristics** → Professionalism, training/experience, low value of rewards.
Iowa State University (Kurt Lewin and Associates)

**Autocratic Leader**
Tends to centralize authority, rely on legitimate reward or coercive power to manage subordinates.

**Democratic Leader**
Delegates authority to others, encourages participation, and relies on expert and referent power to manage subordinates.

**Sources of leadership Influence or Power**

- **Legitimate Power**
  Stems from a formal management position and authority granted

- **Reward Power**
  From leader's authority to reward others

- **Coercive Power**
  From leader's authority to punish

- **Expert Power**
  From leader's special knowledge or skill in tasks performed by subordinates

- **Referent Power**
  Results from respect and admiration, and desire to emulate the leader.

**Followers Reactions to Each Source of Power**

- **Commitment by Followers**
  Workers share leader's point of view and enthusiastically carry out instructions (expert and referent power most likely generate commitment)

- **Compliance by Followers**
  Workers will obey orders but may personally disagree and not necessarily be enthusiastic (legitimate and reward power most likely generate follower compliance)

- **Resistance by Followers**
  Workers will deliberately try to avoid carrying out instructions and attempt to disobey orders (coercive power)

**Personal Characteristics of Leaders**

- **Physical characteristics**
  Activity, Energy

- **Social Background**
  Mobility

- **Intelligence & Ability**
  Judgment, decisiveness, knowledge & fluency of speech.

- **Personality**
  Alertness, originality, creativity, personal integrity, ethical conduct, self-confidence.

- **Work-Related Charac.**
  Achievement drive, desire to excel, drive for responsibility, responsibility in pursuit of objectives, task orientation.

- **Social Characteristics**
  Ability to enlist cooperation, cooperativeness, popularity, prestige, sociability, interpersonal skills, social participation, tact, diplomacy.

**Bass and Avolio's Approach to Leadership**

- **Transformational Leadership**
  Transformational leaders behave in ways to achieve superior results by employing one or more of the Four I's: idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration.

- **Transactional Leadership**
  Transactional leadership occurs when the leader rewards the follower depending on performance. It is divided in: contingent reward, and the active and passive forms of management-by-exception.

- **Laissez-Faire**
  The Laissez-Faire style is the avoidance or absence on leadership, and it is the most inactive, and ineffective style.
APPENDIX C

Leadership Patterns Found in the Literature
Leadership Patterns Found in the Literature

**Leader's Behavior Style**
Adapted from:
- Blake & Mouton (1985)
- Cartwright & Zander (1960)
- Fiedler (1967)
- Group Dynamic Studies
- Hersey, Blanchard, & Johnson (1996)
- Likert (1961)
- Michigan Studies
- Niehouse (1998)
- Ohio State Studies
- Path-goal theory
- Tannenbaum & Schmidt (1973)

**People or Relationship Orientation**
The leader is sensitive to subordinates, respects their ideas, and feelings, mutual trust, friendly, open communication, teamwork, oriented to subordinate's welfare. Supportive and open behavior toward subordinates. Attention to the needs of people, friendly organization atmosphere.

**Production or Task Orientation**
The leader is task oriented, direct work activities toward goal attainment. Give Instructions, spend time planning, deadlines, schedules. In favor of meeting schedules, keeping costs low, and achieving production efficiency. Efficiency in operations results from arranging conditions of work in such a way that human elements interfere to a minimum degree.

**Diagnosing**
Cognitive or cerebral competency, where the leader understands what the situation is now and what is expected in the future

**Adapting**
Behavioral competency that involves altering the leader's behavior and other resources available to meet the contingencies of the situation.

**Communicating**
Process competency that involves interacting with others in a way that people understand and accept.

**Physical characteristics**
Such as activity, and energy of the leader.

**Social Background**
Mobility of the leader

**Intelligence & Ability**
Judgment, decisiveness, knowledge and fluency of speech of the leader

**Personality**
Alertness, creativity, personal integrity, ethical conduct, self-confidence, flexibility, credibility, persistent, innovative thinking, initiative, and enthusiasm, analytical, emotional, extroverted, opportunistic, dependable, independent, and methodical.

