Published in final edited form as:
*J Consum Psychol*. 2010 April ; 20(2): 215–220. doi:10.1016/j.jcps.2010.03.002.

# SEM with simplicity and accuracy

**Peter M. Bentler**
University of California, Los Angeles, USA

## Abstract

Professor Iacobucci has provided a useful introduction to the computer program LISREL, as well as to several technical topics in structural equation modeling (SEM). However, SEM has not been synonymous with LISREL for several decades, and focusing on LISREL's 13 Greek matrices and vectors is not the most intuitive way to learn SEM. It is possible today to do model specification via a path diagram without any need for filling in matrix elements. The simplest alternative is based on the Bentler–Weeks model, whose basic concepts are reviewed. Selected additional SEM topics are discussed, including some recent developments and their practical implications. New simulation results on model fit under null and alternative hypotheses are also presented that are consistent with statistical theory but in part seem to contradict those reported by Iacobucci.

For anyone who has been involved in theory and applications of SEM since its early days, it is hard to avoid a sense of déjà vu when reading Professor Iacobucci's (2009) well-formed introduction to LISREL. The approach to modeling using 13 Greek matrices and vectors was state-of-the-art three decades ago, but it certainly is not so today. In fact it can be argued that research in consumer psychology is artificially held back when users who are interested in SEM also are required to become proficient in LISREL's matrix language. Of course, LISREL continues to be a good program for sophisticated users, and it certainly is still used, so Iacobucci's emphasis is certainly justifiable. Yet its Greek language provides an artificial and unnecessary roadblock to becoming proficient in the use of SEM for substantive research. This note reminds potential SEM researchers that simpler alternatives to LISREL are available, e.g., AMOS (Arbuckle, 2006) and EQS (Bentler, 2000–2008). We emphasize the Bentler and Weeks (1980) model that is the foundation of the EQS approach; the model provides conceptual clarity, while the program provides implementation simplicity and cutting-edge statistics. An introduction to SEM based on the Bentler–Weeks model is given by Mulaik (2009) and to EQS is provided by Byrne (2006), while its role in marketing research is discussed by Savalei and Bentler (2006).

In addition to overviewing basic modeling setups, Professor Iacobucci (2010) ably reviews several issues in application of SEM that are relevant to its appropriate use. Because of their critical role in SEM, this note expands on selected statistical topics including the behavior of $\chi^2$ when the model is correct and when it is wrong, model modification, the modeling of correlations and dealing with ordinal variables, evaluation of model nesting and equivalence, handling missing data, and dealing with outliers and bad distributions. Recent developments, especially those in which we participated, are emphasized. Some new simulations are reported whose results tend to be consistent with statistical theory but are partially at odds with findings reported by Iacobucci (2010).


bentler@ucla.edu .

## Simplicity versus complexity in SEM

Many important concepts in SEM, such as the distinction between observed variables (*V*s) and latent factors (*F*s), or setting the scale of latent variables (*F*s and residual errors, *E*s) that are not in the data file, are, of course, crucial to the informed use of SEM. These concepts can, however, be taught without Greek letters. To set up and run a model such as Fig. 5 in Iacobucci (2009) certainly requires understanding basic ideas such as the distinction between *V*s and *F*s, and between one-way (directional) and two-way (correlational) arrows. But sucha diagram —without a single Greek letter—is all the user needs. A program such as EQS will directly implement the model without any need to learn or use the many other matrices discussed by Iacobucci and shown in her Figs. 2, 4, and 6.

There are three main reasons that 13 matrices and vectors are not optimal for teaching or implementing SEM in practice. The first is that there are much simpler ways to think about SEM, especially with the Bentler and Weeks (1980) model. The second is that such simpler thinking automatically allows the researcher to consider a wider range of model types are found in the standard LISREL setup. The third is that good computer programming can relieve the user of routine but troublesome SEM chores and handle them internally in the program. EQS incorporates these methods.[1] At the same time, for the advanced researcher, EQS allows interface to the unlimited world of advanced computing with its R interface (Mair, Wu, & Bentler, 2010).

