Structural Equation Modeling with Amos

Basic Concepts, Applications, and Programming
THIRD EDITION

Barbara M. Byrne



Brief Contents

Acknowled	gments xvi Author xxii
Section I: I	ntroduction
Chapter 1	Structural Equation Modeling: The Basics3
Chapter 2	Using the Amos Program16
Section II:	Single-Group Analyses
Confirmator	y Factor Analytic Models
Chapter 3	Application 1: Testing the Factorial Validity of a Theoretical Construct (First-Order CFA Model)69
Chapter 4	Application 2: Testing the Factorial Validity of Scores from a Measurement Scale (First-Order CFA Model)
Chapter 5	Application 3: Testing the Factorial Validity of Scores from a Measurement Scale (Second-Order CFA Model)
Full Latent \	Variable Model
Chapter 6	Application 4: Testing the Validity of a Causal Structure

Section III:	Multiple-Group Analyses	
Confirmator	y Factor Analytic Models	
Chapter 7	Application 5: Testing Factorial Invariance of Scales from a Measurement Scale (First-Order CFA Model)	227
Chapter 8	Application 6: Testing Invariance of Latent Mean Structures (First-Order CFA Model)	263
Full Latent \	Variable Model	
Chapter 9	Application 7: Testing Invariance of a Causal Structure (Full Structural Equation Model)	2 93
Section IV:	Other Important Applications	
Chapter 10	Application 8: Testing Evidence of Construct Validity: The Multitrait-Multimethod Model	311
Chapter 11	Application 9: Testing Change Over Time: The Latent Growth Curve Model	339
Section V: 0	Other Important Topics	
Chapter 12	Application 10: Use of Bootstrapping in Addressing Nonnormal Data	365
Chapter 13	Application 11: Addressing the Issues of Missing Data	393
References		407
	ex	
Subject Inde	ex	433

Contents

Preface	xvi
Acknowledgments	xxi
About the Author	
Section I: Introduction	
Chapter 1 Structural Equation Modeling: The Basics	
Key Concepts	
What Is Structural Equation Modeling?	3
Basic Concepts	
Latent versus Observed Variables	4
Exogenous versus Endogenous Latent Variables	5
The Factor Analytic Model	5
The Full Latent Variable Model	7
General Purpose and Process of Statistical Modeling	
The General Structural Equation Model	9
Symbol Notation	
The Path Diagram	10
Structural Equations	
Nonvisible Components of a Model	12
Basic Composition	
The Formulation of Covariance and Mean Structures	
Notes	15
Chapter 2 Using the Amos Program	16
Key Concepts	
Model Specification Using <i>Amos Graphics</i> (Example 1)	
Amos Modeling Tools	
The Hypothesized Model	
Drawing the Path Diagram	
Model Specification Using <i>Amos Tables View</i> (Example 1)	
Understanding the Basic Components of Model 1	
The Concept of Model Identification	
1	

x Contents

Model Spec	cification Using Amos Graphics (Example 2)	43
The Hy	pothesized Model	43
Drawin	g the Path Diagram	46
Model Spec	cification Using <i>Amos Tables View</i> (Example 2)	51
	cification Using Amos Graphics (Example 3)	
	pothesized Model	
	g the Path Diagram	
	he Amos Default Color for Constructed Models	
	cification Using Amos Tables View (Example 3)	
_ , , , , , , , , , , , , , , , , , , ,		
Section II:	Single-Group Analyses	
Confirmator	y Factor Analytic Models	
,		
Chapter 3	Application 1: Testing the Factorial Validity of a Theoretical Construct (First-Order CFA Model)	69
Key Concer	ots	
	nesized Model	
	1: Self-concept is a 4-Factor Structure	
	vith Amos Graphics	
	Specification	
	pecification	
	tion of Estimates	
	ext Output: Hypothesized 4-Factor Model	
	Summary	
Model V	Variables and Parameters	82
	Evaluation	
	ter Estimates	
	as a Whole	
	Misspecification	
	nalyses	
	2: Self-concept is a 2-Factor Structure	
	d Amos Text Output: Hypothesized 2-Factor Model.	
Hypothesis	3: Self-concept is a 1-Factor Structure	110
Modeling v	vith Amos Tables View	110 111
	VIII / 111105 1110105 VIEW	
1 NOTES		113
Chapter 4	Application 2: Testing the Factorial Validity of	
	Scores from a Measurement Scale (First-Order	
	CFA Model)	
	ots	
Modeling v	vith Amos Graphics	115

