Latent Variable Analysis With Categorical Outcomes: Multiple-Group And Growth Modeling In Mplus

Bengt Muthén
University of California, Los Angeles

Tihomir Asparouhov
Muthén & Muthén *

Mplus Web Notes: No. 4
Version 5, December 9, 2002

*The research of the first author was supported under grant K02 AA 00230 from NIAAA. The email address is bmuthen@ucla.edu.

Abstract

This note describes latent variable modeling with categorical outcomes in several groups and for longitudinal data. Different parameterizations are discussed as well as issues of identification. A comparison is made between formulating the modeling in terms of conditional probabilities versus using a latent response variable formulation. Two parameterizations used in Mplus are described, including a new parameterization introduced in Version 2.1, May 2002. Differences between binary outcomes and polytomous outcomes are discussed. The LISREL approach is also presented and compared to the Mplus approaches. It is shown that the Mplus approach avoids the LISREL restriction of across-group or across-time invariance of all thresholds parameters, making it possible to study (partial) non invariance also in the thresholds. The techniques are illustrated by factor analysis of antisocial behavior items and by Monte Carlo simulation examples of multiple-group factor analysis and growth modeling, showing good chi-square testing and estimation performance at rather low sample sizes.

1 Introduction

This note contains a technical discussion of parameterization and modeling in latent variable analysis of binary and ordered polytomous outcomes. Multiple-group and longitudinal settings are discussed together because they have analogous considerations regarding invariance and noninvariance of parameters. Two Mplus parameterizations are presented. A new approach introduced recently in Mplus allows for invariance testing of residual variances. The LISREL approach to multiple-group and longitudinal modeling is also presented and compared to the Mplus approaches. For technical references, see Muthén (1979, 1983, 1984, 1996), Muthén and Christofferson (1981), and Technical Appendix 2 of the Mplus User's Guide (Muthén & Muthén, 1998-2002).

Section 2 presents single-group cross-sectional modeling. Section 3 discusses an example, analyzing a factor model for antisocial behavior items. Section 4 presents the Mplus approach for multiple-group modeling. Section 5 extends the analysis of the antisocial behavior items to a gender comparison. Section 6 presents Mplus growth modeling. Section 7 presents the LISREL approach to multiple-group and growth modeling and Section 8 compares Mplus and LISREL approaches. Section 9 presents Monte Carlo simulation examples using Mplus for multiple-group and growth modeling.

2 Parameterization In Single-Group, Cross-Sectional Studies

Latent variable models for categorical outcomes can be presented in two ways, directly postulating a conditional probability model or deriving a conditional probability model from a linear model for latent response variables, where the observed outcomes are obtained by categorizing the latent response variables. It is shown that the two formulations give equivalent results. The discussion clarifies that the latent response variables are a convenient conceptualization, but that it is not necessary that the data have been generated by categorizing latent response variables.

Below, the univariate and bivariate probability expressions are considered. For simplicity, a factor model with a single factor η is considered.

2.1 Conditional Probability (CP) Formulation

Consider an ordered polytomous y_i variable with categories $c = 0, 1, 2, \ldots, C-1$ for individual i. Consider the standard proportional-odds model (Agresti, 1990, pp. 322-324), expressing the probability of being in one of the highest categories,

$$P(y_i \geq c|\eta_i) = F[\alpha_c + \beta\, \eta_i], \tag{1}$$

where F is typically chosen as a normal or logistic distribution function (cf. Muthén & Muthén, 1998-2002; Technical Appendix 1). This formulation includes a binary y with only one α parameter. With a polytomous y, the model assumes parallel probability curves for the events $y \geq c$, i.e. only intercepts α change over those probability expressions, not the slope β.

Item Response Theory uses an equivalent conditional probability formulation written slightly differently,

$$P(y_i \geq c|\eta_i) = F[a\ (\eta_i - b_c)], \tag{2}$$

where a is called the item discrimination and b_c the item difficulty.

The marginal distribution for y is obtained by integrating (1) over η, typically using a normality assumption for η.

