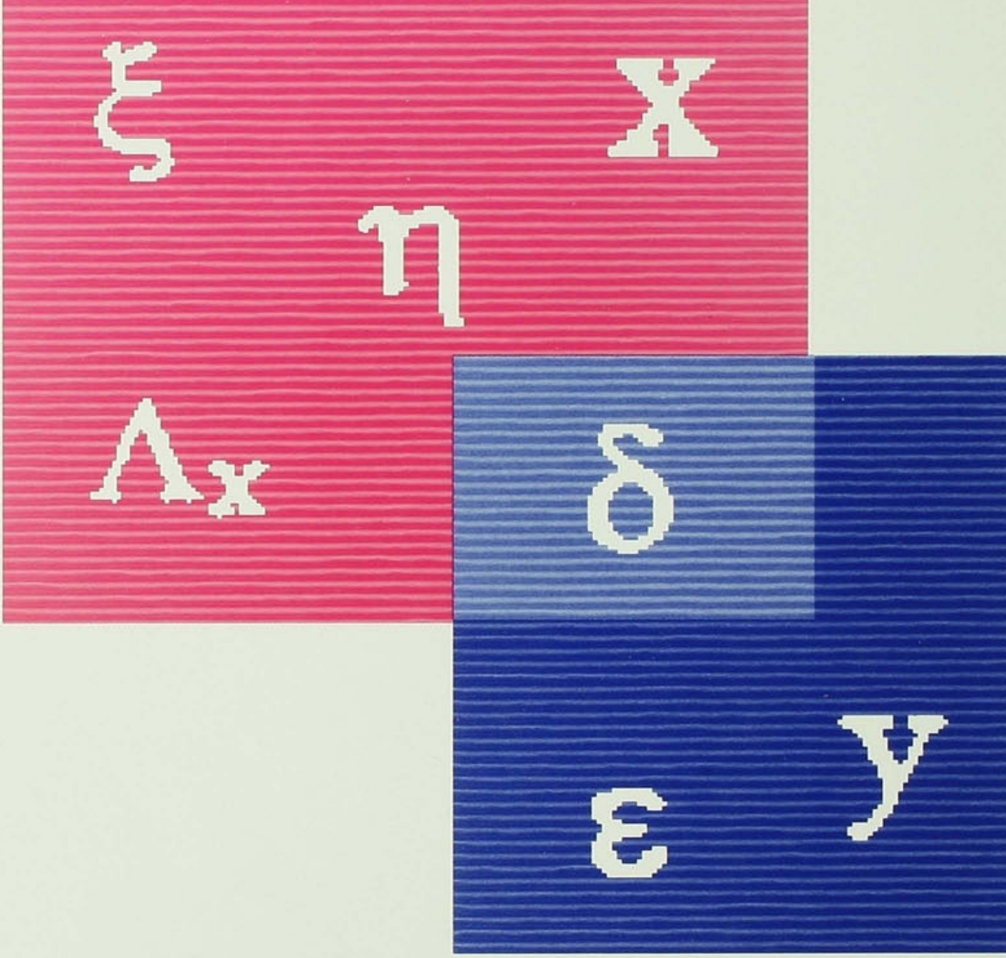
STRUCTURAL EQUATION MODELING with LISREL, PRELIS, and SIMPLIS:

Basic Concepts, Applications, and Programming



Barbara M. Byrne

Structural Equation Modeling With LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications, and Programming

MULTIVARIATE APPLICATIONS BOOK SERIES

I am delighted and honored to be introducing the latest book from Barbara Byrne, a remarkable scholar and methodological spokesperson. As the second book in Lawrence Erlbaum Associates' recent Multivariate Application book series, *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS* describes intricate methodology clearly and understandably. What's more, Dr. Byrne leaves readers assured that they can also tackle the nuances and complexities of structural equation modeling with ease and facility. This kind of book epitomizes the very goals and substance that were intended for this series.

Dr. Byrne's background is aptly suited for turning out the kind of high-quality scholarship that has become her trademark. Her research focuses on structural equation modeling, largely applied to self-concept data. Over the last 15 years, Barbara's laudatory and prolific record includes several major books, dozens of research publications, and numerous presentations and invited addresses. In quantitative terms alone, Barbara's work is impressive.

A qualitative review of her scholarship is equally compelling. Few people could match Barbara's talent for describing complex methodology in completely understandable words. Barbara's writings on structural equation modeling are probably on the shelves of every academic institution and laboratory in North America. Whereas many, if not most, researchers focus their work on a limited academic audience, Barbara prefers to write in a manner that is accessible to a wide array of researchers, methodologists, teachers, practitioners, and students. I have used her structural equation modeling books in a graduate class that I teach and they have been unequivocal hits!

In closing, I am pleased to welcome *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS*, a book that conveys Dr. Byrne's erudite knowledge in a thoroughly lucid and comprehensible style.

Lisa L. Harlow, Ph.D., Series Editor
Department of Psychology
University of Rhode Island
10 Chafee Rd., Suite 8
Kingston, RI 02881-0808

Phone: 401-874-4242
Fax: 401-874-5562
E-Mail: LHarlow@uriacc.uri.edu

Other books in the series:

Harlow/Mulaik/Steiger • What If There Were No Significance Tests?

Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications, and Programming

Barbara M. Byrne
University of Ottawa

LEA LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
1998 Mahwah, New Jersey London

Lawrence Erlbaum Associates, Inc., Publishers
10 Industrial Avenue
Mahwah, New Jersey 07430

Cover Design by Kathryn Houghtaling Lacey

Library of Congress Cataloging-in-Publication Data

Byrne, Barbara M.
Structural equation modeling with LISREL, PRELIS, and SIMPLIS : basic concepts, applications, and programming / Barbara M. Byrne.
p. cm.
Includes bibliographical references and index.
ISBN 0-8058-2924-5 (alk. paper)
1. Multivariate analysis. 2. Social sciences—Statistical methods. I. Title.
QA278.B97 1998
519.5'35—dc21 97-49492
CIP

Books published by Lawrence Erlbaum Associates are printed on acid-free paper, and their bindings are chosen for strength and durability.

Printed in the United States of America
10 9 8 7 6 5

Contents

Preface

The intent of the present book is to illustrate the ease with which various features of the LISREL 8 program and its companion preprocessor program, PRELIS 2, can be implemented in addressing research questions that lend themselves to structural equation modeling (SEM). Specifically, the purpose is threefold: (a) to present a nonmathematical introduction to basic concepts associated with SEM, (b) to demonstrate basic applications of SEM using both the DOS and Windows versions of LISREL 8, as well as both the LISREL and SIMPLIS lexicons, and (c) to highlight particular features of the LISREL 8 and PRELIS 2 programs that address important caveats related to SEM analyses.

This book is intended neither as a text on the topic of SEM, nor as a comprehensive review of the many statistical functions available in the LISREL 8 and PRELIS 2 programs. Rather, the intent is to provide a practical guide to SEM using the LISREL approach. As such, the reader is "walked through" a diversity of SEM applications that include both factor analytic and full latent variable models, as well as a variety of data management procedures. Ideally, this volume serves best as a companion book to the LISREL 8 (Joreskog & Sorbom, 1996b) and PRELIS 2 (Joreskog & Sorbom, 1996c) manuals, as well as to any statistics textbook devoted to the topic of SEM.

The book is divided into four major sections. In Part I, I introduce basic concepts associated with SEM in general, and the LISREL approach to SEM in particular (chap. 1). Also detailed and illustrated in this chapter are the fundamental equations and matrices associated with the LISREL program. Chapter 2 focuses solely on the LISREL 8 and PRELIS 2 programs. Here, I detail the key elements needed in building files related to both programs. I also review the relatively new SIMPLIS language and illustrate the structuring of input files based on this lexicon.

Part II focuses on applications involving single-group analyses; these include three first-order confirmatory factor analytic (CFA) models, one second-order CFA model, and one full latent variable model. Accompanying each of these applications is at least one illustrated use of the PRELIS 2 program. The first-order CFA applications demonstrate testing for the validity of the theoretical structure of a construct (chap. 3), the factorial structure of a measuring instrument (chap. 4), and multiple traits assessed by multiple methods within the framework of a multitrait–multimethod (MTMM) design (chap. 6); unique to the present book are applications of both the General CFA and Correlated Uniqueness MTMM models. The second-order CFA model bears on the factorial structure of a measuring instrument (chap. 5). The final single-group application tests for the validity of a hypothesized causal structure (chap. 7). The PRELIS 2 examples demonstrate the procedures involved in creating a smaller subset of data (chap. 3), a covariance matrix (chap. 4), and polychoric and asymptotic matrices (chap. 5). Also illustrated are the descriptive statistics available for the screening of data (chaps. 6 & 7).

In Part III, I present applications related to multigroup analyses. Specifically I show how to test for measurement and structural invariance related to a theoretical construct (chap. 8), a measuring instrument (chap. 9), and a causal structure (chap. 11). Finally, in chapter 10, I outline basic concepts associated with latent means structures and demonstrate testing for their invariance across groups.

Part IV contains the final chapter of the book, chapter 12, which is devoted to the new Windows versions of both the LISREL 8 and PRELIS 2 programs. For purposes of demonstrating assorted aspects of the Microsoft Windows environment, the final application of SEM involves a two-wave panel study. The chapter begins by first describing the various drop-down menus, and then by detailing each of the options made available to the user. Of particular importance is the diversity of options associated with the Path Diagram menu; a broad array of these schematic displays are illustrated and described. Finally, a further example is provided in demonstrating the use of PRELIS 2 to formulate a subset of data.

In writing a book of this nature, it was essential that I have access to a number of different data sets that lent themselves to various applications. To facilitate this need, all examples presented throughout the book are drawn from my own research; related journal references are cited for readers who may be interested in a more detailed discussion of theoretical frameworks, aspects of the methodology, and/or substantive issues and findings. In summary, each application in the book is accompanied by the following:

- a statement of the hypothesis to be tested;
- a schematic representation of the model under study;
- a full explanation bearing on related LISREL 8 and PRELIS 2 input files;
- a full explanation and interpretation of selected LISREL 8 and PRELIS 2 output files;
- at least one model input file based on the SIMPLIS language;
- the published reference from which the application was drawn;
- published references related to disciplines other than psychology.

It is important to emphasize that, although all applications are based on data that are of a social/psychological nature, they could just as easily have been based on data representative of the health sciences, leisure studies, marketing, or a multitude of other disciplines; my data, then, serve only as one example of each application. Indeed, I strongly recommend seeking out and examining similar examples as they relate to other subject areas. To assist in this endeavor, I have provided a few references at the end of each application chapter (chaps. 3–12).

Because many readers are familiar with my earlier book (referenced here as *The Primer*), which also addressed the basic concepts of SEM and demonstrated applications based on the LISREL program (Byrne, 1989), it seems prudent that I outline the many features of the present volume that set it apart from its predecessor. Indeed, there are several major differences between the two publications. First, discussion of important issues related to SEM (e.g., statistical identification, goodness-of-fit, post hoc model fitting) are substantially more extensive in the present volume than in *The Primer*. Second, whereas *The Primer* focused solely on first-order CFA models, the present text additionally includes second-order CFA and full SEM models. Specifically, in addition to the six CFA applications illustrated in the earlier book, this new book includes applications of (a) a second-order CFA model, (b) a full SEM model, (c) invariance related to a full SEM model, and (d) a full SEM panel model. Third, in contrast to *The Primer*, in which no mention was made with respect to categorical data, the present volume provides both a discussion of the related issues and an application wherein the ordinality of the data is taken into account (chap. 5). Fourth, in contrast to the example of an MTMM model presented in *The Primer*, the application shown in the present text additionally includes the Correlated Uniqueness MTMM model. Fifth, whereas examples in *The Primer* were based on the LISREL V and VI versions of the program, the present book is based on LISREL 8.14, the most current version as this volume goes to press. Furthermore, to illustrate the various features of the LISREL 8 Windows version, one application (chap. 12) is conducted solely within the Windows environment. Sixth, in light of the development of both the PRELIS program and the SIMPLIS language, subsequent to the writing of *The Primer*, the present text, as detailed here, provides considerable description and various examples related to each. Seventh, to address any concerns that all applications in the text are based on psychoeducational data, references bearing on similar applications pertinent to other disciplines (as noted previously) are provided at the end of each application chapter (chaps. 3–12). Finally, with the exception of two applications (chaps. 8 & 9), all examples presented in the current text differ from those in *The Primer*. Whereas the application presented in chapter 10 is based on the same data as that used in the earlier book, the procedure for testing the invariance of latent means structures using LISREL 8 is substantially different.

Although there are now several structural equation modeling texts available, the present book distinguishes itself from the rest in a number of ways. First, it is the only book to demonstrate, by application to actual data, a wide range of confirmatory factor analytic and latent variable models drawn from published studies, along with a detailed explanation of each input and output file. Second, it is the only book to incorporate applications based on both the LISREL 8 and PRELIS 2 programs, and to include example files based on the SIMPLIS language. Third, it is the first book to present an application of a structural equation model using the LISREL 8 program in the Windows environment. Fourth, it is the only book to literally "walk" the reader through (a) the various matrices associated with the LISREL approach to structural equation modeling, (b) the model specification, estimation, and modification decisions and processes associated with a variety of applications, and (c) the various drop-down menus and diagrammer features associated with the Windows version of LISREL 8 and PRELIS 2. Fifth, it is the first book to demonstrate an application of structural equation modeling to the Correlated Uniqueness Multitrait–multimethod model (chap. 5). Finally, the present book contains the most extensive and current description and explanation of goodness-of-fit indices to date.

ACKNOWLEDGMENTS

As with any project of this nature, there are many people to whom I owe a great deal of thanks. First and foremost, I am truly indebted to Larry Erlbaum who provided me with an absolutely peerless team of editorial, artistic, and marketing experts. This extremely talented, efficient, and personable cadre of individuals literally performed miracles in getting this book through the production process as if there were no tomorrow. Thank you all for a wonderful experience!

On a more personal basis, I wish to thank: Anne Monaghan for keeping me updated on a regular basis regarding the progression of this volume through the many stages of production, and for always being true to her word regarding time lines. If Anne advised me that I could expect to receive certain materials on a specific date, they did indeed arrive!; to Sharon Levy, who not only understood, but also shared my love of bright colors, for making sure that the cover of my present book met that criterion; to Joe Petrowski for spreading the word that this volume is now available; and to Art Lizza for overseeing the entire production of the book.

I wish also to extend thanks to my three reviewers: Karl Jöreskog, for "setting me straight" on the appropriate language and expression of particular mathematical formulae; Bruce Thompson, for his very thorough reading of the text and his welcome suggestions for expanded explanations regarding particular topics; an anonymous reviewer for very helpful comments and suggestions bearing on the clarity of text.

I am most grateful to Darlene Worth for her invaluable assistance in the production of all figures related to the various applications presented in this book, as well as to the students in my 1997 (winter semester) Causal Modeling course, for their careful attention to typographical errors, and textual ambiguities and inconsistencies related to the initial draft of the manuscript.

As always, I am truly indebted to my husband, Alex, for his continued support and understanding of the incredible number of hours that my computer and I necessarily spend together on a project of this sort. At this point in time, however, I think that he has given up trying to modify my compulsive behavior that will inevitably lead me into taking on yet another project of similar dimension.

Finally, for my many readers who have inquired from time to time about my yellow lab Amy, my constant companion throughout the writing of my last three books, it is with much sadness that I report that, at 12 1/2 years of age, and following my initial submission of the present manuscript, she decided that a good long sleep was in order. I miss her dearly!

PART ONE

Introduction

CHAPTER 1
Structural Equation Models: The Basics

CHAPTER 2
Using LISREL, PRELIS, and SIMPLIS

CHAPTER
1

Structural Equation Models: The Basics

Structural equation modeling (SEM) is a statistical methodology that takes a confirmatory (i.e., hypothesis-testing) approach to the multivariate analysis of a structural theory bearing on some phenomenon. Typically, this theory represents "causal" processes that generate observations on multiple variables (Bentler, 1988). The term **structural equation modeling** conveys two important aspects of the procedure: (a) that the causal processes under study are represented by a series of structural (i.e., regression) equations, and (b) that these structural relations can be modeled pictorially to enable a clearer conceptualization of the theory under study. The hypothesized model can then be tested statistically in a simultaneous analysis of the entire system of variables to determine the extent to which it is 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.