### A one-paragraph introduction to SEM

Even when not using the program diagram to create a model, any model can be easily specified in EQS by writing out some equations and variance/covariance specifications. First, a path diagram is scanned to find *dependent* variables—those variables that have at least one one-way arrow aiming at them. Any model then consists of a series of regression-type equations, one for each dependent variable. If there are 3 (or 10 or 30) dependent variables, the model setup requires 3 (or 10 or 30) equations. Every one-way arrow corresponds to a $b$ or $\beta$ coefficient in an equation. Such coefficients are free parameters (denoted by "*" in EQS) unless fixed for identification or by theory. So, if a diagram has two substantive arrows aiming at a variable plus an arrow from a residual, the hypothesis can be specified by an equation. An example is V3=*V1+*V2+E3, which is simply a regression equation with two predictors and a residual. If there are 5 arrows aiming at V3, there have to be 5 terms on the right-hand side of the equation. No complicated matrices are needed to expand this idea to slightly varied, yet important, types of equations, especially (1) V3=*F1+*F2+E3, for a measurement equation that relates a dependent *V* to its explanatory latent variables; and (2) F3=*F1+*F2+D3, for predicting a latent factor from other factors (so as not to confuse the error E3 in V3, with the residual in F3, the latter is called D3, for disturbance). These two aspects cover all possible equations specified via matrices in LISREL. Furthermore, it is immediately obvious that we could go beyond the LISREL model by also allowing mixed equations when they make sense, such as (3) V3=*F1 +*V4+E3, where both *F*s and *V*s predict a *V*, or (4) F3=*F1+*V4+D3, where both *F*s and *V*s predictan *F*. Then to complete the model setup, the *independent* variables—those that do not have one-way arrows aiming at them and hence are never on the left side of any equation—have variances and possibly covariances (two-way arrows) as free parameters that need specification. Depending on the model, any *V*, *F*, *E*, or *D* thus could be an independent variable. In EQS, a covariance is specified by the variables involved, and a statement such as V1, V2=* or E5,E7=* sets the covariance as a free parameter. In EQS, then, there is a 1:1 correspondence between a model diagram and its equation/covariance specification. Hence, given an equation/covariance specification, a unique diagram of connections also can be created.

[1]Bentler acknowledges a financial interest in EQS and its distributor, Multivariate Software.

The above paragraph covers all the models allowed by LISREL, and even more. It covers all of the models diagrammed in Iacobucci (2009,2010). It is simply not necessary for researchers on consumer psychology to become familiar with the complicated Greek equation setup discussed by Iacobucci (2009). Of course, there is more to the Bentler–Weeks model than summarized above, e.g., the matrix form that generates a model-reproduced covariance matrix. Although this also is a simpler form than that of LISREL, it is not needed to set up, run, and obtain complete results for a SEM analysis. The matrix details are relegated to the program internals (and, of course, to a manual).

### A typical problem avoided with the simpler approach

It was already noted that matrices, such as shown in Figs. 2, 4, and 6 of Iacobucci (2009), are not needed in the approach summarized above. In addition, a Greek label is not needed for variables in diagrams such as shown in Figs. 1, 3, and 5. However, perhaps the most typical problem avoided with the EQS approach is to not confuse latent factors with observed variables as is often done when using LISREL. This problem is illustrated in Iacobucci's (2010) Fig. 6. The diagram shows a latent factor *D* and its observed indicator D1,[2] and Iacobucci rightly notes that since the error in D1 is taken to be zero and D1=1.0D, it follows that the factor actually is not a factor at all. The factor *D* is identical to the observed variable D1, that is, D1≡*D*, nothing more and nothing less. Hence, the path diagram is not a direct representation of what the model states and invites confusion by implying that a factor exists in this part of the model. The diagram could simply show C→D1→E without the confusing implication that a factor *D* is involved. That is precisely how the model would be specified in EQS, but since LISREL does not allow factor→variable→factor paths, one has to pretend that D1 actually is a factor when it is not. Of course, technically, what LISREL does is correct, since the identity between D1 and *D* is part of the specification. But a casual glance at Fig. 6 incorrectly implies that *D* is a latent variable.

