Factor Analysis and Structural Equation Modeling

Sultan Altikriti and Claudia N. Anderson
School of Criminal Justice, University of Cincinnati, Cincinnati, Ohio, USA

The purpose of structural equation modeling (SEM) is twofold. Specifically, structural equation models simultaneously combine confirmatory factor analysis (CFA) and path analysis. CFA entails using multiple, measured variables to create "unobserved" measures, or latent constructs. Path analysis estimates the structural relationships among observed measures and latent constructs. More broadly, this allows researchers to explore complex causal relationships between observed variables and latent constructs. Briefly then, SEM includes (i) a measurement model, which creates latent constructs from observed variables, and (ii) a structural model, which assesses relationships between variables. This entry presents a brief and nontechnical overview of SEM in two major parts. The first part elaborates on the conceptual underpinnings of measurement and structural models. The second reviews the specific uses and goals of SEM in social science research.

Measurement Model

Many constructs of interest are not directly observable. For example, socioeconomic status (SES), happiness, and motivation are not directly observable, and consequently not readily measured (Byrne 2012). The purpose of a measurement model is to create measures of unobservable constructs using existing, observed variables called indicators. For example, income, education, and net worth are all indicators that can be quantified. These indicators can be postulated to embody the broader construct of SES. Generally, factor analysis capitalizes on the covariation among indicators tapping into a related, broader construct. To elaborate on the SES example, there is likely covariation among income, education, and net worth. Consequently, the measurement model portion of an SEM estimates a latent variable – a hypothetical construct – with these underlying, covarying indicators.

The ability to create measures of latent constructs has additional benefits, one being increased validity of the measure of interest. The use of multiple, related indicators allows researchers to capture a construct more comprehensively (i.e. content validity) and, in theory, more accurately than any single measure (Anastasi and Urbina 1997). In other words, creating a latent construct from several related indicators better represents multidimensional and difficult-to-quantify constructs. For example, income alone may be insufficient in capturing the essence of socioeconomic status; however, incorporating education and net worth better taps into the broader notion of SES.

The Encyclopedia of Research Methods in Criminology and Criminal Justice, Volume II, First Edition.
Edited by J.C. Barnes and David R. Forde.

Latent constructs also mitigate concerns stemming from measurement error. All measures comprise two forms of measurement error: random and systematic (Holland 1986). Random error decreases reliability and inflates standard errors, attenuating potential relationships and reducing their precision. Systematic error, meanwhile, has a more insidious consequence: the introduction of bias. The use of multiple observed indicators to create a latent construct allows for the estimation and correction of the error variances (Saris and Revilla 2016). Thus, confirmatory factor analysis permits a methodical process of creating latent constructs involving the validation of the construct and the underlying measures. In other words, factor analysis assesses the assumption that the indicators covary, and that variation in the latent construct accounts for covariation among the underlying measures, with the goal of representing several indicators through fewer latent constructs.

Although related, exploratory factor analysis (EFA) is distinct from confirmatory factor analysis. The main distinction is that EFA estimates all intercorrelations among indicators. Relying on empirical research and theory, CFA models have a priori specified paths from latent constructs to the observed measures. Unlike EFA, which explores intercorrelations among all the variables, using CFA restricts nonspecified paths to zero and estimates the remaining paths. CFA requires an exact number of hypothesized factors *and* the arrangement of indicators across factors to be specified prior to the analyses. For example, a researcher would specify an "SES" factor by telling the software that income, education, and net worth constitute the same factor. In EFA, an unrestricted exploratory process unfolds where indicators are estimated for an undefined number of latent factors. Placing these restrictions in CFA allows for a unique solution to create the latent construct. Two questions should be answered before conducting a CFA: (i) which indicators are caused by which factors, and (ii) which indicators are unrelated to which factors?