Several aspects of SEM set it apart from the older generation of multivariate procedures (see Fornell, 1982). First, as noted earlier, it takes a confirmatory, rather than an exploratory, approach to the data analysis (although aspects of the latter can be addressed). Furthermore, by demanding that the pattern of intervariable relations be specified a priori, SEM lends itself well to the analysis of data for inferential purposes. By contrast, most other multivariate procedures are essentially descriptive by nature (e.g., exploratory factor analysis), so that hypothesis testing is difficult, if not impossible. Second, whereas traditional multivariate procedures are incapable of either assessing or correcting for measurement error, SEM provides explicit estimates of these parameters. Finally, whereas data analyses using the former methods are based on observed measurements only, those using SEM procedures can incorporate both unobserved (i.e. latent) and observed variables.

Given these highly desirable characteristics, SEM has become a popular methodology for nonexperimental research, where methods for testing theories are not well developed and ethical considerations make experimental design unfeasible (Bentler, 1980). Structural equation modeling can be utilized very effectively to address numerous research problems involving nonexperimental research; in this book, I illustrate the most common applications (e.g., chapters 3, 4, 7, 9, and 11), as well as some that are less frequently found in the substantive literatures (e.g., chapters 5, 6, 8, 10, and 12). Before showing you how to use the LISREL program (Jöreskog & Sörbom, 1993b), along with its companion package PRELIS (Jöreskog & Sörbom, 1993c) and second language option SIMPLIS, however, it is essential that I first review key concepts associated with the methodology. We turn now to their brief explanation.

BASIC CONCEPTS

Latent Versus Observed Variables

In the behavioral sciences, researchers are often interested in studying theoretical constructs that cannot be observed directly. These abstract phenomena are termed **latent variables**, or **factors**. Examples of latent variables in psychology are self-concept and motivation; in sociology, powerlessness and anomie; in education, verbal ability and teacher expectancy; in economics, capitalism and social class.

Because latent variables are not observed directly, it follows that they cannot be measured directly. Thus, the researcher must operationally define the latent variable of interest in terms of behavior believed to represent it. As such, the unobserved variable is linked to one that is observable, thereby making its measurement possible. Assessment of the behavior, therefore, constitutes the direct measurement of an observed variable, albeit the indirect measurement of an unobserved variable (i.e., the underlying construct).

It is important to note that the term **behavior** is used here in the very broadest sense to include scores on a particular measuring instrument. Thus, observation may include, for example, self-report responses to an attitudinal scale, scores on an achievement test, in vivo observation scores representing some physical task or activity, coded responses to interview questions, and the like. These measured scores (i.e., measurements) are termed **observed** or **manifest** variables; within the context of SEM methodology, they serve as **indicators** of the underlying construct that they are presumed to represent. Given this necessary bridging process between observed variables and unobserved latent variables, it should now be clear why methodologists urge researchers to be circumspect in their selection of assessment measures. Although the choice of psychometrically sound instruments bears importantly on the credibility of all study findings, such selection becomes even more critical when the observed measure is presumed to represent an underlying construct.[1]

The Factor-Analytic Model

The oldest and best-known statistical procedure for investigating relations between sets of observed and latent variables is that of **factor analysis.** In using this approach to data analyses, the researcher examines the covariation among a set of observed variables in order to gather information on their underlying latent constructs (i.e., factors). There are two basic types of factor analyses: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). We turn now to a brief description of each.

EFA is designed for the situation where links between the observed and latent variables are unknown or uncertain. The analysis thus proceeds in an exploratory mode to determine how, and to what extent the observed variables are linked to their underlying factors. Typically, the researcher wishes to identify the minimal number of factors that underlie (or account for) covariation among the observed variables. For example, suppose a researcher develops a new instrument designed to measure five facets of physical self-concept (e.g., Health, Sport Competence, Physical Appearance, Coordination, Body Strength). Following the formulation of questionnaire items designed to measure these five latent constructs, he or she would then

[1]Throughout the remainder of the book, the terms **latent**, **unobserved**, or **unmeasured** variable are used synonymously to represent a hypothetical construct or factor; the terms **observed**, **manifest**, and **measured** variable are also used interchangeably.

conduct an EFA to determine the extent to which the item measurements (the observed variables) were related to the five latent constructs. In factor analysis, these relations are represented by **factor loadings.**[2] The researcher would hope that items designed to measure health, for example, exhibited high loadings on that factor, albeit low or negligible loadings on the other four factors. This factor analytic approach is considered to be exploratory in the sense that the researcher has no prior knowledge that the items do, indeed, measure the intended factors. (For extensive discussions of EFA, see Comrey, 1992; Gorsuch, 1983; McDonald, 1985; Mulaik, 1972).

In contrast to EFA, CFA is appropriately used when the researcher has some knowledge of the underlying latent variable structure. Based on knowledge of the theory, empirical research, or both, he or she postulates relations between the observed measures and the underlying factors a priori, and then tests this hypothesized structure statistically. For example, based on the example cited earlier, the researcher would argue for the loading of items designed to measure sport competence self-concept on that specific factor, and not on the health, physical appearance, coordination, or body strength self-concept dimensions. Accordingly, a priori specification of the CFA model would allow all sport competence self-concept items to be free to load on that factor, but restricted to have zero loadings on the remaining factors. The model would then be evaluated by statistical means to determine the adequacy of its goodness of fit to the sample data. (For more detailed discussions of CFA, see e.g., Bollen, 1989b; Hayduk, 1987; Long, 1983a.)

In summary, therefore, the factor analytic model (EFA or CFA) focuses solely on how, and the extent to which the observed variables are linked to their underlying latent factors. More specifically, it is concerned with the extent to which the observed variables are generated by the underlying latent constructs and thus strength of the regression paths from the factors to the observed variables (the factor loadings) are of primary interest. Although interfactor relations are also of interest, any regression structure among them is not considered in the factor analytic model. Because the CFA model focuses solely on the link between factors and their measured variables, within the framework of SEM, it represents what has been termed a **measurement model.**

---

[2]Although the term **loading** is commonly used in the CFA literature, Thompson and Daniel (1996) have argued that the more specific term of **structure coefficient** should be used.

The Full Latent Variable Model

In contrast to the factor analytic model, the full latent variable (LV) model allows for the specification of regression structure among the latent variables. That is, the researcher can hypothesize the impact of one latent construct on another in the modeling of causal direction. This model is termed **full** (or **complete**) because it comprises both a measurement model **and** a structural model; the **measurement model** depicting the links between the latent variables and their observed measures (i.e., the CFA model), and the **structural model** depicting the links among the latent variables themselves.

A full LV model that specifies direction of cause from one direction only is termed a **recursive model**; one that allows for reciprocal or feedback effects is termed a **nonrecursive model**. Only applications of recursive models are considered in this book.

General Purpose and Process of Statistical Modeling

Statistical models provide an efficient and convenient way of describing the latent structure underlying a set of observed variables. Expressed either diagrammatically, or mathematically via a set of equations, such models explain how the observed and latent variables are related to one another.

Typically, a researcher postulates a statistical model based on his or her knowledge of the related theory, on empirical research in the area of study, or some combination of both. Once the model is specified, the researcher then tests its plausibility based on sample data that comprise all observed variables in the model. The primary task in this model-testing procedure is to determine the goodness of fit between the hypothesized model and the sample data. As such, the researcher imposes the structure of the hypothesized model on the sample data, and then tests how well the observed data fit this restricted structure. (A more extensive discussion of this process is presented in Chapter 3.) Because it is highly unlikely that a perfect fit will exist between the observed data and the hypothesized model, there will necessarily be a discrepancy between the two; this discrepancy is termed the **residual**. The model-fitting process can therefore be summarized as:

Data = Model + Residual

where:

Data represent score measurements related to the observed variables as derived from persons comprising the sample
Model represents the hypothesized structure linking the observed variables to the latent variables, and in some models, linking particular latent variables to one another
Residual represents the discrepancy between the hypothesized model and the observed data

In summarizing the general strategic framework for testing structural equation models, Jöreskog (1993) distinguished among three scenarios that he termed **strictly confirmatory** (SC), **alternative models** (AM), and **model generating** (MG). In the first instance (the SC scenario), 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. From the results of this test, the researcher either rejects or fails to reject the model; no further modifications to the model are made. In the **AM** case, the researcher proposes several alternative (i.e., competing) models, all of which are grounded in theory. Following analysis of a single set of empirical data, he or she selects one model as most appropriate in representing the sample data. Finally, the MG scenario represents the case where the researcher, having postulated and rejected a theoretically derived model on the basis of its poor fit to the sample data, proceeds in an exploratory (rather than confirmatory) fashion to modify and reestimate the model. The primary focus here is to locate the source of misfit in the model and to determine a model that better describes the sample data. Jöreskog noted that, although respecification may be either theory- or data-driven, the ultimate objective is to find a model that is both substantively meaningful and statistically well fitting. He further posited that despite the fact that "a model is tested in each round, the whole approach is model generating, rather than model testing" (p. 295).

Of course, even a cursory review of the empirical literature will clearly show the **MG** situation to be the most common of the three scenarios, and for good reason. Given the many costs associated with the collection of data, it would be a rare researcher indeed who could afford to terminate his or her research on the basis of a rejected hypothesized model! As a consequence, the SC scenario is not commonly found in practice. Although the **AM** approach to modeling has also been a relatively uncommon practice, at least two important papers on the topic (e.g., MacCallum, Roznowski, & Necowitz, 1992; MacCallum, Wegener, Uchino, & Fabrigar, 1993) have recently precipitated more activity with respect to this analytic strategy.

Statistical theory related to these model-fitting processes can be found (a) in texts devoted to the topic of SEM (e.g., Bollen, 1989b; Hayduk, 1987; Loehlin, 1992; Long, 1983b; Mueller, 1996; Saris & Stronkhurst, 1984; Schumacker & Lomax, 1996), (b) in edited books devoted to the topic (e.g., Bollen & Long, 1993; Hoyle, 1995b; Marcoulides & Schumacker, 1996), and (c) in methodologically oriented journals such as *British Journal of Mathematical and Statistical Psychology, Journal of Educational Statistics, Multivariate Behavioral Research, Psychological Methods, Psychometrika, Sociological Methodology, Sociological Methods & Research*, and *Structural Equation Modeling: A Multidisciplinary Journal*.

THE GENERAL LISREL MODEL

As with communication in general, one must first acquire an understanding of the language before being able to interpret the message conveyed. So it is with SEM. Although the researcher wishing to use SEM procedures now has several other computer programs from which to choose (e.g., AMOS—Arbuckle, 1995; EQS—Bentler, 1995; LISCOMP—Muthén, 1988; CALIS—SAS Institute, 1992; RAMONA—Browne, Mels, & Coward, 1994; SEPATH—Steiger, 1994), the LISREL program is the most longstanding and widely distributed. (For an extended list and annotated bibliography, see Austin & Calderón, 1996.) Indeed, it has served as the prototype for all subsequently developed programs. Nonetheless, each of these programs is unique in the command language it uses in model specification. In this regard, LISREL stands apart from the other programs in its "bilingual" capabilities. That is, in lieu of using the original Greek language traditionally associated with statistical models, the researcher can opt to specify models using everyday language, as made possible by the SIMPLIS command language. Although I describe the SIMPLIS notation in chapter 2, and provide example input files throughout, the format of this book follows my earlier one (Byrne, 1989) in focusing on the original LISREL notation.

To fully comprehend the nature of both CFA and full LV models within the framework of the LISREL program, it is helpful if we first examine a generalized model structure. For didactic purposes only, I have termed this model a **General LISREL model**; later in the chapter, I refer to other types of LISREL models. By taking this general model approach, I can then decompose the model into its component parts, which, I believe, will reduce much of the mystery associated with model specification using the LISREL command language. We turn now to this decomposition process.

Basic Composition

As noted earlier, the general LISREL model can be decomposed into two submodels: a measurement model, and a structural model. The **measurement model** defines relations between the observed and unobserved variables. In other words, it provides the link between scores on a measuring instrument (i.e., the observed indicator variables) and the underlying constructs they are designed to measure (i.e., the unobserved latent variables). The measurement model, therefore, specifies the pattern by which each measure loads on a particular factor. It also describes the measurement properties (reliability, validity) of the observed variables (Jöreskog & Sörbom, 1989). The **structural model** defines relations among the unobserved variables. Accordingly, it specifies which latent variable(s) directly or indirectly influences (i.e., "causes") changes in the values of other latent variables in the model.

One necessary requirement in working with LISREL is that in specifying full structural equations models the researcher must distinguish between latent variables that are exogenous, and those that are endogenous. **Exogenous** latent variables are synonymous with independent variables; they "cause" fluctuations in the values of other latent variables in the model. Changes in the values of exogenous variables are not explained by the model. Rather, they are considered to be influenced by other factors external to the model. **Endogenous** latent variables are synonymous with dependent variables and, as such, are influenced by the exogenous variables in the model, either directly, or indirectly. Fluctuation in the values of endogenous variables is said to be explained by the model because all latent variables that influence them are included in the model specification.

The Link Between Greek and LISREL

In the Jöreskog tradition, the command language of LISREL is couched in the mathematical statistical literature, which commonly uses Greek letters to denote parameters. Despite the availability of SIMPLIS, as noted earlier, it behooves the serious user of LISREL to master the original language. Thus, a second requirement in learning to work with the LISREL program is to become thoroughly familiar with the various LISREL matrices and the Greek letters that represent them. What follows now is a description of these basic matrices.

Consistent with matrix algebra, matrices are represented by upper-case Greek letters, and their elements, by lower case Greek letters; the elements represent the parameters in the model. By convention, observed measures are represented by Roman letters. Accordingly, those that are exogenous are termed **X-variables**, and those that are endogenous are termed **Y-variables**.

As becomes more evident later in the section dealing with the LISREL CFA model, the **measurement model** may be specified either in terms of LISREL exogenous notation (i.e., X-variables), or in terms of its endogenous notation (i.e., Y-variables). Each of these measurement models can be defined by two matrices and two vectors as follows:[3] one regression matrix relating the exogenous (Λ_X) (or endogenous [Λ_Y]) LVs to their respective observed measures, one vector of latent exogenous (ξs) (or endogenous [ηs]) variables, and one vector of measurement errors related to the exogenous (δs) (or endogenous [εs]) observed variables.

The **structural model** can be defined by two matrices and three vectors. These include one matrix of coefficients relating exogenous LVs to endogenous LVs (Γ), one matrix of coefficients relating endogenous LVs to other endogenous LVs (B), one vector of latent exogenous variables (ξs), one vector of latent endogenous variables (ηs), and one vector of residual errors associated with the endogenous LVs (ζs).[4]

Taken together, the general LISREL model can be captured by the following three equations:

Measurement Model for the X-variables:

$x = \Lambda_X\xi + \delta \qquad (1.1)$

Measurement Model for the Y-variables:

$y = \Lambda_Y\eta + \varepsilon \qquad (1.2)$

---

[3]A **matrix** represents a series of numbers written in rows and columns; each number in the matrix is termed an element. A **vector** is a special matrix case; a column vector has several rows, albeit one column, whereas a row vector has several columns and only one row.

[4]These residual terms represent errors in the equation in the prediction of endogenous factors from exogenous factors; they are typically referred to as **disturbance terms** to distinguish them from errors of measurement associated with the observed variables.

Structural Equation Model:

$$\eta = B\eta + \Gamma\xi + \zeta \qquad (1.3)$$

In an attempt to clarify the nature of the various components of these equations, each is now described in more detail. Because many readers may feel a little overwhelmed by the use of symbols and letters in these descriptions, I wish first to explain the basic rule by which the subscript numbers are governed. By convention, matrices are defined according to their number of rows and columns; the number of rows is always specified first. This row by column description of a matrix is termed the "order" of the matrix. We turn now to a more specific definition for each of these terms.

