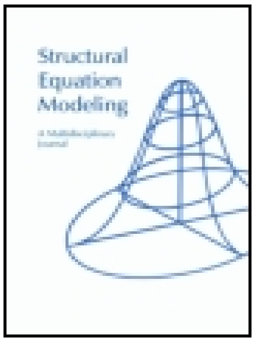

Structural Equation Modeling

ISSN: 1070-5511 (Print) 1532-8007 (Online) Journal homepage: https://www.tandfonline.com/loi/hsem20

Modeling Strategies: In Search of the Holy Grail

Kenneth A. Bollen

To cite this article: Kenneth A. Bollen (2000) Modeling Strategies: In Search of the Holy Grail, Structural Equation Modeling, 7:1, 74-81, DOI: 10.1207/S15328007SEM0701_03

To link to this article: https://doi.org/10.1207/S15328007SEM0701_03

Published online: 19 Nov 2009.

STRUCTURAL EQUATION MODELING, *7*(1), 74–81

Modeling Strategies: In Search of the Holy Grail

Kenneth A. Bollen
Department of Sociology
Carolina Population Center
University of North Carolina at Chapel Hill

Modeling strategies are subject to debate for virtually all statistical procedures. Witness the sharp disagreements over stepwise regression, the interpretation of clusters in cluster analysis, or the identification of outliers and influential points. The largely objective basis of statistical algorithms does not remove the need for human judgment in their implementation. So it is not surprising that the use of structural equation models is subject to disputes over the best way to formulate and test models. Though I must admit considerable scepticism about whether it is possible to have a single generic strategy that would prove optimal over all substantive areas and types of structural equation models, articles and discussions such as those of Hayduk and Glaser (2000) and Mulaik (1998) are very helpful in that they bring out the merits and limits of the alternative procedures.

The options that are the focus of their discussions are (a) the *one-step* procedure and (b) the *four-step* approach. In a sense, we could see the one-step procedure as there at the birth of contemporary structural equation models. One of the attractive features of structural equation models was the ability to *simultaneously* model the latent variable and the measurement models in one step.[1] Defenders of this original practice include Hayduk (1987, 1996) and Fornell and Yi (1992). More recently, Mulaik (1998) has advocated a four-step method that is the subject of the Hayduk and Glaser (2000) article.

Requests for reprints should be sent to Kenneth A. Bollen, Department of Sociology, CB 3210 Hamilton, University of North Carolina, Chapel Hill, NC 27599–3210. E-mail: bollen@email.unc.edu

[1]I depart from the current practice of referring to the latent variable model as the "structural" model. Such a labeling suggests that the measurement model is not structural and that only the latent variable model involves structural parameters (Bollen, 1989, p.11). Using the term *latent variable model* can avoid this confusion.

Hayduk and Glaser (2000) singled out the four-step procedure for their critique. The procedure is described in theirs and other pieces, so to conserve space I will not repeat the steps here. Hayduk and Glaser's central critique is that the procedure is incapable of determining the correct number of factors, and that the wrong number of factors can throw off the whole process. They also question the use of the chi-square test statistic and the root mean squared error of approximation (RMSEA) in the four-step process. This comment focuses on the modeling strategy, ignoring the issues of the chi-square test statistic and the RMSEA.

I readily agree with Hayduk and Glaser's critique that the four-step method depends on, but cannot guarantee, the correct number of factors. But I would add two things. The first is that *both of the preceding modeling procedures suffer the same limitation.* Neither can tell us whether we have the right number of factors. More specifically, the one-step procedure that Hayduk and Glaser recommended has this same limitation as does the four-step. So, if we accept the critique and ask for the way out, the one-step method does not provide a solution. Though I see this as an important limitation, I do not see why the four-step modeling procedure should be singled out.

Second, the number-of-factors issue is a particularly difficult one when we consider that there are some chi-square equivalent models that have a different number of factors. For instance, in research that colleagues and I were engaged in, we examined the reported distance to a family-planning facility in a rural sample of villages in Tanzania (Bollen, Mroz, Speizer, & Ngallaba, 1996). Three female and 3 male informants were independently asked the distance to a specific facility, and this was done for each village. Altogether we had 6 informant estimates of distance for each village with different informants for each village. Figure 1a represents a three-factor model to describe these data. The common perception of distance influences all 6 informants' estimates of the distance. In addition, there are two method factors, one for perceptual influences of female informants and another introduced by using male informants. Because the ordering of the informants was arbitrary and the informants changed from village to village, there is no reason to expect the factor loadings to differ from one, and so all loadings are constrained to one.[2] The model in Figure 1b has only two factors. The two models in Figure 1 have identical fits. So here we have an instance where there is a different number of factors, but the empirical fit of the models is identical.

