Structural Equation Models*

David Knoke

University of Minnesota

To appear in *Encyclopedia of Social Measurement*,

edited by Kimberly Kempf-Leonard. San Diego, CA: Academic Press

* I thank George W. Bohrnstedt, Francisco J. Granados, and three anonymous reviewers for their comments on previous drafts.

OUTLINE

GLOSSARY

confirmatory factor analysis A multivariate equation model with one or more unobserved common factors describing or explaining the relationships among empirical measures.

random error Unpredictable error resulting in normally distributed variation around a measure's true value.

reliability The extent to which different operationalizations of the same concept produce consistent results.

structural equation model A multivariate equation model combining relations among unobserved constructs with links to empirical indicators.

validity The degree to which the operationalizations of a variable accurately reflect the concept that they purport to measure.

STRUCTURAL EQUATION MODELS (SEM) are a family of analysis methods that represent translations of a series of hypothesized cause-effect relationships among variables; for making quantitative estimates of model parameters and their standard errors; for assessing the overall fit of a model to data; and for determining the equivalences of model parameters across several samples. The techniques for analyzing multivariate relationships among systems of equations build directly on multiple regression, exploratory factor analysis, and path models. Although SEM methods can be applied to complex problems, such as nonrecursive models that estimate reciprocal causal effects, space constraints allow only a basic exposition.

I. Reliability and Validity Issues

An important advantage of structural equation models lies in their capacity to combine empirical observations with relations among unobserved constructs into a single integrated system. Measurement theory seeks to represent a latent (unobserved) construct with one or more observable indicators (operational measures or variables) that accurately capture a theoretically intended concept. Two desirable properties of any empirical measure are high levels of reliability and validity. Reliability indicates the extent to which different operationalizations of the same concept produce consistent results. Reliability refers to the replication of measurement results under the same conditions; a perfectly reliable instrument must generate identical scores when the re-measurement conditions are unchanged. Alternative or multiple measures are reliable indicators of the same construct to the extent that they correlate highly. Validity is the degree to which the operationalizations of a variable accurately reflect the concept that they purport to measure. Many validity issues concern how well or poorly a particular instrument, whether consisting of a single or multiple empirical indicators, represents its intended latent concept. To be valid, a measure must demonstrate at least moderate reliability. In the extreme, if a measure has zero reliability, its validity would be attenuated relative to a more reliable measure. Multiple indictors may vary in their validity as measures of the unobserved concept they are intended to measure. Some measures may be very reliable but not valid; that is, an instrument might very precisely measure a particular phenomenon yet be invalid for some purposes. For example, individual height is very reliably measured, yet is worthless as an indicator of a person's physical health. A multiple-item health battery is less reliably measured, yet is far more valid for measuring physical health. Unfortunately, researchers never obtain perfect measurements in the real world; that is, every empirical measure is subject to some degree of measurement error. Measurement theory is therefore also a theory about how to estimate magnitudes and sources of errors in empirical observations.

Measurement reliability assumes random errors. If random error occurs when a measure is repeated several times on the same cases under the same conditions, then the resulting variations in scores form a normal distribution about the measure's true value. The standard error of that distribution represents the magnitude of the measurement error: the larger the standard error, the lower the measure's reliability. In classical test theory, the observed score (X) of respondent *i* on a measuring instrument (such as an aptitude test score or a survey item) arises from two hypothetical unobservable sources: the respondent's "true score" and an error component:

$$\mathbf{True_i} \longrightarrow \mathbf{X_i} \longleftarrow \mathbf{Error_i}$$

$$X_i = T_i + \varepsilon_i$$

On the assumption of random error, the error term is assumed to be uncorrelated with the true score. Both sources make unique contributions to the observed score variance in a population: $\sigma^2_X = \sigma^2_T + \sigma^2_\varepsilon$. The ratio of true score to observed score variances is defined as the reliability of measure *X*:

$$\rho_X = \frac{\sigma^2_T}{\sigma^2_X} = 1 - \frac{\sigma^2_\varepsilon}{\sigma^2_X}$$

This formula demonstrates that reliability ranges between 0 and 1: if the entire observed variance is error, $\rho_X = 0$; but if no random error exists, then $\rho_X = 1$. Rearranging the reliability formula also reveals that the true score variance equals the observed score variance times the reliability: $\sigma^2_T = \rho_X \sigma^2_X$. Similarly, for two parallel measures (i.e., items having equal variances), the true score variance can be estimated as the product of their correlation $(\rho_{X_1X_2})$ times either measure's variance; that is, $\sigma^2_T = \rho_{X_1X_2} \sigma^2_X$. Hence, reliability equals the correlation of two parallel measures, $\rho_X = \rho_{X_1X_2}$ while the correlation between a true score and its indicator equals the square root of the reliability: $\rho_{TX_1} = \sqrt{\rho_X}$. The measurement theory principles summarized in this section are encompassed within structural equation models, and used in the next section on the confirmatory factor analytic approach to modeling the relationships between observed indicators and latent constructs.