The Measurement Model

x is a q x 1 vector of observed exogenous variables
y is a p x 1 vector of observed endogenous variables
ξ is an n x 1 vector of latent exogenous variables
η is an m x 1 vector of latent endogenous variables
δ is a q x 1 vector of measurement errors in x
ε is a p x 1 vector of measurement errors in y
Λ_x (lambda-x) is a q x n regression matrix that relates n exogenous factors to each of the q observed variables designed to measure them.
Λ_y (lambda-y) is a p x m regression matrix that relates m endogenous factors to each of the p observed variables designed to measure them.

The Structural Model

Γ (gamma) is an m x n matrix of coefficients that relates the n exogenous factors to the m endogenous factors.
B (beta) is an m x m matrix of coefficients that relates the m endogenous factors to one another.
ζ (zeta) is an m x 1 vector of residuals representing errors in the equation relating η and ξ.

A summary of these matrices, together with their Greek and program notations, is presented in Table 1.1.

It is of further importance to note that, in the specification of any LISREL model, the following minimal assumptions are presumed to hold:

ε is uncorrelated with η
δ is uncorrelated with ξ
ζ is uncorrelated with ξ
ζ, ε, and δ are mutually uncorrelated.

Table 1.1

Summary of Matrices, Greek Notation, and Program Codes

Greek Letter	Matrix	Matrix Elements	Program Code	Matrix Type
Measurement Model				
Lambda-X	Λ_x	λ_x	LX	Regression
Lambda-Y	Λ_y	λ_y	LY	Regression
Theta delta	Θ_δ	θ_δ	TD	Var/cov
Theta epsilon	Θ_ε	θ_ε	TE	Var/cov
Structural Model				
Gamma	Γ	γ	GA	Regression
Beta	B	β	BE	Regression
Phi	Φ	ϕ	PH	Var/cov
Psi	Ψ	ψ	PS	Var/cov
Xi (or Ksi)	---	ξ	---	Vector
Eta	---	η	---	Vector
Zeta	---	ζ	---	Vector

Var/cov = variance-covariance

To comprehend more fully how the Greek lettering system works, and how variables in a specified model may be linked to one another within the context of the LISREL program, let us turn to the pictorial presentation of a simple hypothetical model illustrated in Fig. 1.1. Schematic representations of a models such as this one are termed **path diagrams** because they provide a visual portrayal of relations that are assumed to hold among the variables under study.

The LISREL Path Diagram

Visible Components

By convention, in the schematic presentation of structural equation models, measured variables are shown in boxes and unmeasured variables in circles (or ellipses). Thus, in reviewing the model shown in Fig. 1.1, we see that there are two latent variables (ξ_1, η_1), and five observed variables (X_1–X_3;

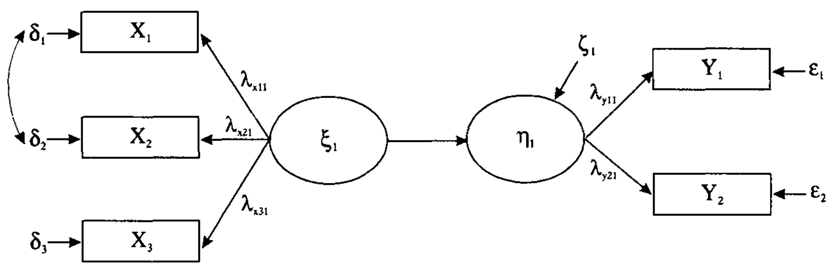


FIG. 1.1. A LISREL Structural Equation Model.

Y_1–Y_2); the Xs and Ys function as **indicators** of their respective underlying latent factors. Associated with each observed variable is an error term (δ_1–δ_3; ε_1–ε_2), and with the factor being predicted (η_1), a disturbance (i.e., residual) term (ζ_1). It is worth noting that, technically speaking, both measurement and residual error terms represent unobserved variables and, thus, they quite correctly could be enclosed in circles. For example, although the disturbance terms (only) are often enclosed in the presentation of path diagrams using the EQS program; (Bentler, 1995), it is not customary to enclose either of these terms in LISREL path diagrams.

In addition to the aforementioned symbols that represent variables, certain others are used in path diagrams to denote hypothesized processes involving the entire system of variables. In particular, one-way arrows represent structural regression coefficients and thus indicate the impact of one variable on another. In Fig. 1.1, for example, the unidirectional arrow pointing toward the η_1 implies that the exogenous factor ξ_1 "causes" the endogenous factor η_1.[5] Likewise, the three unidirectional arrows leading from the exogenous factor (ξ_1) to each of the three observed variables (x_1, x_2, x_3), suggest that these score values (representing, say, item or subscale scores of an assessment measure) are influenced by ξ_1. In LISREL, these regression paths are represented by λ; consistent with the exogenous and endogenous specification of the measurement models, these regression coefficients are labeled λ_x and λ_y, respectively. As such, they represent the magnitude of

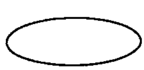
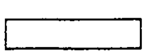
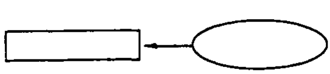
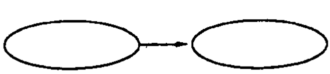
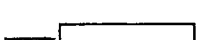
[5]In this text, a cause is a direct effect of a variable on another within the context of a complete model. Its magnitude and direction are given by the partial regression coefficient. If the complete model contains all relevant influences on a given dependent variable, its causal precursors are correctly specified. In practice, however, models may omit key predictors, and may be misspecified, so that it may be inadequate as a causal model in the philosophical sense.

expected change in the observed variables for every change in the related latent variable (or factor). The ordering of the subscripts related to these coefficients is keyed to the direction of the arrow. That is, one works from the variable to which the arrow is pointing, to the source variable. Thus λ_{31}, for example, represents the regression of x_3 on ξ_1. For sake of clarity in the schematic presentation of all subsequent models, however, these "x" and "y" subscripts will not be explicitly labeled; their designation is implicit in the specification of the particular Λ matrix modeled.

The sourceless one-way arrows pointing from the δs and εs indicate the impact of random measurement error on the observed Xs and Ys respectively, and from ζ_1, the impact of error in the prediction of η_1. Finally, curved two-way arrows represent covariances or correlations between pairs of variables. Thus, the bidirectional arrow linking δ_1 and δ_2, as shown in Fig. 1.1, implies that measurement error associated with X_1 is correlated with that associated with X_2. These symbols are summarized in Table 1.2.

Table 1.2

Symbols Associated with Structural Equation Models in LISREL

Symbol	**Representation**
	• Unobserved (latent) Factor
	• Observed Variable
	• Path coefficient for regression of observed variable onto unobserved factor
	• Path coefficient for regression of one factor onto another
	• Residual error (disturbance) in prediction of unobserved factor
	• Measurement error associated with observed variable

As noted in the initial paragraph of this chapter, in addition to lending themselves to pictorial description via a schematic presentation of the causal processes under study, structural equation models can also be represented by a series of regression (i.e., structural) equations. Because (a) regression equations represent the influence of one or more variables on another, and (b) this influence, conventionally in SEM, is symbolized by a single-headed arrow pointing **from** the variable of influence **to** the variable of interest (i.e., the dependent variable), we can think of each equation as summarizing the impact of all relevant variables in the model (observed and unobserved) on one specific variable (observed or unobserved). Thus, one relatively simple approach to formulating these equations is to note each variable that has one or more arrows pointing toward it, and then record the summation of all such influences for each of these dependent variables.

To illustrate this translation of regression processes into structural equations, let us turn again to Fig. 1.1. We can see that there are six variables with arrows pointing towards them; five represent observed variables (X_1–X_3; Y_1–Y_2), and one represents an unobserved variable (or factor; η_1). Thus, we know that the regression functions symbolized in the model shown in Fig. 1.1 can be summarized in terms of six separate regression equations:

$$
\begin{aligned}
\eta_1 &= \xi_1 + \zeta_1 \\
x_1 &= \lambda_{x11} + \delta_1 \\
x_2 &= \lambda_{x12} + \delta_2 \qquad (1.4) \\
x_3 &= \lambda_{x13} + \delta_3 \\
y_1 &= \lambda_{y11} + \varepsilon_1 \\
y_2 &= \lambda_{y12} + \varepsilon_2
\end{aligned}
$$

Nonvisible Components

Although, in principle, there is a one-to-one correspondence between the schematic presentation of a model and its translation into a set of structural equations, it is important to note that neither one of these model representations tells the whole story; some parameters critical to the estimation of the model, are not explicitly shown and thus may not be obvious to the novice structural equation modeler. For example, in both the path diagram and the equations presented earlier, there is no indication that the variances of the independent variables are parameters in the model; indeed, such parameters are essential to all structural equation models. Thus the researcher must be mindful of this inadequacy of path diagrams when specifying his or her model input.

Considering the other side of the coin, however, it is equally important to draw your attention to the specified nonexistance of certain parameters in the model. For example, in Fig. 1.1, we detect no curved arrow between ε_1 and ε_2, which suggests the lack of covariance between the error terms associated with the observed variables Y_1 and Y_2. Similarly, there is no hypothesized covariance between η_1 and ζ_1; absence of this path addresses the common, and most often necessary assumption that the predictor (or exogenous) variable is in no way associated with any error arising from the prediction of the criterion (or endogenous) variable.

The core parameters of concern in structural equation models are the regression coefficients, and the variances and covariances of the independent variables.[6] However, given that sample data comprise observed scores only, there needs to be some internal mechanism whereby the data are transposed into parameters of the model. This task is accomplished via a mathematical model representing the entire system of variables. Such representation systems can, and do, vary with each SEM computer program. (For a comparison of the LISREL and EQS representation systems, for example, see Bentler, 1988.) Because adequate explanation of the way in which the LISREL representation system operates demands knowledge of the program's underlying statistical theory, the topic goes beyond the aims and intent of the present volume. Thus, readers interested in a comprehensive explanation of this aspect of the LISREL approach to the analysis of covariance structures are referred to the following texts (Bollen, 1989b; Hayduk,1987; Saris & Stronkhorst, 1984) and monographs (Jöreskog & Sörbom, 1979; Long, 1983b).

Finally, an important corollary of SEM is that the variances and covariances of dependent (or endogenous) variables, whether they be observed or unobserved, are **never** parameters of the model; rather, they remain to be explained by those parameters that represent the independent (or exogenous) variables in the model. In contrast, the variances and covariances of independent variables are important parameters that need to be estimated.

THE LISREL CONFIRMATORY FACTOR-ANALYTIC MODEL

In an earlier section of this chapter entitled "Basic Concepts," I presented a brief comparison of exploratory and confirmatory factor analyses. In describ-

---

[6]Other parameters may also include the means and intercepts (see chapter 10).

ing the CFA model, I noted that it focuses solely on relations between the observed variables and their underlying factors. 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; as noted earlier, it represents only a portion of the general LISREL model.

Specification of a LISREL CFA model requires the researcher to choose between one that is exogenously oriented and one that is endogenously oriented. Although the choice is purely an arbitrary one, once the decision is made within the framework of one or the other, the specification of all parameters must be consistent with the chosen orientation.[7] In other words, in working with a CFA model using LISREL, the researcher must specify either an all-X (i.e., exogenous) or an all-Y (i.e., endogenous) model. Although some researchers prefer to work within the all-Y model (see e.g., Marsh, 1987), most work within the framework of the all-X model.

For sake of didactic purposes in clarifying this important distinction, let us now examine Fig. 1.2, in which the same model presented in Fig. 1.1 has been demarcated into measurement and structural components.

Considered separately, the elements modeled within each rectangle in Fig. 1.2 represent two CFA models—one exogenous (all-X model), and one endogenous (all-Y model). The enclosure of the two factors within the ellipse represents a structural model and, thus, is not of interest in CFA research. The CFA all-X model represents a one-factor model (ξ_1) measured by three observed variables (X_1–X_3), whereas the CFA all-Y model represents a one-factor model (η_1) measured by two observed variables (Y_1–Y_2). In either case, the regression of the observed variables on the factor, and the variances of both the errors of measurement and the factor, as well as the error covariance, are of primary interest. Although both CFA models described in Fig. 1.2 represent first-order factor models, second- and higher order CFA models can also be analyzed using LISREL. First-order factor models are those in which correlations among the observed variables (i.e., the raw data) can be described by a smaller number of latent variables (i.e., factors), each of which may be considered to be one level, or one unidimensional arrow away from the observed variables (see Fig. 1.2); these factors are termed

[7]Technically speaking, because there is no specification of causal relations among the factors in CFA models, there is no specific designation of variables as either exogenous or endogenous. Relatedly, ζ, the residual term associated with the prediction of one factor from another, will necessarily be zero thereby making the all-X and all-Y models equivalent. This, then accounts for the arbitrariness of the decision.

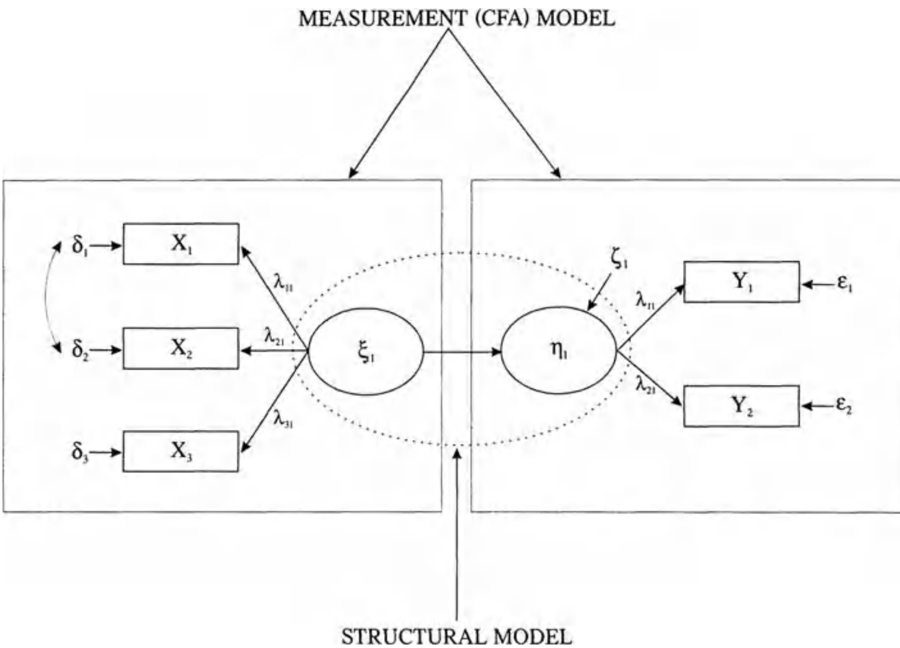


FIG. 1.2. A LISREL Structural Equation Model demarcated into its Measurement and Structural Components.

primary or **first-order factors**. Second-order factor models are those in which correlations among the first-order factors, in turn, can be represented by a single factor, or at least a smaller set of factors. Relatedly, one can think of these higher order factors as being two levels, or two unidimensional arrows away from the observed variables; hence the term **second-order factor**. Likewise, if correlations among the second-order factors can be represented by yet another factor (or smaller set of factors), a third-order model exists, and so on.

Although Kerlinger (1984) has deemed such hierarchical CFA models to be less commonly found in the literature, second-order CFA models are beginning to appear more frequently in the methodological journals. Discussion and application of CFA models in the present book are limited to first- and second-order models only. (For a more comprehensive discussion and explanation of first- and second-order CFA models, see Bollen, 1989b; Kerlinger, 1984.) We turn now to an explanation of the first-order CFA model within the framework of the LISREL program.