**Work-Related Characteristics**
Conflict resolution skills, technical skills, expertise, planning skills, organizational skills, entrepreneurial skills, administrative skills, management support building skills, and resource allocation skills.

**Social Characteristics**
Ability to enlist cooperation, cooperativeness, popularity, prestige, sociability, interpersonal skills, social participation, tact, diplomacy, communication skills, persuasiveness, and inspiration.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Follower's Style
Adapted from:
• Bass & Avolio (1994)
• Evans (1970)
• Kelley (1992)
• Owens (1998)

Exemplary Followers
Independent and critical thinkers, separate from the leader or the group. They take initiative, go beyond the job, think for themselves, give constructive criticism, and are willing to stand up to leaders.

Alienated Followers
Followers that think independently and critically, but they are not very active in taking initiative. They see themselves as victims and sarcastically criticize the leader's efforts, holding back own effort.

Conformist Followers
The opposite of alienated followers, conformists score high on the active engagement scale but low on independent thinking. They find comfort in structure and in having someone above them.

Pragmatist Followers
Followers that moderately think independently and critically, and moderately take initiative. They manipulate others and the organization to their benefit. They emerge when organization is unstable.

Passive Followers
Followers that in a low degree think independently and critically, and lowly take initiative. They are lazy, incompetent, and unmotivated. They take action only when the boss gives instructions (mindless)

Include such factors as productivity, satisfaction, autonomy, participation, needs, and motivations, ability to communicate clearly, job experience, innovative thinking, initiative and enthusiasm, discipline, and control techniques.

Psychological Readiness
Willingness to take responsibility, achievement motivation, self-confidence, and commitment to do what needs to be done

Job Readiness
Job knowledge, competence, expertise of a subordinate to fulfill an assigned task

Commitment by Followers
Followers share leader's point of view and enthusiastically carry out instructions (expert and referent power most likely generate commitment).

Compliance by Followers
Followers will obey orders but may personally disagree and not necessarily be enthusiastic (legitimate and reward power most likely generate follower compliance)

Resistance by Followers
Followers will deliberately try to avoid carrying out Instructions and attempt to disobey orders (coercive power)
Organization

Organizational Subsystems
Adapted from:
• Carlson (1996)
• Hersey, Blanchard, & Johnson (1996)
• Kerzner (1997)
• Owens (1998)
• Path-Goal Theory

Sources of Influence or Power
Adapted from:
• Carlson (1996)
• Daft (1994)
• Fiedler (1967)
• Hersey, Blanchard, & Johnson (1996)
• Kerzner (1997)
• Owens (1998)
• Sergiovanni & Starratt (1998)

Indicators of Organizational Health
Adapted from:
• Daft (1994)
• Carlson (1996)
• Hersey, Blanchard, & Johnson (1996)
• Owens (1998)
• Sergiovanni & Starratt (1998)

External Environment
Adapted from:
• Kerzner (1997)
• Hersey, Blanchard & Johnson (1996)

Administrative
Authority, organizational structure, size, and responsibility within the organization.

Economic/Technological
The work to be done and the cost effectiveness of that work within the specific organizational goals.

Information/Decision Making
Emphasizes key decisions and their informational needs to keep the system working.

Human/Social
Motivation and needs of the members of the organization and on the leadership provided or required.

Legitimate Power
Stems from a formal management position and authority granted

Reward Power
From leader's authority to reward others

Coercive Power
From leader's authority to punish

Expert Power
From leader's special knowledge or skill in tasks performed by subordinates

Referent Power
Results from respect and admiration, and desire to emulate the leader.

Goal Focus
Extent to which people in the organization understand and accept goals.