Bivariate counterparts to the conditional probability expression (1) introduces no further modeling issue because the y responses are assumed conditionally independent given η. Bivariate marginal probabilities, not conditioning on η, are again obtained by integrating over η using a normality assumption.

2.2 Latent Response Variable (LRV) Formulation

The LRV formulation considers a continuous latent response variable y^* that expresses the amount of understanding, attitude, or illness required to respond in a certain category. This acknowledges that a more fine-grained measurement could have been attempted. Consider for simplicity the 1-factor model for the continuous latent response variable y_i^* for individual i,

$$y_i^* = \nu + \lambda\ \eta_i + \epsilon_i, \tag{3}$$

where ν is an intercept parameter, λ is a factor loading, η is a factor variable, and ϵ is a residual. The expectation and variance of y^* are

$$\mu^* = \nu + \lambda\ \alpha, \tag{4}$$
$$\sigma^* = \lambda^2\ \psi + \theta, \tag{5}$$

where α is the mean of η, ψ is the variance of η, and θ is the variance of the residual ϵ. Here, y^* is related to the observed ordered polytomous variable y via threshold parameters τ as

$$y = c,\ if \quad \tau_c < y^* \leq \tau_{c+1} \tag{6}$$

for categories $c = 0, 1, 2, \ldots, C-1$, where $\tau_0 = -\infty$, $\tau_C = \infty$. This leads to the conditional probability expression,

$$P(y \geq c|\eta) = F[-(\tau_c - \nu - \lambda\ \eta)\ \theta^{-1/2}], \tag{7}$$

where F is typically chosen as a normal or logistic distribution function depending on the distributional assumption for ϵ.

Consider the standardization $E(\eta) = \alpha = 0$ so that $\mu^* = \nu$. Because y^* is a latent variable, its metric is not determined. It is therefore common to standardize to $\nu = 0$ and $\sigma^* = 1$. This may be viewed as θ not being a free parameter to be estimated, but a parameter that is obtained as the remainder

$$\theta = 1 - \lambda^2 \, \psi. \tag{8}$$

This standardization results in one particular metric for the τ and λ parameters of (7), but other metrics are also possible and have the same fit to data. Defining the scaling factor $\mathbf{\Delta}$ corresponding to the inverted latent response variable standard deviation

$$\mathbf{\Delta} = 1/\sqrt{\sigma^*}, \tag{9}$$

one obtains a more general form of (8) where the scaling factor $\mathbf{\Delta}$ can be fixed at other values than one, resulting in

$$\theta = \mathbf{\Delta}^{-2} - \lambda^2 \, \psi. \tag{10}$$

Standardizing to $\theta = 1$ instead of $\sigma^* = 1$,

$$\mathbf{\Delta} = 1/\sqrt{\lambda^2 \, \psi + 1}, \tag{11}$$

which gives yet another metric for the τ and λ parameters.

Given a uni- and bi-variate normality assumption for the y^* variables, the LRV formulation leads to the univariate and bivariate marginal probability expressions in the binary case

$$P(y_j = 1) = \int_{\mathbf{\Delta}_j(\tau_j - \mu_j^*)}^{\infty} \phi_1(y_j^*) \, dy_j^*, \tag{12}$$

and

$$P(y_j = 1, y_k = 1) = \int_{\mathbf{\Delta}_j(\tau_j - \mu_j^*)}^{\infty} \int_{\mathbf{\Delta}_k(\tau_k - \mu_k^*)}^{\infty} \phi_2(y_j^*, y_k^*) \, dy_k^* \, dy_j^*, \tag{13}$$

where ϕ_1 denotes a univariate standard normal density, ϕ_2 denotes a bivariate normal density with unit variances, zero means, and correlation coefficient,

$$Corr(y_j^*, y_k^*) = \mathbf{\Delta}_j \, \sigma_{jk}^* \, \mathbf{\Delta}_k, \tag{14}$$