First-Order CFA Model

As demonstrated in the previous section, structural equation models can be depicted both diagrammatically and as a series of equations. Furthermore, the equation format can be expressed either in matrix form or as a series of regression statements. Using a simple two-factor model, let us now reexamine the LISREL CFA model within the framework of each of these two formats; we focus here on a first-order factor model. To help you in acquiring a thorough understanding of all matrices and their related elements, and because examples of both are found in the literature, this model will be expressed first as an all-X model, and then as an all-Y model.

The All-X CFA Model

Suppose that we have a two-factor model of self-concept. Let the two factors be academic self-concept (ASC) and social self-concept (SSC). Suppose that each factor has two indicator variables.[8] Let the two measures of ASC be scores from the Academic Self-Concept subscale of the Self Description Questionnaire–I (SDQASC; Marsh, 1992) and the Scholastic Competence subscale of the Self-Perception Profile for Children (SPPCASC; Harter, 1985). Let the two measures of SSC be scores from the Peer Relations subscale of the SDQ–I (SDQSSC) and the Social Acceptance subscale of the SPPC (SPPCSSC). This model is presented schematically in Fig. 1.3.

Shown here then, is a two-factor model representing the constructs of ASC and SSC, with each factor having two indicator variables. These two observed measures for ASC are SDQASC and SPPCASC; for SSC they are SDQSSC and SPPCSSC. Associated with each observed measure is an error term. The curved two-headed arrow indicates that the two factors, ASC and SSC, are correlated.

Now let us translate this model into LISREL notation and reexamine it both diagrammatically and as a set of equations. We will focus first on the all-X CFA model and then turn our attention to the all-Y CFA model. Fig. 1.4, therefore, represents a schematic translation of the hypothesized model of self-concept with LISREL notation specific to the all-X formulation.

---

[8]It is now widely recommended that at least three observed variables should be used as indicators of the underlying constructs (see, e.g., Anderson & Gerbing, 1984; Bentler & Chou, 1987; Wothke, 1993). However, for didactic purposes in the examination of a simple model, I have limited the number of indicators to two.

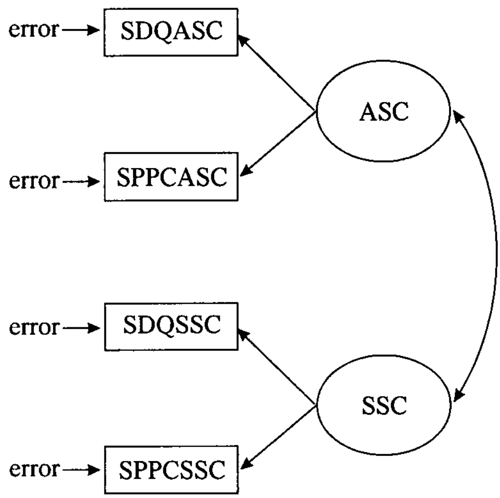


FIG. 1.3. Hypothesized First-order CFA Model.

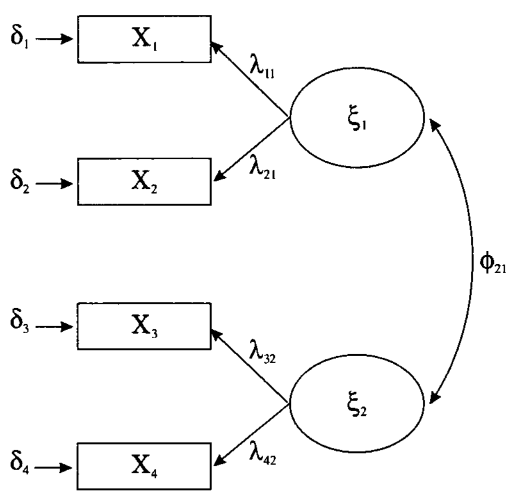


FIG. 1.4. Hypothesized First-order CFA Model with LISREL All-X Model Notation. From *A Primer of LISREL: Basic Applications and Programming for Confirmatory Factor Analytic Models* (p. 11), by B. M. Byrne, 1989, New York: Springer-Verlag. Copyright 1989 by Springer-Verlag. Reprinted with permission.

In reviewing Fig. 1.4, we see that the factors ASC and SSC are represented by ξ_1 and ξ_2, respectively; their correlation is represented by ϕ_{21}. The parameters λ_{11} and λ_{21} represent the regression of X_1 (SDQASC) and X_2 (SPPCASC) on ξ_1 (ASC), and λ_{32} and λ_{42}, the regression of X_3 (SDQSSC) and

X_4 (SPPCSSC) on ξ_2 (SSC), respectively. Finally, δ_1 to δ_4 represent errors of measurement associated with X_1 to X_4, respectively.

We can now decompose the LISREL model in Fig. 1.4 into its related regression equations as:

$$\begin{aligned} x_1 &= \lambda_{11}\xi_1 + \delta_1 \\ x_2 &= \lambda_{21}\xi_1 + \delta_2 \\ x_3 &= \lambda_{32}\xi_2 + \delta_3 \\ x_4 &= \lambda_{42}\xi_2 + \delta_4 \end{aligned} \qquad (1.5)$$

This set of equations can be captured by a single expression that describes relations among the observed variables (xs), the latent variables (ξs), and the errors of measurement (δs); shown here as Equation 1.6, this expression represents the general LISREL factor analytic model (for the all-X CFA model):

$$x = \Lambda_x\xi + \delta \qquad (1.6)$$

The parameters of this model are Λ_X, Φ, and Θ_δ where:

- Λ_X represents the matrix of regression coefficients related to the ξs (described earlier).
- Φ (phi) is an n x n symmetrical variance-covariance matrix among the n exogenous factors.
- Θ_δ (theta-delta) is a symmetrical q x q variance-covariance matrix among the errors of measurement for the q exogenous observed variables.

Consistent with the model in Fig. 1.4 and the series of regression statements in Equation 1.5, this general factor analytic model can be expanded as:

$$\begin{array}{ccccccc} x & & \Lambda_x & & \xi & & \delta \\ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} & = & \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ 0 & \lambda_{32} \\ 0 & \lambda_{42} \end{bmatrix} & & \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} & + & \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix} \end{array} \qquad (1.7)$$

In reviewing Equation 1.7 which represents the regression equations in matrix format, it is now easy to follow the direct transposition of parameters from the model depicted in Fig. 1.4. Further elaboration of the regression matrix (Λ_x), however, is in order. Accordingly, I wish to note four features of this matrix that may be helpful to you in your understanding of these equations. **First**, consistent with the number of factors (ξs) in the model, there are two columns—one representing each factor. **Second**, recall that in interpreting the double subscripts, the first one represents the row, and the second one the column. From a statistical perspective, λ_{11} for example, indicates an element in the first row and first column of the matrix; from a substantive perspective, it represents the regression coefficient for the first observed variable (X_1) on Factor 1 (ξ_1); likewise, λ_{21} represents the second indicator (X_2) of Factor 1. Both of these parameters therefore appear in column 1; in contrast, those representing observed variables X_3 and X_4, the indicators of Factor 2 (ξ_2), are shown in column 2. **Third**, the zeros are fixed values indicating that, for example, X_1 and X_2 are specified to load on Factor 1 and *not* on Factor 2; the reverse holds true for X_3 and X_4. In sum, then, the λs represent parameters to be estimated, whereas the 0s represent parameters whose values have been fixed to zero a priori. **Finally**, the Λ matrix is often termed the **factor-loading matrix** because it portrays the pattern by which each observed variable is linked to its respective factor; it is classified as a "full" matrix in terms of LISREL language. Information bearing on types of matrices is important in the specification of models in LISREL, and constitutes a topic that we review in Chapter 2.

As presented in Equation 1.7, both the factors (ξs) and the measurement error terms (δs) were expressed as vectors. Recall, however, that in addition to the regression coefficients, the core parameters in any structural equation model also include the variances and covariances of the independent variables. Although it is easy to see that the ξs in the model represent independent latent variables, this designation for the δs is likely less obvious (and really not of any consequence, other than to make a point here); technically speaking, given that they impact on the observed variables (i.e., the arrow points **from** the error term **to** the observed variable), they may be regarded as independent variables. Overall, then, we will want to estimate the variances for both factors, as well as for the measurement errors.[9] In addition, because the two factors are related, their covariance also needs to be estimated.

[9]Typically, the specification of structural equation models assumes no relations among the errors of measurement a priori. Such parameters, however, can often lead to major misspecification of a model; their presence is determined from the application of post hoc model-fitting procedures, a topic that is discussed at various points throughout the book.

To illustrate these variance–covariance parameters more explicitly, it is helpful to further expand the matrix equation shown in Equation 1.7 to include all elements in the related variance–covariance matrices. As noted earlier, variances and covariances of the ξs are represented in the Φ matrix; those for the δs are represented in the Θ_δ matrix. Both variance–covariance matrices are classified as being symmetric. This expansion is presented in Equation 1.8:

$$\underset{\mathbf{x}}{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}} = \underset{\Lambda_x}{\begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ 0 & \lambda_{32} \\ 0 & \lambda_{42} \end{bmatrix}} \underset{\Phi}{\begin{bmatrix} \phi_{11} & \\ \phi_{21} & \phi_{22} \end{bmatrix}} + \underset{\Theta_\delta}{\begin{bmatrix} \theta_{11} & & & \\ 0 & \theta_{22} & & \\ 0 & 0 & \theta_{33} & \\ 0 & 0 & 0 & \theta_{44} \end{bmatrix}} \tag{1.8}$$

As with all variance–covariance matrices, the variances are represented on the diagonals and the covariances on the off-diagonals. Thus, ϕ_{11} and ϕ_{22} represent the variance of Factors 1 (ξ_1) and 2 (ξ_2), respectively; ϕ_{21} represents their covariance. Likewise, θ_{11}–θ_{44} represent the variance of each error term associated with the four observed variables (X_1–X_4).[10] Because the model specified no error covariances, each of the off-diagonal elements in the Θ_δ matrix is fixed to zero. On the other hand, had the presence of error covariances (commonly termed **correlated errors**) been known a priori, then the appropriate off-diagonal parameter would have been specified as free (i.e., estimable).

The All-Y CFA Model

Now let us review this same two-factor model expressed as an all-Y LISREL first-order CFA model. Presented first is the schematic presentation of the model (Fig. 1.5), followed by the set of four regression statements (Equation 1.9) and the general LISREL factor model (for the all-Y CFA model; Equation 1.10). The parameters of this model are Λ_y, Ψ, and Θ_ε where:

[10]As with the x/y designation of the λ parameters noted earlier, the subscript for each θ parameter could be preceded by a "δ" to distinguish it from the θ parameters in the Θ matrix. However, in the interest of clarity and neatness, these additional symbols are not included.

Λ_y represents the matrix of regression coefficients related to the η's (described earlier)

Ψ (psi) is an *m* x *m* symmetrical variance-covariance matrix among the *m* residual errors for the *m* endogenous factors.

Θ_ε (theta-epsilon) is a symmetrical *p* x *p* variance-covariance matrix among the errors of measurement for the *p* endogenous observed variables.

The regression equations expressed in matrix format are presented in Equation 1.11, followed by the expanded matrix equation (Equation 1.12):

$$\begin{aligned} y_1 &= \lambda_{11}\eta_1 + \varepsilon_1 \\ y_2 &= \lambda_{21}\eta_1 + \varepsilon_2 \\ y_3 &= \lambda_{32}\eta_2 + \varepsilon_3 \\ y_4 &= \lambda_{42}\eta_2 + \varepsilon_4 \end{aligned} \qquad (1.9)$$

$$Y = \Lambda_y\eta + \varepsilon \qquad (1.10)$$

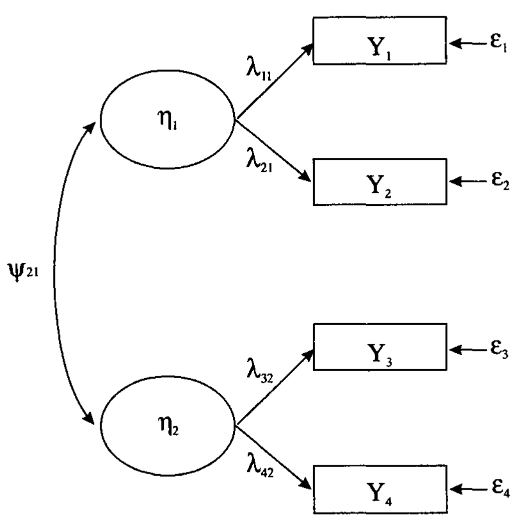


FIG. 1.5. Hypothesized First-order CFA Model with LISREL All-Y Model Notation. From *A Primer of LISREL: Basic Concepts, Applications, and Programming* (p. 14), by B. M. Byrne, 1989, New York: Springer-Verlag. Copyright 1989 by Springer-Verlag. Reprinted with permission.

$$\begin{array}{cccc} Y & \Lambda_y & \eta & \varepsilon \\ \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} & \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ 0 & \lambda_{32} \\ 0 & \lambda_{42} \end{bmatrix} & \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} & \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix} \end{array} \quad (1.11)$$

$$\begin{array}{ccccccc} Y & & \Lambda_y & \Psi & & \Theta_\varepsilon \\ \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} & = & \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ 0 & \lambda_{32} \\ 0 & \lambda_{42} \end{bmatrix} & \begin{bmatrix} \psi_{11} & \\ \psi_{21} & \psi_{22} \end{bmatrix} & + & \begin{bmatrix} \theta_{11} & & & \\ 0 & \theta_{22} & & \\ 0 & 0 & \theta_{33} & \\ 0 & 0 & 0 & \theta_{44} \end{bmatrix} \end{array} \quad (1.12)$$

Chances are you were doing just fine until you reached Equation 1.12, at which point you likely wondered where on earth the Psi (Ψ) matrix came from. In fact, the variance–covariance matrix for the residual terms (ζs) is represented in the Ψ matrix. However, because the LISREL CFA model does not include a causal path between the two factors (e.g., $\xi_1 \rightarrow \eta_1$), the residual term is nonexistant (i.e., equal to a value of zero). Thus, variances and covariances for the η factors in the all-Y model are estimated in the ψ matrix.

One Final Model

To finalize our examination of the first-order CFA model, and to allow me to introduce you to the issue of statistical identification, let us now take a brief look at another slightly more complex model that incorporates four factors, each of which has three indicator variables. A schematic presentation of this model is shown in Fig. 1.6.

Considered as an all-X model, this CFA structure comprises four self-concept (SC) factors— academic SC (ASC; ξ_1), social SC (SSC; ξ_2), physical SC (PSC; ξ_3), and emotional SC (ESC; ξ_4). Each SC factor is measured by three observed variables (x_1–x_3; x_4–x_6; x_7–x_9; x_{10}–x_{12}, respectively). The reliability of each of these indicators is influenced (as indicated by the $\rightarrow$s) by random measurement error (δ_1–δ_3; δ_4–δ_6; δ_7–δ_9; δ_{10}–δ_{12}, respectively). Each of these observed variables is regressed onto its respective factor (λ_{11}–λ_{31}; λ_{42}–λ_{62}; λ_{73}–λ_{93}; $\lambda_{10,4}$–$\lambda_{12,4}$, respectively). Finally, the four factors are shown to be intercorrelated (ϕ_{21}, ϕ_{31}, ϕ_{41}, ϕ_{32}, ϕ_{42}, ϕ_{43}).