An immediate advantage of the Bentler–Weeks model is to provide a clear-cut definition of what a latent variable model is. It is one in which the dimensionality of the independent variables (their covariance matrix) is larger than the dimensionality of the observed variables (Bentler, 1982). See Bollen (2002) for additional viewpoints. Principal components and PLS, as noted by Iacobucci, are not latent variable models.

### Formative constructs

Measurement models in the LISREL tradition and implemented via Bentler–Weeks require that latent factors generate theobserved variables, and not the other way around as is proposed in formative measurement. Although we fully agree with Iacobucci that this is the appropriate way to specify measurement models, we recently developed a two step approach to provide an identified approach to formative constructs. First, a formative construct is operationalized with multiple observed variables. Second, these are embedded as reflective indicators of a latent factor in any standard SEM model (Treiblmaier, Bentler, & Mair, 2008).

## Additional statistical topics

Although simplicity is important, accuracy in model testing and evaluation is even more important. Hence in this section we discuss a few important SEM topics that extend the overview provided by Professor Iacobucci. An overview of statistical issues in SEM is given by Hayashi, Bentler, and Yuan (2008).

[2]This notation has nothing to do with the EQS usage of *D* as disturbance.

## Test statistics and sample size

Iacobucci discusses the behavior of the $\chi^2$ test statistic when the SEM model is true. Under the null hypothesis that a model holds, and when other modeling assumptions are met, in large samples the expected value of a test statistic is its degrees of freedom (df). Hence the first diagram in Fig. 1 of Iacobucci (2010) describing the behavior of the $\chi^2_{ML}$ model test with increasing sample size for a simulated 2-factor model is puzzling, since the line should be approximately flat and centered at about 8, the df associated with her model. We repeated her simulation and obtained the results in Table 1, showing mean results across 2000 replications of four statistics (including RMSEA not studied by Iacobucci) at various sample sizes. We found the mean $\chi^2_{ML}$ to be quite stable and near its expected value at all sample sizes, so we suspect that a setup or reporting error crept into Iacobucci's figure. Note also that convergence problems, such as error variances tending to go negative, occur frequently at small sample sizes, and that while the descriptive fit indices show some bias at smaller sample sizes, the bias is much smaller than observed by Iacobucci and largely disappears as convergence becomes dependable. Actually, these problems would disappear at much smaller *N*s if the correlations among variables in the model had more structure than in this example (see below).

## Test statistics with incorrect models

It is also useful to consider what happens when the model is false, i.e., incorrect in some respects. Then any decent goodness of fit test becomes arbitrarily large as sample size *N* increases, reflecting increasing power to reject the null hypothesis. This can be seen in the case of maximum likelihood (ML) with test statistic $T_{ML}=NF_{\min}$ (see Iacobucci, 2009, Appendix I where $T_{ML}=X^2$ and $F_{\min}$ is the part in brackets, evaluated at the optimal estimate). Here $F_{\min}$ approaches some nonzero number, but *N* goes to infinity, and thus $T_{ML}$ will get very large. As a result, almost any model with sufficiently large *N* will almost surely be rejected. Models with very bad fit to the population covariance matrix (large $F_{\min}$) may be rejected even at relatively smaller *N*. Large df also work against the investigator. For example, in a model with 30 variables there will be 30(31) / 2 = 465 sample covariances to model, and if a model has 200 free parameters, there are 465–200=265 df. Each df represents one way of being wrong about the model (e.g., a missing path). In such a situation, even the best a priori theory is likely to be incorrect in some ways and perhaps in many ways, implying almost certain statistical model rejection. For simulation and theoretical work see Curran, Bollen, Paxton, Kirby and Chen (2002) and Yuan, Hayashi, and Bentler (2007), respectively.

