Preface

Structural equation modeling (SEM) has come of age. As recently as the 1980s, SEM was perceived by many students and researchers in the social and behavioral sciences as virtually intractable—esoteric notation, difficult-to-use computer programs, and little published guidance targeted to would-be users with basic graduate-level training in statistical methods. The traditional LISREL notation system is now more familiar to many students and researchers, and alternative, more intuitive systems have been developed. Whereas there was once only LISREL for mainframe computers, there now are multiple computer programs for implementing SEM that run on desktop computers using syntax that does not require knowledge of matrix algebra. And one could now fill a shelf with textbooks and edited volumes devoted to SEM and SEM-related topics. A statistical approach that once was accessible only to social and behavioral scientists with advanced coursework in statistical methods and strong computing skills is now part of the methodological mainstream.

Despite the growing literature on SEM targeted to graduate students and researchers, there has, to date, been no single resource that offers broad and deep coverage of both the mechanics of SEM and specific SEM strategies and applications. This handbook is that resource. If offers comprehensive coverage of SEM, beginning with background issues, continuing through statistical underpinnings and steps in implementation, then moving into basic and advanced applications of SEM. In a single volume, it offers virtually complete coverage of SEM and its use.

The book is intended for advanced graduate students and postgraduate researchers with graduate-level training in applied statistical methods that include multiple regression analysis and at least basic coverage of factor analysis. The structure of the book, described below, is designed to lead readers from basic, foundational material through coverage of the increasing number of modeling approaches and model types for which SEM is appropriate. As such, the book could serve as the primary textbook for a graduate-level course on SEM. Alternatively, it could serve as a resource for students and researchers who have completed their statistical training but need to know more about how SEM works and how it could be used in their work. In either case, the goal is to provide coverage at a level suitable for graduate students and postgraduate researchers who have had basic statistical training typical of the social and behavioral sciences.

To that end, the authors, of whom many are at the forefront of developments related to the topic about which they have written, were challenged with producing focused chapters that balance sophistication and accessibility. The level of sophistication necessarily varies but, generally, increases from early to later chapters. Some chapters in the last part of the book cover highly specialized applications at a level that assumes a solid grasp of the statistical underpinnings of SEM. Yet, even in these chapters, the authors have provided conceptually oriented descriptions and revealing examples. Many of the chapters offer fully explicated analyses, including access to data and syntax files for readers interested in trying their hand at reproducing the authors' results. (These can be accessed at the website for the *Handbook*: *www.handbookofsem.com*.) The result is a set of chapters that provide up-to-date, accessible, and practical coverage of the full array of SEM topics.

The 40 chapters are arrayed in five parts designed to move the reader from foundational material through the statistical underpinnings and practicalities of using SEM, to basic and advanced applications. The chapters in Part I provide important background, beginning with a historical account of key advances and including material on path diagrams, latent variables, causality, and simulation methods. Part II is the "nuts-and-bolts" portion of the book, comprising chapters on assumptions, specification, estimation, statistical power, fit, model modification, and equivalent models. Also included is a chapter on the use of categorical data in SEM. Part III, a practically oriented "how-to" portion of the book, covers preparing data, managing missing data, bootstrapping, choosing computer software, and writing the SEM research report. Parts IV and V cover the many types of models and data for which SEM is appropriate. Part V includes chapters on "basic" applications—those that have been in use for the longest period of time and/or serve as building blocks for newer, more complex or specialized applications. These include confirmatory factor analysis; models of mediation and moderation; models of longitudinal data; models focused on means; models for the construction and development of measurement scales; and models for evaluating measurement equivalence for different populations. Part V includes a dozen chapters that showcase the newest and most specialized SEM models and modeling strategies. Some chapters focus on the use of SEM to model data generated by relatively new methods such as brain imaging, genotyping, and geocoding. Others cover strategies for more general types of data that pose particular challenges but offer unique opportunities; these include multilevel data, categorical measurement data, longitudinal growth data, data from intensive longitudinal assessments, dyadic data, and data from heterogeneous samples for which the source of heterogeneity is not observed. Also included in Part V are chapters on emerging strategies—Bayesian methods and automated model specification.

Together, these parts form a coherent whole that provides comprehensive, in-depth, coverage of SEM in a style appropriate for advanced graduate students and researchers in the social and behavioral sciences.

CHAPTER 1

Introduction and Overview

Rick H. Hoyle

Structural equation modeling (SEM) is a growing family of statistical methods for modeling the relations between variables. Although the data from which these relations are modeled and estimated are observed, models may include variables that are unobserved, or latent. For this reason, SEM has been referred to as latent variable modeling. The primary data for most uses of SEM are covariances, which explains why SEM has also been referred to as covariance structure modeling. And the intent of many uses of SEM is to estimate causal effects between variables, explaining why SEM is sometimes referred to as causal modeling. Regardless of the label, the family of methods referred to as SEM in this handbook is a comprehensive, flexible, and increasingly familiar approach to hypothesis testing and modeling in the social and behavioral sciences.

Unlike more widely used statistical methods such as analysis of variance, multiple regression analysis, and factor analysis, SEM is not yet fully developed. Although the core capabilities of SEM have been well established since the early 1970s and generally accessible to researchers since the early 1980s, new capabilities are being developed and incorporated into computer programs for SEM analyses with regularity (see Matsueda, Chapter 2, this volume, for an informative history of SEM). These emerging capabilities, coupled with powerful and intuitive computer programs for implementing them, have spurred phenomenal growth in the amount and diversity of SEM usage. This handbook is a response to that growth. Our goal is to provide detailed coverage of SEM, beginning with foundational concerns and moving through an impressive array of modeling possibilities.

In this opening chapter, I offer an introduction to SEM that also serves as an overview of the remainder of the handbook. I begin by discussing the relation between SEM and statistical methods with which many readers new to SEM will be familiar. I then provide a brief description of the basic logic of SEM as it typically is used in the social and behavioral sciences. The heart of the chapter is the presentation of an implementation framework that serves as both context for the remainder of the chapter and an outline of the first three parts of the handbook. In the final section of the chapter I succinctly describe types of data and models for which SEM can be profitably used, and point the reader to chapters in the fourth and fifth parts of the book that offer detailed descriptions and demonstrations.