In visually representing these relationships, SEM diagrams depict single indicators as rectangles and latent factors as ellipses. Pattern coefficients indicate the effect of each factor on the indicators (also called factor loadings). Visually, these are represented via a unidirectional arrow from the latent factors to the underlying indicators. These are termed effect (or reflective) indicators, which are understood to be caused by the latent construct (Bollen and Bauldry 2011). For example, intelligence causes verbal and mathematical IQ scores, where changes in intelligence are expected to cause changes in the underlying measures. The factors are presumed to cause the underlying indicators. This is not always the case. Causal (or formative) indicators are the opposite of effect indicators. These indicators are presumed as causes of the latent factor. For example, SES does not cause income or education. Rather, income and education determine SES, so the directionality of the arrows is reversed – from the indicators to the latent factors. It should be noted that few CFAs are strictly confirmatory. Typically, CFA models are modified if initial analyses indicate poor factor loadings or fit.

Structural Model

The structural portion of an SEM relies on a method of multiple regression where systems of relationships are specified between multiple independent and dependent variables (Bag 2015). In general, three types of variables are identified. Endogenous variables are outcomes – or dependent variables – with a specified causal factor within the path model. Causal factors are termed exogenous variables. Exogenous variables are free to vary and do not have variables influencing them specified within the model. Lagged endogenous variables specified in a path analysis play both roles: a causal factor that influences another variable and an outcome that is influenced by a preceding variable.

An advantage of using path analysis is that the system of relationships can be represented either in equation form or visually through diagrams. In visually representing the relationships between variables in the model, directional arrows are used to denote a relationship. Covariances between variables are delineated by a curved double-headed arrow. This indicates a nondirectional relationship where two variables are assumed to covary. Straight, one-headed arrows are specified from a variable (exogenous) believed to cause another variable (endogenous). A variable with an incoming arrow and outgoing arrow is lagged endogenous as it is both caused and causal. Specification of directional arrows should be grounded in theory, temporal order, and other factors (see Davis 1985 for a review on conditions of causality). Statistically, these specifications represent a system of regression estimates (Hox and Bechger 2015).

Building on the logic of simultaneous "systems" of relationships, direct and indirect effects can be estimated. Decomposing total effects into direct and indirect effects can provide evidence of mediation between variables. Regardless of the methodology – SEM or otherwise – truly causal claims of mediation are not possible without experimental designs. Therefore, in path analysis, and in SEM in general, hypothesized mediation effects must be supported by theory, logic, and proper specification of the model. Even if all three elements are met, causal claims should be viewed with caution (see Davis 1985 for a thorough introduction of causal flow in path analysis).

Structural Equation Model

Models are recognized as full SEMs when they comprise both a measurement model (CFA) and a structural model (path analysis). In other words, when at least one latent variable is included in a path analysis, the analysis is termed an SEM. Within an SEM, the measurement model serves to create the latent constructs while the structural model estimates the paths specified between them (Byrne 2012). The benefit of using SEM is that the unobservable latent constructs can be measured simultaneously within the specified system of relationships in the path analysis.

Kline (2015) outlines six steps to conducting an SEM. The first involves the specification of the initial model. Relying on theory and postulated hypotheses, a researcher must specify relationships between the variables in the model. In other words, the initial model should represent the expected relationships based on theory, although later refinement and justification may follow, as in step five. Once the SEM has been fully specified, the model parameters – relationships between variables – must be estimated. To estimate these parameters, the model must be "identified." SEMs with parameters that can be uniquely determined are considered identified. To achieve this, there must be a balance between known parameters (variances and covariances) and the unknown parameters to be estimated by the structural and measurement models. An identified model includes as many known parameters as unknown parameters. Ideally, a model is overidentified, where the known parameters exceed the unknown parameters, permitting tests of model fit. An underidentified model cannot be estimated due to having more unknown than known parameters. Identification is represented by the model degrees of freedom. Although it is beyond the scope of the current entry to detail identification, most SEM software will report identification information in the output.

Once a model has been identified, or, preferably, overidentified, step three is to curate the sample data and prepare it for analysis. Step four involves the estimation of the model and evaluation of model fit, with three potential outcomes: (i) if the model exhibits adequate fit, then proceed to step six; (ii) if the model fit is not adequate and the model can be justifiably respecified, then step five is to respecify the model to improve fit; (iii) if the model cannot be justifiably respecified, no model is retained. Step six involves interpreting and reporting the results of the model, including the fit statistics and parameter estimates.