II. Confirmatory Factor Analysis

Factor analysis refers to a family of statistical methods that represent the relationships among a set of observed variables in terms of a hypothesized smaller number of latent constructs, or common factors. The common factors presumably generate the observed variables' covariations (or correlations, if all measures are standardized with zero means and unit variances). In confirmatory factor analysis (CFA) a researcher posits an *a priori* theoretical measurement model to describe or explain the relationship between the underlying common factors and the empirical measures. Then the analyst uses statistical fit criteria to assess the degree to which the sample data are consistent with the posited model; that is, to ask whether the results confirm the hypothesized model.

Figure 1 hypothesizes that observed measures of three harmful and four beneficial effects of psychiatric medicines (X_1 to X_7) load on separate but correlated latent factors (ξ_1 and ξ_2), labeled "Psychmed1" and "Psychmed2." Data for the estimates come from 1,070 respondents in the 1998 General Social Survey (for more details, see Knoke, Bohrnstedt and Mee 2002). The seven λ_i are the factor loadings of each observed variable on the two common factors, and the seven δ_i are the observed variables' unique error terms. This diagram implies that the latent constructs are responsible for the covariation among the observed variables. Each observed score is a linear combination of its shared unobserved factor plus its unique error term. We can also see these relationships by writing the implied measurement equation for the first and seventh indicators: $X_1 = \lambda_1 \xi_1 + \delta_1$ and $X_7 = \lambda_7 \xi_2 + \delta_7$. Note the similarity of each factor analytic equation to classical test theory's representation of an observed score as a sum of a true score plus an error term. (Figure 1 about here.)

Figure 1 assumes that all seven error terms are uncorrelated with both factors and among themselves (although alternative models allow such specifications). Hence, the only sources of an indicator's variance are its common factor ξ and its unique error term:

$$\sigma^2_{X_i} = \lambda^2_i \, \sigma^2_{\xi_k} + \Theta^2_{\delta_i}$$

where $\Theta^2_{\delta_i}$ signifies the variance of the error in X_i. Because ξ_k is unobserved, its variance is unknown. And because it is unknown, we may assume it to be a standardized variable with a variance equal to 1.0. Therefore,

$$\sigma^2_{X_i} = \lambda^2_i + \Theta^2_{\delta_i}$$

Again, note that this formula closely resembles the classical test theory in which the variance of a measure equals the sum of two components—the true score variance plus the error variance. When both components are standardized, their sum must equal 1.0. A CFA model exhibits another similarity to the classical test theory. The reliability of indicator X_i is defined as the squared correlation between a factor and the indicator (if that indicator loads on only one factor). This value is the proportion of variation in X_i that is statistically "explained" by the common factor (the "true score" in classical test theory) that it purports to measure:

$\rho_{X_i} = \rho^2_{\xi X_i} = \lambda^2_i$. Finally, the covariation between any two indicators in a multiple-factor model is the expected value of the product of their two factor loadings times the correlation between the factors. Because the error terms are uncorrelated with the factor and with each other, simplifies to: $\sigma_{X_i X_j} = \lambda_i \, \lambda_j \, \varphi^2_{\xi_k \xi_l}$.

As noted above, an unobserved common factor has no definite scale, meaning that both the origin and the unit of measurement are arbitrary. Researchers usually fix a factor's origin by assuming it has a mean of zero. The measurement unit must be scaled one of two ways: (1) by fixing the unobserved factor's variance to unity; or (2) by forcing the factor loading of one indicator (λ_i), called the reference indicator, to take a specific value (typically by setting it equal to 1.00). This latter procedure forces the factor's true score variance to equal the reliable portion of the reference indicator's variance.

The CFA example for psychiatric medicine effects in Figure 1 illustrates the second technique for setting the two factor scales by constraining the factor loadings of the X_2 and X_5 indicators equal to 1.00. Although the seven estimated loadings are all positive, the two factors have a negative covariation (-0.30). This inverse relationship is not surprising, given the substantive wordings of the seven GSS items, respectively emphasizing harmful or beneficial effects. Because respondents generally do not regard psychiatric medicines as simultaneously harmful and beneficial, a negative covariation occurs between the latent constructs represented by these two sets of empirical indicators.