where σ_{jk}^* is the covariance between the two latent response variables y_j^* and y_k^*,

$$Cov(y_j^*, y_k^*) = \lambda_j \, \psi \, \lambda_k. \tag{15}$$

2.2.1 LRV Model Identification

From an identification point of view, it is important to note that only standardized quantities enter the probability expressions of (12) and (13) in that both the integration limits and the correlation have been scaled by $\mathbf{\Delta}$. From these two expressions the identification status of the model can be determined (cf Muthén, 1979). LRV models that are not identified in terms of uni- and bi-variate probabilities are not identified when also using higher-order terms. This is a function of the model and holds irrespective of estimation method. To see this, consider a multivariate normal $\mathbf{y}^*$ which is $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and has threshold vector $\boldsymbol{\tau}$. The same $\mathbf{y}$ distribution is obtained considering $\boldsymbol{\Delta}\ \mathbf{y}^*$, which is $N(\boldsymbol{\Delta}\ \boldsymbol{\mu}, \boldsymbol{\Delta}\ \boldsymbol{\Sigma}\ \boldsymbol{\Delta})$, with thresholds $\boldsymbol{\Delta}\ \boldsymbol{\tau}$. This implies that the maximum number of parameters that can be identified is $p(p-1)/2 + r$, where p is the number of variables and r is the total number of thresholds summed over all variables. With multiple groups, this number of parameters is multiplied by the number of groups.

In the standardization $\mu^* = 0$, $\sigma^* = 1$, (12) shows that τ is identified as the corresponding z score. Given the threshold parameters, (13) shows the identification of the correlation coefficient. In line with factor analysis for continuous outcomes, fixing one λ to one, the remaining parameters λ and ψ are identified in terms of these correlations.

It may be noted that even when the loadings λ are held equal or fixed for all items, as when testing equality of loadings across items, the residual variances θ are not separately identifiable. This is seen from (14) and (15). A change in a θ value can be absorbed in the factor variance ψ to give the same correlation.

2.3 Comparison Of The Two Formulations

It is clear from (1), (7), and (8) that the CP and LRV formulations are equivalent in terms of relating y and η to each other, with parameters related as

$$\alpha_c = \frac{-\tau_c}{\sqrt{\theta}}, \tag{16}$$

$$\beta = \frac{\lambda}{\sqrt{\theta}}, \tag{17}$$

or using the IRT parameterization,

$$b_c = \frac{\tau_c}{\lambda}, \tag{18}$$

$$a = \frac{\lambda}{\sqrt{\theta}}. \tag{19}$$

It is seen from (17) and (19) that an increased residual variance θ gives rise to a flatter conditional probability curve and therefore attenuates the strength of the relationship between y and η. It is clear, however, that (in a single-group analysis) θ is not separately identifiable from the other parameters, motivating a standardization such as (8). Note that if the standardization of (8) is used, the degree of attenuation depends both on the loading λ and the factor variance ψ. For related discussions, see e.g. Muthén (1979), Muthén (1988), and Muthén, Kao and Burstein (1991). IRT estimates are obtained from the Mplus solution by using the transformation shown in (18) and (19). Standard errors can be obtained via the Delta method (see MacIntosh & Hashim, 2000).

A researcher may take two different viewpoints in terms of modeling, (A) focusing on the CP formulation and its parameters α and β (or b and a) without introducing y^*, or (B) focusing on the LRV formulation and its parameters τ_c, ν, λ, ψ, and θ. The two viewpoints have somewhat different consequences in multiple-group and longitudinal analysis.

Under (A), the researcher is only interested in relating y to η and the α and β (b and a) parameters of the CP formulation are the only ones relevant. The CP model may be derived via the LRV formulation using y^* and related parameters, but this is only used as a pedagogical vehicle to motivate the CP formulation and only the resultant (16), (17) (or (18), (19)) functions of the LRV parameters are relevant.

Under (B), the researcher believes that the LRV formulation has actually generated the data, so that y^* is a substantively meaningful variable and all the LRV parameters are in principle meaningful to the extent that they can be identified.