SEM IN RELATION TO OTHER STATISTICAL MODELS

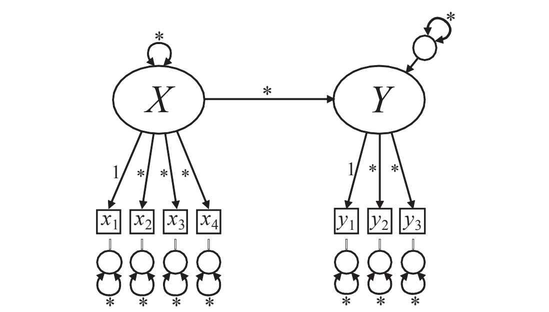
As a linear model concerned with accounting for the relations between variables, SEM is not unrelated to narrower and more familiar statistical models such as analysis of variance (ANOVA), multiple regression analysis, and principal factor analysis. Indeed, any of these analyses could be accomplished, and would yield identical results, using SEM. As such, SEM can be described as a generalization, integration, and extension of these familiar models.

Consider, for example, tests involving means. In the most limited case, a single mean estimated from a sample is compared against a population value, often zero, and the difference tested for significance. This test can be usefully generalized to the situation in which both means are estimated from samples, which may be independent or dependent; alternatively, the means may come from two observations of the same sample. The same comparison could be made using ANOVA, which offers the additional benefit of allowing for both more than two means and means generated by more than one factor. The number of levels a factor might reasonably take on in ANOVA is relatively small, making it unsuitable for independent variables measured on a continuous or quasi-continuous scale such as survey items. Multiple regression analysis can accommodate both traditional ANOVA factors and quantitative measures that take on many values; thus, it has all the capabilities of ANOVA and more. Although both ANOVA and multiple regression analysis can accommodate multiple dependent variables, they are limited in how the relations between those variables are specified. Furthermore, a variable can be an independent or dependent variable, but not both. SEM can accommodate both analytic situations. For instance, a set of variables might be used to predict a pair of outcomes that are correlated, uncorrelated, or related in such a way that one is regressed on the other. In the latter case, one of the dependent variables is also an independent variable in that it is used to predict the other dependent variable.

An alternative path to SEM that highlights additional capabilities begins with the zero-order correlation coefficient, which indexes the commonality between two variables. The degree to which that commonality can be attributed to a common influence can be evaluated using partial correlation analysis, assuming the putative influence has been measured. In the case of three or more variables, this logic can be extended to consider common influences that are not measured using factor analysis. The traditional factor analysis model is referred to as exploratory factor analysis (EFA) because those influences, even in the presence of well-developed hypotheses, cannot be specified a priori. More an inconvenience than a limitation is the fact that an infinite number of factor scores can be derived from the parameters (factor loadings and uniquenesses) estimated by EFA (Steiger & Schönemann, 1978). Finally, EFA requires that uniquenesses be uncorrelated. Within SEM, factors have traditionally been referred to as latent variables and are modeled in a more flexible, mathematically defensible manner that allows for a wide array of models that could not be evaluated using EFA. Applications of SEM that focus exclusively on the relations between latent variables and their indicators are referred to as restricted factor analysis or, more commonly, confirmatory factor analysis (CFA). Both labels are apt because it is the restrictions that CFA requires that make it confirmatory (i.e., subject to statistical testing). Conditional on appropriate restrictions (illustrated below), CFA permits specification and testing of a wide array of factor models.

Although each of these generalizations of basic statistical models is impressive in its own right, it is the integration of the two that constitutes the core strength of SEM. The traditional approach to integrating multiple regression analysis and factor analysis involves factoring a set of indicators of one or more predictors and outcomes, generating factor scores (which, as noted, are indeterminate) or creating unit-weighted composites of the highest-loading indicators, then using these variables as predictors or outcomes. SEM allows for these two components of the analytic strategy to be done simultaneously; that is, the relations between indicators and latent variables and the relations between latent variables are evaluated in a single model.

This integration of regression analysis and factor analysis is depicted in three ways in Figure 1.1. The model in question is one in which an outcome, *Y*, is regressed on a predictor, *X*. *Y* is operationally defined by three indicators, y_1, y_2, and y_3, and *X* is operationally defined by four indicators, x_1, x_2, x_3, and x_4. These indicators could be survey items, total scores on different instruments designed to measure *X* and *Y*, behavioral observations, physical characteristics, or some combination of these and other fallible indicators of the constructs. Regardless of how the values on these indicators were generated, it is assumed that x_1 to x_4 share in common their reflection of construct *X* but not



$$
\begin{array}{ll}
v_1 = 1F_1 + e_1 & e_1,e_1 = * \\
v_2 = *F_1 + e_2 & e_2,e_2 = * \\
v_3 = *F_1 + e_3 & e_3,e_3 = * \\
v_4 = *F_1 + e_4 & e_4,e_4 = * \\
v_5 = 1F_2 + e_5 & e_5,e_5 = * \\
v_6 = *F_2 + e_6 & e_6,e_6 = * \\
v_7 = *F_2 + e_7 & e_7,e_7 = * \\
F_2 = *F_1 + d_2 & F_1,F_1 = * \\
 & d_2,d_2 = *
\end{array}
$$

$$
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda_{21} \\ \lambda_{22} \\ \lambda_{23} \end{bmatrix} [\xi_1] + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix} \qquad \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda_{21} \\ \lambda_{31} \end{bmatrix} [\eta_1] + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}
$$

$$
[\eta_1] = [0][\eta_1] + [\gamma_{11}][\xi_1] + [\zeta_1]
$$

$$
[\varphi_{11}] \qquad [\psi_{11}]
$$

FIGURE 1.1. Alternative depictions of a model.

Y and, conversely, y_1 to y_3 reflect construct *Y* but not *X*. In order to estimate the effect of *X* on *Y* using regression analysis, composite scores would need to be produced, perhaps by summing x_1 to x_4 and y_1 to y_3 or, if the indicators were on different scales, standardizing scores and taking a mean. As illustrated in the top panel of Figure 1.1, the regression portion of the model involves only latent variables (i.e., factors), designated by ovals. These are unobserved forms of *X* and *Y* that reflect the commonality among observed indicators of them, designated by squares. Variance in each indicator is attributable to two unobserved sources: one of the latent variables of interest, *X* or *Y*, and uniqueness, or specificity, designated by the small circles. The straight lines indicate directional effects, and the sharply curved lines indicate variances. The asterisks designate parameters to be estimated. These include factor loadings, uniquenesses, a regression coefficient, a disturbance (regression error of prediction), and the variance of *X*. This approach to depicting a model is called a path diagram, about which detailed information is provided by Ho, Stark, and Chernyshenko (Chapter 3, this volume).