The Me	asuring Instrument under Study	116
The Hypotl	nesized Model	117
	l Amos Output: The Hypothesized Model	
	Evaluation	
	nalyses	
	l Amos Output: Model 2	
	l Amos Output: Model 3	
	l Amos Output: Model 4	
	rison with Robust Analyses Based on the	
	ra–Bentler Scaled Statistic	145
Modeling v	vith Amos Tables View	146
Chapter 5	Application 3: Testing the Factorial Validity of	
-	Scores from a Measurement Scale (Second-Order	i
	CFA Model)	
Key Conce	ots	149
	nesized Model	
Modeling v	vith Amos Graphics	152
Selected	d Amos Output File: Preliminary Model	155
Selected	d Amos Output: The Hypothesized Model	161
	Evaluation	
	Based on Continous Versus Categorical Data	
	ical Variables Analyzed as Continuous Variables	
	ical Variables Analyzed as Categorical Variables	
The Amos	Approach to Analysis of Categorical Variables	170
	Bayesian Estimation?	
	tion of Bayesian Estimation	
	vith Amos Tables View	
Note		182
Full Latent	Variable Model	
Chapter 6		
	Causal Structure	
	ots	
	nesized Model	
	vith Amos Graphics	
	ation of Indicator Variables	
Confirm	natory Factor Analyses	189

xii Contents

Selecte	d Amos Output: Hypothesized Model	197
	odel Assessment	
Post Hoc A	nalyses	203
	d Amos Output: Model 2	
Mo	odel Assessment	203
Selecte	d Amos Output: Model 3	204
Mo	odel Assessment	204
	d Amos Output: Model 4	
Mo	odel Assessment	205
	d Amos Output: Model 5	
Mo	odel Assessment	206
Selecte	d Amos Output: Model 6	208
	odel Assessment	
	e Issue of Model Parsimony	
	d Amos Output: Model 7 (Final Model)	
Mo	odel Assessment	210
	eter Estimates	
Modeling	with Amos Tables View	219
Notes		221
Chapter 7	Application 5: Testing Factorial Invariance of Scales from a Measurement Scale (First-Order	
	CFA Model)	227
Key Conce	pts	227
	Multigroup Invariance	
	neral Notion	
The Tes	sting Strategy	230
The Hypot	hesized Model	230
	shing Baseline Models: The General Notion	231
	shing the Baseline Models: Elementary and	
	ndary Teachers	
	with Amos Graphics	
	of Steps in Testing Multigroup Invariance	
	ng for Configural Invariance	238
	ected Amos Output: The Configural Model	
	(No Equality Constraints Imposed)	240
	ng for Measurement and Structural Invariance: The	
Speci	fication Process	244

III. Testing for Measurement and Structural Invariance:	
Model Assessment	252
Testing For Multigroup Invariance: The Measurement	
Model	253
Model Assessment	
Testing For Multigroup Invariance: The Structural	
Model	261
Notes	
Chapter 8 Application 6: Testing Invariance of Latent Mean	
Structures (First-Order CFA Model)	263
Key Concepts	263
Basic Concepts Underlying Tests of Latent Mean Structures	
Estimation of Latent Variable Means	
The Hypothesized Model	
The Baseline Models	
Modeling with Amos Graphics	
The Structured Means Model	
Testing for Latent Mean Differences	
The Hypothesized Multigroup Model	
Steps in the Testing Process	
Selected Amos Output: Model Summary	
Selected Amos Output: Goodness-of-fit Statistics	
Selected Amos Output: Parameter Estimates	
Notes	
Full Latent Variable Model	
1 W 2 M 2 M 1 W 1 W 1 X 1 X 1 X 1 X 1 X 1 X 1 X 1 X	
Chapter 9 Application 7: Testing Invariance of a Causal	
Structure (Full Structural Equation Model)	293
Key Concepts	
Cross-Validation in Covariance Structure Modeling	
Testing for Invariance across Calibration/Validation Samples	
The Hypothesized Model	
Establishing a Baseline Model	
Modeling with Amos Graphics	
Testing for the Invariance of Causal Structure Using the	
Automated Multigroup Approach	303
Selected Amos Output: Goodness-of-fit Statistics	
for Comparative Tests of Multigroup Invariance	305