Model Estimators

Unlike traditional regression models, which rely on patterned variation between two or more variables, SEM relies on the correspondence of the specified model with the underlying data it is tested against. Where regression models use raw data to predict outcomes, SEM relies on the variance–covariance matrix of the data. Because of this, different estimators are often used. In assessing the variance–covariance structure of a model, there are several available estimators. The most popular method for assessing continuous data that meet distributional assumptions of multivariate normality is Maximum Likelihood (ML) estimation (Kline 2015). ML relies on having overidentified models that are compared for the best fit to estimate the unknown parameters specified in the model (Bag 2015). Software packages provide alternative estimation strategies such as weighted least squares (WLS) and other asymptotically distribution-free estimators (see Bag 2015; Kline 2015) if the data do not meet the assumption of normality or the outcome variables are not continuous. These alternative estimators have their own limitations, however, such as onerous demands on sample size.

Fit

After a model has been specified and the parameters estimated, the results of the analysis can be compared to the underlying data variance–covariance matrix. The congruence between the hypothesized model and the underlying data provides what are called goodness-of-fit statistics. These convey how well the structure in the model represents the sample data. Adequate fit statistics lend support for the plausibility of the specified model, whereas inadequate fit suggests that the model is not tenable (Byrne 2012). This is integral to evaluating hypothesized models and rejecting poorly specified models.

Fit statistics can generally be categorized into two bins: global and local fit. Global fit statistics – also called simultaneous or full-information – evaluate the entire model, resulting in a single statistic to quantify the fit of the proposed model. Local fit tests provide more precise points of potential specification concerns (Thoemmes et al. 2018). This entry briefly reviews several common tests of global fit (see Thoemmes et al. 2018 for a dedicated review of local fit).

The global χ^2 statistic is often used to evaluate an SEM. This statistic is calculated based on the discrepancy between the variances and covariances in the specified model and the those of the sample data used (Kline 2015). This is considered an accept-support test, where the resulting statistic provides a p-value that can be used to assess the model, where failing to reject the null (e.g., $p \geq .05$) implies support for the model. Researchers should view the χ^2 statistic with caution, however, as its reliance on sample size can result in smaller p-values with larger samples, in turn making it more difficult to find support for a hypothesized model.

Supplementary global fit statistics are often reported along with the χ^2 statistic. These include (but are not limited to) the comparative fit index (CFI) and root mean square error of approximation (RMSEA), which rely on cut-off values to provide support for the model. Although the calculation of these statistics is beyond the scope of the current entry, the resulting estimates are often provided in the output of the software used to conduct an SEM. Common cut-off criteria used by researchers to support a proposed model are CFI ≤ .95 and RMSEA ≤ .06 (see Hu and Bentler 1999 for a thorough review of fit statistics).

Uses in Criminology

Creating Measures

Measurement has plagued the field of criminology for some time, making valid and comprehensive measures scarce (Cullen et al. 2019). Using CFA, latent factors can be created to represent intended constructs. This opens avenues for researchers to use latent factors either from primary data or from theoretically sound indicators in existing secondary data. Of course, researchers should be judicious in creating these measures, ensuring that they meet validity concerns. The same principle can be applied to a single measure collected over multiple time periods to create latent intercepts and slopes. These latent variables can then be used to assess change over time using latent growth curve analysis.

Testing Causal Models

SEM can be used, as multivariate analyses are often used, to assess research questions regarding the relationship between directly or indirectly observable independent and dependent variables. One key goal of SEM is to assess the validity of a specified causal model or theory. Given this, SEM is typically characterized as a confirmatory rather than an exploratory technique. SEMs are tasked with finding support for a specified model's correspondence with underlying data. In other words, does the theorized causal model match the data? Criminological theory and attendant hypothesized causal relationships can be supported or disconfirmed with this technique. At the same time, SEMs aid researchers in examining mechanisms proposed by theory by parsing out direct and indirect effects (i.e. mediation) of purported causal variables on a particular outcome (Gunzler et al. 2013).

Integrating Theory

Criminology is averse to theoretical disconfirmation. Rather, theories are retrofitted and brought back for more testing (Pratt et al., 2006). These theories often lack the conceptual breadth necessary to explain complex behavior such as deviance or offending (Weisburd & Piquero, 2008). SEM can bolster existing theories, or support more comprehensive ones, by integrating distinct theories into one cohesive causal structure. Integrating fundamentally compatible theories under one framework using SEM could improve prediction and uncover more complex causal paths to offending.

