A NOTE ON ESTIMATING SECOND-ORDER INTERACTIONS AND QUADRATICS IN LATENT VARIABLES

Robert A. Ping, Jr., Wright State University, U.S.A.

ABSTRACT

Because little is k0nown about interactions involving second-order latent variables (LVs) (i.e., LVs that have other LVs as "indicators") in structural equation models, the paper suggests a specification for an interaction between a second-order LV and a first-order LV. Second-order constructs were proposed by Jöreskog (1970, *Biometrika*). These second-order LVs have first-order LVs as their "indicators." Each of these first-order "indicator" LVs has observed indicators as usual. Second-order LVs have received attention recently (e.g., Gerbing, Hamilton and Freeman 1994, *J. Mgt.*). Estimating these LVs, although not difficult, is not a straightforward task in LISREL, EQS, etc. In addition, there is no guidance for estimating an interaction involving a second-order LV (e.g., XZ in

$$Y = \beta_1X + \beta_2Z + \beta_3XZ + \zeta_Y \,, \qquad (1$$

where Z is a second-order construct, β_1 through β_3 are unstandardized "regression" or structural coefficients, and ζ_Y is the estimation or prediction error, also termed the structural disturbance term). Equation 1 can be factored to produce a coefficient of Z due to the interaction XZ, i.e.,

$$Y = \beta_0 + \beta_1X + (\beta_2 + \beta_3X)Z + \zeta_Y \,. \qquad (2$$

Alternatively, Equation 1 can be re-factored to produce a coefficient of X due to the interaction XZ (i.e., $(\beta_1 + \beta_3Z)X)$. The amount of interaction between X and Z in their association with Y in Equation 2 (also termed X's moderation of the Z-Y association) is the strength (i.e., the magnitude) and the direction (i.e., the sign) of the coefficient of Z, $(\beta_2 + \beta_3X)$, in Equation. Because X takes on a range of values in the study, $(\beta_2 + \beta_3X)$ takes on a range of values (and a range of significances). The possibilities for specifying a second-order by first-order interaction, XZ for example, where X is a first-order LV and Z is a second-order LV with 3 first-order LVs, Z_1, Z_2 and Z_3, are considerable, but most of them are impractical. However, "indicator LVs" of XZ, XZ_i, could be specified with the observed indicators $x{:}z_1 = (x_1+x_2+...+x_m)(z_{1,1}+z_{1,2}+...+z_{1,n})$, $x{:}z_2 = (x_1+x_2+...+x_m)(z_{2,1}+z_{2,2}+...+z_{2,p})$, and $x{:}z_3 = (x_1+x_2+...+x_m)(z_{3,1}+z_{3,2}+...+z_{3,q})$. These indicators have fixed loadings and measurement error variances (see Ping 1995, *JMR*), and their observed values $x{:}z_i$ can be computed for each case in a data set. Alternatively, the second-order construct Z could be re-specified as a first-order construct by replacing Z_1 by the sum of its indicators, and doing the same for Z_2 and Z_3. This re-specification of a second-order construct as a first-order construct using sums of indicators has been reported (e.g., Dwyer and Oh 1987). The corresponding XZ interaction would then be a first-order by first-order interaction with the indicator $x{:}z = (x_1+x_2+...+x_m)(\Sigma z_{1,i}+\Sigma z_{2,i}+\Sigma z_{3,i})$, where $\Sigma z_{j,i}$ is the sum of the terms of Z_j. This indicator also has fixed loadings and measurement error variances (see Ping 1995), and its observed values (i.e., $(x_1+x_2+...+x_m)(\Sigma z_{1,i}+\Sigma z_{2,i}+\Sigma z_{3,i})$) can be computed for each case in a data set. A variation of this approach would be to specify Z as a first-order construct by replacing Z_1 by its factor score, and doing the same for Z_2 and Z_3. The resulting XZ interaction would then be a first-order by first-order interaction with the indicator $x{:}z = (x_1+x_2+...+x_m)(\Sigma\omega_{1,i}d_i+\Sigma\omega_{2,i}d_i+\Sigma\omega_{3,i}d_i)$, where d_i is the ith indicator in the measurement model corresponding to the structural model of interest (i.e., x_1, x_2, ... x_m, $z_{1,1}$, $z_{1,2}$, ... $z_{1,q}$, $z_{2,1}$, $z_{2,2}$, ... $z_{2,q}$, $z_{3,1}$, $z_{3,2}$, ... $z_{1,q}$, and any other indicators of the exogenous and endogenous variables in the model--in this case there are none), $\omega_{1,i}$ is the factor score weight or coefficient for Z_1 and indicator d_i, $\omega_{2,i}$ is the factor score weight/coefficient for Z_2 and indicator d_i, etc. (i.e., $\Sigma\omega_{1,i}d_i$ is the full factor score for Z1, etc.). This x:z indicator also has fixed loadings and measurement error variances (see Ping 1995), and its observed values (i.e., $(x_1+x_2+...+x_m)(\Sigma\omega_{1,i}d_i+\Sigma\omega_{2,i}d_i+\Sigma\omega_{3,i}d_i)$) can be computed in each case. For pedagogical purposes a real-world data was reanalyzed. A survey involving the first-order LVs U, V and W, the second-order LV T with 3 first-order "indicator" LVs, T_1, T_2 and T_3, and the interaction UxT, produced more than 200 usable responses. T was specified initially as a second-order construct (i.e., its proper specification) and the model (not shown) was estimated. Then the model was re-estimated with T re-specified as a first-order construct by replacing its first-order indicator T_1 by the sum of the indicators of T_1, and doing the same for T_2 and T_3. Finally the model was re-estimated with T re-specified as a first-order construct by replacing T_1 by its factor score, and doing the same for T_2 and T_3. Based on the reproduced covariance matrices (and several other criteria) of the measurement models containing T, U, V and W with the various specifications for T, a second-order interaction may be adequately specified in this survey by replacing the second-order construct's "indicators" (i.e., T_1, T_2 and T_3) by the factor-score for each "indicator" (i.e., the factor score for T_1, the factor score for T_2, etc.) and specifying $u{:}t = (u_1+u_2+...+u_m)(\Sigma\omega_{1,i}d_i+\Sigma\omega_{2,i}d_i+\Sigma\omega_{3,i}d_i)$, where $\Sigma\omega_{1,i}di$ is the factor score for T_1, etc. Nevertheless, because Maximum Likelihood (common) factor scores are known to be approximate, simulations are required to demonstrate that factor scores provide unbiased estimates of "indicator" constructs (although it is widely believed that factor scores can be used to adequately represent constructs). Simulations are also required to demonstrate that a first-order by second-order interaction specified using a Ping (1995) single-indicator with a factor scored specification for T produces unbiased estimates (although this is likely because factor score indicators do not violate the assumptions underlying the Ping 1995 technique any more or less than any other indicators).