CFA solutions can represent relationships in both unstandardized and standardized forms. Because a structural equation model consists of both structural and measurement levels of analysis, standardization may be done separately at each level: (1) the standardized solution scales the factors to have standard deviations of one, but leaves the observed variables in their original metrics; (2) the completely standardized solution transforms the standard deviations of both latent and observed variables to unity. Figure 2 displays the completely standardized solution for the two-factor psychiatric medicine model. The correlation between the two latent factors is –0.57, indicating that they share 32.5 percent of their variation ($r^2 = (-0.57)(-0.57) = 0.325$). Unlike the factor model in Figure 1, the completely standardized solution does not require constraining any indicators to have loadings equal to one. Hence, their magnitudes can easily be compared to assess the indicators' relative importance. Further, in both standardizations, the sum of a squared factor loading plus its square error term equals 1.00, showing that all the variation in an observed indicator is determined by these two sources. For example, the first indicator in Figure 2 has a standardized factor loading of 0.63 and an error term of 0.78; the sum of their squared values is $(0.63)^2 + (0.78)^2 = 0.397 + 0.608 = 1.00$, within rounding. (Figure 2 about here.)

III. Assessing Model Fit to Data

Parameter significance and overall correspondence between the data and the model's parameters are two important concerns of model testing. SEM computer programs estimate standard errors for all free parameters in confirmatory factor analysis or structural equation models. Thus, analysts can test the null hypothesis that a particular parameter is zero in the population, using appropriate one- or two-tailed t-tests, or Z-tests depending on sample size. All the factor loadings, residuals, variances, and covariances in the CFA model in Figure 1 are significant at $p < .05$. However, testing the significance of individual parameters cannot reveal whether the model as a whole fits the sample data.

Statistical tests for overall model fit involve a comparison of two variance-covariance matrices: (1) the observed matrix (**S**) of covariances among the *K* empirical indicators in the sample data; and (2) the expected matrix ($\mathbf{\Sigma(\theta)}$) of covariances among the same *K* indicators, computed from the model's estimated parameters ($\boldsymbol{\theta}$). An SEM program fits a model to the data by minimizing a fit function $F[\boldsymbol{S}, \boldsymbol{\Sigma}(\theta)]$. (Iterative maximum likelihood estimation is the default procedure of most programs, but alternative methods may be more appropriate for some model specifications, such as generalized least squares or weighted least squares.) The fit function involves discrepancies between the observed and predicted matrices: $F[S, \Sigma(\theta)] = \ln |\Sigma| - \ln |S| + \text{tr}(S\Sigma^{-1}) - p$; where $|\Sigma|$ and $|S|$ are determinants of each matrix, "tr" indicates "trace," the sum of the diagonal elements of the matrix, and p is the number of observed variables in the model. The fit function is always nonnegative and equals zero only if a perfect fit occurs; that is, if $\boldsymbol{S} - \Sigma = 0$. For a large sample N, multiplying $F[\boldsymbol{S}, \boldsymbol{\Sigma(\theta)}]$ by $(N - 1)$ yields a test statistic distributed approximately as a χ^2 with degrees of freedom equal to $d = [k(k + 1)/2] - t$, where t is the number of estimated parameters. Because the minimum fit function χ^2 test statistic increases proportional to sample size, (N), obtaining low chi-square values with large samples often proves difficult. The CFA model in Figure 1, based on $N = 1{,}070$ cases, has $\chi^2 = 24.1$ for df = 13 ($p = .03$), indicating that the model does not fit the data perfectly.

SEM computer programs print numerous goodness-of-fit indexes that can be used to assess overall model fit. Many indices are normed within a 0 to 1 range, with higher values reflecting better fits, but others have arbitrary metrics. Some fit indexes are functions of sample size, like chi-square, while others vary with degrees of freedom. For example, the widely used root mean square error of approximation (RMSEA) measures the mean of the squared discrepancies between observed and predicted matrices per degree of freedom. Small RMSEA values (< .05) indicate a "close fit". The RMSEA for the CFA model in Figure 1 is 0.029, indicating that the model fits the data quite well. One useful classification system distinguishes absolute, relative, and adjusted goodness-of-fit measures (Maruyama 1998:240-41). Absolute indexes assess whether a specific model leaves appreciable unexplained variance. Relative fit indexes compare the specific model to possible baseline or null models estimated using the same data. Adjusted measures ask how well the model combines both fit and parsimony, taking into account the degrees of freedom used in the model specification. Space limitations prevent a detailed review of these alternative indexes' merits. Analysts remain divided about criteria for selecting fit index and evaluating good fit. Several major points of consensus have emerged (Bollen and Long 1993:6-8): (1) a strong substantive theory is the best guide to assessing model fit; (2) chi-square should not be the sole basis for determining fit; (3) analysts should not rely on a single measure of overall fit; (4) other model components should be taken into account, such as equation R-squares and magnitudes of coefficient estimates; and (5) rather than "attempt to assess a single model's fit in some absolute sense," several models should be examined for plausible alternative structures.