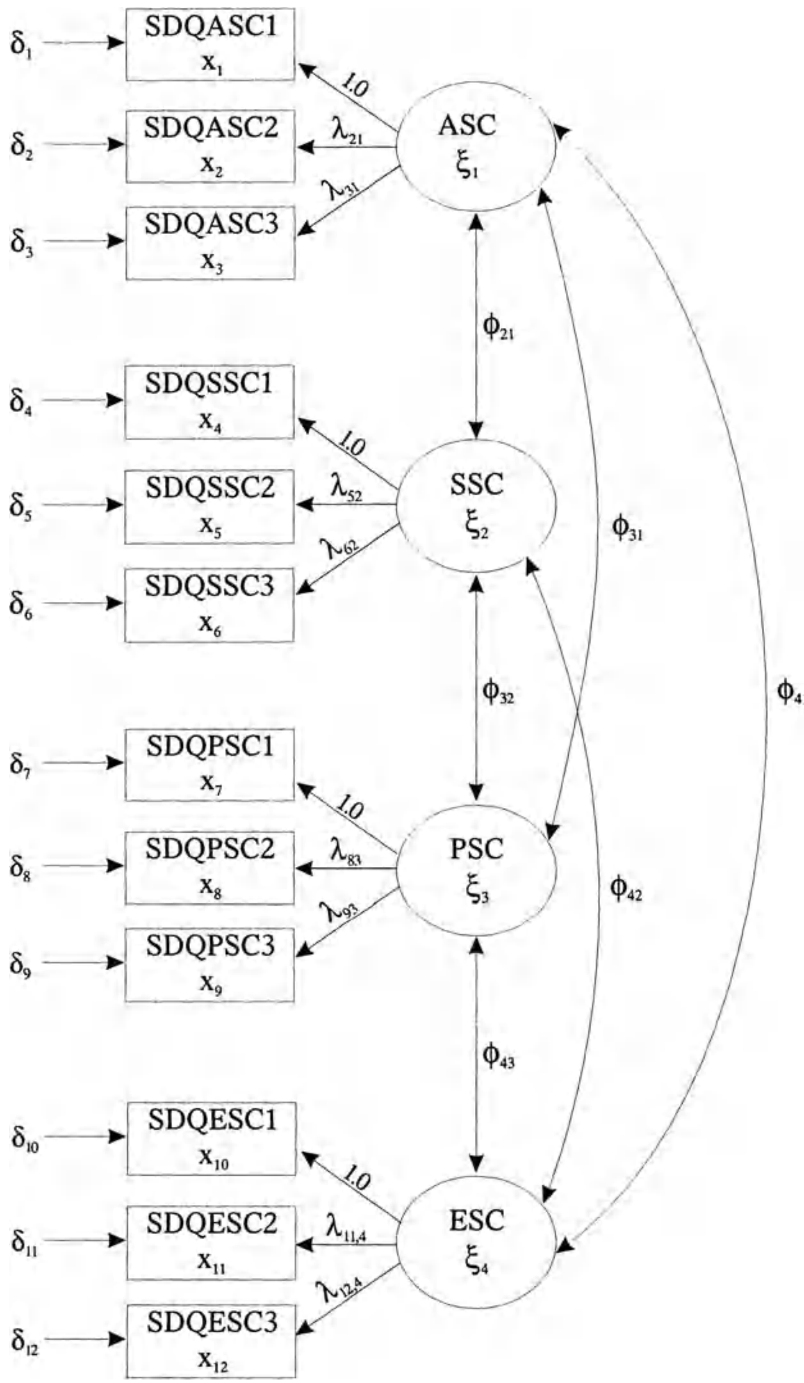


FIG. 1.6. Hypothesized First-order CFA Model with LISREL Notation.

As noted earlier, the key parameters to be estimated in a CFA model are the regression coefficients (i.e., factor loadings), the factor and error variances, and in some models (as is the case here), the factor covariances. One extremely important caveat, in working with structural equation models, is to **always** tally the number of parameters in the model to be estimated prior to running the analyses. This information is critical to your knowledge of whether or not the model that you are testing is statistically identified, a topic to which I turn next.

As a prerequisite to the discussion of identification, however, let us first count the number of parameters to be estimated for the model portrayed in Fig. 1.6. From a review of the figure, we can ascertain that there are 12 regression coefficients (λs), 16 variances—12 error variances (δs) and 4 factor variances (ϕs), and 6 covariances (ϕ_{21}–ϕ_{43}). However, you will also note that the first of each set of λ parameters has been fixed to a value of 1.00 (they are therefore not to be estimated). The rationale underlying this constraint is tied to the issue of statistical identification. In total, then, there are 30 parameters to be estimated for the CFA model depicted in Fig. 1.6. Let us now turn to a brief discussion of this important concept.

The Issue of Statistical Identification

Statistical identification is a complex topic that is difficult to explain in nontechnical terms. Although a thorough explanation of the identification principle exceeds the scope of the present book, it is not critical to the reader's understanding and use of the book. Nonetheless, because some insight into the general concept of the identification issue will undoubtedly help you to better understand why, for example, particular parameters are specified as having certain fixed values, I attempt now to give you a brief, nonmathematical explanation of only one aspect of this concept—the measurement issue relative to the latent variables. Essentially, I address the so-called "*t*-rule," one of several tests associated with identification. Readers are urged to consult the following texts for a more comprehensive treatment of the topic: Bollen (1989b); Hayduk (1987); Long (1983a, 1983b); Saris and Stronkhorst (1984).

In broad terms, the issue of identification focuses on whether or not there is a unique set of parameters consistent with the data. This question bears directly on the transformation of the variance–covariance matrix of observed variables (the data) into the structural parameters of the model under study. If a unique solution for the values of the structural parameters can be found, the model is considered to be identified, and the parameters are therefore estimable and the model testable. If, on the other hand, a model cannot be identified, it indicates that many sets of very different parameter estimates could fit the data equally well and, in this sense, any one set of values would be arbitrary.

Structural models may be just-identified, overidentified, or underidentified. A **just-identified model** is one in which there is a one-to-one correspondence between the data and the structural parameters. That is to say, the number of data variances and covariances equals the number of parameters to be estimated. However, despite the capability of the model to yield a unique solution for all parameters, the just-identified model is not scientifically interesting because it has no degrees of freedom and therefore can never be rejected. An **overidentified model** is one in which the number of estimable parameters is less than the number of data points (i.e., variances, covariances of the observed variables). This situation results in positive degrees of freedom that allow for rejection of the model, thereby rendering it of scientific use. The aim in SEM, then, is to specify a model such that it meets the criterion of overidentification. Finally, an **underidentified model** is one in which the number of parameters to be estimated exceeds the number of variances and covariances. As such, the model contains insufficient information (from the input data) for the purpose of attaining a determinate solution of parameter estimation; that is, an infinite number of solutions are possible for an underidentified model.

Reviewing the CFA model in Fig. 1.6 again, let us now determine how many data points with which we have to work (i.e., how much information do we have with respect to our data?). As noted earlier, these constitute the variances and covariances of the observed variables; with p variables, there are $p(p + 1)/2$ such elements. Because there are 12 observed variables, this means that we have $12(12 + 1)/2 = 78$ data points. Prior to this discussion of identification, we determined a total of 30 unknown parameters. Thus, with 78 data points and 30 parameters to be estimated, we have an overidentified model with 48 degrees of freedom. It is important to point out, however, that the specification of an overidentified model is a necessary, but not sufficient condition to resolve the identification problem. Indeed, the imposition of constraints on particular parameters can sometimes be beneficial in helping the researcher to attain an overidentified model. An example of such a contraint is illustrated in chapter 5 with the application of a second-order CFA model.

Linked to the issue of identification is the requirement that every latent variable have its scale determined. This requirement arises because these variables are unobserved and therefore have no definite metric scale; this can be accomplished in one of two ways: The **first** approach is tied to specification of the measurement model whereby the unmeasured latent variable is mapped onto its related observed indicator variable. This scaling requisite is satisfied by constraining to some nonzero value (typically 1.0), one factor loading parameter (λ) in each set of loadings designed to measure the same factor. This constraint holds for both the independent (ξ) and dependent (η) latent variables. In other words, in reviewing Fig. 1.6, this means that for one of the three regression paths leading from each SC factor to its set of observed indicators, some fixed value should be specified; this observed variable is termed a **reference** variable.[11] The **second** way to establish the scale of the latent factors is to standardize them directly thereby making their units of measurement equal to their population standard deviations. This approach eliminates the need for a reference variable in the Λ matrix; all factor loading parameters are therefore freely estimated. Jöreskog and Sörbom (1993a, 1993b) claimed that standardization of the latent variables is the more useful and convenient way of assigning the units of a measurement to the latent variables. Thus, by default, LISREL 8 automatically standardizes both the independent and dependent latent factors, if no reference variable is specified for the measurement model. (This default represents a new feature in LISREL 8; only standardization of the independent latent factor, ξ, was possible with LISREL 7.) If, on the other hand, a reference variable is assigned to a particular latent factor, its variance will automatically be estimated by the program. It is important to note, however, that for reasons linked to statistical identification, it is not possible to estimate all factor loading regression coefficients **and** the variance related to the same factor; one must either constrain one parameter in a set of regression coefficients **or** the factor variance.

With respect to the model shown in Fig. 1.6, the scale has been established by constraining to a value of 1.0, the first λ parameter in each set of observed variables. From the schema presented in Fig. 1.6, we can now summarize the structural regression portion of the model as a series of equations:

$$\begin{array}{lll} x_1 = 1.00\,\xi_1 + \delta_1 & x_2 = \lambda_{21}\,\xi_1 + \delta_2 & x_3 = \lambda_{31}\,\xi_1 + \delta_3 \\ x_4 = 1.00\,\xi_2 + \delta_4 & x_5 = \lambda_{52}\,\xi_2 + \delta_5 & x_6 = \lambda_{62}\,\xi_2 + \delta_6 \\ x_7 = 1.00\,\xi_3 + \delta_7 & x_8 = \lambda_{83}\,\xi_3 + \delta_8 & x_9 = \lambda_{93}\,\xi_3 + \delta_9 \\ x_{10} = 1.00\,\xi_4 + \delta_{10} & x_{11} = \lambda_{11,4}\,\xi_4 + \delta_{11} & x_{12} = \lambda_{12,4}\,\xi_4 + \delta_{12} \end{array} \tag{1.13}$$

[11]Although the decision as to which parameter to constrain is purely an arbitrary one, the measure having the highest reliability is recommended, if this information is known; the value to which the parameter is constrained, is also arbitrary.

Second-Order CFA Model

In our previous factor analytic model, we had four factors (ASC, SSC, PSC, ESC) that operated as independent variables; each could be considered to be one level, or one unidirectional arrow away from the observed variables, and therefore served as first-order factors. However, it may be the case that the theory argues for some higher level factor that is considered accountable for the lower order factors. Basically, the number of levels, or unidirectional arrows that the higher order factor is removed from the observed variables determines whether a factor model is labeled as being second-order, third-order or some higher order; as noted earlier, only a second-order model will be considered here. Accordingly, let us examine the pictorial representation of this model in Fig. 1.7.

Although this model has essentially the same first-order factor structure as the one shown in Fig. 1.6, it differs in that a higher order general self-concept (GSC) factor is hypothesized as accounting for, or explaining all of the covariances among the first-order factors. As such, general SC is termed the second-order factor. It is important to take particular note of the fact that general SC does not have its own set of measured indicators; rather, it is linked indirectly to those measuring the lower order factors. Let us now take a closer look at the parameters to be estimated for this second-order model.

Because LISREL assigns different notation to parameters that are estimated in all-X versus all-Y measurement models, the factorial structure portrayed in Fig. 1.7 is readily distinguishable from the one shown in Fig. 1.6, despite the fact that the first-order structure is exactly the same. Indeed, several distinctive features of this second-order model set it apart from the first-order model shown in Fig. 1.6. Let us now take a closer look at these characteristics.

It is likely that the first thing you noticed about the model in Fig. 1.7 was the change in notation for the model parameters. This modification reflects the fact that the four SC factors (ASC, SSC, PSC, ESC) are now dependent, rather than independent variables in the model because they are presumed to be explained by the higher order factor of general SC; the latter therefore serves as the one and only independent factor (ξ_1). Relatedly, because the first-order SC factors operate as dependent variables, the notation of their observed indicators and measurement errors is consistent with an all-Y model.

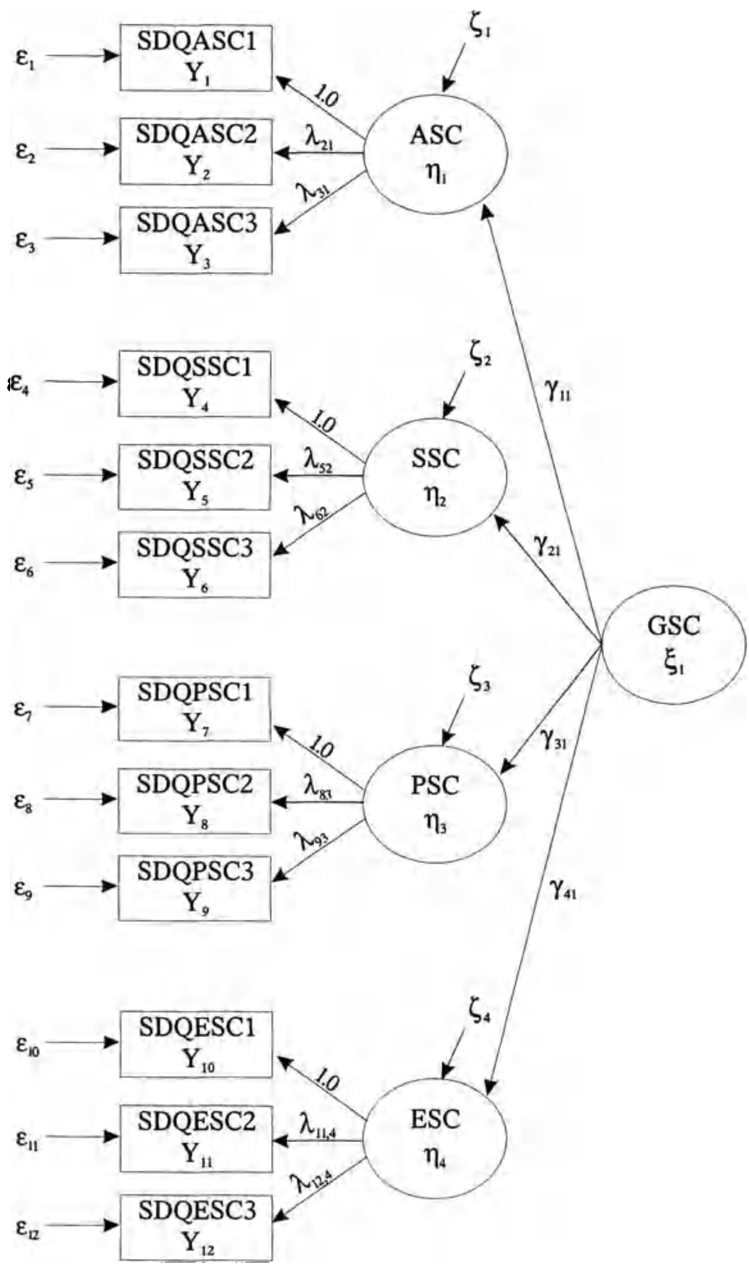


FIG. 1.7. Hypothesized Second-order CFA Model with LISREL Notation.

Relative to the altered parameter notation, a second feature of this model is the presence of single-headed arrows leading from the second-order factor (ξ_1) to each of the first-order factors (η_1–η_4). These regression paths (γ_{11}, γ_{21}, γ_{31}, γ_{41}) represent second-order factor loadings, and all are freely estimated.[12] Recall, however, that for reasons linked to the statistical identification issue, a constraint must be placed either on one of the regression paths, or on the variance of an independent factor; both parameters cannot be estimated. Because the impact of general SC on each of the lower-order SC factors is of primary interest in second-order CFA models, the variance of the higher-order factor is necessarily constrained to equal 1.0.

A third aspect of this second-order model that perhaps needs amplification is the initial appearance that the first-order factors operate as both independent and dependent variables. This, however, is not so; variables can serve either as independent or as dependent variables in a model, but not both.[13] Because the first-order factors function as dependent variables, this necessarily means that their variances and covariances are no longer estimable parameters in the model; indeed, such variation is presumed to be accounted for by the higher order factor. In comparing Figs. 1.6 and 1.7, then, you will note that there are no longer two-headed curved arrows linking the first-order SC factors, thereby indicating that neither their factor covariances nor variances are to be estimated.

Finally, the prediction of each of the first-order factors from the second-order factor is presumed not to be without error. Thus, a residual error term is associated with each of the lower level factors (ζ_1–ζ_4).