Communication Adequacy
Vertical, horizontal, internal and external communication, and its easiness.

Innovativeness
Tendency to devise new procedures and goals, to grow, and develop over time.

Morale
This is exhibited as feelings of well-being and satisfaction.

Adaptation or Change
Organizations should be able to change, correct, and adapt to the environment.

Conflict Management Adequacy
Includes mechanisms for sensing and solving problems with minimum strain (e.g., avoidance, domination, accommodation, compromise, and collaboration).

Remote Environment
Includes economic, social, political, technological, and ecological factors.

Industry Environment
Includes barriers to entry into the industry, and factors such as competition, and suppliers.

Operating Environment
Includes such factors as competitors, creditors, customers, labor force, and suppliers.
Task Structure or Complexity
Adapted from:
• Bass & Avolio (1994)
• Fiedler (1967)
• Kerzner (1997)
• Path-Goal Theory
• Stinson-Johnson Model

Clear or Structured Task
It has specific instructions on what the leader and their followers should do. Tasks performed by the group are defined, involve specific procedures, and have explicit goals.

Ambiguous or Unstructured Task
It has no prescribed operating procedures, and it cannot be delineated specifically. Tasks performed by the group are not clearly defined, they do not involve specific procedures, nor have explicit goals.

Skill Variety
The number of diverse activities, and the number of skills used to perform a task.

Task Identity
Degree to which an employee performs a task with a recognizable beginning and ending.

Task Significance
Degree to which the task is perceived as important and having impact on the company or consumers.

Autonomy
Degree to which the employee has freedom, discretion, and self-determination in planning and carrying out tasks.

Feedback
Extent to which doing the task provides information back to the employee about his or her performance.

Additive Task
Task in which individual contributions of members are added together.

Conjunctive Task
Require that each member carry out the entire assignment parallel with other members.

Disjunctive Task
Require a team choice among alternatives: for example, whether to work overtime.

Task Demands
Adapted from:
• Bass & Avolio (1994)
• Kerzner (1997)
• Lewis (1999)

Task Obstacles
Adapted from:
• Daft (1994)
• Kerzner (1997)

Some of the task obstacles are: unstable economy, shortages, resources (e.g., cost, time), increased complexity or size of the task, technological changes, societal concerns, consumerism, ecology, and quality of work.

Core Dimensions
Adapted from:
• Daft (1994)
• Hackman & Oldham (1980)

Skill Variety
The number of diverse activities, and the number of skills used to perform a task.

Task Identity
Degree to which an employee performs a task with a recognizable beginning and ending.

Task Significance
Degree to which the task is perceived as important and having impact on the company or consumers.

Autonomy
Degree to which the employee has freedom, discretion, and self-determination in planning and carrying out tasks.

Feedback
Extent to which doing the task provides information back to the employee about his or her performance.

Additive Task
Task in which individual contributions of members are added together.

Conjunctive Task
Require that each member carry out the entire assignment parallel with other members.

Disjunctive Task
Require a team choice among alternatives: for example, whether to work overtime.

Task Demands
Adapted from:
• Bass & Avolio (1994)
• Kerzner (1997)
• Lewis (1999)

Task Obstacles
Adapted from:
• Daft (1994)
• Kerzner (1997)

Some of the task obstacles are: unstable economy, shortages, resources (e.g., cost, time), increased complexity or size of the task, technological changes, societal concerns, consumerism, ecology, and quality of work.
APPENDIX D

Instrument Duplication Permission
April 15, 2002.

To Whom It May Concern:

Liliana Rodriguez Campos has permission to: (1) use the partial data set requested; (2) reproduce The Educational Leadership Survey to attach to HSIRB application, and (3) reproduce The Educational Leadership Survey as an appendix, with the understanding that I will be appropriately referenced.

