**Visualization Strategies to Evaluate SEMs**

Whereas visual model specification using path diagrams or directed acyclic graphs (DAGs) is ubiquitous in SEM applications, there is less emphasis on post-estimation visualization strategies (Bollen & Arminger, 1991; Raykov & Penev, 2014; Yuan & Hayashi, 2010). SEM practitioners generally rely on global fit tests and approximate fit indices when evaluating model adequacy rather than visual inspection (Bollen & Arminger, 1991; Yuan & Hayashi, 2010). Poor model fit in SEM can be the result of misspecification of the causal structure and/or violation of model assumptions. Consequently, many SEM scholars have emphasized the importance of diagnostic procedures evaluating the tenability of model assumptions and indicators or poor local fit. However, these procedures are rarely reported in applied papers and, unlike global fit indices, are not typically included by default in output provided by SEM applications.

There is sparse literature describing visual strategies for evaluating the adequacy of SEMs relative to the ponderous literature discussing the merits of global fit indices. However, visual procedures for exist for evaluating the tenability of model assumptions, diagnosing causal misspecifications, and selecting the best model from a group of competitors. Importantly, evidence of violations of modeling assumptions (e.g., non-normally distributed case residuals) could be a signal of causal misspecification.

**Visualization using aggregate statistics.** A common approach for identifying indicators of poor local model fit is to plot the distribution of the residual covariances/correlations (e.g., using stem-and-leaf plots or histograms) to aid in identifying specific model implications that are most inconsistent with the observed data (Bollen, 1989; Bollen & Arminger, 1991). Residual covariances may signal problems that are otherwise masked in global fit indices. Distributional plots of residual covariances are particularly helpful for identifying aberrant residual covariances in complex models in which the residual covariance matrix has many elements.

**Visualization using factor score estimates and individual case residuals.** Global fit indices and residual covariances are important for detecting model misspecification; however, because they are based on aggregate statistics, they do not identify aberrant cases that may influence model parameter estimates and fit. Bollen and Arminger (1991) demonstrated methods of calculating raw and standardized individual case residuals (ICRs) representing differences between observed and model-predicted case values for outcome variables. They used stem-and-leaf, index, and histogram plots to help locate outlying and influential observations. This technique requires estimation of factor scores (e.g., using regression-based or Bartlett methods). Pek and MacCallum (2011) demonstrated how case diagnostic procedures commonly used in GLMs (e.g., Mahalanobis distance, generalized Cook’s *D*, and *DFBETAs*) can be applied to SEMs to detect influential cases. Similarly, Flora and colleagues applied these diagnostic procedures and others specifically to factor analysis models (Flora et al., 2012). Yuan and Hayashi (2010) used visualizations of Mahalanobis distance metrics to identify high-leverage cases and outliers . Visualization procedures showing case influence on model fit (e.g., likelihood differences) and parameter estimations (e.g., generalized Cook’s *D*) have been implemented in open-source R packages, including fa.outlier (Chalmers & Flora, 2015) and influence.SEM (Pastore & Altoe, 2018). These packages use both statistical tests and graphs to aid in identifying cases with strong influence on model fit and parameter estimates.

Muthén and Asporouhov (2017) demonstrated how estimated factor scores and ICRs can be used to detect specific model misspecifications. First, they showed that plots of estimated factor scores for a latent outcome variable against observed predictor variables can be used to detect unspecified nonlinear effects of the predictor on the latent outcome. Second, they used ICR scatterplots to detect violations of local independence in a latent factor model. When two reflective indicators (Y1 and Y2) of a latent factor (η) had an unmodeled common cause, they showed that plotting the indictors’ ICRs against one another could help identify the non-zero residual covariance, prompting modelers to consider an alternate model. Finally, they demonstrated in a latent factor model how plotting predicted values for a reflective indicator () against the observed indicator values (*Y*) could uncover unmodeled heterogeneity that could be better captured using a mixture model.

Raykov and Penev (2014) demonstrated how ICRs can be used to aid in model selection in the context of latent growth curve modeling. When comparing linear and quadratic growth curve models for the same data, for example, they showed that a scatterplot of the ICRs for the quadratic model vs. ICRs for a linear model can help the modeler see which model best minimizes ICRs, leading to selection of the appropriate model. In the context of growth mixture modeling, Chang et al. (2005) showed how visualization of empirical Bayes residuals (e.g., Q-Q and trajectory plots) can aid in determining the appropriate number of classes, an adequate shape of within-class growth trajectories, and missing confounders.

A potentially limitation of SEM visual model-evaluation procedures using individual factor score estimates and ICRs is that they rely on factor score estimates. It is well-known that individual latent factor scores cannot be uniquely determined (Grice, 2001; Rigdon et al., 2019; Steiger, 1996). In cases where factors are highly indeterminant (e.g., factors with few indicators only weakly predicted by the latent factor), different factor score estimation methods can yield highly discrepant values, potentially even estimates that are negatively correlated (Grice, 2001). Therefore, the conclusions drawn from visual diagnostic strategies using individual scores could vary depending on which factor score estimation strategy is employed.