Two additional ways of depicting the same model are shown in the remainder of Figure 1.1. In the middle panel, the model is outlined in a series of equations and "double-label" terms (Bentler & Weeks, 1980). In this notational scheme, observed variables are designated as *v* and latent variables as *F*, uniquenesses as *e*, and disturbances as *d*. As in the path diagram, parameters to be estimated are denoted by asterisks. In the left column are equations; the first seven are measurement equations, which specify the relations between indicators and latent variables, and the last one is a structural equation, which specifies the directional relation between the latent variables. In the right column are variances corresponding to the uniquenesses, the latent predictor, and the disturbance. The scheme draws its name from the approach to designating parameters. The double-label format is evident for the variances. The same format could be used for the equations; for instance, the factor loading for v_2 on F_1 could be written $v_2,F_1 = *$, and the regression coefficient for the structural equation could be written $F_2,F_1 = *$. Note that every asterisk in the path diagram has a counterpart in the double-label depiction.

In the bottom panel of Figure 1.1, the model is depicted using matrix notation, sometimes referred to as LISREL notation in recognition of its use in the original computer program for implementing SEM (Jöreskog & Sörbom, 1999). In this scheme, observed variables are denoted as *x* (independent) or *y* (dependent), whereas variables defined in the model and parameters are denoted by Greek letters. ξ_1 corresponds to *X* and F_1 in the previous panels, and η_1 corresponds to *Y* and F_2. The parameters are now differentiated, with λ corresponding to factor loadings, and δ and ε variances of uniquenesses. The regression coefficient is denoted by γ. The remaining parameters are the variance of ξ_1 and the disturbance, denoted by ϕ and ψ, respectively. As with the scheme illustrated in the middle panel, equations corresponding to the measurement and structural components of the model are distinguished. Unlike the earlier scheme, a further distinction is made between the measurement equations for the independent (ξ_1) and dependent (η_1) latent variables.

In many applications of SEM, the observed variables are assumed to be measured on a continuous scale, and any latent variables are assumed to be continuous as well. Yet variables often are measured coarsely (e.g., 5- or 7-point response scales) and sometimes categorically (e.g., *true–false*), raising questions as to the appropriateness of standard SEM approaches to estimation and testing (see Edwards, Wirth, Houts, & Xi, Chapter 12, this volume). Fortunately, much of the recent expansion of SEM is attributable to the development of models, estimators, and fit statistics for categorical data. This suggests a third path of generalization from simpler models to the general and integrated models for which SEM is appropriate. At the most elemental level is the simple cross-tabulation of categorical variables. More general models for modeling categorical data include logistic regression analysis and latent class analysis. The integration of latent class analysis and factor analysis yields factor mixture modeling and the possibility of categorical latent variables. These latent variables are unobserved categorical variables that reflect homogeneous groups, or classes, of individuals from a population that is heterogeneous with reference to the parameters in a factor model.

To this point, my description of SEM and comparison with other statistical models has focused on modeling the relations between variables. In some cases, however, the hypothesis of interest requires modeling patterns of means or means of latent variables. SEM is appropriate for testing such hypotheses; however, doing so requires moving beyond covariance structure modeling, which typifies most uses of SEM, to models that include a mean structure. This addition allows for the expansion of models such as the one shown in Figure 1.1 to include intercepts in the measurement, and structural equations and means of the latent variables. It also allows for modeling individual patterns of means over time, capturing variability in latent growth variables. When these variables are examined in relation to latent variables that reflect the commonality among sets of indicators (e.g., *X* and *Y* in Figure 1.1), the model includes three components—measurement and structural equations, which together constitute the covariance structure, and the mean structure. The full generality and flexibility of SEM would be evident in a model that includes all three components, and both continuous and categorical observed and latent variables.

BASIC LOGIC AND APPLICATION

The chapters in the second and third parts of this book provide detailed coverage of the basic logic and general application of SEM. I offer a summary in this introductory chapter both as an overview of the contents of Parts II and III, and as background for the chapters that follow in Part I of this handbook.

A fundamental difference between SEM and more familiar statistical models such as ANOVA and multiple regression analysis is the target of parameter estimation. In typical applications of multiple regression analysis, for example, the regression coefficients are estimated using ordinary least squares (OLS). The coefficients define a regression line that minimizes the average squared distance between the individual data points (the target) and the line. Residuals index the degree to which the estimated line misses each data point, that is, the degree of error in predicting the observed data from those estimated by the model. The goal of estimation in SEM is the same: Find values of the parameters that best account for the observed data given a substantively interesting model. A major difference, however, is what constitutes the observed data, or target. In the prototypic application of SEM—as, for example, the model shown in Figure 1.1—the data are the observed covariances between the variables and their variances. The goal of estimation, typically by the maximum likelihood method, is to find values for the parameters that, given the model, maximize the likelihood of the observed data. Stated differently, as with OLS regression, the goal is to minimize the difference between the observed and estimated data, but the observed and estimated data in SEM are variances and covariances. Thus, the residuals are the differences between the observed variances and covariances, and those estimated by the model given the data.

Returning to the model depicted in Figure 1.1, the data are the seven variances of the observed variables plus the 21 covariances between them (easily calculated as $p(p + 1)/2$, where p is the number of observed variables). As with the casewise observed data in OLS regression, the degrees of freedom available for model testing are derived from the number of data points. Unlike degrees of freedom for casewise data, the number of degrees of freedom available for model testing is equal to the total number of data points—28 in this case. As with tests involving casewise data, the number of degrees of freedom for a given test is the number of available degrees of freedom minus the number of parameters to be estimated. Referring either to the top or middle panel of Figure 1.1 and counting asterisks, there are 15 parameters to be estimated, leaving 13 degrees of freedom for tests of model fit. Working from the matrix notation in the lower panel, the same outcome is reached by counting the λ's, δ's, ε's, γ's, ϕ's, and ψ's.

Models such as the one shown in Figure 1.1 are specified by researchers; that is, there is no default model provided by SEM software for covariance matrices based on seven observed variables. A given specification offers a putative explanation for the pattern of observed covariances and reflects the researcher's hypotheses about those relations; it also reflects certain technical constraints necessary to ensure the model can be estimated. When the parameters in a model are estimated from data, they can be used in combination with the data to produce an estimated covariance matrix equivalent to estimated scores on the outcome variable in OLS regression. The difference between the estimated and observed matrices is the residual matrix, which is implicated directly or indirectly in various tests and indices of fit. Generally speaking, a model fits the data when the elements of the residual matrix are uniformly near zero. Models initially specified by researchers often result in one or more residual covariances that are nontrivially different from zero, meaning they are not adequately explained by the model given the data. In such cases, models often are respecified, estimated, and tested, the equivalent of post hoc comparisons in ANOVA. When support is obtained for either an a priori or respecified model, it is interpreted and presented. Each of these steps in the use of SEM is discussed and illustrated in chapters in Parts II and III of this volume. In the next section of this chapter, I present a framework that integrates the steps involved in the implementation of SEM and offer a brief description of each one.