2.4 Factor Analysis with Covariates

The modeling discussed above can be extended to analysis with covariates (cf Muthén, 1979, 1989). The covariates may influence the factors, and therefore the indicators indirectly, or may influence the indicators directly. Both the CP and LRV formulations can be used to describe such modeling. Using the LRV formulation, the model in (3) is extended as

$$y_i^* = \nu + \lambda\, \eta_i + \kappa\, x_i + \epsilon_i, \tag{20}$$

where κ is the direct effect of the covariate x. In addition, the factor is related to x as

$$\eta_i = \gamma\, x_i + \zeta_i. \tag{21}$$

Using the CP formulation, it follows for a binary item that

$$P(y_i = 1|\eta_i, x_i) = F[-(\tau - \lambda\, \eta_i - \kappa\, x_i)\theta^{-1/2}], \tag{22}$$

so that $\tau - \kappa\, x_i$ can be seen as a new threshold value for the item, a threshold that varies across x values. This implies that the inclusion of direct effects can be used to study item bias, or differential item functioning (DIF), with respect to covariates, e.g. dummy covariates representing groups. If the direct effect is significant, the item shows DIF. This approach has been used in a variety of substantive applications, e.g. Gallo, Anthony and Muthén (1994) and Muthén, Tam, Muthén, Stolzenberg and Hollis (1993).

This approach to studying DIF is a useful first, exploratory step in investigating across-group noninvariance of items. Although the DIF only concerns the threshold parameters, an item showing threshold DIF may also show noninvariance with respect to loadings. The approach consists of two steps. First, a model without any direct effects ($\kappa = 0$) is estimated. "Modification indices" are studied to check the need for including direct effects (in the current Mplus version, first-order derivatives obtained by TECH2 can be used for this purpose[1]).

3 Factor Analysis of Antisocial Behavior

The Mplus Web Note section (see www.statmodel.com) for this web note (Mplus Web Note #4) contains data on antisocial behavior items from the National Longitudinal Survey of Youth (NLSY). NLSY contains 17 antisocial behavior (ASB) items collected in 1980 when respondents were between the ages 16 and 23. The ASB items assessed the frequency of various behaviors during the last year. A sample of $7,326$ respondents is analyzed. Due to a very low proportion of the high-frequency category, the items are dichotomized as 0/1 with 0 representing never in the last year.

The exploratory factor analysis output in the Mplus Web Note section shows that at least 3 factors are clearly defined. Here, the analysis will focus on one of the factors for simplicity. This is the third factor in the 3-factor solution and may be labelled as property offense, being measured well by the 8 items: property, shoplift, stealing less than 50 dollars, stealing more than 50 dollars, conning someone, auto theft, breaking into a building, and stealing goods.

The confirmatory factor analysis output in the Mplus Web Note section shows the results from a 1-factor model for the 8 property offense items. The model appears to fit the data well. The con and auto items have the lowest loadings. In IRT terms, (19) shows that this also implies that these two items have the lowest item discrimination values.

The MIMIC output in the Mplus Web Note section shows a factor analysis of the 8 items using gender as a covariate (male = 0, female = 1). It is seen that females have a lower factor value than males ($\gamma < 0$ in (21)). The model may, however, be misspecified due to gender non-invariance of items. The largest modification index is for the item shoplift (-0.037).

Letting the gender covariate directly influence the shoplift item improves the model fit and gives a significant direct effect ($\kappa = 0.360$ in (20) and (22)). For females the shoplift item has a lower threshold $\tau - \kappa$ in (22), which implies that compared to males, the female conditional probability curve of (22) is shifted to the left. The probability curves can be computed[2] and plotted for different factor values η. The interpretation is that for given property offense factor value, females are more likely to admit to the shoplift item than males.

[1] Note that these modification indices are not scale free, so that values cannot be compared across covariates in different scale, only across items for a given covariate.