IV. Structural Equation Models

This section extends confirmatory factor analysis models to models with two or more latent variables having multiple indicators. Structural equation models combine factor analysis principles with path analysis and other path modeling methods in specifying a set of linear equations representing hypothesized relations among latent constructs and their multiple indicators. Structural equation models consist of two interrelated components, a measurement model and a structural model. The measurement model, which specifies how the latent constructs are indicated by their observed indicators, describes these indicators' measurement properties (reliabilities and validities) and is analogous to CFA. The structural equation model specifies causal relationships among the latent variables, describes their direct and indirect effects, and allocates explained and unexplained variance of the dependent constructs.

The observed indicators are partitioned into exogenous variables whose variation is predetermined outside the model, and endogenous variables whose variation is explained within the model. In the matrix algebra notation popularized by Jöreskog and Sörbom (1996), a generic system of structural equations is denoted by $\boldsymbol{\eta} = \boldsymbol{\beta\eta} + \boldsymbol{\Gamma\xi} + \boldsymbol{\zeta}$, where $\boldsymbol{\eta}$ is a vector of unobserved endogenous variables, $\boldsymbol{\xi}$ is a vector of unobserved exogenous variables, $\boldsymbol{\varsigma}$ is a vector of unobserved errors, and the $\boldsymbol{\beta}$ and $\boldsymbol{\Gamma}$ the matrices of structural parameters to be estimated. The measurement model is specified by two equations: $\mathbf{Y} = \Lambda_{\mathbf{Y}}\eta + \varepsilon$ and $\mathbf{X} = \Lambda_{\mathbf{X}}\xi + \delta$, where **Y** and **X** are vectors of the observed endogenous and exogenous indicators; the two $\boldsymbol{\Lambda}$ parameter matrices specify how the observed indicators are linked to the unobserved constructs (equivalent to factor loadings in a CFA model); and the $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$ vectors contain the error terms of the indicators.

A SEM implies a covariance structure for the observed variables. Estimating the model assumes empirical data from a random sample of N cases for which all the indicators have been measured. A computer program then uses an iterative algorithm to fit the specified SEM to the sample covariance matrix (**S**) of the indicators. The program simultaneously estimates the free parameters of both the structural and measurement models, estimates standard errors for each parameter, and calculates various goodness-of-fit indexes for the whole model. Several SEM computer programs perform these computations; for example, LISREL, AMOS, EQS, MPLUS, and SAS CALIS. Most programs no longer require analysts to specify their models in formal matrix algebra language, but use simple programming instructions to denote the hypothesized relations among latent and observed variables. Some SEM programs allow a researcher to draw a diagram on a computer screen, then translate it into software commands that estimate the model parameters.

Diagrams are indispensable tools for conceptualizing and interpreting a SEM. In the simple example in Figure 3, the structural model depicts an exogenous "Political Ideology" construct causing variation in an endogenous "Federal Help" construct. The measurement model consists of two observed indicators of Politics (conservative political views and party identification) and four indicators of Help (attitudes against the federal government's responsibility for solving social problems: not helping with poverty; not helping with any problems; not helping with medical bills; not helping African-Americans. See Knoke et al. 2002 for more details about these measures). Parameter estimates for the completely standardized solution were computed by LISREL from a covariance matrix computed for 1,594 respondents in the 1998 GSS. The overall model fit is very good ($\chi^2 = 15.7$, 8 df, $p = .052$; other fit indexes have quite acceptable values) and the individual parameter estimates are all highly significant. The estimated structural parameter (0.63) means that a difference of one standard deviation in political ideology is associated with a three-fifths standard deviation difference in attitude towards the federal government's role in solving social problems. Given that all indicators were measured with conservative responses scored high and liberal responses scored low, the positive sign means that more ideologically conservative respondents favor more individualistic solutions to social problems (i.e., less federal government involvement). (Figure 3 about here.)

SEM programs allow an explicit statistical test of the hypothesis that two or more parameters are equal in the population. Constraining a pair of parameters to be equal (rather than letting them freely take differing values) requires estimating only one parameter instead of two. As a result, one degree of freedom is then available to assess whether constrained and unconstrained models' chi-square statistics differ at a chosen alpha-level. If no significant difference occurs between the two models, then the more parsimonious version with equal parameters (i.e., the model with fewer unconstrained parameters) would be preferred. In the model in Figure 3, the standardized parameters of the four Help indicators seem to have roughly similar magnitudes (ranging between 0.55 and 0.69). Some alternative models that specified equal loadings for several pairs of indicators didn't produce significantly worse fits to the data. However, a model that hypothesized equal loadings for HELPOOR and HELPBLK was rejected (χ^2 difference of 24.9 – 15.7 = 9.2 for 1 df, $p < .01$). Another model, hypothesizing equal factor loadings for the first three indicators (HELPPOOR, HELPNOT, HELPSICK), didn't produce a significantly worse fit compared to the model with no equality constraints (χ^2 difference of 19.2 – 15.7 = 3.5 for 2 df). The three Help indicators each had estimated parameters equal to 0.76, while the HELPBLAK indicator had a lower value (0.66).