IMPLEMENTATION FRAMEWORK

Despite its flexibility and generality, in practice SEM is nearly always implemented following the same series of discrete steps. In this section, I present an implementation framework that positions these steps in relation to each other, providing context for processing material in the remainder of the book. For each step I provide an overview and refer to the relevant chapters. The framework, shown in diagram form in Figure 1.2, comprises four steps—specification, estimation, evaluation of fit, and interpretation and reporting—that are always followed, and a fifth step—respecification—that is included in most implementations of SEM. Because they are important considerations for how the steps are implemented, I also include the related concerns of data acquisition/preparation and identification; these are shown in Figure 1.2 as boxes connected by dashed lines to one or more of the primary steps in implementation.

SEM can be used with different intents, and it is useful to review them here as context for the presentation of the implementation framework. Specifically, Jöreskog (1993) described three common intents when using SEM. Although somewhat rare in practice, SEM can be used with strictly confirmatory intent. In such cases, a single a priori model is specified and evaluated. Either it provides an acceptable account of the data or it

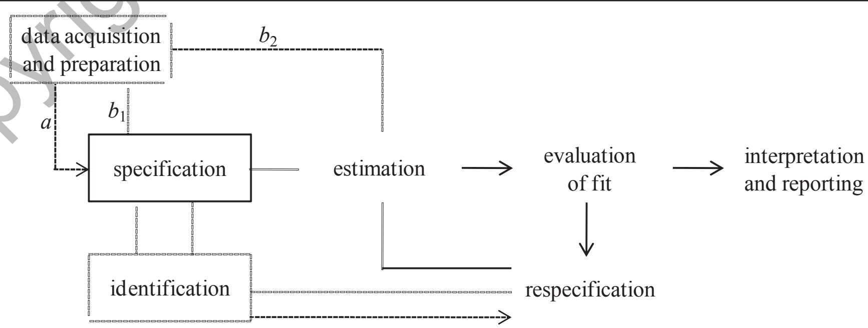


FIGURE 1.2. Steps in the implementation of SEM.

does not. No attempts are made at modifying the model or comparing it with alternative models. Alternatively, the researcher's intent may include both evaluating the fit of a model in an absolute sense and comparing it with alternative models that reflect competing theoretical accounts or offer a more parsimonious account of the data. When SEM is used with the intent of comparing alternative models, these models should be specified a priori and, when possible, specified in such a way that direct statistical comparisons can be made. Finally, the intent of an SEM analysis might be the generation of a model for subsequent evaluation in a strictly confirmatory or alternative models analysis. Although an initial model must be specified, it either follows from results of prior analyses (e.g., multiple regression analysis, factor analyses) of the same data, or offers a sufficiently poor account of the data that it must be modified or abandoned. Many uses of SEM begin with strictly confirmatory or alternative model comparison intent, but they become exercises in model generation when a priori models do not meet fit criteria. At the other extreme, it is possible to begin with a commitment to no particular model and use data mining strategies to generate models (see Marcoulides & Ing, Chapter 40, this volume). With these distinctions in mind, I now turn to an overview of the implementation framework displayed in Figure 1.2.

Specification

The use of SEM always begins with the specification of a model. A *model* is a formal statement of the mechanisms assumed to have given rise to the observed data. Those mechanisms reflect the substantive hypotheses that motivated the analysis, as well as characteristics of the sample and research design. As discussed later in this section, the model also includes features that ensure unique values can be obtained for the parameters to be estimated.

As shown in Figure 1.2, specification can take place either before or after data are acquired and prepared for analysis. The dashed line labeled *a* corresponds to the situation in which specification follows data collection, whereas the line labeled b_1 corresponds to the situation in which data collection follows specification then, as indicated by line b_2, leads to estimation. Again using the model depicted in Figure 1.1 as an example, a researcher might have access to a set of data that includes x_1 to x_4 and y_1 to y_3. These may be data she collected herself but did not collect with this specific model in mind, or data she acquired from a secondary source (e.g., U.S. Census data). Note that in this situation the options for specification are constrained by the contents of a set of data that were not collected with the researcher's model in mind. In such cases, multiple indicators might not be available, precluding the specification of latent variables; the spacing of longitudinal data might not be ideal for the mechanisms being modeled; or in any number of other ways the data might limit the researcher's ability to specify the model she ideally would test. For this reason, the preferred approach is the acquisition of data that allow for estimation and testing of a model that comprises all that the researcher believes relevant to the process being studied. Referring again to Figure 1.1, the model depicted there, when specified before the data are acquired, serves as a guide to data collection or the selection of a secondary data source. Only a data set that includes x_1 to x_4 and y_1 to y_3 would be suitable for the analysis.

The specific actions and concerns in specification are the same whether a model is specified before or after the acquisition and preparation of data. In terms of actions, specification involves designating the variables, relations among the variables, and the status of the parameters in a model. With regard to designating variables, the decisions are which observed variables to include and which latent variables, if any, to model (see Bollen & Hoyle, Chapter 4, this volume, for a detailed treatment of latent variables). Having decided which observed and latent variables to include in the model, the researcher must then decide which variables are related and, for those that are related, whether the relation is nondirectional or directional. Finally, the status of parameters in a model must be specified. In general, a parameter can be specified as either fixed or free. Fixed parameters are those whose values are set by the researcher and, therefore, not estimated. For instance, in the model shown in Figure 1.1, the loading of x_1 on X (or F_1 or ξ_1) is fixed to 1. Less apparent is the fact that the loadings of x_1 to x_4 on Y and the loadings of y_1 to y_3 on X are fixed to 0. Otherwise the loadings are free parameters and will be estimated from the data. I provide detailed information about these actions in Chapter 8 (this volume).