Sincerely,

[Signature]

Dr. Jianping Shen
Professor
APPENDIX E

Human Subjects Institutional Review Board Permission
Date: July 18, 2001

To: Jianping Shen, Principal Investigator
    Liliana Rodriguez-Campos, Student Investigator for dissertation

From: Mary Lagerwey, Chair

Re: HSIRB Project Number 01-06-20

This letter will serve as confirmation that your research project entitled "The Structure of the Leadership Construct: A Test of Factorial Invariance Using Structural Equation Modeling" has been approved under the exempt category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note that you may only conduct this research exactly in the form it was approved. You must seek specific board approval for any changes in this project. You must also seek reapproval if the project extends beyond the termination date noted below. In addition if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: July 18, 2002
APPENDIX F

The Educational Leadership Survey
October 22, 1997

The Educational Leadership Survey

This survey is administered to students enrolled in the EDLD program and graduates of the Department of Educational Leadership who currently work in K-12 settings. Please respond to each of the following questions. The information will be used to examine factors related to the recruitment and retention of educational administrators. This study is voluntary. Responses will remain anonymous. Thank you for your assistance.

### Demographic Information

<table>
<thead>
<tr>
<th>Gender</th>
<th>Marital Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Male</td>
<td>A. Married</td>
</tr>
<tr>
<td>B. Female</td>
<td>B. Single</td>
</tr>
<tr>
<td></td>
<td>C. Divorced</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Asian</td>
<td>A. 21-26</td>
</tr>
<tr>
<td>B. African American</td>
<td>27-30</td>
</tr>
<tr>
<td>C. Hispanic/Latino</td>
<td>31-35</td>
</tr>
<tr>
<td>D. Native American</td>
<td>36-40</td>
</tr>
<tr>
<td>E. White</td>
<td>E. 41-45</td>
</tr>
<tr>
<td>F. Other</td>
<td>F. 46-55</td>
</tr>
<tr>
<td></td>
<td>G. 56 and above</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of years in teaching/education</th>
<th>I am currently employed as</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. 0-1</td>
<td>A. Teacher in public school</td>
</tr>
<tr>
<td>B. 2-4</td>
<td>B. Teacher in private school</td>
</tr>
<tr>
<td>C. 5-7</td>
<td>C. Teacher in charter school</td>
</tr>
<tr>
<td>D. 8-10</td>
<td>D. Administrator in public school</td>
</tr>
<tr>
<td>E. 11-15</td>
<td>E. Administrator in private school</td>
</tr>
<tr>
<td>F. 16-20</td>
<td>F. Administrator in charter school</td>
</tr>
<tr>
<td>G. Above 20</td>
<td>G. I am working outside of education</td>
</tr>
<tr>
<td></td>
<td>H. Other (Please specify)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>If you are a teacher or administrator, at which building level do you work?</th>
<th>Type of School District</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Elementary (K-6 or similar configuration)</td>
<td>A. Urban</td>
</tr>
<tr>
<td>B. Middle School/Junior High</td>
<td>B. Suburban</td>
</tr>
<tr>
<td>C. High School (9-12)</td>
<td>C. Rural/Small District</td>
</tr>
<tr>
<td>D. Other (please specify)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Approximate student population in your school district</th>
<th>If I plan to seek a position in educational administration within</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. less than 300</td>
<td>A. 1 year</td>
</tr>
<tr>
<td>B. 301-750</td>
<td>B. 2-3 years</td>
</tr>
<tr>
<td>C. 751-1200</td>
<td>C. 4-5 years</td>
</tr>
<tr>
<td>D. 1201-1600</td>
<td>D. 6-8 years</td>
</tr>
<tr>
<td>E. 1601-2000</td>
<td>E. No plans to seek a position in educational administration</td>
</tr>
<tr>
<td>F. 2001-3000</td>
<td>F. I am an administrator</td>
</tr>
<tr>
<td>G. 3001-4000</td>
<td></td>
</tr>
<tr>
<td>H. 4001-7000</td>
<td></td>
</tr>
</tbody>
</table>