V. Model Identification and Modification Strategies

For a model to be estimable, both its measurement and structural equation portions must be identifiable. An SEM or CFA model is identified if every unknown parameter has a unique value that can be estimated by fitting the model to the data. A model is "underidentified," and not estimable, if the number of unknown parameters to be estimated exceeds the available degrees of freedom (the number of indicator variances and covariances). In such instances, a model can be respecified to assure identification by constraining sufficient numbers of the unknown parameters to fixed values (typically set to zero). "Just identified" models, with precisely as many unknown parameters as available degrees of freedom, always produce trivially perfect fits but may provide useful baseline estimates against which to test other models with positive degrees. "Overidentified" models, with positive degrees of freedom, reveal whether the model specifications reasonably represent relationships in both the measurement and structural models. For a complicated SEM, a researcher's *a priori* identification of all parameters may become difficult because fulfilling all the formal requirements to assure identification can often be quite complicated (see Kenny, Kashy and Bolger 1998). SEM computer programs usually ascertain whether an hypothesized model is not identified, if they cannot calculate unique estimates with standard errors for the unknown parameters. Nevertheless, model and parameter identification remain relevant concerns because SEM computer programs may occasionally produce solutions for unidentified models. Model modification strategies comprise additional important concerns. Unless a model fits the data well, researchers seldom fit a single hypothesized model to their data, then stop after making the decision to accept or reject that specification without proposing any alternative. The more common practice involves model generation strategy, an exploratory approach that incrementally respecifies parameters and fits a series of alternative CFA or SEM to the same data. The analyst's ultimate objective is to find an overidentified model that fits the data at an acceptable level (using the various goodness-of-fit indexes), while also yielding plausible and meaningful interpretations of the estimated parameters. Exploratory modifications of a tentative initial model should not rely entirely on statistical criteria, but also take into account existing theory and empirical knowledge about a substantive area. Important procedures for locating sources of model misspecification include examining: parameter estimates for unrealistic values or anomalous signs inconsistent with theoretical expectations; assessing squared multiple correlations (R^2) for each equation for evidence of weak or nonlinear relations; inspecting residuals, standardized residuals, model modification indices to pinpoint expected parameter changes and fit improvement, for example, by correlating error terms. By repeating these steps for successively modified models, analysts may obtain a final version that fits the sample data reasonably well and provides a plausible interpretation.

Unfortunately, a final SEM or CFA model with an improved fit to the data is unlikely to be the "true" or "best-fitting" model, in the sense that its successive improvements involved capitalizing on chance covariation in the sample data. Instead, it is probably one of several alternative models of equivalent overall fit that approximate the unknown true population SEM. A more robust approach to SEM generation cross-validates the modified model results with an independent sample. Alternatively, researchers can randomly split a sufficiently large sample in half, using the first subsample to estimate the modified model and the second subsample to cross-validate that specification (Cudeck and Browne 1983). Recent developments in automated algorithms, such as TETRAD and TABU, assist researchers in their search for true model specifications and parameter estimates (Richardson and Sprites 1999).

Space limitations prevent examination of several important procedures and issues in CFA and SEM, including: hierarchical linear modeling (HLM); comparisons of parameters across populations; second-order factor structures (factors of factors); multitrait-multimethod models (alternative procedures for measuring several constructs); noncontinuous (ordinal or dichotomous) indicators; nonlinear and nonrecursive relations among variables; missing data; multicollinearity; correlated errors; longitudinal data collection designs; and statistical power in SEM. Interested readers can further their understanding by consulting basic and advanced texts (e.g., Bollen 1989; Marcoulides and Schumaker 1996; Maruyama 1998) and articles in such journals as *Psychological Methods*, *Sociological Methods & Research*, and *Structural Equation Modeling*.

VI. Strengths and Limitations of Structural Equation Models

Both CFA and SEM methods, implemented in a variety of computer packages, provide researchers with powerful data analysis tools. Applied judiciously, these methods have important advantages over traditional multivariate methods, such as linear regression, that assume no errors in observed measures. If a SEM model is true, then its structural parameters linking the latent constructs take into account the biases of less reliably and validly measured indicators. But, the price paid for these advantages is susceptibility to erroneous parameter estimates and model fits if analysts misspecify the true measurement and structural relationships. For example, covariation in cross-sectional data offers no clues to asymmetric or reciprocal causation; even the temporal sequences among repeated measures in longitudinal panel designs is not an infallible guide to causal order. (See Cliff 1983 and Breckler 1990 for additional critiques of SEM methods.) Because SEM methods by themselves do not enable researchers to distinguish among many alternative models with statistically equivalent fits, analysts face heavy requirements to apply logic and theory jointly to distinguish incredible from plausible alternative model specifications (MacCallum, Wegener, Uchino and Fabrigar 1993). The protean qualities of SEM methods should spur researchers to work harder at improving their theoretical understanding of the social processes they seek to explain.