As demonstrated earlier, a specified model is expressed formally using a system of notation coupled with either a set of equations or a diagram. Historically, each computer program for conducting SEM analyses accepted only one means of depicting a model. For example, early versions of the LISREL program required specification using matrix notation. Early versions of the EQS program required equations and double-label notation. The first program designed specifically for use on desktop computers, AMOS, accepted either line-by-line code or path diagrams constructed using built-in drawing capability. These and other commercial programs such as Mplus now allow for model specification using multiple means, including one or more of the means illustrated in Figure 1.1 as well as program-specific shorthand coding schemes. These programs are reviewed by Byrne (Chapter 19, this volume). In Chapter 20, Fox, Byrnes, Boker, and Neale review two open-source programs for SEM analyses implemented in the R environment.

A key concern in specification is identification. Each parameter in a specified model must be identified and, if all parameters are identified, the model is an identified model. A parameter is identified when it takes on a single value given the model and observed data. Parameters can be identified in two ways. The most straightforward and direct means of identifying a parameter is to fix its value. Because a fixed parameter can, by definition, assume no other value, it is identified. Free parameters are identified if there is but one estimated value for them that satisfies the estimation criterion and is obtained when the data are used to solve relevant structural and measurement equations. In some models, there is more than one way to obtain the estimate for a free parameter from these equations. As long as all such computations produce the same estimate, the parameter is overidentified. If a single value for a given parameter cannot be obtained through estimation, the parameter is unidentified and, as a result, the model is unidentified. Although a few straightforward rules of thumb offer some assurance that a model is identified, the only way to ensure identification is to show mathematically that a single value can be obtained for each parameter in all ways it might be expressed as a function of other parameters in the model. As illustrated in Figure 1.2, identification is linked to (re)specification. Ideally, prior to estimation, researchers would verify that all parameters are identified. That said, it bears noting that not all identification problems are related to specification. Parameter estimates near zero and highly correlated parameters can result in empirical underidentification, which can only be detected by attempting estimation. Kenny and Milan (Chapter 9, this volume) offer a relatively nontechnical treatment of identification.

An additional concern related to specification is the statistical power of tests of model fit. The model that best reflects the researcher's hypotheses about the mechanisms that gave rise to the data may be perfectly captured in the specification with all parameters identified, but the likelihood of finding support for the model or specific parameters in the model given the specification and data is too low to justify the analysis. The statistical power of SEM analyses is affected by multiple factors (e.g., degrees of freedom, sample size, correlations between parameters) that may vary from one fit index or statistical test to the next. The role of degrees of freedom—which derive, in part, from model specification—in the statistical power of SEM analyses argues for the consideration of statistical power as part of model specification. Detailed treatment of statistical power in the SEM context is provided by Lee, Cai, and MacCallum (Chapter 11, this volume).

Estimation

Once a model has been specified, its parameters identified, and the data prepared for analysis (see Malone & Lubansky, Chapter 16, this volume, for recommendations), the implementation moves to estimation. The goal of estimation is to find values for the free parameters that minimize the discrepancy between the observed covariance matrix and the estimated, or implied, covariance matrix given the model and the data. The means by which parameter estimates are derived depends on which of a number of possible estimation methods are used. Examples are maximum likelihood, unweighted least squares, generalized least squares, weighted least squares, and asymptotically distribution-free estimators (see Lei & Wu, Chapter 10, this volume, for detailed coverage of estimation and estimation methods). By far the most commonly used method of estimation is maximum likelihood, the default in most SEM computer programs. Because the validity of model evaluation rests most fundamentally on the integrity of estimates, a critical concern for researchers is whether maximum likelihood estimation is appropriate given their data and model. If it is not, then a decision must be made as to which alternative estimator overcomes the limitations of maximum likelihood without introducing additional concerns about the integrity of estimates. The key assumptions and how they are evaluated are discussed by Kline (Chapter 7, this volume). The robustness of different estimators to violations of assumptions often is determined by simulation studies, the logic and interpretation of which are covered by Bandalos and Gagné (Chapter 6, this volume). An approach to constructing indices of fit and statistical tests of parameters when assumptions are not met is described and illustrated by Hancock and Liu (Chapter 18, this volume).

Most estimation methods, including maximum likelihood, are iterative. They begin with a set of start values for the free parameters. These values are, in effect, used along with the fixed parameter values to solve the equations that define the model and produce an implied covariance matrix. The degree of discrepancy between the observed and implied covariance matrices is reflected in the value of the fitting function, the computation of which varies from one estimator to the next. The goal of estimation is, through iterative updating of parameter estimates (beginning with the start values), to minimize the value of the fitting function, which takes on a value of zero when the observed and implied covariance matrices are identical. Because the start values are nothing more than guesses at the values of the free parameters, the starting point often is a wide discrepancy between the observed and implied covariance matrices reflected in a relatively large value of the fitting function. The first few iterations typically result in substantial reductions in the discrepancy between the two matrices and relatively large declines in the value of the fitting function. When the value of the fitting function can be minimized no further through updates to the parameter estimates, the process is said to have converged on a solution. Often convergence is achieved in 10 or fewer iterations, though complex models or estimation situations in which start values are highly discrepant from the final estimates may require more. Unidentified models and models estimated from ill-conditioned data typically do not converge, forcing the researcher to revisit the model specification or data evaluation and preparation. Although convergence is necessary for evaluation of model fit, the number of iterations required for convergence has no relevance for that evaluation.

Evaluation of Fit

Although a set of parameter estimates obtained from suitable data for an identified model are those estimates that minimize the discrepancy between the observed and implied covariance matrices, that discrepancy may be relatively large or small. That is, the fixed and estimated parameters may imply a covariance matrix that is sufficiently similar to the observed covariance matrix to support an inference that the model fits the data; or it may imply a covariance matrix in which one or more values are sufficiently discrepant from the observed data that the model does not fit the data. In an SEM analysis, the evaluation of fit concerns whether the specified model offers an acceptable account of the data or should be rejected (if the intent is strictly confirmatory) or respecified (if the original or reconsidered intent is model generation). How this evaluation is done and a decision reached remains a topic of research and debate among methodologists.

A useful starting point for considering how decisions about fit are made is the so-called χ^2 test. In reality, the value typically labeled χ^2, under conditions rather typical of SEM analyses, is a poor approximation. Moreover, the statistical test, when it is legitimate, is of a hypothesis that few researchers would venture: that the specified model fully accounts for the observed data (i.e., there is no discrepancy between the observed and implied covariance matrices). Nonetheless, it is prototypic of goodness-of-fit tests, the goal of which is to find no difference between the observed data and data implied by a model.