As a first step in determining whether this second-order model is identified, we now sum the number of parameters to be estimated; as such we have: 8 first-order regression coefficients (λ_{21}, λ_{31}, λ_{52}, λ_{62}, λ_{83}, λ_{93}, $\lambda_{11,4}$, $\lambda_{12,4}$), 4 second-order regression coefficients (γ_{11}–γ_{44}), 12 measurement error variances (ε_1–ε_{12}), and 4 residual error terms (ζ_1–ζ_4), for a total of 28. Given that there are 78 pieces of information in the sample variance–covariance matrix, we conclude that this model is overidentified with 50 degrees of freedom.

---

[12]As noted earlier, with respect to the first-order regression coefficients, the first number of the subscript represents the dependent variable whereas the second one represents the independent variable.

[13]In SEM, once a variable has an arrow pointing at it, thereby targeting it as a dependent variable, it maintains this status throughout the analyses.

Before leaving this identification issue, however, a word of caution is in order. With complex models in which there may be several levels of latent variable structures, it is wise to visually check each level separately for evidence that identification has been attained. For example, although we know from our initial CFA model that the first-order level is identified, it is quite possible that the second-order level may indeed be underidentified. Because the first-order factors function as indicators of (i.e., the input data for) the second-order factor, identification is easy to assess. In the present model, we have four factors, thereby giving us 10 ([4 x 5]/2) pieces of information from which to formulate the parameters of the higher order structure. According to the model depicted in Fig. 1.7, we wish to estimate 8 parameters, thus leaving us with 2 degrees of freedom, and an overidentified model. However, suppose that we only had three first-order factors. We would then be left with a just-identified model as a consequence of trying to estimate six parameters from six (3[3 + 1])/2) pieces of information. In order for such a model to be tested, additional constraints would need to be imposed. Finally, let us suppose that there were only two first-order factors; we would then have an underidentified model because there would be only three pieces of information, albeit four parameters to be estimated. Although it might still be possible to test such a model, given further restrictions on the model, the researcher would be better advised to reformulate his or her model in light of this problem (see Rindskopf & Rose, 1988).

Considering the model in Fig. 1.7, let us now write the series of regression statements that summarize its configuration. As such, we need to address two components of the model—the higher order factor structure (represented by a structural model in the LISREL sense), and the lower order factor structure (represented by the measurement model in the LISREL sense). These are:

$$\eta_1 = \gamma_{11}\xi_1 + \zeta_1 \quad \eta_2 = \gamma_{21}\xi_1 + \zeta_2 \quad \eta_3 = \gamma_{31}\xi_1 + \zeta_3 \quad \eta_4 = \gamma_{41}\xi_1 + \zeta_4$$

$$\begin{array}{lll}
y_1 = 1.00\,\eta_1 + \varepsilon_1 & y_2 = \lambda_{21}\,\eta_1 + \varepsilon_2 & y_3 = \lambda_{31}\,\eta_1 + \varepsilon_3 \\
y_4 = 1.00\,\eta_2 + \varepsilon_4 & y_5 = \lambda_{52}\,\eta_2 + \varepsilon_5 & y_6 = \lambda_{62}\,\eta_2 + \varepsilon_6 \\
y_7 = 1.00\,\eta_3 + \varepsilon_7 & y_8 = \lambda_{83}\,\eta_3 + \varepsilon_8 & y_9 = \lambda_{93}\,\eta_3 + \varepsilon_9 \\
y_{10} = 1.00\,\eta_4 + \varepsilon_{10} & y_{11} = \lambda_{11,4}\,\eta_4 + \varepsilon_{11} & y_{12} = \lambda_{12,4}\,\eta_4 + \varepsilon_{12}
\end{array} \tag{1.14}$$

Just as we were able to capture the essence of the regression statements describing the first-order CFA model in a single matrix equation (see Equation 1.10), so we can for this second-order CFA model. Consistent with the two levels of factor structure in this model, however, the matrix equation necessarily comprises two separate statements. We turn first to the higher order structure which can summarized as:

$$\eta = \Gamma\xi + \zeta^{14} \qquad (1.15)$$

Because Equation 1.15 representing the structural model is somewhat tricky due to the change in notation related to the ζs, let us review its expansion in two stages—expressed first as a set of vectors, and expressed second in its expanded matrix form. We turn now to the equation in vector form.

$$\begin{array}{ccccccc} \eta & & \Gamma & \xi & & \zeta & \\ \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{bmatrix} & = & \begin{bmatrix} \gamma_{11} \\ \gamma_{21} \\ \gamma_{31} \\ \gamma_{41} \end{bmatrix} & \begin{bmatrix} \xi_1 \end{bmatrix} & + & \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \end{bmatrix} & (1.16) \end{array}$$

Before rewriting this equation to show the expansion of the ζ vector, it is worthwhile to note two important points. **First**, as shown earlier for the Λ matrix, Γ is also a regression matrix and, as such, can be classified as a full matrix. However, within the context of a higher order CFA model, it serves as a special case where there is only one ξ, albeit more than one η. As a consequence, its elements remain described in vector form. To understand better how this works, recall that the order of the Γ matrix is m x n, where m represents the number of endogenous (ηs) factors, and n, the number of exogenous (ξs) factors; in addition, m x n represents a row $\times$ column orientation. Thus, with respect to the model shown in Fig. 1.7, we know that the Γ matrix will be described by four rows (the ηs) and one column (the ξ) thereby resulting in a vector.

The **second** point worthy of note pertains to the residual disturbance term which, as shown on the diagram (Fig. 1.7), and in Equation 1.16, is represented by the Greek letter zeta (ζ). This variable, however, is estimated in

[14]Had the model involved a third level (i.e., a third-order CFA model), then the equation would involve the regression of the second-order factors on the third-order factor, which would then have involved the Beta matrix; the equation would then be: $\eta = B\eta + \Gamma\xi + \zeta$. However, this expansion goes beyond the intent of the present book.

the Psi (Ψ) matrix; thus, the elements of this matrix are ψs, and *not* ζs. We turn now to Equation 1.17 showing the expanded Ψ matrix:

$$
\begin{array}{ccccccc}
\eta & & \Gamma & \xi & & \Psi & \\
\begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} & = & \begin{bmatrix} \gamma_{11} \\ \gamma_{21} \\ \gamma_{31} \\ \gamma_{41} \end{bmatrix} & \begin{bmatrix} \xi_1 \end{bmatrix} & + & \begin{bmatrix} \psi_{11} & & & \\ 0 & \psi_{22} & & \\ 0 & 0 & \psi_{33} & \\ 0 & 0 & 0 & \psi_{44} \end{bmatrix} & (1.17)
\end{array}
$$

Turning to the lower order factor structure, we see first the summary matrix equation (Equation 1.18), followed by its expansion (Equation 1.19) showing the elements of each matrix. However, due to space restrictions, and because the error variance-covariance matrix ($\Theta\varepsilon$) was illustrated for the first-order CFA model, measurement error in this equation is represented here only in vector form (ε). By now, this equation should look familiar to you because it follows the same pattern as the first-order CFA structure based on an all-Y model:

$$Y = \Lambda_y \eta + \varepsilon \qquad (1.18)$$

$$
\begin{array}{cccccccc}
Y & & \Lambda y & \eta & & \varepsilon & \\
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \\ y_{11} \\ y_{12} \end{bmatrix} & = & \begin{bmatrix} \lambda_{11} & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 \\ \lambda_{31} & 0 & 0 & 0 \\ 0 & \lambda_{42} & 0 & 0 \\ 0 & \lambda_{52} & 0 & 0 \\ 0 & \lambda_{62} & 0 & 0 \\ 0 & 0 & \lambda_{73} & 0 \\ 0 & 0 & \lambda_{83} & 0 \\ 0 & 0 & \lambda_{93} & 0 \\ 0 & 0 & 0 & \lambda_{10,4} \\ 0 & 0 & 0 & \lambda_{11,4} \\ 0 & 0 & 0 & y_{12,4} \end{bmatrix} & \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} & + & \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \\ \varepsilon_{11} \\ \varepsilon_{12} \end{bmatrix} & (1.19)
\end{array}
$$

In summary, and for sake of completion, I wish to note that the general LISREL equation for a second-order CFA model derives from a combination of both Equation 1.15 (representing the higher order structural model), and Equation 1.18 (representing the lower-order measurement model). The product of this combination, then, is:

$$y = \Lambda_Y (\Gamma\xi + \zeta) + \varepsilon \qquad (1.20)$$

THE LISREL FULL STRUCTURAL EQUATION MODEL

In contrast to a first-order CFA model, which comprises only a measurement component, and a second-order CFA model, for which the higher order level is represented by a reduced form of a structural model, the full structural equation model encompasses both a measurement and a complete structural model. Accordingly, the full model embodies a system of variables whereby latent factors are regressed on other factors as dictated by theory, as well as on the appropriate observed measures. In other words, in the full SEM, certain latent variables are connected by one-way arrows, the directionality of which reflects hypotheses bearing on the causal structure of variables in the model. For a clearer conceptualization of this model, let us examine the relatively simple structure presented in Fig. 1.8.

The structural component of this model represents the hypothesis that a child's self-confidence (SCONF) derives from his or her self-perception of overall social competence (SSC; social SC) which, in turn, is influenced by the child's perception of how well he or she gets along with family members (SSCF), as well as his or her peers at school (SSCS). The measurement component of the model shows each of the SC factors to have three indicator measures, and the self-confidence factor to have two.

Turning first to the structural part of the model, we can see that there are four factors; the two independent factors (ξ_1, ξ_2) are postulated as being correlated with each other, as indicated by the curved two-way arrow joining them, (ϕ_{21}), but they are linked to the other two factors by a series of regression paths, as indicated by the unidirectional arrows. Because the factors SSC (η_1) and SCONF (η_2) have one-way arrows pointing at them, they are easily identified as dependent variables in the model. Residual errors associated with the regression of η_1 on both ξ_1 and ξ_2, and the regression of η_2 on η_1, are captured by the disturbance terms ζ_1 and ζ_2 , respectively.

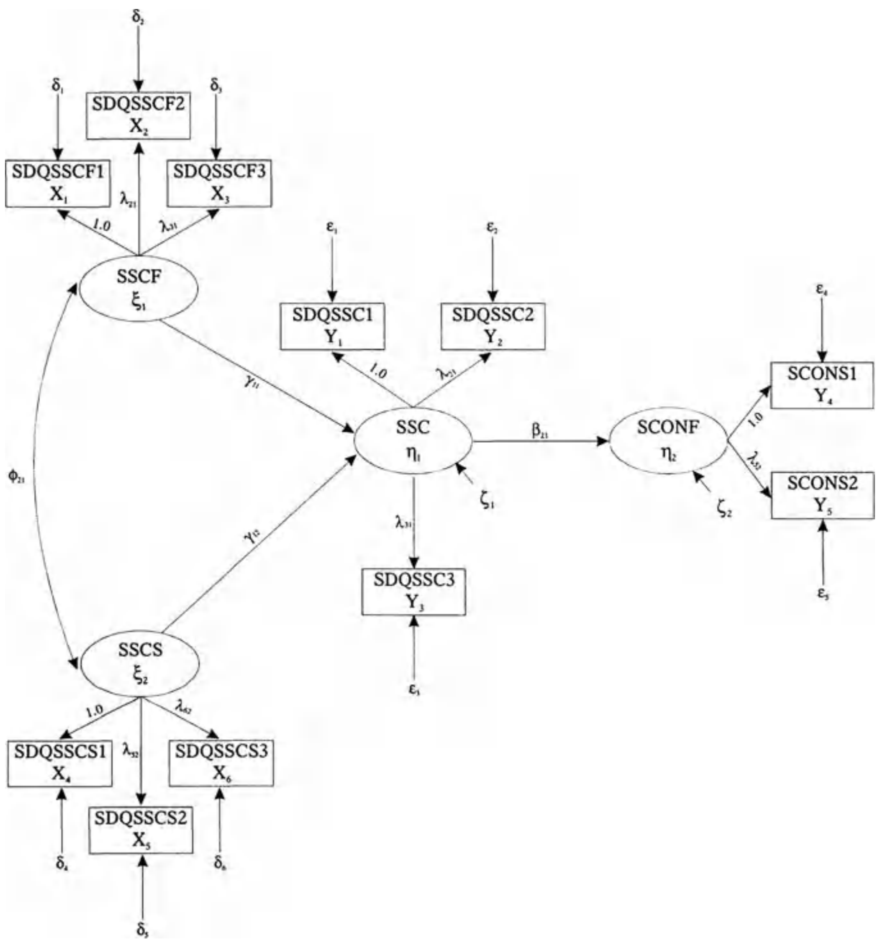


FIG. 1.8. Hypothesized Full Structural Equation Model with LISREL Notation.

Finally, because one path from each of the two independent factors (ξ_1, ξ_2) to their respective indicator variables is fixed to 1.0, their variances can be freely estimated; variances of the dependent variables (η_1, η_2), however, are not parameters in the model.

By now, you likely feel fairly comfortable in interpreting the measurement portion of the model, and thus, substantial elaboration is not necessary here. As usual, associated with each observed measure is an error term, the variance of which is of interest. (Because the observed measures technically operate as dependent variables in the model, as indicated by the arrows pointing towards them, their variances are not estimated.) Finally, to establish the scale for each unmeasured factor in the model (and for purposes of identification), the first of each set of regression paths is fixed to 1.0; note again that path selection for the imposition of this constraint was purely arbitrary.

For this, our last example, let us again determine if we have an identified model. Given that we have 11 observed measures, we know that we have 66 (11[11 + 1]/2) pieces of information from which to derive the parameters of the model. Counting up the unknown parameters in the model, we see that we have 24 parameters to be estimated: 7 measurement regression paths (4 λ_xs and 3 λ_ys; the factor loadings); 3 structural regression paths (γ_{11}, γ_{12}, β_{11}); 11 error variances (6 δs and 5 εs); 2 residual error variances (ζ_1, ζ_2); 1 covariance (ϕ_{21}). We therefore have 42 (66 – 24) degrees of freedom and, as a consequence, an overidentified model.

Our final task is to translate this full model into a set of equations. With the structural model being summarized first, followed by the measurement model, these equations are:

$$\eta_1 = \gamma_{11}\xi_1 + \gamma_{12}\xi_2 + \zeta_1; \qquad \eta_2 = \beta_{21}\eta_1 + \zeta_2;$$

$$\begin{array}{lll} x_1 = 1.00\,\xi_1 + \delta_1; & x_2 = \lambda_{21}\,\xi_1 + \delta_2; & x_3 = \lambda_{31}\,\xi_1 + \delta_3; \\ x_4 = 1.00\,\xi_2 + \delta_4; & x_5 = \lambda_{52}\,\xi_2 + \delta_5; & x_6 = \lambda_{62}\,\xi_2 + \delta_6; \\ y_1 = 1.00\,\eta_1 + \varepsilon_1; & y_2 = \lambda_{21}\,\eta_1 + \varepsilon_2; & y_3 = \lambda_{31}\,\eta_1 + \varepsilon_3; \\ y_4 = 1.00\,\eta_2 + \varepsilon_4; & y_5 = \lambda_{52}\,\eta_2 + \varepsilon_5; & \end{array} \tag{1.21}$$

Because (a) by now the equations representing the measurement part of the model likely present no problem to you, and (b) the beta matrix (B) is new to you, further elaboration is limited to the structural portion of the model (i.e., the structural model). As such, the two structural statements in Equation 1.21 can be summarized more compactly as:

$$\eta = B\eta + \Gamma\xi + \zeta \tag{1.22}$$

Equation 1.22 therefore represents the general LISREL model for a full structural equation model. Now let us focus on the decomposition of this equation. Recall that B represents a *m* x *m* regression matrix that relates the *m* endogenous factors (ηs) to one another. Thus, because there are two

Application 1: Testing the Factorial Validity of a Theoretical Construct (First-Order CFA Model)

Business: Nelson, J. E. , Duncan, C. P. , & Kiecker, P. L. (1993). Toward an understanding of the distraction construct in marketing. *Journal ofBusiness Research* , 26,201–221.

