Structural Equation Modeling: Some Issues

A paper by Lei and Wu covers some basic issues concerning the use of a special statistical method known as structural equation modeling (SEM). That paper makes the following points, among others:

1. “One of the primary advantages of SEM (vs. other applications of GLM) is that it can be used to study relationships among latent constructs that are indicated by multiple measures. It is also applicable to both experimental and non-experimental data, as well as cross-sectional and longitudinal data…” [p. 33]

2. “SEM takes a confirmatory (hypothesis testing) approach to the multivariate analysis of structural theory, one that stipulates causal relations among multiple variables. The causal pattern of intervariable relations within the theory is specified a priori. The goal is to determine whether a hypothesized theoretical model is consistent with the data collected to reflect this theory. The consistency is evaluated through model-data fit, which indicates the extent to which the possible network of relations among variables is plausible…” [pp. 34-35]

3. “SEM is a large sample technique…A general rule of thumb is that the minimum sample size should be no less than 200 (preferably no less than 400 especially when observed variables are not multivariate normally distributed) or 5-20 times the number of parameters to be estimated, whichever is larger…” [p.36]

4. “In SEM, a variable can serve both as a source variable (called an exogenous variable, which is analogous to an independent variable) and a result variable (called an endogenous variable, which is analogous to a dependent variable) in a chain of causal hypotheses…Furthermore, feedback loops among variables…are permissible in SEM, as are reciprocal effects (e.g., learning motivation and achievement affect each other). In path analyses [a precursor in history to SEM] observed variables [not latent constructs] are treated as if they are measured without error [unlike SEM], which is an assumption that does not likely hold in most social and behavioral sciences. When observed variables contain error, estimates of path coefficients may be biased in unpredictable ways, especially for complex models…” [p.34]

5. “A sound model is theory based. Theory is based on findings in the literature, knowledge in the field, or one’s educated guesses, from which causes and effects among variables within the theory are specified. Models are often easily conceptualized and communicated in graphical forms…a directional arrow (→) is universally used to indicate a hypothesized causal direction. The variables to which variables are pointing are commonly termed endogenous variables…and the variables having no arrows pointing to them are called exogenous variables…” [p.35]

6. “A properly specified structural equation model often has some fixed parameters and some free parameters to be estimated from the data…Once model parameters have been estimated, one would like to make a dichotomous decision, either to retain or reject the hypothesized model. This is essentially a statistical hypothesis-testing problem, with the null hypothesis being that the model under consideration fits the data. [So, in SEM, a modeler who fails to reject the null hypothesis has a “finding” of sorts to report, unlike most other modeling situations.] [p.36]


Additional abstracts of research can be viewed at our website at:

[Abstract by Willard Hom, Director of Research & Planning, System Office, California Community Colleges, 5/21/08]