Bibliography

Bollen, K. (1989). *Structural Equations with Latent Variables*. Wiley, New York.

Bollen, K. and Long, J. (1993). *Testing Structural Equation Models*. Sage Publications, Thousand Oaks, CA.

Breckler, S. (1990). Applications of Covariance Structure Modeling in Psychology: Cause for Concern? *Psychological Bulletin* **107**, 260-273.

Cliff, N. (1983). Some Cautions Concerning the Application of Causal Modelling Methods. *Multivariate Behavioral Research* **18**, 115-126.

Cudeck, R. and Browne, M. (1983). Cross-Validation of Covariance Structures. *Multivariate Behavioral Research* **18**, 147-167.

Jöreskog, K. and Sörbom, D. (1996). *LISREL 8: User's Reference Guide*. Scientific Software International, Chicago.

Kenny, D. Kashy, D. and Bolger, N. (1998). Data Analysis in Social Psychology. In *Handbook of Social Psychology* (D. Gilbert, S. Fiske and G. Lindzey, eds.) 4th edition, Vol. 1, pp. 233-265. McGraw-Hill, Boston, MA.

Knoke, D., Bohrnstedt, G. and Mee, A. 2002. *Statistics for Social Data Analysis*. 4th Ed.: F.E. Peacock, Itasca, IL.

Marcoulides, G. and Schumaker, R. eds. (1996). *Advanced Structural Equation Modeling: Issues and Techniques*. Erlbaum, Mahwah, NJ.

Maruyama, G. (1998). *Basics of Structural Equation Modeling*. Sage Publications, Thousand Oaks, CA.

MacCallum, R., Wegener, D., Uchino, B. and Fabrigar, L. (1993). The Problem of Equivalent Models in Applications of Covariance Structure Analysis. *Psychological Bulletin* **114**, 185-199.

Richardson, T. and Spirtes, P. (1999). Automated Discovery of Linear Feedback Models. In *Computation, Causation, and Discovery* (C. Glymour and G. Cooper, eds.) Pp. 253-304. MIT Press, Cambridge, MA.

