

Latent Variable Structural Equation Modeling: A Flexible Tool for Establishing Validity and More

Drew Altschul

University of Edinburgh & Scottish Primate Research Group

Petrinovich highlighted many salient issues in the behavioral and social sciences that are of concern to this day, such as insufficient attention to construct validity. Structural equation modeling, particularly with regard to latent variables, is introduced and discussed in this context. Though conceptual issues remain, analytic and statistical techniques have made immense strides in the past three decades since the article was written, and properly used, offer solutions to many problems Petrinovich identified.

Although originally written in the 1990s, the target article by Petrinovich (2021) stands alongside more recent commentaries on philosophy of science and the behavioral sciences, much like Paul Meehl's (1967, 1990) work still stands up today. However, while research culture has been slower to change, methods and analysis have come a long way. In particular, powerful statistical analyses have become possible thanks to improvements in computing power (and affordability of computers) as well as increased availability of free software over the past three decades (see Vazire and Gelman's commentary for more on the recent and relevant history). To further Petrinovich's aims, robust and approachable statistical techniques are necessary.

Structural equation modeling (SEM) is a multivariate statistical framework that can be used to assess unobserved or latent variables. SEMs have two parts. The measurement model defines latent variables through the measurement of multiple observed variables, which is analogous to confirmatory factor analysis. In the structural model, latent variables can be linked to one another (or to other observed variables) through conventional correlations and regressions, analogous to path modeling. SEMs are a form of network modeling, and lend themselves to visual representation via path diagrams, and intuitive syntax (Rosseel, 2011). Petrinovich discusses latent variables throughout. He mentions SEM once, near the end of the commentary, as an example of a method for causal modeling, which I will discuss further along.

SEM is highly flexible and can be directly applied to many of the problems Petrinovich raises. Petrinovich then asks if there is a cure to the inadequacies endemic to the behavioral sciences, and though SEM is not a panacea, it is a powerful analytic tonic.

Petrinovich advocates for using multitrait-multimethod (MTMM) matrices to establish construct validity. In its most basic form a MTMM matrix is a correlation matrix of measures, some tapping the construct of interest, and some the tapping other constructs that ought to be discriminable from the construct of interest (Campbell & Fiske, 1959).

Examining the MTMM matrix of these measures allows one to evaluate convergent and discriminant validity of the construct of interest. Petrinovich notes that these measures are “of some latent attribute”, thus a researcher ought to use best practice for latent variables and confirmatory factor analysis (see, for example, Kenny & Kashy, 1992; Marsh et al., 1992) when performing such an analysis. SEM allows us to model each measure as more than a simple sum score; as a variable formed statistically from its indicators (e.g., questions, items), while formally modeling measurement error (Eid et al., 2008; Widaman, 1985). Once the latent variable (or variables) for each measure have been modeled and deemed reliable and valid, finding correlations between these variables is straightforward. SEM has numerous advantages over correlational and regression analyses, such as allowing the researchers to model traits and method effects as separate factors (e.g., Pohl & Steyer, 2010) as well as controlling for the systematic overestimation (Westfall & Yarkoni, 2016) and underestimation (Riemann et al., 1997) of correlations when measurement error is not effectively controlled for.

Once latent variables are established and validated in SEM, the researcher can use these variables in additional analyses. SEM has advantages over a piecemeal approach, wherein a researcher might use different software packages for MTMM analysis, multiple regression, and mediation. Performing all analyses in SEM allows the researcher to examine fit across models while monitoring changes in coefficients as model complexity increases, all while making sure that construct quality is maintained in successive models.

Like all statistical models, SEM performance is incumbent on fit. SEM fit has been researched extensively (Chen et al., 2008; Hu & Bentler, 1999), and best practice places a reasonably high bar on what is considered acceptable fit. As with other statistical models, SEM can produce point predictions; one use of which is in cross validation. Cross validation, otherwise known as out-of-sample testing, is ubiquitous in other methods of analysis, such as machine learning (Browne, 2000). The quality of machine learning models cannot be adequately evaluated without testing an already fitted model on data that were not used to fit the model. Cross validation and the principle of testing by prediction can be applied to any statistical model (Salganik et al., 2020). This is another concept Petrinovich draws attention to that has begun to see widespread adoption in the behavioral and social sciences. SEM is well-suited to producing predictions using out-of-sample data that can cross validate a model.