Turn over page - go to Page 2.
The level at which I currently work or intend to seek an administrative position is
A. Elementary (K-6 or similar configuration)
B. Middle School/Junior High
C. High School (9-12)
D. Other

Position I have obtained or will most likely seek for my first administrative position
A. Assistant Principal
B. Principal
C. Curriculum Director
D. Business Manager
E. Personnel Director
F. Other

Highest position I intend to seek
A. Assistant principal
B. Principal
C. Curriculum Director
D. Business Manager
E. Personnel Director
F. Assistant Superintendent
G. Superintendent
H. Other (please specify)

Degree I am presently seeking
A. M.A., M. Ed.
B. Ed.S. (Educational Specialist)
C. Ed.D. or Ph.D.
D. Other

The highest academic degree that I hope to attain is
A. M.A., M. Ed.
B. Ed.S. (Educational Specialist)
C. Ed.D. or Ph.D.

Major assigned teaching area
A. Art
B. Business Ed.
C. English
D. Home Economics
E. Industrial Ed.
F. Math
G. Music
H. Physical Ed.
I. Science
J. Social Studies
K. Elementary
L. Other

Where are you now in your preparation program?
A. I am just getting into the program
B. I am about halfway through the program
C. I am three quarters of the way through the program
D. I am near completion of the program

Satisfaction with my present position would be described as
A. Very satisfied
B. Satisfied
C. Dissatisfied
D. Extremely dissatisfied

Factors Influencing Applying for and Resigning from an Administrative Position

Please respond to each statement below in both columns by circling the number that best relates to your perspective. The first column focuses on factors that might affect your applying for an administrative position. In the second column, please consider reasons that might cause you to leave an administrative position.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Applying for Administrative Position</th>
<th>Leaving an Administrative Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>19. Location of district</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>20. Size of the district</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
</tbody>
</table>

Go to Page 3.
<table>
<thead>
<tr>
<th>Statement</th>
<th>Not Important</th>
<th>Neutral</th>
<th>Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reputation of the district</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Length of the administrative contract (e.g., 1-3 years)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Time required for supervising extra-curricular activities</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Finding a suitable position for my spouse</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Wealth and prosperity of the district</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Diversity of the job responsibilities</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>A position in an urban area</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>A position in a suburban area</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Proximity of a school district to a metropolitan area</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Salary must be commensurate with responsibilities</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Administrator’s daily rate (salary) significantly exceeds that of a teacher</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Length of contract days (e.g., 220-260 days)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Relationship between the board, administration, and teachers</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Reputation of the superintendent</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Encouragement from my professors about the reputation of the district</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Only seeking position in current district</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Nature of the work - e.g., relationship with parents and students</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Political nature of administration</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Stress of the position</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Community support</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Impact of administrative position on my home life</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Ability to make a difference in the lives of students</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Recognition for accomplishing important goals</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Autonomy in a leadership position</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Status within the educational community</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Professional advancement</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>To work with adults</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Encouragement by my family and others</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Quality of life in the community (housing, cultural activities, recreation, etc.)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Poor working conditions (paperwork, long hours, little time, and freedom, etc.)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Emotional aspects (stress, boredom, frustration, burnout, lack of fulfillment, etc.)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Lack of respect for educators</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>No chance for advancement</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Labor related issues</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Turn over page - go to Page 4.
<table>
<thead>
<tr>
<th>Statement</th>
<th>Least Important</th>
<th>Most Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of work day</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Waiting until children are grown</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>School board micro-management in district</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Personal safety</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Lack of support for administrators</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>To be a change agent (e.g., to develop new programs and procedures)</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
</tbody>
</table>