A less clear-cut instance in which the number of factors differs but the fit is the same occurs when we have correlated errors of measurement. If we replace a pair of correlated errors with a new latent variable with paths set to one and that

[2]There is no statistically significant difference in fit when only the scaling indicators are set to 1 and the other factor loadings are free.

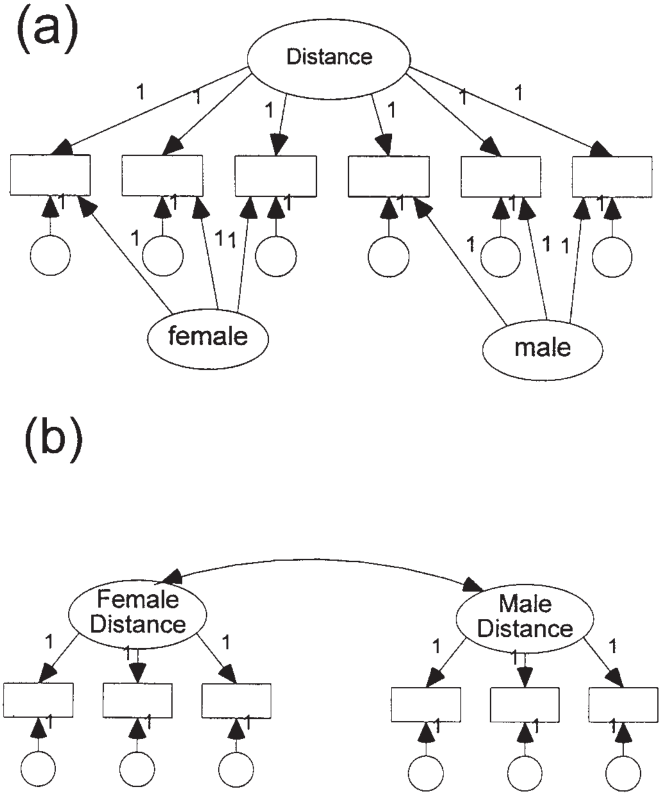

FIGURE 1 Models of informant distance estimates with identical fit and different number of latent variables.

is uncorrelated with the other latent variables, we get two models with identical fit. Figures 2a and 2b provide an example. In Figure 2a we have a single latent variable with a pair of correlated errors. Figure 2b has two latent variables. The fit of both models is identical. We can complicate the model by adding more latent variables, indicators, and pairs of correlated errors. By replacing each pair of correlated errors with a latent variable with two paths set to one, we can alter the number of latent variables, often without changing the fit.

I do not know how common it is to have models that have a different number of factors but that have identical fits, but the existence of any shows a limitation of all empirically based searches for the proper number of factors. If we loosen the restriction from having identical fit to having fits that are close, we encompass a much broader range of situations in which the empirical procedures will not make clean distinctions between models with different numbers of factors. The conclu-

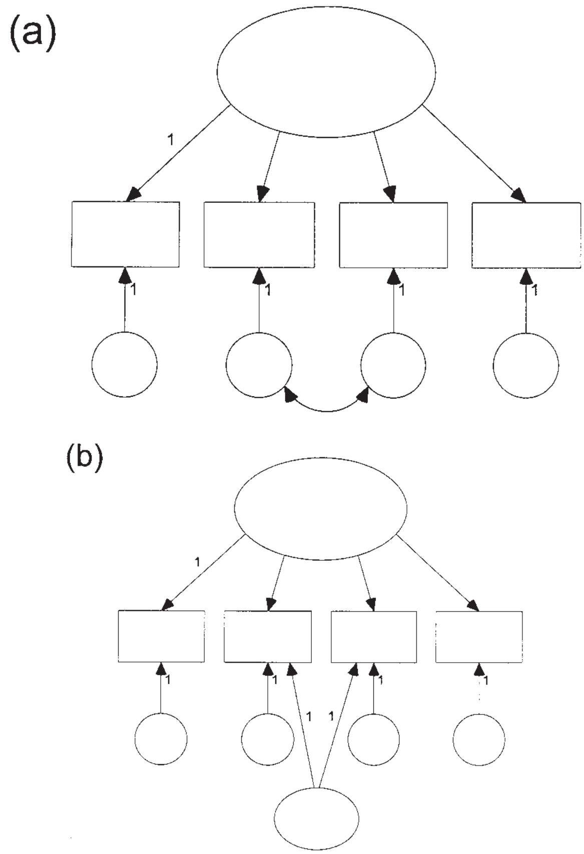

FIGURE 2 Replacement of correlated errors by latent variable.

sion is that a researcher can use a four-step or a one-step procedure and still be led to the wrong model.