SEM as originally conceived operates in a null hypothesis significance testing (NHST) framework, despite NHST’s inadequacies noted by Petrinovich and others (Nickerson, 2000). However, SEM has been redeveloped to function in a Bayesian framework, and takes full advantage of what Bayesian data analysis offers (Muthén & Asparouhov, 2012). Bayesian analysis and its strengths are beyond the scope of this commentary, but the benefits of such an approach (Tejedor, 2017) address many of Petrinovich’s concerns. For instance, Petrinovich (via

Oakes, 1986) calls for “the development of rich causal theories capable of producing non-trivial statistical predictions.” As mentioned, SEMs are a form of path modeling, and a restricted subset of path models can be plotted as directed acyclic graphs (DAG), which are a modern staple of causal reasoning. DAGs ask a researcher to specify all factors, both those of interest and confounders, as well as the relationships among these factors. Crucially, a DAG must have no closed loops of feedback among variables, which is what makes them “acyclic”. Creating DAGs is the first step in developing a causal analysis, which can then be carried forward into a formal structural causal model (Pearl, 2009), of which SEM is one of the most common.

In summary, SEM is a widely applicable, intuitive framework for statistical inference, particularly when unobserved variables are at work. The future of the behavioral sciences depends on the adoption of such robust analytic and statistical procedures as well as more rigorous conceptual thinking. SEM provides a path forward for researchers at many levels, across many disciplines.

References

- Browne, M. W. (2000). Cross-validation methods. *Journal of Mathematical Psychology, 44*, 108-132.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.
- Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods & Research, 36*, 462-494.
- Eid, M., Nussbeck, F. W., Geiser, C., Cole, D. A., Gollwitzer, M., & Lischetzke, T. (2008). Structural equation modeling of multitrait-multimethod data: Different models for different types of methods. *Psychological Methods, 13*, 230-253.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1-55.
- Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin, 112*, 165-172.
- Marsh, H. W., Byrne, B. M., & Craven, R. (1992). Overcoming problems in confirmatory factor analyses of MTMM data: The correlated uniqueness model and factorial invariance. *Multivariate Behavioral Research, 27*, 489-507.
- Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science, 34*, 103-115.
- Meehl, P. E. (1990). Appraising and amending theories: The strategy of Lakatosian defense and two principles that warrant it. *Psychological Inquiry, 1*, 108-141.

- Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods, 17*, 313-335.
- Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods, 5*, 241-301.
- Oakes, M. (1986). An evaluation of the major school of statistical inference. *Statistical Inference: A Commentary for the Social and Behavioural Sciences, 6*, 118-148.
- Pearl, J. (2009). *Causality: Models, reasoning, and inference*. 2nd ed. Cambridge University Press.
- Pohl, S., & Steyer, R. (2010). Modeling Common Traits and Method Effects in Multitrait-Multimethod Analysis. *Multivariate Behavioral Research, 45*, 45-72.
- Riemann, R., Angleitner, A., & Strelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using self- and peer report NEO-FFI scales. *Journal of Personality, 65*, 449-475.
- Rosseel, Y. (2011). *lavaan: An R package for structural equation modeling and more Version 0.4-9 (BETA)*. Ghent University.
- Salganik, M. J., Lundberg, I., Kindel, A. T., Ahearn, C. E., Al-Ghoneim, K., Almaatouq, A., ... & McLanahan, S. (2020). Measuring the predictability of life outcomes with a scientific mass collaboration. *Proceedings of the National Academy of Sciences of the United States of America, 117*, 8398-8403.
- Shadish, W. R., & Sullivan, K. J. (2012). Theories of causation in psychological science. In: H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 1. Foundations, planning, measures, and psychometrics* (pp. 23-52). American Psychological Association.
- Tejedor, J. P. (Ed.). (2017). *Bayesian Inference*. IntechOpen.
- Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PLOS One, 11*, e0152719.
- Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement, 9*, 1-26.