We also did a simulation in which the data is generated by Iacobicci's 2-factor model and parameters, but the hypothesized model is incorrectly specified as a 1-factor model with 9 df. Table 2 shows the results. Now all the fit indices at all sample sizes indicate that the model has problems, with the $\chi^2_{ML}$ becoming substantially larger than df=9 at the larger sample sizes. It does not become extremely large because with these data most models will obtain a small $F_{\min}$. This occurs because Iacobucci's 6 variables are hardly correlated, with 60% of the correlations being .1 or less and the largest correlation being .26. In such a situation $\chi^2_{ML}$ will not have much power to reject any model structure unless *N* is really huge. On the other hand, if the indicators of a given factor had correlated .8, and the two factors correlated lowly, power to reject the model would have been high and $\chi^2_{ML}$ would be large even at smaller sample sizes. For recent work on power in SEM, see Kim (2005).

## Model modification

Of course, high power to reject models due to large sample sizes motivate the use of fit indices like CFI (Bentler, 1990) that provide another meaningful metric for evaluating degree of misspecification. But it also motivates model modification. If a data set cost a fortune to collect, is it the really best course of action to stop all further data analyses when one's very best theory is found to be not perfect? That is the implication of Iacobucci's discouragement of model modification. We don't want to be "statistics devils," but our view is that modification indices (as in econometrics, called Lagrange Multiplier or LM tests in EQS) are very useful, perhaps even necessary, to help isolate unanticipated ways in which a model can be improved. In EQS, LM tests are always used in an a priori way to evaluate equality restrictions, such as those imposed for invariance testing in multiple samples. Iacobucci does not like the 1-parameter at a time approach to searching for large unexpected misspecifications as used in LISREL, but in EQS, there is also a multivariate LM test that takes into account parameter dependencies. Of course, LM test based improvements based on blind search have to make sense, are subject to capitalization on chance in small samples, and will not necessarily find the "true" modeling errors if these exist (e.g., MacCallum, Roznowski, & Necowitz, 1992). Nonetheless they almost always provide useful hints on model inadequacies. Honesty in reporting any such post-hoc search is, of course, required, since readers deserve to know about the necessity for cross-validation. Actually, sometimes a researcher can do such cross-validation on their own data if the original sample is large enough to be split into exploratory and confirmatory samples.

## Correlations and ordinal variables

SEM is almost universally presented as a covariance structure method rather than as a correlation structure method. Actually, the original invention of path analysis with latent variables involved the use of correlations (e.g., Wright, 1923), and in fact, most simple SEM models are only concerned with correlational implications. Exploratory factor analysis had always been (and continues to be) done with correlation input, but ever since Jöreskog (1969) it has been assumed that covariances are critical to SEM modeling. Although this was true historically since the statistics in SEM were originally derived from the distribution of sample covariances, this is no longer true. Modern programs such as EQS give correct statistics for correlation structures. Of course, there are situations where variances matter, such as when testing invariance of factor structures across groups. For a summary on the role and limitations of correlation structures, see Bentler (2007) or Bentler and Savalei (2010).

A closely related topic is that of ordinal variables. In many situations, it can be assumed that an observed binary or ordinal variable represents an underlying latent continuous variable. For example, Likert scales with few response options are probably best thought of as ordinal and not interval scaled, while the underlying attribute being measured is likely to be continuous and, in many cases, normally distributed. In such situations the observed product-moment correlations are biased (usually downward) and tetrachoric and polychoric correlations among ordered variables, and polyserial correlations between continuous and ordinal variables, provide more appropriate estimates of the underlying correlations. These correlations are then considered to be based on some interesting SEM (e.g., Muthén, 1984). In EQS, the combined methodology is based on Lee, Poon, and Bentler (1995), and in future versions, also will be based on maximum pairwise likelihood (Liu & Bentler, 2009).

## Model nesting and equivalence

Usually it is easy to see when models are nested, because the more restricted model is a version of the more general model in which some parameters have been set to zero. It is not so easy when models are only covariance matrix nested, where the graphs of the models may be quite different but their implied covariance matrices are subsets of each other (Bentler & Bonett, 1980). Iacobucci mentions nested models in the context of chisquare difference tests, but nesting is also assumed for fit indices such as CFI that by default are computed as if the model of uncorrelated variables is a special case of any given structural model. Unfortunately, as noted by Widaman and Thompson (2003), sometimes such nesting is not true and then the CFI makes no sense. Bentler and Satorra (2009) propose a simple way to evaluate model nesting as well as model equivalence. Equivalence is an often overlooked property of a model, but it is important because a quite different model structure may lead to identical statistical fit but a different interpretation (Hershberger, 2006). For example, model X→Y→Z (residuals omitted) will have identical fit as model X←Y←Z.