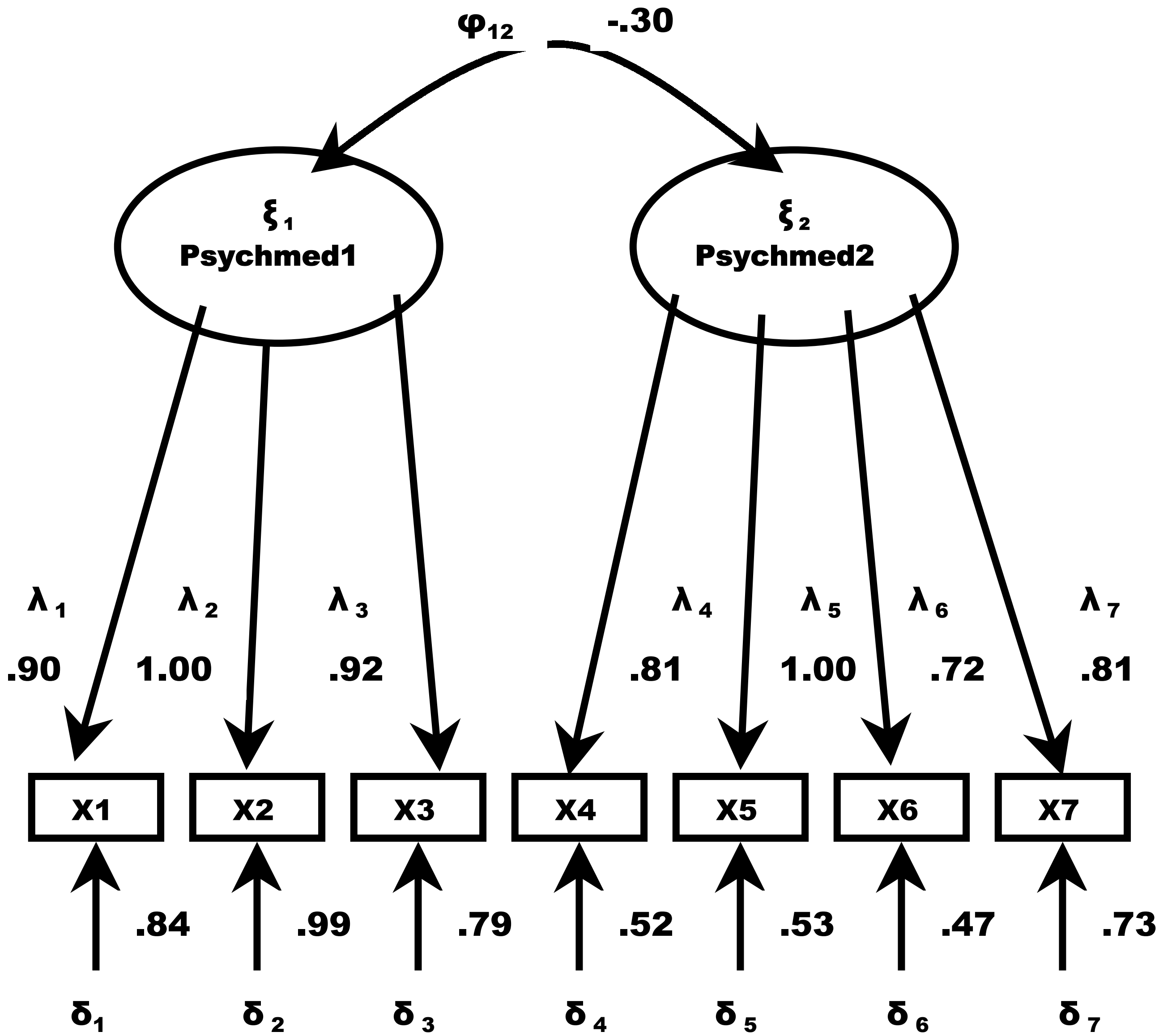

Figure 1. A Two-Factor Confirmatory Factor Analysis Model with Seven Psychiatric Medicine Indicators