Another concern I have with the four-step modeling strategy is its heavy reliance on traditional factor analysis techniques. For instance, the starting point in the four-step method is an exploratory factor analysis in which the number of factors is specified in advance, but correlated errors are not permitted. This type of structure is too confining and restricts the options that a researcher might consider. Correlated errors due to similar question wording or overlapping content are not that unusual. How they would manifest themselves in the first step of the four-step analysis is ambiguous.

Other relations also could be difficult to capture from a traditional factor analysis in the first part of the four-step procedure. For instance, in a recent analysis, Juan Diez Medrano and I sought to measure perceived attachment to Spain among a probability sample of Spaniards (Bollen & Medrano, 1998). Morale and Sense of Belonging were the two substantive dimensions of attachment that were measured with three indicators apiece. However, our final model also included response bias effects such that the response to one indicator influenced the response to the indicator that came immediately after it in the questionnaire. Figure 3 has a path diagram of this model where the arrows directly connecting the indicators represent the response bias effects. It is not clear how such a model would fair if we took a four-step approach to it.[3] We found the two-factor model without the response bias effects lacking in fit, but adding factors does not seem to be the solution. Rather, having each indicator except the first being affected by the indicator that precedes it on the questionnaire was a more plausible structure that led to a good fit. There are other models that do not neatly fit into a four-step modeling strategy, and pursuing such a strategy might push the researcher in the wrong direction.

Is the one-step method the solution? Unfortunately, no. The one-step procedure shares with the four-step technique the inability to determine the correct number of factors. Another major limitation of the one-step approach is that in most cases an initial model will not have adequate fit. Testing the whole model in a single step makes locating the source of the poor fit extremely difficult. Given the practical problem that our initial models often have specification errors, the one-step strategy provides insufficient guidance on where to turn after the model fit is found wanting.

OTHER MODELING STRATEGIES

The critiques of the four-step and the one-step strategies to fitting structural equation models leads to the question of whether there are other alternatives to fitting models. Hayduk and Glaser (2000) and Mulaik (1998) largely restrict their discussion to (a) the one-step modeling strategy and (b) the four-step procedure. To these we could add (c) the *separate factor analysis* method, (d) the *two-step* procedure, and (e) the *jigsaw piecewise* technique. In the 1970s, Burt (1973, 1976) raised the possibility of breaking the full model into separate factor analyses and fixing the parameter estimates from these separate analyses when estimating the full model. Fornell and Yi (1992) illustrated how separate factor analysis could individually fit even if the indicators were inadvertently placed with the wrong factor and how this could lead to a mistaken impression of a

[3]If we stuck with Mulaik's rule of having four indicators per latent variable, we could not apply the four-step procedure to this example. Though I would agree that more indicators are better than fewer indicators, there are practical considerations that often prevent collecting so many measures per latent variable.

poorly fitting latent variable structure. Between the extremes of treating each factor separately and handling the whole model at once were several efforts that treated the whole measurement model first and then turned to the latent variable model in a two-step procedure (e.g., Anderson & Gerbing, 1988; Herting & Costner, 1985; James, Mulaik, & Brett, 1982). This procedure is the predecessor of the four-step procedure. Critiques by Fornell and Yi (1992) and Hayduk (1996) are available.

The *jigsaw piecewise technique* is a term that I use to refer to a strategy that lies between those already mentioned. As the name suggests, the modeling strategy is somewhat like a jigsaw puzzle, where we fit pieces of the model individually and then together until we find a coherent whole. It is similar to the separate factor analysis method in that pieces of the model are examined prior to the whole model. However it differs from the separate factor analysis in that part of the analysis is to see if the fit still is reasonable as you assemble the pieces at each stage and you do *not* fix the coefficient estimates when assembling the full model.

A hypothetical illustration would be a model with four interrelated latent variables, each measured with at least four indicators. In the jigsaw piecewise, we could fit one factor separately, ensuring that it fits, then turn to each of the other

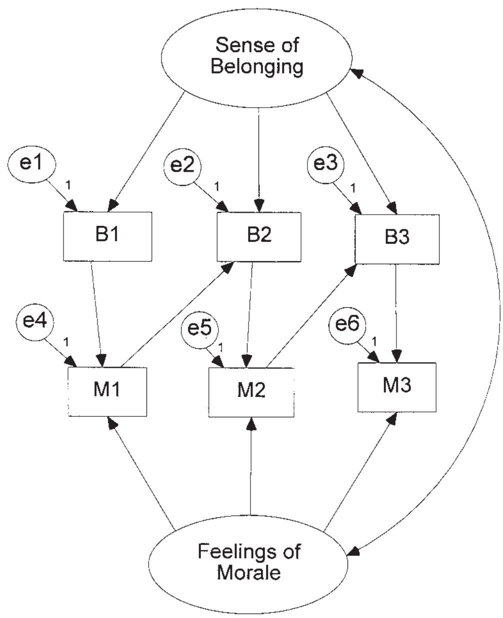

FIGURE 3 Measurement model of attachment to Spain. From "Who Are the Spaniards? Nationalism and Identification in Spain," by K. A. Bollen and J. D. Medrano, 1998, *Social Forces, 77,* p. 602. Copyright 1998 by *Social Forces*. Reprinted with permission.

factors and check their fit.[4] But unlike the separate factor analysis strategy, we would then put together two or three of the factors at a time to make sure that the expanded model still fits. This piecing together of factors is an important way to check whether the fit of the separate models was obscuring spurious or suppressor relations that were missed by treating the factors separately.