Medicine: Lobel, M. & Dunkel-Schetter, C. (1990). Conceptualizing stress to study effects on health: Environmental, perceptual, and emotional components. *Anxiety Research* , 3,213–230.

Sociology: Shu, X. , Fan, P. L. , Li, X. , & Marini, M. M. (1996). Characterizing occupations with data from the Dictionary of occupational titles. *Social Science Research* , 25, 149–173.

Application 2: Testing the Factorial Validity of Scores From a Measuring Instrument (First-Order CFA Model)

Education: O'Grady, K. E. (1989). Factor structure of the WISC-R. *Multivariate Behavioral Research* , *24* , 177–193.

Kinesiology: Pelletier, L. G. , Vallerand, R. J. , Briere, N. M. , Tuson, K. M. , & Blais, M. R. (1995). The Sport Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in sport. *Journal of Sport and Exercise Psychology* , *17* , 35–53.

Medicine: Coulton, C. J. , Hyduk, C. M. , & Chow, J. C. (1989). An assessment of the Arthritis Impact Measurement Scales in 3 ethnic groups. *Journal of Rheumatology* , *16* , 1110–1115.

Application 7: Testing for Invariant Factorial Structure of Scores from a Measuring Instrument (First-order CFA Model)

Education: Dauphinee, T. L. , Schau, C. , & Stevens, J. J. (1997). Survey of Attitudes Towards Statistics: Factor structure and factorial invariance across women and men. *Structural Equation Modeling: A Multidisciplinary Journal* , 4, 129–141

Kinesiology: Li, F. , Harmer, P. , & Acock, A. (1996). The Task and Ego Orientation in Sport Questionnaire: Construct equivalence and mean differences across gender. *Research Quarterly for Exercise and Sport* , 68, 228–238.

Medicine: Cole, D. A. , Peeke, L. G. , & Ingold, C. (1996). Charac-terological and behavioral self-blame in children: Assessment and development considerations. *Development and Psychopathology* , 8, 381–397.

Application 8: Testing for Invariant Latent Mean Structures

Education: Rosen, M. (1995). Gender differences in structure, means, and variances of hierarchically ordered ability dimensions. *Learning and Instruction* , *5* , 37–62.

Education: Marsh, H. W. , & Grayson, D. (1990). Public/Catholic differences in the high school and beyond data: A multi-group structural equation modeling approach to testing mean differences. *Journal of Educational Statistics* , *15* , 199–235.

Gerontology: Schaie, K. W. , Maitland, S. B. , Willis, S. L. , & Intrieri, R. C. (1997). Longitudinal invariance of adult psychometric ability factor structures across seven years. *Psychology and Aging* , *12* , 1–18.

Kinesiology: Li, F. , Harmer, P. , & Acock, A. (1996). The Task and Ego Orientation in Sport Questionnaire: Construct equivalence and mean differences across gender. *Research Quarterly for Exercise and Sport* , *68* , 228–238.

References

Abramson, L. Y. , Metalsky, G. I. , & Alloy, L. B. (1988). The hopelessness theory of depression: Does the research test the theory? In L. Y. Abramson (Ed.), *Social cognition and clinical psychology: A synthesis* (pp. 33–65). New York: Guilford.

Aitchison, J. , & Silvey, D. C. (1958). Maximum likelihood estimation of parameters subject to restraints. *Annals of Mathematical Statistics* , *29* , 813–828.

Aish, A. M. , & Jöreskog, K. G. (1990). A panel model for political efficacy and responsiveness: An application of LISREL *7* with weighted least squares. *Quality and Quantity* , *19* , 716–723.

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika* , 52, 317–332.

Aldrich, J. H. , & Nelson, F. D. (1984). *Linear probability, logit, andprobit models* . Beverly Hills, CA: Sage.

Anderson, J. C. , & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika* , *49* , 155–173.

Anderson, J. C. , & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin* , *103* , 411–423.

Arbuckle, J. L. (1995). *AMOS user's guide* . Chicago: Smallwaters.

Atkinson, L. (1988). The measurement-statistics controversy: Factor analysis and subinterval data. *Bulletin of the Psychonomic Society* , *26* , 361–364.

Austin, J. T. , & Calderón, R. F. (1996). Theoretical and technical contributions to structural equation modeling: An updated annotated bibliography. *Structural Equation Modeling: A Multidisciplinary Journal* , 105–175.

Babakus, E. , Ferguson, C. E., Jr. , & Jöreskog, K. G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research* , *24* , 222–228.

Bacharach, S. B. , Bauer, S. C. , & Conley, S. (1986). Organizational analysis of stress: The case of elementary and secondary schools. *Work and Occupation* , *13* , 7–32.

Bagozzi, R. P. (1993). Assessing construct validity in personality research: Applications to measures of self-esteem. *Journal of Research in Personality* , *27* , 49–87.

Bagozzi, R. P. , & Yi, Y. (1990). Assessing method variance in multitrait-multimethod matrices: The case of self-reported affect and perceptions at work. *Journal of Applied Psychology* , *75* , 547–560.

Bagozzi, R. P. , & Yi, Y. (1983). Multitrait-multimethod matrices in consumer research: Critique and new developments. *Journal of Consumer Psychology* , *2* , 143–170.

Bandalos, D. L. (1983). Factors influencing cross-validation of confirmatory factor analysis models. *Multivariate Behavioral Research* , *28* , 351–374.

Beek, A.T. , Ward, C. H. , Mendelson, M. , Mock, J. , & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry* , *4* , 561–571.

Benson, J. (1987). Detecting item bias in affective scales. *Educational and Psychological Measurement* , *47* , 55–67.

Benson, J. , & Bandalos, D. L. (1992). Second-order confirmatory factor analysis of the Reactions to Tests Scale with cross-validation. *Multivariate Behavioral Research* , *27* , 459–487.

Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology* , *31* , 419–456.

Bentler, P. M. (1988). Causal modeling via structural equation systems. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 317–335). New York: Plenum.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin* , *107* , 238–246.

Bentler, P. M. (1992). On the fit of models to covariances and methodology to the *Bulletin. Psychological Bulletin* , *112* , 400–404.

Bentler, P. M. (1995). *EQS: Structural equations program manual* . Encino, CA: Multivariate Software Inc.

Bentler, P. M. , & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin* , *88* , 588–606.

Bentler, P. M. , & Bonett, D. G. (1987). This week's citation classic. *Current Contents, Institute for Scientific Information* , *9* , 16.
Bentler, P. M. , & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research* , *16* , 78–117.
Bentler, P. M. , & Wu, E. J. C. (1995). *EQS for Windows: User's guide* . Encino, CA: Multivariate Software Inc.
Biddle, B. J. , & Marlin, M. M. (1987). Causality, confirmation, credulity, and structural equation modeling. *Child Development* , *58* , 4–17.
Bollen, K. A. (1989a). A new incremental fit index for general structural models. *Sociological Methods & Research* , *17* , 303–316.
Bollen, K. A. (1989b). *Structural equations with latent variables* . New York: Wiley.
Bollen, K. A. , & Barb, K. H. (1981). Pearson's *r* and coursely categorized measures. *American Sociological Review* , *46* , 232–239.
Bollen, K. A. , & Long, J. S. (Eds.). (1993). *Testing structural equation models* . Newbury Park, CA: Sage.
Boomsma, A. (1982). The robustness of LISREL against small sample sizes in factor analysis models. In H. Wold & K. Jöreskog (Eds.), *Systems under indirect observation* (pp. 149–173). New York: Elsevier North-Holland.
Boomsma, A. (1985). Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation. *Psychometrika* , *50* , 229–242.
Bozdogan, H. (1987). Model selection and Akaike's information criteria (AIC): The general theory and its analytical extensions. *Psychometrika* , *52* , 345–370.
Breckler, S. J. (1990). Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin* , *107* , 260–271.
Brookover, W. B. (1962). *Self-Concept of Ability Scale* . East Lansing, MI: Educational Publication Services.
Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (Ed.), *Topics in applied multivariate analysis* (pp. 72–141). Cambridge, England: Cambridge University Press.
Browne, M. W. (1984a). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology* , *37* , 62–83.
Browne, M. W. (1984b). The decomposition of multitrait-multimethod matrices. *The British Journal of Mathematical and Statistical Psychology* , *37* , 1–21.
Browne, M. W. , & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research* , *24* , 445–455.
Browne, M. W. , & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 445–455). Newbury Park, CA: Sage.
Browne, M. W. , Mels, G. , & Coward, M. (1994). *Path analysis: RAMONA: SYSTAT for DOS: Advanced applications (version 6)* . Evanston, IL: SYSTAT.
Byrne, B. M. (1986). Self-concept/academic achievement relations: An investigation of dimensionality, stability, and causality. *Canadian Journal of Behavioural Science* , *18* , 173–186.
Byrne, B. M. (1988a). Adolescent self-concept, ability grouping, and social comparison: Reexamining academic track differences in high school. *Youth and Society* , *20* , 46–67.
Byrne, B. M. (1988b). Measuring adolescent self-concept: Factorial validity and equivalency of the SDQ III across gender. *Multivariate Behavioral Research* , *24* , 361–375.
Byrne, B. M. (1988c). The Self Description Questionnaire HI: Testing for equivalent factorial validity across ability. *Educational and Psychological Measurement* , *48* , 397–406.
Byrne, B. M. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models* . New York: Springer-Verlag.
Byrne, B. M. (1991). The Maslach Inventory: Validating factorial structure and invariance across intermediate, secondary, and university educators. *Multivariate Behavioral Research* , *26* , 583–605.
Byrne, B. M. (1993). The Maslach Inventory: Testing for factorial validity and invariance across elementary, intermediate, and secondary teachers. *Journal of Occupational and*

Organizational Psychology , *66* , 197–212.
Byrne, B. M. (1994a). Burnout: Testing for the validity, replication, and invariance of causal structure across elementary, intermediate, and secondary teachers. *American Educational Research Journal* , *31* , 645–673.
Byrne, B. M. (1994b). Testing for the factorial validity, replication, and invariance of a measuring instrument: A paradigmatic application based on the Maslach Burnout Inventory. *Multivariate Behavioral Research* , *29* , 289–311.
Byrne, B. M. (1995). Strategies in testing for an invariant second-order factor structure: A comparison of EQS and LISREL. *Structural Equation Modeling. A Multidisciplinary Journal* *2* , 53–72.
Byrne, B. M. (1996). *Measuring selfconcept across the lifespan: Issues and instrumentation* . Washington, DC: American Psychological Association.
Byrne, B. M. (in press). The nomological network of teacher burnout: A literature review and empirically validated model. In M. Huberman & R. Vandenberghe (Eds.), *Understanding and preventing teacher burnout: A sourcebook of international research and practice* . Cambridge, England: Cambridge University.
Byrne, B. M. , & Baron, P. (1993). The Beck Depression Inventory: Testing and cross-validating an hierarchical factor structure for nonclinical adolescents. *Measurement and Evaluation in Counseling and Development* , *26* , 164–178.
Byrne, B. M. , & Baron, P. (1994). Measuring adolescent depression: Tests of equivalent factorial structure for English and French versions of the Beck Depression Inventory. *Applied Psychology: An International Review* , *43* , 33–47.
Byrne, B. M. , Baron, P. , & Balev, J. (1996). The Beck Depression Inventory: Testing for its factorial validity and invariance across gender for Bulgarian adolescents. *Personality and Individual Differences* , *21* , 641–651.
Byrne, B. M. , Baron, P. , & Balev, J. (in press). The Beck Depression Inventory: A cross-validated test of second-order structure for Bulgarian adolescents. *Educational and Psychological Measurement* .
Byrne, B. M. , Baron, P. , & Campbell, T. L. (1993). Measuring adolescent depression: Factorial validity and invariance of the Beck Depression Inventory across gender. *Journal of Research on Adolescence* , *3* , 127–143.
Byrne, B. M. , Baron, P. , & Campbell, T. L. (1994). The Beck Depression Inventory (French version): Testing for gender-invariant factorial structure for nonclinical adolescents. *Journal of Adolescent Research* , *9* , 166–179.
Byrne, B. M. , Baron, P. , Larsson, B. , & Melin, L. (1995). The Beck Depression Inventory: Testing and cross-validating a second-order factorial structure for Swedish nonclinical adolescents. *Behaviour Research and Therapy* , *33* , 345–356.
Byrne, B. M. , Baron, P. , Larsson, B. , & Melin, L. (1996). Measuring depression for Swedish nonclinical adolescents: Factorial validity and equivalence of the Beck Depression Inventory Across Gender. *Scandinavian Journal of Psychology* , *37* , 37–45.
Byrne, B. M. , & Bazana, P. G. (1996). Investigating the measurement of social and academic competencies for early/late preadolescents and adolescents: A multitrait-multimethod analysis. *Applied Measurement in Education* , *9* , 113–132.
Byrne, B. M. , & Goffin, R. D. (1993). Modeling MTMM data from additive and multiplicative covariance structures: An audit of construct validity concordance. *Multivariate Behavioral Research* , *28* , 67–96.
Byrne, B. M. , & Shavelson, R. J. (1986). On the structure of adolescent self-concept. *Journal of Educational Psychology* , 78, 474–481.
Byrne, B. M. , & Shavelson, R. J. (1987). Adolescent self-concept: Testing the assumption of equivalent structure across gender. *American Educational Research Journal* , *24* , 365–385.
Byrne, B. M. , & Shavelson, R. J. (1996). On the structure of social self-concept for pre-, early, and late adolescents. *Journal of Personality and Social Psychology* , *70* , 599–613.
Byrne, B. M. , Shavelson, R. J. , & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin* , *105* , 456–466.
Byrne, B. M. , & Worth Gavin, D. A. (1996). The Shavelson model revisited: Testing for the structure of academic self-concept across pre-, early, and late adolescents. *Journal of Educational Psychology* , *88* , 215–228.

Campbell, D. T. , & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin* , *56* , 81–105.

Carlson, M. , & Mulaik, S. A. (1993). Trait ratings from descriptions of behavior as mediated by components of meaning. *Multivariate Behavioral Research* , *28* , 111–159.

Chou, C. P. , & Bentler, P. M. (1995). Estimates and tests in structural equation modeling. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 37–55). Thousand Oaks, CA: Sage.

Chou, C. P. , Bentler, P. M. , & Satorra, A. (1991). Scaled test statistics and robust standard errors for non-normal data in covariance structure analysis: A Monte Carlo study. *British Journal of Mathematical and Statistical Psychology* , *44* , 347–357.

Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research* , *18* , 115–126.

Cohen, J. (1994). The earth is round (*p* < .05). *American Psychologist* , *49* , 997–1003.

Cole, D. A. , & Maxwell, S. E. (1985). Multitrait-multimethod comparisons across populations: A confirmatory factor analytic approach. *Multivariate Behavioral Research* , *20* , 389–417.

Comrey, A. L. (1992). *A first course in factor analysis* . Hillsdale, NJ; Lawrence Erlbaum Associates.

Cudeck, R. (1988). Multiplicative models and MTMM matrices. *Journal of Educational Statistics* , *13* , 131–147.

Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin* , *105* , 317–327.

Cudeck, R. , & Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research* , *18* , 147–167.

Cudeck, R. , & Henly, S. J. (1991). Model selection in covariance structures analysis and the “problem” of sample size: A clarification. *Psychological Bulletin* , *109* , 512–519.

Curran, P. J. , West, S. G. , & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods* , *1* , 16–29.

Dillon, W. R. , Kumar, A. , & Mulani, N. (1987). Offending estimates in covariance structure analysis: Comments on the causes of and solutions to Heywood cases. *Psychological Bulletin* , *101* , 126–135.

Fornell, C. (1982). *A second generation of multivariate analysis Vol. 1: Methods* . New York: Praeger.

Francis, D. J. , Fletcher, J. M. , & Rourke, B. P. (1988). Discriminant validity of lateral sensorimotor tests in children. *Journal of Clinical and Experimental Neuropsychology* , *10* , 779–799.