Relatively early in the history of SEM, the χ^2 goodness-of-fit test fell into disfavor as a test of the absolute fit of a specified model. The earliest alternatives were indices that reflected the improvement of a specified model over a model that assumed no relations between the variables (i.e., the independence, or null, model). In some cases these values were standardized so that their values ranged from 0 to 1, with higher values indicating greater improvement of the specified model relative to the model that offered no account of the relations between variables. A drawback to these comparative fit indices is that because they do not follow a known probability distribution, they cannot be used to construct formal statistical tests. As such, their use is governed by rules of thumb, typically involving the designation of a criterion value that must be exceeded for a model to be considered acceptable.

Because of the critical importance of the decision to accept or reject a specified model, the development of new statistics and fit indices has continued. The most promising of these follow a known probability distribution, focus on absolute rather than comparative fit, evaluate the hypothesis of approximate rather than perfect fit, and account for the complexity of the model. West, Taylor, and Wu (Chapter 13, this volume) review a wide range of fit statistics and indices, and offer recommendations for using them to judge the adequacy of a specified model or to choose between alternative models.

Two additional aspects of evaluating fit bear mention. If the intent of an SEM analysis is strictly confirmatory or model generating, then the strategy I have described is appropriate for judging model adequacy. If, however, the analysis involves the comparison of alternative models, this strategy is not appropriate. Ideally the models to be compared are *nested*; that is, one model is produced by changing the status of one or more parameters in the other. In the same way that hierarchical multiple regression models can be formally compared by testing change in F or R^2, nested models in SEM can be compared by testing change in χ^2. The comparison of alternative models that are not nested is more informal and less precise but may be necessary when the alternatives cannot be specified to be nested.

Beyond these tests of overall model adequacy are tests of the estimated parameters. These typically are tested for difference from zero using a test that is comparable to the test of coefficients in multiple regression analysis.

Respecification

Referring back to Figure 1.2, the evaluation of fit can send the researcher in one of two directions—to interpretation and reporting or to respecification (the exception being when SEM is used with strictly confirmatory intent, in which case respecification is not an option). Although interpretation and reporting is the desired direction, often the evaluation of fit does not produce support for the specified model and any alternatives, sending the researcher in the direction of respecification. Note that respecification requires a reconsideration of identification, then a return to estimation and evaluation of fit. Once a researcher engages in respecification, regardless of his original intent, the goal has shifted to model generation.

Decisions about how a model might be modified to improve its fit are based on specification searches, the goal of which is to find sources of misspecification among the fixed and free parameters in the initially specified model. Specification searches can be manual, which involves a visual inspection of the residual matrix in search of subjectively large residuals, or automated, which involves the use of a statistical algorithm that evaluates the incremental improvement in fit if each fixed parameter is freed or each free parameter is fixed. In Chapter 14, this volume, Chou and Huh offer a detailed treatment of specification searching, including discussion of how modified models should be interpreted given their post hoc nature.

Interpretation and Reporting

When the evaluation of fit yields support for a model—either the originally specified model or a modified version of it—the researcher moves to the final step in the implementation framework. Given the technical challenges associated with specification, estimation, and evaluation of fit, it is perhaps surprising that many of the criticisms leveled at SEM have focused on the interpretation and reporting of results. For that reason the researcher who uses SEM must take special care in interpreting results and reporting information about the analysis and results.

With regard to interpretation, the primary concerns are the basis for the model, the meaning of particular parameters in the model, and the degree to which the model is unique in accounting for the observed data. Generally speaking, the basis for the model can be either a priori, as in models that reflect theoretical models or form a set of interrelated hypotheses that perhaps derive from multiple theories, or post hoc, as in models that include modifications to the initially specified model or have their basis in exploratory analyses of the same data to which they were fit. The former affords stronger conclusions and allows for more straightforward interpretation based primarily on the concepts and their interrelations. The latter requires qualifying with reference to the means by which the model was derived or modified. A second interpretational issue concerns the meaning of certain parameters in the model. Specifically, I refer to parameters associated with directional paths and the degree to which they can be interpreted as reflecting causal effects. In this regard, the prevailing wisdom among methodologists has moved from a willingness to view tests of parameters as tests of causal effects in the 1960s and 1970s, to an increasing reluctance to interpret parameters in this way beginning in the 1980s and continuing into the early 2000s. As detailed by Pearl (Chapter 5, this volume), there is evidence of a move away from such conservative interpretation of directional effects to a view that, when properly justified, parameters can be interpreted as tests of causal effects even when the design is cross-sectional and the data are correlational. Finally, an issue that has received too little attention from researchers who use SEM, despite repeated expressions of concern by methodologists (e.g., Breckler, 1990; MacCallum, Wegener, Uchino, & Fabrigar, 1993), is the degree to which the model accepted by the researcher is the only model that offers an acceptable account of the data. Of particular concern are equivalent models, models that yield fit statistics and indices that are identical to those of the accepted model but include paths that directly contradict those in the accepted model. Means of detecting such models and the interpretational issues they raise are treated by Williams (Chapter 15, this volume). The degree to which the researcher can successfully manage these interpretational concerns often determines whether the findings are taken seriously by other researchers.

Beyond these interpretational concerns is a more mundane set of concerns that focuses on what is to be included in research reports describing SEM analyses and results. Given the flexibility of SEM and the multiple approaches to estimation and evaluation of fit, the research report must include information that generally is not expected in reports of ANOVA, multiple regression, or factor analysis. At the most basic level, the reader needs full information regarding the model specification, including the full array of fixed and free parameters and an accounting for degrees of freedom. Additional information includes the estimation method used and the outcome of evaluating its assumptions, the information to be consulted in order to evaluate fit, and the specific criteria that distinguish a model that offers an acceptable account of the data from one that does not. Information about missing data, if any, and how they are managed in the analysis is important, particularly given the fact that some approaches to managing missing data affect model specification (e.g., inclusion of auxiliary variables; see Graham & Coffman, Chapter 17, this volume, for information about methods for addressing missing data in SEM analyses). Once this background information has been provided, the researcher must decide which parts of the large amount of statistical information generated by an SEM analysis to report and how to report them. General guidelines for reporting statistical results and suggestions related to specific types of models are provided by Boomsma, Hoyle, and Panter (Chapter 21, this volume).

This general framework captures the primary steps in any implementation of SEM, regardless of the type of model or data under study. In the final major section of the chapter, I describe the various types of models and the types of data for which they would be appropriate. Variations on each type are discussed in detail and illustrated in Parts IV and V of this handbook.