### Reasons for Enrollment in Educational Leadership Courses

<table>
<thead>
<tr>
<th>Reason you have enrolled in educational leadership courses</th>
<th>1 2 3 4 5 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>To learn more about leadership</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>To provide options in case I change my mind about becoming an administrator</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>The courses are interesting</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>My friends or colleagues took EDLD courses</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>The courses are easier than in other departments</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Only to satisfy requirements for Master's degree</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>To satisfy licensing requirements</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Pay raise associated with completion of course work</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Other (Please specify)</td>
<td></td>
</tr>
</tbody>
</table>

### Benefits of the Educational Leadership Program

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Least Important</th>
<th>Most Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>To understand schooling as enculturation into a democracy</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>To understand the importance of continuing school renewal</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>To understand issues concerning education as a real profession</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>To develop knowledge and skills in group dynamics and group process</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>To use the principles of school law in the development and implementation of policy and procedure</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>To have a working knowledge of general curriculum in terms of scope and sequence</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>To establish effective school/community relations</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>To maintain a school climate which promotes growth, learning, and excellence among staff and students</td>
<td>1 2 3 4 5 6</td>
<td>1 2 3 4 5 6</td>
</tr>
</tbody>
</table>

Go to Page 5.
<table>
<thead>
<tr>
<th>Question</th>
<th>Least Helpful</th>
<th>Most Helpful</th>
</tr>
</thead>
<tbody>
<tr>
<td>To what extent have you been helped through courses taken in the</td>
<td></td>
<td></td>
</tr>
<tr>
<td>educational leadership program?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To manage a school promoting fair and humane discipline and control</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>techniques</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To use different leadership styles as needed</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To foster collaborative group decision-making</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To encourage input and evaluation from staff, faculty, students, and</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>the community about school programs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To use theories of organizational development and change process to</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>implement change</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To establish a schedule for staff and faculty within a school</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To plan and integrate special activities</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To select staff and faculty appropriate to school and community needs</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To establish a system for staff and faculty supervision and evaluation</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>for the purposes of school improvement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To collaboratively plan and implement a staff development program</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To assess needs for school improvement and plan appropriate intervention</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To implement district policies and procedures</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To use knowledge of school law to plan and intervene in student/faculty/parent situations</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To plan and manage a school budget</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To report school financial matters to community and faculty</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To develop and conduct staff development workshops or other learning activities for teachers</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To give teachers timely and constructive feedback for instructional improvement</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To manage special programs within the school (e.g., guidance,</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>special education, gifted/talented, ESL, multi-ethnic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To articulate the mission of the school to staff and community</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To analyze, synthesize, and report data on student achievement to</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>school and community</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To analyze a teaching episode</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>To diagnose and prescribe an intervention strategy for improving the</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>instructional practices of the teacher</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Academy Questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Worst Time</th>
<th>Best Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>I would be willing to attend a three-day academy workshop.</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>Nominated by my superintendent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Received graduate credit for participation</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>Conference and registration fees paid by school district</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>I receive release time from teaching/current responsibilities</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>The month and time most likely to increase chances of my participation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>June</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>July</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>August</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>Weekend during school year</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>Weekdays during the school year</td>
<td>1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>I am willing to participate in the academy at anytime.</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

End of Survey
Thank You for Your Participation
APPENDIX G

Desirable Knowledge/Skill Base for Educational Leaders Survey
You are invited to participate in a research project entitled "The Structure of the Leadership Construct: A Test of Factorial Invariance Using Structural Equation Modeling" designed to study people's perception about leadership, being conducted by Dr. Jianping Shen and Liliana Rodriguez-Campos from Western Michigan University, Department of Teaching, Learning, and Leadership. This research is being conducted as part of the dissertation requirements for Liliana Rodriguez-Campos. If you are currently a school teacher, you are invited to answer the 30-item questionnaire.

