A comparison of estimates from Mplus with WLSMV and Gibbons' program with ML in the bi-factor analysis of QOL data

Mplus did not converge when a method factor loading of the first, Life as a whole, item was estimated. Therefore, I fixed its value close to your estimate (0.001) in following report. Because only one digit of estimate was reported, sensitivity analysis was conducted by varying its value from 0.0006 to 0.0014. Estimates were not sensitive to the ranging values.

The first and the second block in Table 1 (in Excel file) are estimates from Gibbons' and Mplus programs respectively. Their differences are generally small.

Transformation and rescaling of parameters in Mplus

Scale: Because the scales, the standard deviation of the latent response variable, y^*, are different in Gibbons et al. (2007) and my result of Mplus (theta parameterization with residual variance is 1), rescaling is necessary. The scales in Gibbons et al. (2007) and my Mplus are 1 and $\sqrt{a_{1j}^2 + a_{kj}^2 + 1}$ respectively. Denote parameters in Gibbons et al. (2007), γ_{jt} s, α_{1j} s, α_{kj} and, an intercept, and in my Mplus c_j^* s, d_{jt}^* s, a_{1j} s, and a_{kj} s . Denote the scale in my Mplus $\sigma_j \equiv \sqrt{a_{1j}^2 + a_{kj}^2 + 1}$.

Rating scale parameterization: Because the rating scale parameters (c_j s and d_{jt} s) cannot be estimated directly in Mplus, equivalent parameters (c_j^* s and d_{jt}^* s) are estimated. A restriction of $d_1^* = 0$ instead of $\sum_{t=1}^{m-1} d_t = 0$ is imposed.

Transformation of my Mplus to Gibbons et al. (2007) is two step, 1) my Mplus parameters are transformed to the rating scale parameters, then 2) the rating scale parameters are rescaled.

Implementing the rating scale model in Mplus

Rating scale model: I estimated thresholds (d_t^* in my notation) but not distances of adjacent thresholds (d_t , the category parameters) because the category parameters can not be estimated directly in Mplus. The first thresholds parameters were fixed at 0, $d_{1,j}^* = 0$, and each threshold is constrained to be the same over items, $d_{t,j}^* - d_{t,j'}^* = 0$ $j \neq j'$ and for $t = 2...m-1$ (I omit item subscript *j* for the d and d* in my presentation in other places). Then I estimated an intercept parameter, c_j^* , for each item (note c_j^* is different from "intercept parameter" in Gibbons et al. (2007)). Using a relationship of $c_j + d_t = -c_j^* + d_t^*$ and the constraint of $\sum_{t=1}^{m-1} d_t = 0$, the rating scale parameters are obtained from the Mplus parameters,

$$c_j = \frac{1}{m-1}\sum_{t=1}^{m-1}\left(-c_j^* + d_t^*\right)$$

and

$$d_t = \left(c_j + d_t\right) - c_j = \left(c_j^* + d_t^*\right) - \frac{1}{m-1}\sum_{t=1}^{m-1}\left(-c_j^* + d_t^*\right).$$

Fake latent variables: Because all thresholds were specified (saturated), Mplus did allow to specify the intercept term. I created fake latent variables (mj) which are identical to latent response variables (yj) (mj by yj@1, for j=1…35 in Mplus). Then estimate means of the fake latent variables ([mj] in Mplus) specifying their variances are 0 (mj@0 in Mplus). Because Mplus automatically estimates covariances among latent variables, we have to fix covariates among the 35 fake latent variables (e.g., m1 with m2@0 in Mplus) and between 9 factors and 35 fake variables (e.g., m1 with f1@0 in Mplus) to be 0.