TYPES OF MODELS

A covariance matrix to be modeled using SEM, especially a large matrix, affords a wide array of modeling possibilities, constrained only by features of the sampling strategy, the research design, and the hypotheses or patterns the researcher is willing to entertain. In fact, an infinite number of models is possible with even a few observed variables (e.g., Raykov & Marcoulides, 2001). Of course, not all models that might be specified and estimated are plausible or interesting. The point is that SEM allows for the study of a wide array of models using a single comprehensive and integrative statistical approach. In the remainder of this section, I describe a sample of the models for which SEM is well-suited; references are provided to relevant chapters in the latter half of the book. Although these models do not sort cleanly into a small number of categories, for efficiency, I present them in relatively homogeneous groups based on the type of data and hypotheses for which they are appropriate.

Models Focused Primarily on Latent Structure

The variables implicated in many research questions cannot be directly observed in pure form, if at all. Rather, they must be inferred from fallible indicators such as administrative records, observer ratings, self-reports, or the status of some biological characteristic such as heart rate or changes in blood volume in selected regions of the brain. A means of separating variance in these indicators attributable to the variable of interest from variance attributable to other factors is to gather data on multiple indicators that share in common only their reflection of the variable of interest. This commonality—the latent variable—is assumed to be a relatively pure reflection of the variable of interest, free of the error and idiosyncrasies of the individual indicators (see Bollen & Hoyle, Chapter 4, this volume, for further details and discussion of other types of latent variables). This notion of commonality-as-latent-variable is familiar to many researchers as the basic premise of EFA. In the

SEM context, it is the basic logic and building block for a large number of models.

The most straightforward model concerned primarily with the latent structure of a set of indicators is the first-order factor model. The two factors in the model depicted in Figure 1.1 are first-order factors that account for the commonality among the seven indicators. Unlike EFA, indicators are assigned a priori to factors and, ordinarily, each indicator is assumed to reflect only one factor. This prototypic model can be used to test a wide array of hypotheses such as whether the factors are correlated and, if so, whether they are distinguishable; whether each item is, in fact, a reflection of only one factor; whether the loadings are equal; and whether subsets of the uniquenesses are correlated. The basic first-order model and extension of it are discussed by Brown and Moore (Chapter 22, this volume).

If the model includes enough first-order factors, the researcher might choose to explore the latent structure of the first-order factors. In the same way that the commonality among indicators can be attributed to a smaller number of latent variables, it is possible that the commonality among first-order factors can be attributed to a smaller number of second-order factors. The classic example is Thurstone's use of EFA to argue for the presence of seven primary (i.e., first-order) mental abilities but later to concede that a single (i.e., second-order) thread, presumably general intelligence, ran through them (Ruzgis, 1994). With enough first-order factors, it is possible to have multiple second-order factors and the possibility of modeling one or more third-order factors.

Another class of models concerned primarily with the latent structure of a set of indicators includes models with "subfactors," which are additional first-order factors that explain commonality in subsets of indicators that may span the factors of interest (Rindskopf & Rose, 1988). Returning to Figure 1.1 and referencing the second equation in the middle panel, note the model implies that variance in v_2 is attributable only to F_1 (X) and a uniqueness. Imagine that v_2, v_4, and v_6 were negatively worded and for that reason assumed to share a source of commonality not captured by F_1 and F_2. In order to account for this commonality, a subfactor, F_3, could be specified that influences v_2, v_4, and v_6, in which case the equation for v_2 becomes $*F_1 + *F_3 + e_2$. The inclusion of subfactors can be used strategically to tease apart trait and method variance, as in multitrait–multimethod models (Marsh & Grayson, 1995), or trait and state variance, as in trait–state models (see Cole, Chapter 34, this volume). These models, as well as first- and higher-order models, can be estimated for indicators that are continuous or categorical. The specific concerns of models that include categorical indicators are discussed by Bovaird and Koziol (Chapter 29, this volume).

Regardless of the specific model of latent structure, the question of whether a single model applies to all members of a given population may be of interest. (The same question may be asked of any model, regardless of type.) There are two approaches to studying model equivalence. When the subpopulations for which the model is to be compared can be distinguished by an observed variable (e.g., gender, ethnicity), then multigroup modeling may be used (Sörbom, 1974). In multigroup modeling, a model is estimated separately for different groups subject to constraints placed on individual parameters or groups of parameters. For instance, the loadings in a factor model might be constrained to be equal across groups and compared to a model in which they are free to vary as a means of evaluating the equivalence of the loadings. This approach is described and illustrated by Millsap and Olivera-Aguilar (Chapter 23, this volume). It is also possible that a given model does not describe the data for all members of the population but the variable that defines homogeneous subgroups in terms of parameter values is not observed. In such cases, factor mixture modeling can be used to estimate a categorical latent variable that indexes subgroup membership (Lubke & Muthén, 2005).

Models Focused Primarily on Directional Effects

A second type of model is concerned primarily with the estimation of the directional relations between variables, which may be latent or observed. The most basic model of this type is equivalent to the multiple regression model, in which the relations between a set of potentially correlated predictor variables and a single outcome are estimated. In this simplest structural model, all variables are observed, and there are no directional relations between the predictor variables. SEM allows for the extension of this basic model in three primary ways: (1) Any of the variables may be observed or latent; (2) there may be multiple outcomes among which there are directional relations; and (3) there may be directional relations between predictors. The first extension is illustrated in our example model, in which latent variable X predicts latent variable Y. The second and third extensions are somewhat redundant because they allow for models in which variables are both predictor and outcome. In fact, it is possible to have a model in which only one of many variables is only a predictor, with all other variables serving as predictors with reference to some variables in the model and outcomes with reference to others.

This point is evident in a relatively straightforward but highly useful model: the model that includes an indirect, or mediated, effect. Imagine that we add a variable, *Z*, to the model depicted in Figure 1.1. This variable is presumed to mediate the effect of *X* on *Y*. To evaluate this hypothesis, *Z* is positioned between *X* and *Y* with a directional path running from *X* to it and from it to *Y*. Thus, *Z* is both an outcome and a predictor. This particular model, the topic of Cheong and MacKinnon (Chapter 25, this volume), has received considerable attention from methodologists and is widely used in some research literatures.

Discussions of statistical mediation often compare and contrast it with statistical moderation—the qualification of a direct effect by another variable. Moderation is tested by interaction terms, which are routinely included in ANOVAs, less frequently considered in multiple regression analyses, and rarely included in models analyzed using SEM. In part, the relative neglect of interaction terms in SEM analyses may be attributed to the complexity of specifying interactions involving latent variables. Recent developments regarding strategies for modeling latent interactions have resulted in specification alternatives that significantly reduce the complexity of specification and estimation. These strategies are reviewed and demonstrated by Marsh, Wen, Nagengast, and Hau (Chapter 26, this volume).