xiv Contents

Section IV: Other Important Applications

Chapter 10 Application 8: Testing Evidence of Construct	
Validity: The Multitrait-Multimethod Model	311
Key Concepts	311
The Correlated Traits-Correlated Methods Approach to	
MTMM Analyses	313
Model 1: Correlated Traits-Correlated Methods	315
Model 2: No Traits-Correlated Methods	
Model 3: Perfectly Correlated Traits-Freely Correlated Methods	
Model 4: Freely Correlated Traits-Uncorrelated Methods	
Testing for Evidence of Convergent and Discriminant	
Validity: MTMM Matrix-level Analyses	327
Comparison of Models	
Evidence of Convergent Validity	
Evidence of Discriminant Validity	
Testing for Evidence of Convergent and Discriminant Validity:	
MTMM Parameter-level Analyses	328
Examination of Parameters	
Evidence of Convergent Validity	
Evidence of Discriminant Validity	
The Correlated Uniquenesses Approach to MTMM Analyses	
Model 5: Correlated Uniqueness Model	
Notes	
Chapter 11 Application 9: Testing Change Over Time:	
The Latent Growth Curve Model	339
Key Concepts	
Measuring Change in Individual Growth over Time: The General	
Notion	341
The Hypothesized Dual-domain LGC Model	
Modeling Intraindividual Change	
Modeling Interindividual Differences in Change	
Testing Latent Growth Curve Models: A Dual-Domain Model	
The Hypothesized Model	
Selected Amos Output: Hypothesized Model	
Testing Latent Growth Curve Models: Gender as a Time-invariant	
Predictor of Change	354
Notes	361

Section V: Other Important Topics

Chapter 12 Application 10: Use of Bootstrapping in	
Addressing Nonnormal Data	365
Key Concepts	
Basic Principles Underlying the Bootstrap Procedure	
Benefits and Limitations of the Bootstrap Procedure	
Caveats Regarding the Use of Bootstrapping in SEM	
Modeling with Amos Graphics	
The Hypothesized Model	
Characteristics of the Sample	
Applying the Bootstrap Procedure	
Selected Amos Output	
Parameter Summary	
Assessment of Normality	
Parameter Estimates and Standard Errors	
Note	
Chapter 13 Application 11: Addressing the Issues	
of Missing Data	393
Key Concepts	
Basic Patterns of Missing Data	
Common Approaches to Handling Incomplete Data	
Ad Hoc Approaches to Handling Missing Data	
(Not recommended)	396
Theory-based Approaches to Handling Missing Data	
(Recommended)	300
The Amos Approach to Handling Missing Data	
Modeling with Amos Graphics	
The Hypothesized Model	
Selected Amos Output: Parameter and Model Summary	401
Information	404
Selected Amos Output: Parameter Estimates	
Selected Amos Output: Farameter Estimates Selected Amos Output: Goodness-of-fit Statistics	
Note	
INOIC	400
References	407
Author Index	428
Subject Index	433

the observed data fit this restricted structure. Because it is highly unlikely that a perfect fit will exist between the observed data and the hypothesized model, there will necessarily be a differential between the two; this differential is termed the *residual*. The model-fitting process can therefore be summarized as follows:

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, which he termed strictly confirmatory (SC), alternative models (AM), and model generating (MG). In the strictly confirmatory 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 alternative models 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 model generating 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, in this instance, is to locate the source of misfit in the model and to determine a model that better describes the sample data. Jöreskog (1993) 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" (Jöreskog, 1993, 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 case 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 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, 1989a; Kline, 2011; Loehlin, 1992; Long, 1983b; Raykov & Marcoulides, 2000; Saris & Stronkhurst, 1984; Schumacker & Lomax, 2004); (b) in edited books devoted to the topic (e.g., Bollen & Long, 1993; Cudeck, du Toit, & Sörbom, 2001; Hoyle, 1995a; Marcoulides & Schumacker, 1996); and (c) in methodologically oriented journals such as *British Journal of Mathematical and Statistical Psychology, Journal of Educational and Behavioral Statistics, Multivariate Behavioral Research, Psychological Assessment, Psychological Methods, Psychometrika, Sociological Methodology, Sociological Methods & Research, and Structural Equation Modeling.*

The General Structural Equation Model

Symbol Notation

Structural equation models are schematically portrayed using particular configurations of four geometric symbols—a circle (or ellipse), a square (or rectangle), a single-headed arrow, and a double-headed arrow. By convention, circles (or ellipses;) represent unobserved latent factors, squares (or rectangles;) represent observed variables, single-headed arrows () represent the impact of one variable on another, and double-headed arrows (() represent covariances or correlations between pairs of variables. In building a model of a particular structure under study, researchers use these symbols within the framework of four basic configurations, each of which represents an important component in the analytic process. These configurations, each accompanied by a brief description, are as follows:

Path coefficient for regression of an observed variable onto
an unobserved latent variable (or factor)
Path coefficient for regression of one factor onto another
factor
→ Measurement error associated with an observed variable
Residual error in the prediction of an unobserved factor

The Path Diagram

Schematic representations of models are termed *path diagrams* because they provide a visual portrayal of relations that are assumed to hold among the variables under study. Essentially, as you will see later, a path diagram depicting a particular SEM model is actually the graphical equivalent of its mathematical representation whereby a set of equations relates dependent variables to their explanatory variables. As a means of illustrating how the above four symbol configurations may represent a particular causal process, let me now walk you through the simple model shown in Figure 1.1, which was formulated using *Amos Graphics* (Arbuckle, 2015).

In reviewing the model shown in Figure 1.1, we see that there are two unobserved latent factors—math self-concept (MSC) and math achievement (MATH)—and five observed variables—three of which are considered to measure MSC (SDQMSC; APIMSC; SPPCMSC), and two to measure MATH (MATHGR; MATHACH). These five observed variables function as indicators of their respective underlying latent factors.

Associated with each observed variable is an error term (err1–err5), and with the factor being predicted (MATH), a residual term (resid1)²; there is an important distinction between the two. Error associated with observed variables represents *measurement error*, which reflects on their adequacy in measuring the related underlying factors (MSC; MATH). Measurement error derives from two sources: random measurement error (in the psychometric sense) and error uniqueness, a term used to describe error variance arising from some characteristic that is considered to be specific (or unique) to a particular indicator variable. Such error often represents *nonrandom* (or *systematic*) measurement error. *Residual* terms represent error in the prediction of endogenous factors from exogenous factors. For example, the residual term shown in Figure 1.1 represents error in the prediction of MATH (the endogenous factor) from MSC (the exogenous factor).

It is worth noting that both measurement and residual error terms, in essence, represent unobserved variables. Thus, it seems perfectly reasonable

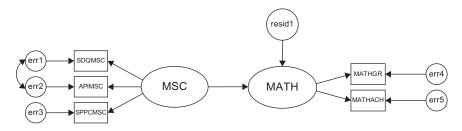


Figure 1.1 A general structural equation model.

that, consistent with the representation of factors, they too should be enclosed in circles. For this reason, then, Amos path diagrams, unlike those associated with most other SEM programs, model these error variables as circled enclosures by default.³