Gerbing, D. W. , & Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research* , *11* , 572–580.

Gerbing, D. W. , & Anderson, J. C. (1993). Monte Carlo evaluations of goodness-of-fit indices for structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 40–65). Newbury Park, CA: Sage.

Goffin, R. D. (1987). *The analysis of multitrait-multimethod matrices: An empirical Monte Carlo comparison of two procedures* . Unpublished doctoral dissertation, University of Western Ontario, London, Canada.

Goffin, R. D. (1993). A comparison of two new indices for the assessment of fit of structural equation models. *Multivariate Behavioral Research* , *28* , 205–214.

Gorsuch, R. L. (1983). *Factor analysis* . Hillsdale, NJ: Lawrence Erlbaum Associates.

Green, D. R. (1975). What does it mean to say a test is biased? *Education and Urban Society* , *8* , 33–52.

Harter, S. (1985). *Manual for the Self–perception Profile for Children: Revision of the Perceived Competence Scale for Children* . Denver, CO: University of Denver.

Harter, S. (1990). Causes, correlates, and the functional role of global self-worth: A lifespan perspective. In R.J. Sternberg & J. Kolligian (Eds.), *Competence considered* (pp. 67–97). New Haven CT: Yale University Press.

Hayduk, L. A. (1987). *Structural equation modeling with LISREL: Essentials and advances* . Baltimore: Johns Hopkins University Press.

Hoelter, J. W. (1983). The analysis of covariance structures: Goodness-of-fit indices. *Sociological Methods & Research* , *11* , 325–344.

Hoyle, R. H. (1995a). The structural equation modeling approach: Basic concepts and fundamental issues. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 1–15). Thousand Oaks, CA: Sage.

Hoyle, R. H. (Ed.). (1995b). *Structural equation modeling: Concepts, issues, and applications* . Thousand Oaks, CA: Sage.

Hu, L. T. , & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Thousand Oaks, CA: Sage.

Hu, L. T. , Bentler, P. M. , & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin* , *112* , 351–362.

James, L. R. , Mulaik, S. A. , & Brett, J. M. (1982). *Causal analysis: Assumptions, models, anddata* . Beverly Hills, CA: Sage.

Jöreskog, K. G. (1971a). Simultaneous factor analysis in several populations. *Psychometrika* , *36* , 409–426.

Jöreskog, K. G. (1971b). Statistical analysis of sets of congeneric tests. *Psychometrika* , *36* , 109–133.

Jöreskog, K. G. (1990). New developments in LISREL: Analysis of ordinal variables using polychoric correlations and weighted least squares. *Quality and Quantity* , *24* , 387–404.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (294–316). Newbury Park, CA: Sage.

Jöeskog, K. G. , & Sörbom, D. (1978). *LISREL IV user's guide* . Chicago: National Educational Resouces.

Jöreskog, K. G. , & Sörbom, D. (1979). *Advances in factor analysis and structural equation models* . New York: University Press of America.

Jöreskog, K. G. , & Sörbom, D. (1988). *LISREL 7: A guide to the program and applications* . Chicago: SPSS Inc.

Jöreskog, K. G. , & Sörbom, D. (1989). *LISREL 7 User's reference guide* . Chicago: Scientific Software Inc.

Jöreskog, K. G. , & Sörbom, D. (1993a). *LISREL 8: Structural equation modeling with the SIMPLIS command language* . Chicago, IL: Scientific Software International.

Jöreskog, K. G. , & Sörbom, D. (1993b). *LISREL 8: User's reference guide* . Chicago: Scientific Software International.

Jöreskog, K. G. , & Sörbom, D. (1993c). *PRELIS 2: User's reference guide* . Chicago: Scientific Software International.

Jöreskog, K. G. , & Sörbom, D. (1996a). *LISREL 8: User's reference guide* . Chicago: Scientific Software International.

Jöreskog, K. G. , & Sörbom, D. (1996b). *PRELIS: User's reference guide* . Chicago: Scientific Software International.

Jöreskog, K. G. , & Sörbom, D. (1996c, April). *Structural equation modeling* . Workshop presented for the NORC Social Science Research Professional Development Training Sessions, Chicago.

Jöreskog, K. G. , & Yang, F. (1996). Nonlinear structural equation models: The Kenny-Judd model with interaction effects. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 57–88). Mahwah, NJ: Lawrence Erlbaum Associates.

Kenny, D. A. (1976). An empirical application of confirmatory factor analysis to the multitrait-multimethod matrix. *Journal of Experimental Social Psychology* , *12* , 247–252.

Kenny, D. A. (1979). *Correlation and causality* . New York: Wiley.

Kenny, D. A. , & Kashy, D. A. (1992). Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin* , *112* , 165–172.

Kerlinger, F. N. (1984). *Liberalism and conservatism: The nature and structure of social attitudes* . Hillsdale, NJ: Lawrence Erlbaum Associates.

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement* , *56* , 746–759.

La Du, T. J. , & Tanaka, J. S. (1989). Influence of sample size, estimation method, and model specification on goodness-of-fit assessments in structural equation modeling.

Journal of Applied Psychology , *74* , 625–636.
Lee, S. Y. , & Bender, P. M. (1980). Some asymptotic properties of constrained generalized least squares estimation in covariance structure models. *South African Statistical Journal* , *14* , 121–136.
Leiter, M. P. (1991). Coping patterns as predictors of burnout: The functions of control and escapist coping patterns. *Journal of Organizational Behavior* , *12* , 123–144.
Little, R. J. A. , & Rubin, D. B. (1987). *Statistical analysis with missing data* . New York: Wiley.
Loehlin, J.C. (1992). *Latent variable models: An introduction to factor, path, & structural analyses* . Hillsdale, NJ: Lawrence Erlbaum Associates.
Long, J. S. (1983a). *Confirmatory factor analysis* . Beverly Hills, CA: Sage.
Long, J. S. (1983b). *Covariance structure models: An introduction to LISREL* . Beverly Hills, CA: Sage.
MacCallum, R. C. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin* , *100* , 107–120.
MacCallum, R. C. (1995). Model specification: Procedures, strategies, and related issues. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Newbury Park, CA: Sage.
MacCallum, R. C. , Browne, M. W. , & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods* , *1* , 130–149.
MacCallum, R. C. , Roznowski, M. , Mar, M. , & Reith, J. V. (1994). Alternative strategies for cross-validation of covariance structure models. *Multivariate Behavioral Research* , *29* , 1–32.
MacCallum, R. C. , Roznowski, M. , & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin* , *111* , 490–504.
MacCallum, R. C. , Wegener, D. T. , Uchino, B. N. , & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin* , *114* , 185–199.
Marcoulides, G. A. , & Schumacker, R. E. (Eds.). (1996). *Advanced structural equation modeling-Issues and techniques* . Mahwah, NJ: Lawrence Erlbaum Associates.
Marsh, H. W. (1987). The hierarchical structure of self-concept and the application of hierarchical confirmatory factor analyses. *Journal of Educational Measurement* , *24* , 17–39.
Marsh, H. W. (1988). Multitrait-multimethod analyses. In J. P. Keeves (Ed.), *Educational research methodology, measurement, and evaluation: An international handbook* (pp. 570–578). Oxford, England: Pergamon.
Marsh, H. W. (1989). Confirmatory factor analyses of multitrait-multimethod data: Many problems and a few solutions. *Applied Psychological Measurement* , *15* , 47–70.
Marsh, H. W. (1992a). *Self Description Questionnaire (SDQ) I: A theoretical and empirical basis for the measurement of multiple dimensions of preadolescent self-concept: A test manual and research monograph* . Macarthur, Australia: Faculty of Education, University of Western Sydney.
Marsh, H. W. (1992b). *Self Description Questionnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple dimensions of late adolescent selfconcepv An interim test manual and research monograph* . Macarthur, Australia: Faculty of Education, University of Western Sydney.
Marsh, H. W. , & Bailey, M. (1991). Confirmatory factor analyses of multitrait-multimethod data: A comparison of alternative models. *Applied Psychological Measurement* , *15* , 47–70.
Marsh, H. W. , Balla, J. R. , & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin* , *103* , 391–410.
Marsh, H. W. , & Grayson, D. (1995). Latent variable models of multitrait-multimethod data. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 177–198). Thousand Oaks, CA: Sage.

Marsh, H. W. , & Hocevar, D. (1985). The application of confirmatory factor analysis to the study of self-concept: First and higher order factor structures and their invariance across age groups. *Psychological Bulletin* , *97* , 562–582.
Maslach, C. , & Jackson, S. E. (1981). *Maslach Burnout Inventory manual* . Palo Alto, CA: Consulting Psychologists Press.
Maslach, C. , & Jackson, S. E. (1986). *Maslach Burnout Inventory manual* (2nd ed.). Palo Alto, CA: Consulting Psychologists Press.
McCarthy, M. (1990). The thin ideal, depression, and eating disorders in women. *Behavior Research and Therapy* , *28* , 205–215.
McDonald, R. P. (1985). *Factor analysis and related methods* . Hillsdale, NJ: Lawrence Erlbaum Associates.
McDonald, R. P. , & Marsh, H. W. (1990). Choosing a multivariate model: Noncentrality and goodness-of-fit. *Psychological Bulletin* , *107* , 247–255.
Mislevy, R. J. (1986). Recent developments in the factor analysis of categorical variables. *Journal of Educational Statistics* , *11* , 3–31.
Mueller, R. O. (1996). *Basic principles of structural equation modeling: An introduction to LISREL and EQS* . New York: Springer-Verlag.
Mulaik, S. A. (1972). *The foundations of factor analysis* . New York: McGraw-Hill.
Mulaik, S. A. , James, L. R. , Van Altine, J. , Bennett, N. , Lind, S. , & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin* , *105* , 430–445.
Muthén, B. O. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika* , *49* , 115–132.
Muthén, B. O. (1988). *LISCOMP: Analysis of linear structural equations using a comprehensive measurement model: User's guide* . Chicago: Scientific Software International.
Muthen, B. , & Christoffersson, A. (1981). Simultaneous factor analysis of dichotomous variables in several groups. *Psychometrika* , *46* , 407–419.
Muthen, B. , & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology* , *38* , 171–189.
Petersen, A. C. , & Taylor, B. (1980). The biological approach to adolescence. In J. Adelson (Ed.), *Handbook of adolescent psychology* (pp. 117–155). New York: Wiley.
Pettegrew, L. S. , & Wolf, G. E. (1982). Validating measures of teacher stress. *American Educational Research Journal* , *19* , 373–396.
Rindskopf, D. (1984). Structural equation models: Empirical identification, Heywood cases, and related problems. *Sociological Methods and Research* , *13* , 109–119.
Rindskopf, D. , & Rose, T. (1988). Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research* , *23* , 51–67.
Rock, D. A. , Werts, C. E. , & Flaugher, R. L. (1978). The use of covariance structures for comparing the psychometric properties of multiple variables across populations. *Multivariate Behavioral Research* , *13* , 403–418.
Rozeboom, W. W. (1960). The fallacy of the null hypothesis significance test. *Psychological Bulletin* , *57* , 416–428.
Saris, W. E. , Satorra, A. , & Sorbom, D. (1987). The detection and correction of specification errors in structural equation models. In C. Clogg (Ed.), *Sociological methodology 1987* (pp. 105–130). San Francisco: Jossey-Bass.
Saris, W. E. , & Stronkhorst, H. (1984). *Causal modelling: nonexperimental research: An introduction to the LISREL approach* . Amsterdam: Sociometric Research Foundation.
SAS Institute Inc . (1992). *The CALIS procedure extended user's guide* . Cary, NC: Author.
Satorra, A. , & Bender, P. M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. *In American Statistical Association 1988 proceedings of the business and economics section* (pp. 308–313). Alexandria, VA: American Statistical Association.
Satorra, A. , & Saris, W. E. (1985). Power of the likelihood ratio test in covariance structure analysis. *Psychometrika* , *50* , 83–90.

Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods* , *1* , 115–129.
Schmitt, N. , & Stults, D. M. (1986). Methodology review: Analysis of multitrait-multimethod matrices. *Applied Psychological Measurement* , *10* , 1–22.
Schumacker, R. E. , & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling* . Mahwah, NJ: Lawrence Erlbaum Associates.
Shavelson, R. J. , Hubner, J. J. , & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research* , *46* , 407–441.
Soares, A. T. , & Soares, L. M. (1979). *The Affective Perception Inventory: Advanced level* . Trumbell, CT: ALSO.
Sobel, M. F. , & Bohrnstedt, G. W. (1985). Use of null models in evaluating the fit of covariance structure models. In N. B. Tuma (Ed.), *Sociological methodology 1985* (pp. 152–178). San Francisco: Jossey-Bass.
Sorbom, D. (1974). A general method for studying differences in factor means and factor structures between groups. *British Journal of Mathematical and Statistical Psychology* , *27* , 229–239.
Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research* , *25* , 173–180.
Steiger, J. H. (1994). *Structural equation modeling with SePath: Technical documentation* . Tulsa, OK: STATSOFT.
Steiger, J. H. , & Lind, J. C. (1980, June). *Statistically based tests for the number of common factors* . Paper presented at the Psychometric Society Annual Meeting, Iowa City, IA.
Stine, R. A. (1989). An introduction to bootstrap methods: Examples and ideas. *Sociological Methods and Research* , *8* , 243–291.
Striegel-Moore, S. H. , Silberstein, L. R. , & Rodin, R. (1986). Towards an understanding of risk factors for bulimia. *American Psychologist* , *41* , 246–263.
Sugawara, H. M. , & MacCallum, R. C. (1993). Effect of estimation method on incremental fit indexes for covariance structure models. *Applied Psychological Measurement* , *17* , 365–377.
Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In J. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp.10–39). Newbury Park, CA: Sage.
Tanaka, J. S. , & Huba, G. J. (1984). Confirmatory hierarchical factor analyses of psychological distress measures. *Journal of Personality and Social Psychology* , *46* , 621–635.
Thompson, B. (1995). Exploring the replicability of a study's results: Bootstrap statistics for the multivariate case. *Educational and Psychological Measurement* , *55* , 84–94.
Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Researcher* , *25* , 26–30.
Thompson, B. (1997). The importance of structure coefficients in structural equation modeling confirmatory factor analysis. *Educational and Psychological Measurement* , *57* , 5–19.
Thompson, B. , & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement* , *56* , 197–208.
Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society* , *54* , 426–482.
Werts, C. E. , Rock, D. A. , Linn, R. L. , & Joreskog, K. G. (1976). Comparison of correlations, variances, covariances, and regression weights with or without measurement error. *Psychological Bulletin* , *83* , 1007–1013.
West, S. G. , Finch, J. F. , & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). Thousand Oaks, CA: Sage.
Wheaton, B. (1987). Assessment of fit in overidentified models with latent variables. *Sociological Methods & Research* , *16* , 118–154.

Wheaton, B. , Muthén, B. , Alwin, D. F. , & Summers, G. F. (1977). Assessing reliability and stability in panel models. In D. R. Heise (Ed.), *Sociological methodology 1977* (pp. 84–136). San Francisco: Jossey-Bass.

Widaman, K.F. (1985). Hierarchically tested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement* , *9* , 1–26.

Williams, L. J. , & Holahan, P.J. (1994). Parsimony-based fit indices for multiple-indicator models: Do they work? *Structural Equation Modeling* , *1* , 161–189.

Wothke, W. (1987, April). *Multivariate linear models of the multitrait-multimethod matrix* . Paper presented at the American Educational Research Association Annual Meeting, Washington, DC.

Wothke, W. (1993). Nonpositive definite matrices in structural modeling. In K. A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 256–293). Newbury Park, CA: Sage.

Wothke, W. (1996). Models for multitrait-multimethod matrix analysis. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 7–56). Mahwah, NJ: Lawrence Erlbaum Associates.

Yung, Y. F. , & Bentler, P. M. (1996). Bootstrapping techniques in analysis of mean and covariance structures. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 195–226). Mahwah, NJ: Lawrence Erlbaum Associates.