A particularly useful class of models focused on directional relations is for data on the same sample at multiple points in time. These models can be distinguished in terms of the intensity of assessment or observation. Traditional longitudinal models involve the collection of data at relatively few points in time (typically two to four) at relatively long time intervals (typically 1–6 months). Intensive longitudinal models involve the collection of data at many time points at short time intervals (occasionally even in a continuous stream). The prototypic model for traditional longitudinal data is the autoregressive model, in which each variable is included in the model at each point in time. This permits estimation of the effect of one variable on another from one wave to the next while controlling for stability of the variables from wave to wave. Autoregressive models are covered by Biesanz (Chapter 27, this volume). When the data collection is more intensive, as in the case of many observations over a short period of time, SEM can be used to model dynamic change as it is observed taking place. One approach to such data is dynamic factor analysis, by which the latent structure of a set of indicators is simultaneously modeled at each time and across the multiple times for which data are available. Use of the basic dynamic factor model is described and demonstrated by Wood (Chapter 33, this volume). In Chapter 35, Ferrer and Song show how this model is extended to the dyadic case.

These longitudinally intensive data, as well as data used for models described in the next section, are clustered; that is, the individual observations of each individual are almost certainly more related to each other than they are to the individual observations of other individuals in the data set. The same concern applies when each individual observation applies to a different individual, but subsets of individuals share an experience (e.g., treatment by one of several health care professionals) or place in an organization (e.g., one of several classrooms or schools) that is not shared by all individuals in the sample. SEM permits modeling of such clustering while retaining all of the flexibility in modeling described in this section of the chapter. Rabe-Hesketh, Skrondal, and Zheng (Chapter 30, this volume) cover a general method for estimating these multilevel models using SEM methods.

Models That Include Means

The goal of most models estimated using SEM, including all those described to this point, is to account for covariances between variables. An additional model type, which may be integrated with the models reviewed thus far, focuses on estimating the pattern of observed means or estimating latent means. These models require as input an augmented matrix that includes an additional line for the variable means. Models fit to such matrices add intercepts to the measurement and structural equations, which allows for the modeling and comparison of means of latent variables, as well as attempts to account for, and perhaps predict, the pattern of means. The additional considerations raised by the inclusion of means and hypotheses involving means that can be evaluated using SEM are covered by Green and Thompson (Chapter 24, this volume).

Particularly useful is a set of models that are longitudinal, multilevel, and focused on modeling means—

latent growth models. These models express as latent variables the variability between individuals in the pattern of means over time. For instance, bonding to school might be assessed annually on four occasions, beginning with the first year of middle school. These assessments are clustered within individual; thus, the model is multilevel. With four time points, both linear and quadratic patterns could be modeled, yielding three latent growth factors: intercept, linear, and quadratic. In multilevel parlance, these factors are Level 2 variables that can be related to other Level 2 (i.e., individual level) latent and observed variables, as described in the previous section. The basics of this modeling approach and variations on it are described by McArdle (Chapter 32, this volume).

To further extend a model that already comprises many of the capabilities SEM affords, a researcher might ask whether there is evidence in the data of distinct subsets of individuals who show evidence of a similar pattern of bonding to school scores across the four time points. Although it is possible that the researcher has anticipated and measured the characteristic that defines these subsets, more often the heterogeneity in growth either is not expected or, if expected, its source is not known. In such cases, growth mixture modeling can be used to model a categorical latent variable that defines subsets of individuals with similar patterns of bonding to school scores. This latent variable is not unlike the latent variables discussed thus far except that its interpretation is not as simple as inferring the source of commonality among its indicators. Rather, it can be correlated with or predicted by other variables, latent or observed, to examine potential explanations for membership in these emergent groups defined by different patterns of bonding to school. Growth mixture modeling is covered by Shiyko, Ram, and Grimm (Chapter 31, this volume).

These different model types can be adapted to a wide array of data and analytic situations. For instance, SEM is increasingly used to model genetic (Franić, Dolan, Borsboom, & Boomsma, Chapter 36, this volume) and imaging (McIntosh & Protzner, Chapter 37, this volume) data. A relatively new use is for modeling spatial data (Wall, Chapter 39, this volume). And, across an array of data types, SEM has proven useful as an integrative approach to measurement scale development and validation (Raykov, Chapter 28, this volume). Across all these data and model types, parameters can be estimated and models selected using Bayesian methods, which are now available in commercial SEM computer programs. An introduction and demonstration of the Bayesian approach to SEM analyses is provided by Kaplan and Depaoli (Chapter 38, this volume).

CONCLUSION

SEM is a comprehensive and flexible approach to modeling the relations among variables in a set. Historically used primarily to model covariances between variables measured on continuous scales, the capabilities of SEM have expanded dramatically to allow modeling of many data types using an array of estimation methods, and to accommodate means, patterns of means, latent interaction terms, categorical latent variables, clustered data, and models tailored to the needs of researchers working with complex data historically not analyzed using sophisticated multivariate methods. Though SEM is not necessary, or even desirable, for every hypothesis test or modeling need, it is unrivaled in its capacity to fulfill many, varied multivariate hypotheses and modeling needs. How this capacity is harnessed and used to full advantage is the topic of the 39 chapters that follow.

REFERENCES

Bentler, P. M., & Weeks, D. G. (1980). Linear structural equations with latent variables. *Psychometrika, 45*, 289–308.

Breckler, S. J. (1990). Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin, 107*, 260–273.

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

Jöreskog, K. G., & Sörbom, D. (1999). *LISREL 8 user's reference guide*. Lincolnwood, IL: Scientific Software International.

Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods, 10*, 21–39.

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 Methods, 114*, 185–199.

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.

Raykov, T., & Marcoulides, G. A. (2001). Can there be infinitely many models equivalent to a given covariance structure model? *Structural Equation Modeling, 8*, 142–149.

Rindskopf, D., & Rose, T. (1988). Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research, 23*, 51–67.

Ruzgis, P. (1994). Thurstone, L. L. (1887–1955). In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 1081–1084). New York: Macmillan.

Sörbom, 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., & Schönemann, P. H. (1978). A history of factor indeterminacy. In S. Shye (Ed.), *Theory construction and data analysis* (pp. 136–178). Chicago: University of Chicago Press.