In addition, the analyst should compare the shifts in the parameter estimates of the model when fitted in pieces versus the estimates from the assembled pieces of the model. Large changes are evidence of model misspecification even if the overall fit measures do not indicate problems. A Hausman (1978) type test for a significant shift in the parameter estimates is a possibility because the parameter estimator for the fully assembled model and the corresponding estimator for the pieces of the model will both be consistent estimators while the estimator for the full model will be asymptotically more efficient under correct model specification. Highlighting those parameters that have the greatest changes could be helpful in locating the sources of the specification errors.

Unfortunately the jigsaw piecewise strategy and the other alternatives I have mentioned in this section do not solve all the problems of the one-step or four-step methods. None overcomes the problems of having the wrong number of factors. None provides unambiguous guidance on what to do in the all-too-common situation of a inadequate model fit.

CONCLUDING COMMENTS

I agree with Hayduk and Glaser's (2000) critique that the four-step strategy cannot detect the correct number of factors, but I would add this same critique to all of the modeling strategies, including the one-step method advocated by Hayduk. To further illustrate the pitfalls of using empirical procedures to detect the correct number of factors, I gave examples of two models with identical fit but with different numbers of factors. Given the identical fit, we cannot rely on empirical techniques to find the correct number of factors. The frequency and the difficulty of the situation is furthered if we consider those cases where models with different numbers of factors have similar rather than identical fit statistics.

Another problem, at least for the four-step strategy, is its heavy reliance on a traditional factor analysis model with diagonal covariance matrices for the unique ("error") variables and few constraints on factor loadings. Such a model does not give us the flexibility that we require in incorporating correlated errors of measurement, a mixture of correlated and uncorrelated latent variables, and direct relations

[4]If there is insufficient number of indicators to identify a factor analysis model, the researcher might need to estimate more than one factor at a time. In addition, if indicators have factor complexity greater than one, then all the factors influencing the indicator should be part of the analysis.

between indicators. In addition, relying on such restrictive factor analysis approaches is likely to blind us to alternative specifications.

The critique of these modeling strategies pushes us back to reliance on our knowledge or hypotheses about the substantive area. It has been my experience that many researchers have more knowledge or "hunches" about the nature of the relationships between the variables in their models than they include in their initial specifications. Embracing set modeling strategies or steps discourages researchers from thinking about specification and creates a dependence on more mechanical procedures that are far from problem free. Given the uncertainty that accompanies any single specification, I would encourage analysts to formulate several alternative model specifications that they can fit to their data. Comparisons of the fit of these models to each other rather than to only the saturated model can allow assessments of relative performance of different specifications. Constructing these specifications is hard work. Though it is tempting to seek a simple formulaic approach to correct specification of models, it does not yet exist.

REFERENCES

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

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Bollen, K. A., & Medrano, J. D. (1998). Who are the Spaniards? Nationalism and identification in Spain. *Social Forces, 77,* 587–621.

Bollen, K. A., Mroz, T., Speizer, I., & Ngallaba, S. (1996). *Measuring perceived distance to family planning facilities in rural Tanzania.* Unpublished manuscript.

Burt, R. S. (1973). Confirmatory factor-analytic structures and the theory construction process (plus corrigenda). *Sociological Methods and Research, 2,* 131–190.

Burt, R. S. (1976). Interpretational confounding of unobserved variables in structural equation models. *Sociological Methods and Research, 5,* 3–51.

Fornell, C., & Yi, Y. (1992). Assumptions of the two-step approach to latent variable modeling. *Sociological Methods and Research, 20,* 291–320.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica, 46,* 1251–1272.

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

Hayduk, L. (1996). *LISREL: Issues, debates, and strategies.* Baltimore: Johns Hopkins University Press.

Hayduk, L., & Glaser, D. N. (2000). Jiving the four-step, waltzing around factor analysis, and other serious fun. *Structural Equation Modeling, 7,* 1–35.

Herting, J. R., & Costner, H. L. (1985). Respecification in multiple indicator models. In H. M. Blalock (Ed.), *Causal models in the social sciences* (2nd ed., pp. 321–393). New York: Aldine.

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

Mulaik, S. A. (1998). Re: Four-step model. SEMNET Discussion List, 20 November 1998, Time: 13:03.