## Missing data

An important topic not discussed by Iacobucci is missing data. This occurs almost universally in real data analysis. Historical approaches such as mean substitution, regression estimation, pairwise computations, and case-wise deletion are not optimal because they can (and often do) lead to biased estimates of parameters, standard errors, and/or test statistics. In recent years, direct or case-wise maximum likelihood has been recommended (e.g., Arbuckle, 1996; Jamshidian & Bentler, 1999) as a much improved solution, since it makes use of all the data and allows missing at random (MAR) missingness rather than the more restrictive missing completely at random (MCAR) mechanism. In MAR data, missingness can depend on features of the observed data. Recently, another approach has been developed (Cai, 2008; Savalei & Bentler, 2009; Yuan & Lu, 2008) that has some advantages over case-wise ML, especially greater stability in small samples and ability to easily use auxiliary variables (Graham, 2003) to reduce bias or variance due to missing data while not impacting a model structure. It is a two-stage ML approach in which saturated ML means and covariances are first computed for model as well as auxiliary variables, if present. The saturated estimates on the model variables are subsequently used as input to a standard ML modeling run. Statistical adjustments take into account the sequential nature of the estimation. A final aspect is that in recent years, multiple imputation (MI) has been recommendedas another way to handle missing data (e.g., Schafer & Graham, 2002). In the context of SEM, however, Yuan, Wallentin, and Bentler (2009) found that MI parameter estimates tend to be less efficient and to have more biases than those of ML, and that overall ML yields better statistical results than MI when there are distributional violations.

## Bad distributions and outliers

Iacobucci suggests to “stick to ML” even when data does not meet the strict multivariate normality assumption of ML. We agree as far as the estimator is concerned, but disagree as far as test statistics and standard errors are concerned. That is, the ML point estimator indeed tends to be remarkably robust to violation of assumptions and can usually be relied upon (although there are special situations when it is problematic, e.g., Browne, MacCallum, Kim, Anderson, & Glaser, 2002). In our experience, real data often have fat tails, i.e., excess skew or kurtosis beyond that allowed under normality. Such excess variability is easily located with Mardia’s normalized coefficient of multivariate kurtosis and generally will produce larger standard errors and inflated $T_{ML}$ test statistics. Of course, reduced variability also can occur, providing a different distortion. Incorrect standard errors will yield incorrect scientific conclusions on the significance of effects, while unnecessarily large test statistics imply that a model is worse than it really is. The bulk of evidence indicates that statistics such as the Satorra and Bentler (1994) scaled (mean corrected) and adjusted (mean/variance corrected) test statistics perform substantially better than $T_{ML}$ in this circumstance as do the Yuan and Bentler (1998a) residual-based statistics. Due to a tendency of the scaled test to lose power with increasing kurtosis (Curran, West, & Finch, 1996) and the reliable behavior of the adjusted statistic (Yuan & Bentler, 2010), the latter is probably preferable. Standard errors are best handled with the so-called sandwich estimator. In the case of missing nonnormal data, the Yuan and Bentler (2000) methodology works most reliably. These methods are implemented automatically in EQS with a simple statement such as “ML, robust.” Recently an improved method was developed for differences between nested scaled tests (Satorra & Bentler, 2009).

Outliers are a slightly different problem. They produce lumpy distributions rather than smooth distributions with lighter or heavier tails than those of the normal distribution. Lumpy distributions are best handled by classical robust statistics based on case-weighting. In the standard setup, each case is weighted 1.0 (equally) in computing means and covariances, but case-robust statistics allow each case to receive a different weight varying from 0 to 1, depending on its closeness to the center of the distribution and perhaps closeness to the model. Outliers or influential cases may be strongly downweighted, say to .2 or .1, or even to 0 thus eliminating them from the sample. Case-weighted robust statistics for SEM (Yuan & Bentler, 1998b) are incorporated into EQS. A recent illustration is on the effects of smoking on various types of cancer, using a classic epidemiological data set (Bentler, Satorra, & Yuan, 2009). A review of alternatives is given by Yuan and Bentler (2007a).