In addition to 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 Figure 1.1, for example, the unidirectional arrow pointing toward the endogenous factor, MATH, implies that the exogenous factor MSC (math self-concept) "causes" math achievement (MATH).4 Likewise, the three unidirectional arrows leading from MSC to each of the three observed variables (SDQMSC, APIMSC, SPPCMSC), and those leading from MATH to each of its indicators, MATHGR and MATHACH, suggest that these score values are each influenced by their respective underlying factors. As such, these path coefficients represent the magnitude of expected change in the observed variables for every change in the related latent variable (or factor). It is important to note that these observed variables typically represent subscale scores (see, e.g., Chapter 8), item scores (see, e.g., Chapter 4), item pairs (see, e.g., Chapter 3), and/or carefully formulated item parcels (see, e.g., Chapter 6).

The one-way arrows pointing from the enclosed error terms (err1–err5) indicate the impact of measurement error (random and unique) on the observed variables, and that from the residual (resid1) indicates the impact of error in the prediction of MATH. Finally, as noted earlier, curved two-way arrows represent covariances or correlations between pairs of variables. Thus, the bidirectional arrow linking err1 and err2, as shown in Figure 1.1, implies that measurement error associated with SDQMSC is correlated with that associated with APIMSC.

Structural Equations

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, 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's turn again to Figure 1.1. We can see that there are six variables with arrows pointing toward them; five represent observed variables (SDQMSC–SPPCMSC; MATHGR, MATHACH), and one represents an unobserved variable (or factor; MATH). Thus, we know that the regression functions symbolized in the model shown in Figure 1.1 can be summarized in terms of six separate equation-like representations of linear dependencies as follows:

MATH = MSC + resid1

SDQMSC = MSC + err1

APIMSC = MSC + err2

SPPCMSC = MSC + err3

MATHGR = MATH + err4

MATHACH = MATH + err5

Nonvisible Components of a Model

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 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 just shown, there is no indication that the variances of the exogenous variables are parameters in the model; indeed, such parameters are essential to all structural equation models. Although researchers must be mindful of this inadequacy of path diagrams in building model input files related to other SEM programs, Amos facilitates the specification process by automatically incorporating the estimation of variances by default for all independent factors.

Likewise, it is equally important to draw your attention to the specified nonexistence of certain parameters in a model. For example, in Figure 1.1, we detect no curved arrow between err4 and err5, which suggests the lack of covariance between the error terms associated with the observed variables MATHGR and MATHACH. Similarly, there is no hypothesized

covariance between MSC and resid1; 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. In the case of both examples cited here, Amos, once again, makes it easy for the novice structural equation modeler by automatically assuming these specifications to be nonexistent. (These important default assumptions will be addressed in Chapter 2, where I review the specification of Amos models and input files in detail.)

Basic Composition

The general SEM 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, then, represents the CFA model described earlier in that it specifies the pattern by which each measure loads on a particular factor. In contrast, the *structural model* defines relations among the unobserved variables. Accordingly, it specifies the manner by which particular latent variables directly or indirectly influence (i.e., "cause") changes in the values of certain other latent variables in the model.

For didactic purposes in clarifying this important aspect of SEM composition, let's now examine Figure 1.2, in which the same model presented in Figure 1.1 has been demarcated into measurement and structural components.

Considered separately, the elements modeled within each rectangle in Figure 1.2 represent two CFA models. The enclosure of the two factors within the ellipse represents a full latent variable model and thus would not be of interest in CFA research. The CFA model to the left of the diagram represents a one-factor model (MSC) measured by three observed variables (SDQMSC–SPPCMSC), whereas the CFA model on the right represents a one-factor model (MATH) measured by two observed variables (MATHGR, MATHACH). In both cases, the regression of the observed variables on each factor and the variances of both the factor and the errors of measurement are of primary interest; the error covariance would be of interest only in analyses related to the CFA model bearing on MSC.

It is perhaps important to note that, although both CFA models described in Figure 1.2 represent first-order factor models, second- and higher-order CFA models can also be analyzed using Amos. Such hierarchical CFA models, however, are less commonly found in the literature