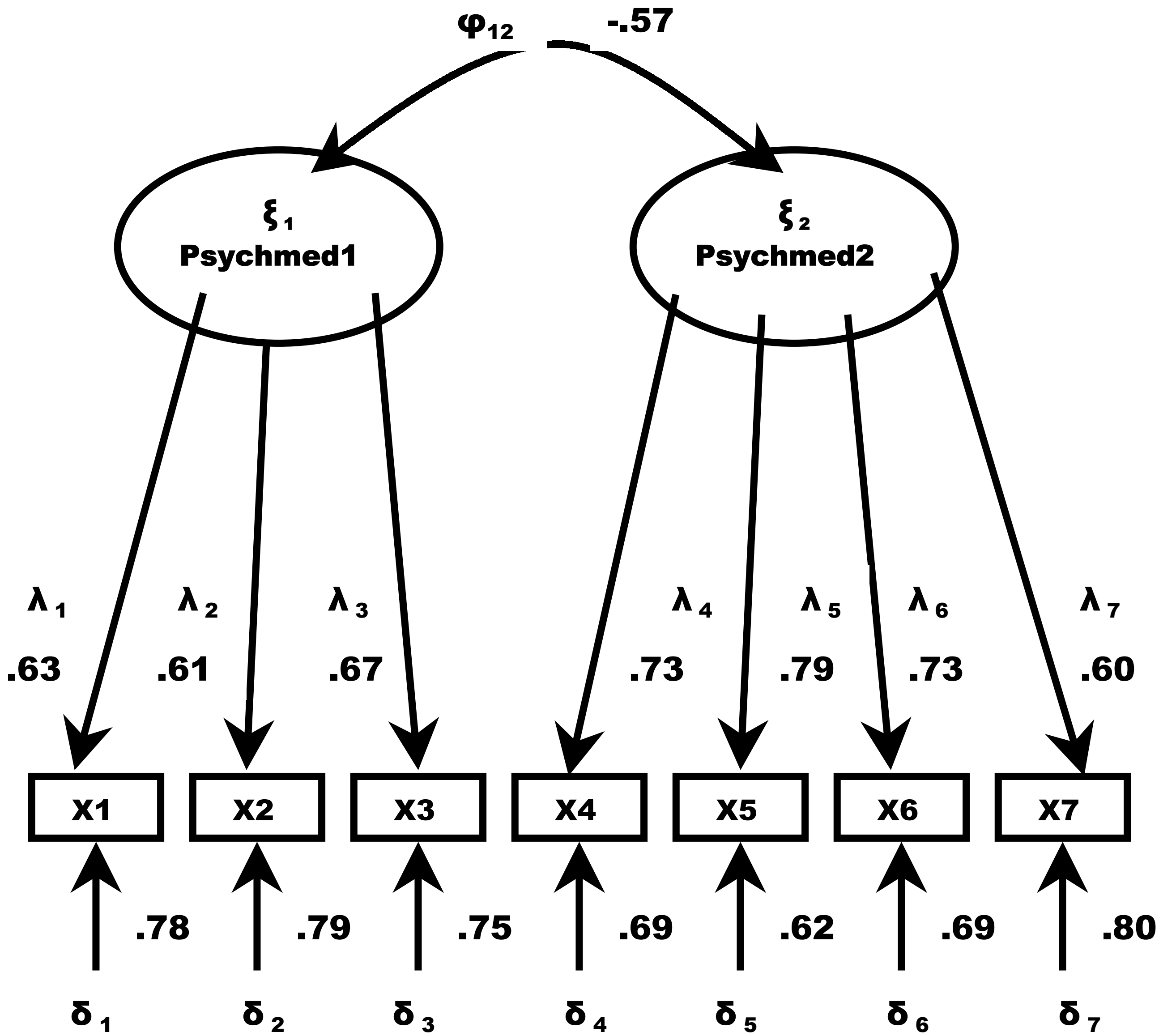

Figure 2. Completely Standardized Solution for a Two-Factor Confirmatory Factor Analysis Model with Seven Psychiatric Medicine Indicators

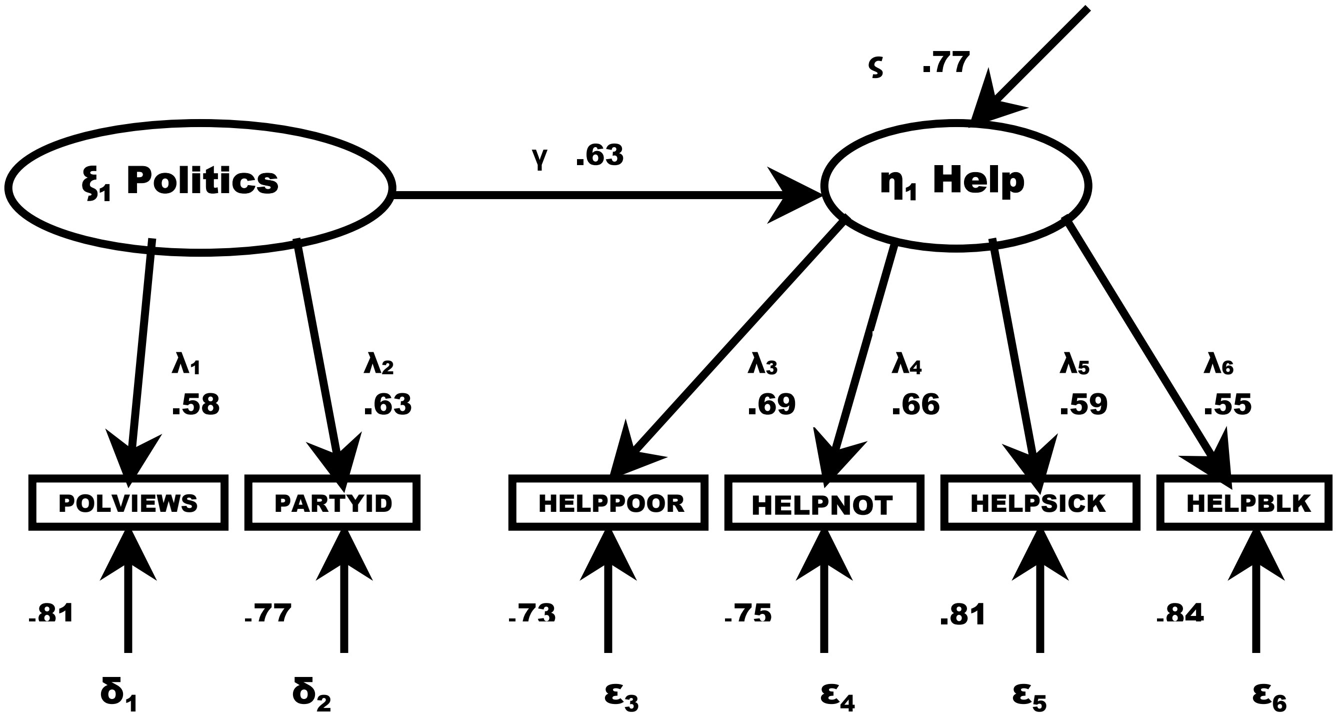

Figure 3. Completely Standardized Solution for a Simple Structural Equation Model

Figure 1. A Two-Factor Confirmatory Factor Analysis Model with Seven Psychiatric Medicine Indicators

Figure 2. Completely Standardized Solution for a Two-Factor Confirmatory Factor Analysis Model with Seven Psychiatric Medicine Indicators

Figure 3. Completely Standardized Solution for a Simple Structural Equation Model