## Conclusion

Professor Iacobucci provided a valuable overview of SEM. However, SEM is simpler than implied in her review. On the other hand, it also has more complex aspects since statistical and scientific accuracy may require the use of specialized methods to handle difficult practical issues.

This commentary, like Professor Iacobucci's overview, has had to be severely restricted in scope due to space limitations. As a result, many exciting new developments could not be included in this review. Among such developments we have been associated with are new approaches to multilevel modeling (Bentler & Liang, 2008; Yuan & Bentler, 2007b), SEM-based reliability estimation that substantially improves on coefficient alpha (e.g., Bentler, 2009a), a new approach to modeling interactions (Mooijaart & Bentler, 2009), a method for evaluating close rather than exact fit of model differences (Li & Bentler, 2009), and new estimation and testing methods for a 1-parameter modeling approach to Guttman data that, like the Rasch model, yields interval scale data (Bentler, 1971, 2009b). More importantly, new developments on SEM by many others continue unabated in the technical literature (e.g., Bauer, 2007); recent reviews can be found in Lee (2007). The most practically relevant of such developments often can be found in the journal *Structural Equation Modeling*.

## Acknowledgments

This research supported was supported by grants 5K05DA000017-32 and 5P01DA001070-35 from the National Institute on Drug Abuse.

## References

Arbuckle, JL. Full information estimation in the presence of incomplete data. In: Marcoulides, GA.; Schumacker, RE., editors. Advanced structural equation modeling: Issues and techniques. Erlbaum; Mahwah, NJ: 1996. p. 243-277.

Arbuckle, JL. AMOS 7 user's guide. SPSS; Chicago: 2006.

Bauer DJ. Observations on the use of growth mixture models in psychological research. Multivariate Behavioral Research 2007;42(4):757–786.

Bentler, PM.; Flamer, GB. An implicit metric for ordinal scales: Implications for assessment of cognitive growth. In: Green, DR.; Ford, MP., editors. Measurement and Piaget. McGraw Hill; New York: 1971. p. 34-63.

Bentler, PM. Linear systems with multiple levels and types of latent variables. In: Jöreskog, KG.; Wold, H., editors. Systems under indirect observation: Causality, structure, prediction. Part I. North-Holland; Amsterdam: 1982. p. 101-130.

Bentler PM. Comparative fit indexes in structural models. Psychological Bulletin 1990;107:238–246. [PubMed: 2320703]

Bentler, PM. EQS 6 structural equations program manual. Multivariate Software; Encino, CA: 2000–2008. www.mvsoft.com

Bentler PM. Can scientifically useful hypotheses be tested with correlations? American Psychologist 2007;62:772–782.

Bentler PM. Alpha, dimension-free, and model-based internal consistency reliability. Psychometrika 2009a;74:137–143. [PubMed: 20161430]

Bentler, PM. Estimation, tests, and extensions of a 1-parameter item scaling model. Paper presented at Society of Multivariate Experimental Psychology, Salishan Lodge OR; 2009b.

Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin 1980;88:588–606.

Bentler, PM.; Liang, J. A unified approach to two-level structural equation models and linear mixed effects models. In: Dunson, D., editor. Random effects and latent variable model selection. Springer; New York: 2008. p. 95-119.

Bentler PM, Satorra A. Testing model nesting and equivalence. Psychological Methods. 2009

Bentler PM, Satorra A, Yuan K-H. Smoking and cancers: Caserobust analysis of a classic data set. Structural Equation Modeling 2009;16:382–390. [PubMed: 20011616]

Bentler, PM.; Savalei, V. Analysis of correlation structures: Current status and open problems. In: Kolenikov, S.; Steinley, D.; Thombs, L., editors. Statistics in the social sciences: Current methodological developments. Wiley; New York: 2010. p. 1-36.