Software

There exists an abundance of software that can be used to conduct SEM. Of these, some of the more popular ones include LISREL, Stata, Mplus, and R. Each software has its benefits and drawbacks (see El-Sheikh et al. 2017 for a dedicated review of SEM software). Given the rapid change and updating of software, however, the reader is cautioned to do their homework on which may be best suited for their individual needs. Interface and features aside, the differences in estimation and results are negligible across software platforms (El-Sheikh et al. 2017).

Conclusion

SEM is a powerful technique that affords researchers several advantages over traditional multiple regression techniques. Chief among these is that (i) SEM allows for the modeling of measurement error and unexplained variance; (ii) unobserved latent variables can be constructed and tested using SEM; and (iii) SEM allows users to test complex processes, including mediation effects and multilevel models (Byrne 2012). In addition to these, SEM allows users to represent the results using visual diagrams rather than traditional equations or tables, potentially increasing its appeal.

SEM modeling, however, is not without drawbacks. To construct an SEM, users must specify each parameter to be estimated prior to any analyses. Further, SEM requires that models be identified or overidentified, which could become a challenge if many parameters are to be estimated. Finally, SEM requires large samples to provide unbiased estimates. Although no simple rules of thumb exist for sample size, increasingly complex models with more variables and parameters only exacerbate this concern (Kline 2015). As with any method of analysis, a researcher must be judicious and thorough in specifying why they used a particular method. It is ultimately up to the individual researcher to choose the most appropriate approach to analyze their data and to adequately justify that decision. For criminologists, thoughtfully applied SEM can be one way to combine creativity and analytical rigor, a pairing essential for unraveling the complexities of human behavior.

References

Anastasi, A. and Urbina, S. (1997). *Psychological Testing*, 7e. Upper Saddle River, NJ: Prentice Hall.

Bag, S. (2015). A short review on structural equation modeling: applications and future research directions. *Journal of Supply Chain Management Systems* 4: 64–69.

Bollen, K.A. and Bauldry, S. (2011). Three Cs in measurement models: causal indicators, composite indicators, and covariates. *Psychological Methods* 16: 265–284.

Byrne, B.M. (2012). *Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming*. New York: Routledge.

Cullen, F.T., Pratt, T.C., and Graham, A. (2019). Why longitudinal research is hurting criminology. *The Criminologist* 44: 2–7.

Davis, J.A. (1985). *The Logic of Causal Order*. Thousand Oaks, CA: Sage Publications.

El-Sheikh, A.A., Abonazel, M.R., and Gamil, N. (2017). A review of software packages for structural equation modeling: a comparative study. *Applied Mathematics and Physics* 5: 85–94.

Gunzler, D., Chen, T., Wu, P., and Zhang, H. (2013). Introduction to mediation analysis with structural equation modeling. *Shanghai Archives of Psychiatry* 25: 390–394.

Holland, P.W. (1986). Statistics and causal inference. *Journal of the American Statistical Association* 81: 945–960.

Hox, J.J. and Bechger, T.M. (2015). An introduction to structural equation modeling. *Family Science Review* 11: 354–373.

Hu, L. and 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.

Kline, R.B. (2015). *Principles and Practice of Structural Equation Modeling*. New York: Guilford Publications.

Pratt, T. C., Cullen, F. T., Blevins, K. R., Daigle, L. E., & Madensen, T. D. (2006). The empirical status of deterrence theory: A meta-analysis. In F. T. Cullen, J. P. Wright, & K. R. Blevins (Eds.), Takingstock: *The status of criminological theory* (pp. 367–395). TransactionPublishers.

Saris, W.E. and Revilla, M. (2016). Correction for measurement errors in survey research: necessary and possible. *Social Indicators Research* 127: 1005–1020.

Thoemmes, F., Rosseel, Y., and Textor, J. (2018). Local fit evaluation of structural equation models using graphical criteria. *Psychological Methods* 23: 27–41.

Weisburd, D. and Piquero, A. R. (2008). How well docriminologists explain crime? Statistical modeling in published studies. *Crime and Justice*, 37, 453–502.