This survey is comprised of 30 items and will take approximately 5 minutes to complete. Your replies will be completely anonymous, so do not put your name anywhere on the form. You may choose to not answer any question and simply leave it blank. If you choose to not participate in this survey, you may either return the blank survey or you may discard it in the box provided. Returning the survey indicates your consent for use of the answers you supply. If you have any questions, you may contact Dr. Jianping Shen at 616-387-3887, Liliana Rodriguez-Campos at 616-344-6091, the Human Subjects Institutional Review Board (616-387-8293) or the vice president for research (616-387-8298).

This consent document has been approved for use for one year by the Human Subjects Institutional Review Board as indicated by the stamped date and signature of the board chair in the upper right corner. You should not participate in this project if the corner does not have a stamped date and signature.
Desirable Knowledge/Skill Base for Educational Leaders

This survey is designed to recognize what type of leadership components are observed in each statement written below. The leadership components are (a) leader, (b) follower, (c) organization, and (d) task.

Please, feel free to circle one or more options.

1. To understand schooling as enculturation into a democracy
   □ Leader □ Follower □ Organization □ Task

2. To understand the importance of continuing school renewal
   □ Leader □ Follower □ Organization □ Task

3. To understand issues concerning education as a real profession
   □ Leader □ Follower □ Organization □ Task

4. To develop knowledge and skills in group dynamics and group process
   □ Leader □ Follower □ Organization □ Task

5. To use the principles of school law in operations and policy development
   □ Leader □ Follower □ Organization □ Task

6. To have a working knowledge of general curriculum in terms of scope and sequence
   □ Leader □ Follower □ Organization □ Task

7. To establish effective school/community relations
   □ Leader □ Follower □ Organization □ Task

8. To maintain a school climate which promotes growth, learning, and excellence
   □ Leader □ Follower □ Organization □ Task

9. To manage a school promoting fair and humane discipline and control techniques
   □ Leader □ Follower □ Organization □ Task

10. To use different leadership styles as needed
    □ Leader □ Follower □ Organization □ Task

11. To foster collaborative group decision-making
    □ Leader □ Follower □ Organization □ Task

12. To encourage staff, faculty, student, and community input on school effectiveness
    □ Leader □ Follower □ Organization □ Task

13. To use theories of organizational development and to implement change
    □ Leader □ Follower □ Organization □ Task

14. To establish a schedule for staff and faculty within a school
    □ Leader □ Follower □ Organization □ Task
15. To plan and integrate special activities
   □ Leader □ Follower □ Organization □ Task

16. To select staff and faculty appropriate to school and community needs
   □ Leader □ Follower □ Organization □ Task

17. To establish and implement a system for staff and faculty supervision and evaluation
   □ Leader □ Follower □ Organization □ Task

18. To collaboratively plan and implement a staff development program
   □ Leader □ Follower □ Organization □ Task

19. To assess needs for school improvement and plan appropriate intervention
   □ Leader □ Follower □ Organization □ Task

20. To implement district policies and procedures
   □ Leader □ Follower □ Organization □ Task

21. To use knowledge of school law to plan and intervene in school situations
   □ Leader □ Follower □ Organization □ Task

22. To plan and manage a school budget
   □ Leader □ Follower □ Organization □ Task

23. To report school financial matters to community and faculty
   □ Leader □ Follower □ Organization □ Task

24. To develop and conduct staff development workshops or other learning activities
   □ Leader □ Follower □ Organization □ Task

25. To give teachers timely and constructive feedback for instructional improvement
   □ Leader □ Follower □ Organization □ Task

26. To manage special programs within the school (e.g., guidance, special education, GT)
   □ Leader □ Follower □ Organization □ Task

27. To articulate the mission of the school to staff and community
   □ Leader □ Follower □ Organization □ Task

28. To analyze, synthesize, and report data on student achievement to school community
   □ Leader □ Follower □ Organization □ Task

29. To analyze a teaching episode
   □ Leader □ Follower □ Organization □ Task

30. To diagnose and prescribe an intervention strategy for improving teaching
   □ Leader □ Follower □ Organization □ Task
BIBLIOGRAPHY


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