Bentler PM, Weeks DG. Linear structural equations with latent variables. Psychometrika 1980;45:289–308.

Bollen KA. Latent variables in psychology and the social sciences. Annual Review of Psychology 2002;53:605–634.

Browne MW, MacCallum RC, Kim C-T, Anderson B, Glaser R. When fit indices and residuals are incompatible. Psychological Methods 2002;7:403–421. [PubMed: 12530701]

Byrne, BM. Structural equation modeling with EQS: Basic concepts, applications, and programming. 2nd ed. Erlbaum; Mahwah, NJ: 2006.

Cai L. SEM of another flavour: Two new applications of the supplemented EM algorithm. British Journal of Mathematical and Statistical Psychology 2008;61:309–329. [PubMed: 17971266]

Curran PJ, Bollen KA, Paxton P, Kirby J, Chen F. The noncentral chi-square distribution in misspecified structural equation models: Finite sample results from a Monte Carlo simulation. Multivariate Behavioral Research 2002;37:1–36.

Curran PJ, West SG, Finch JF. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. Psychological Methods 1996;1:16–29.

Graham JW. Adding missing-data-relevant variables to FIML-based structured equation models. Structural Equation Modeling 2003;10:80–100.

Hayashi, K.; Bentler, PM.; Yuan, K-H. Structural equation modeling. In: Rao, CR.; Miller, J.; Rao, DC., editors. Handbook of statistics, 27: Epidemiology and medical statistics. North-Holland; Amsterdam: 2008. p. 395-428.

Hershberger, SL. The problem of equivalent structural models. In: Hancock, GR.; Mueller, RO., editors. Structural equation modeling: A second course. Information Age; Greenwich, CN: 2006. p. 13-41.

Iacobucci, Dawn. Everything you always wanted to know about SEM (structural equations modeling) but were afraid to ask. Journal of Consumer Psychology 2009;19:673–680.

Iacobucci, Dawn. Structural equations modeling: Fit indices, sample size, and advanced topics. Journal of Consumer Psychology 2010;20:90–98.

Jamshidian M, Bentler PM. ML estimation of mean and covariance structures with missing data using complete data routines. Journal of Educational and Behavioral Statistics 1999;24:21–41.

Jöreskog KG. A general approach to confirmatory maximum likelihood factor analysis. Psychometrika 1969;34:183–202.

Kim K. The relationship among fit indexes, power, and sample size in structural equation modeling. Structural Equation Modeling 2005;12:368–390.

Lee, S-Y. Handbook of latent variable and related models. Elsevier; Amsterdam: 2007.

Lee S-Y, Poon W-Y, Bentler PM. A two-stage estimation of structural equation models with continuous and polytomous variables. British Journal of Mathematical and Statistical Psychology 1995;48:339–358. [PubMed: 8527346]

Li, L.; Bentler, PM. Quantified choice of RMSEAs for evaluation and power analysis of small differences between structural equation models. 2009. Under editorial review

Liu, J.; Bentler, PM. Latent variable approach to high dimensional ordinal data using composite likelihood. 2009. Under editorial review

MacCallum RC, Roznowski M, Necowitz LB. Model modifications in covariance structure analysis: The problem of capitalization on chance. Psychological Bulletin 1992;111:490–504. [PubMed: 16250105]

Mair, P.; Wu, E.; Bentler, PM. EQS goes R: Simulations for SEM using the package REQS. Structural Equation Modeling. 2010. Preprint No. 554. Available from http://preprints.stat.ucla.edu/

Mooijaart A, Bentler PM. An alternative approach for non-linear latent variable models. Structural Equation Modeling. 2009

Mulaik, SA. Linear causal modeling with structural equations. Chapman & Hall/CRC; Boca Raton FL: 2009.

Muthén B. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. Psychometrika 1984;49:115–132.

Satorra, A.; Bentler, PM. Corrections to test statistics and standard errors in covariance structure analysis. In: von Eye, A.; Clogg, CC., editors. Latent variables analysis: Applications for developmental research. Sage; Thousand Oaks, CA: 1994. p. 399-419.

Satorra A, Bentler PM. Ensuring positiveness of the scaled difference chi-square test. Psychometrika. 2009 doi:10.1007/S11336-009-9135-Y.

Savalei, V.; Bentler, PM. Structural equation modeling. In: Grover, R.; Vriens, M., editors. The handbook of marketing research: Uses, misuses, and future advances. Sage; Thousand Oaks, CA: 2006. p. 330-364.

Savalei V, Bentler PM. A two-stage approach to missing data: Theory and application to auxiliary variables. Structural Equation Modeling 2009;16:477–497.

Schafer JL, Graham JW. Missing data: Our view of the state of the art. Psychological Methods 2002;7:147–177. [PubMed: 12090408]

Treiblmaier, H.; Bentler, PM.; Mair, P. A methodology for formative constructs as latent variables. 2008. Under editorial review

Widaman KF, Thompson JS. On specifying the null model for incremental fit indices in structural equation modeling. Psychological Methods 2003;8:16–37. [PubMed: 12741671]

Wright S. The theory of path coefficients: A reply to Niles's criticism. Genetics 1923;8:239–255. [PubMed: 17246011]

Yuan K-H, Bentler PM. Normal theory based test statistics in structural equation modelling. British Journal of Mathematical and Statistical Psychology 1998a;51:289–309. [PubMed: 9854947]

Yuan K-H, Bentler PM. Robust mean and covariance structure analysis. British Journal of Mathematical and Statistical Psychology 1998b;51:63–88. [PubMed: 9670817]

Yuan K-H, Bentler PM. Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. Sociological Methodology 2000:165–200.

Yuan, K-H.; Bentler, PM. Robust procedures in structural equation modeling. In: Lee, S-Y., editor. Handbook of latent variable and related models. North-Holland; Amsterdam: 2007a. p. 367-397.

Yuan, K-H.; Bentler, PM. Multilevel covariance structure analysis by fitting multiple single-level models. In: Xie, Y., editor. Sociological methodology 2007. Vol. vol. 37. Blackwell; New York: 2007b. p. 53-82.

Yuan K-H, Bentler PM. Two simple approximations to the distributions of quadratic forms. British Journal of Mathematical and Statistical Psychology 2010;63:273–291. [PubMed: 19793410]

Yuan K-H, Hayashi K, Bentler PM. Normal theory likelihood ratio statistic for mean and covariance structure analysis under alternative hypotheses. Journal of Multivariate Analysis 2007;98:1262–1282.

Yuan K-H, Lu L. SEM with missing data and unknown population distributions using two-stage ML: Theory and its application. Structural Equation Modeling 2008;43:621–652.

Yuan K-H, Wallentin F, Bentler PM. ML versus MI for missing data with violation of distribution conditions. 2009 Under editorial review.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

**Table 1**

Problem-free convergence and statistic means of all replications for correctly specified model.

| | Sample size | | | | | |
|---|---|---|---|---|---|---|
| | **30** | **50** | **100** | **200** | **500** | **1000** |
| # OK Reps | 490 | 787 | 1373 | 1809 | 1995 | 2000 |
| $\chi^2_{ML}$ | 8.2 | 7.9 | 7.6 | 7.9 | 8.1 | 8.0 |
| SRMR | .10 | .08 | .05 | .04 | .02 | .02 |
| CFI | .75 | .86 | .95 | .97 | .99 | .99 |
| RMSEA | .05 | .04 | .02 | .02 | .01 | .01 |

**Table 2**

Problem-free convergence and statistic means of all replications for misspecified model.

| | Sample size | | | | | |
|---|---|---|---|---|---|---|
| | **30** | **50** | **100** | **200** | **500** | **1000** |
| # OK Reps | 817 | 1079 | 1472 | 1747 | 1952 | 1984 |
| $\chi^2_{ML}$ | 10.5 | 11.4 | 14.5 | 21.0 | 41.0 | 74.3 |
| SRMR | .11 | .09 | .07 | .06 | .06 | .06 |
| CFI | .68 | .75 | .79 | .77 | .75 | .75 |
| RMSEA | .07 | .06 | .07 | .08 | .08 | .08 |