[2] In the Delta parameterization, the residual variance θ of (22) is obtained from the R^2 section of the output, 0.492. In Mplus, the function F is the standard normal distribution function which can be looked up in a table.

The gender non-invariance will be further explored in multiple-group analysis below. The multiple-group analysis allows for more flexibility in representing noninvariance of items, not only with respect to thresholds, but also with respect to loadings and residual variances.

4 The Mplus Approach To Multiple-Group Modeling

This section discusses the Delta and Theta multiple-group parameterizations in Mplus, followed by identification issues.

Considering multiple-group analysis using the LRV formulation of (3) - (5) it is important to not standardize all groups to unit y^* variance since that would hide the across-group variation in y^* variance due to across-group variation in λ, ψ, and θ. The CP formulation of (1) also benefits from a more flexible, analogous parameterization in the case of multiple groups as suggested by the scaling factor $1/\sqrt{\theta}$ in (16), (17). Adding this scaling factor, gives a generalized probit model. Although the y^* variance, or the residual variance θ, is not separately identified in a single group, variance differences across groups are identifiable given threshold and loading invariance and this more flexible parameterization can benefit the analysis.

In Mplus this more flexible parameterization is handled in two alternative parameterizations, using a "Delta approach" or using a "Theta approach". In both approaches, ν in (4) is standardized at 0 (inclusion of ν will be discussed further below).

4.1 The Delta Approach

In the Delta approach (cf. the Mplus User's Guide, page 347; Muthén & Muthén, 1998-2002), consider the scaling factor parameter $\boldsymbol{\Delta}_g$ for group g, where

$$\boldsymbol{\Delta}_g^{-2} = \sigma_g^*, \tag{23}$$

i.e. $\boldsymbol{\Delta}_g$ is the inverted standard deviation for y^* in group g. For two y^* variables j and k, the Delta approach considers the correlation in group g,

$$Corr(y_{gj}^*, y_{gk}^*) = \boldsymbol{\Delta}_{gj}\ \sigma_{gjk}^*\ \boldsymbol{\Delta}_{gk}, \tag{24}$$

where σ_{gjk}^* is the covariance between the two latent response variables, $\sigma_{gjk}^* = \lambda_{gj}\psi_g\lambda_{gk}$. This implies that although correlations between the latent response variables are considered, the across-group variation in λ, ψ, and θ is captured through the corresponding across-group variation in $\boldsymbol{\Delta}_g$. This avoids the well-known problem of analyzing correlations when considering models that are not scale free.

In the Delta approach, θ is not a parameter in the optimization but is obtained as a remainder

$$\theta_g = \mathbf{\Delta}_g^{-2} - \lambda_g^2 \, \psi_g, \tag{25}$$

resulting in (8) for the reference group if $\mathbf{\Delta}_g = 1$. The Delta approach has been found to have some advantages over the Theta approach in model estimation. The Delta approach, however, has the disadvantage that across-group differences in the scaling factors $\mathbf{\Delta}_g$ has three potential sources that are not distinguished: differences in λ; differences in ψ; and differences in θ. This disadvantage is avoided in the Theta approach given below. The Delta parameterization builds on the notion that, drawing on continuous-outcome experiences, residual variances are seldom invariant, and therefore a separate test of this is often less central. What is central for across-group factor comparisons is that thresholds and loadings are invariant to a sufficient degree.

4.2 The Theta Approach

The Theta approach was introduced in Mplus in Version 2.1, May 2002. In the Theta approach, the residual variance θ is a parameter in the optimization and the scaling factor $\mathbf{\Delta}_g$ is obtained as a remainder,

$$\mathbf{\Delta}_g^{-2} = \lambda_g^2 \, \psi_g + \theta_g. \tag{26}$$

This implies that the θ parameters enter into the correlation (24) via the $\mathbf{\Delta}_g$ terms. The Theta approach standardizes to $\theta = 1$ for all variables in a reference group, while estimating the θ parameters in the other groups. To test across-group equality of θ in a comparison group, θ is fixed at 1 in the comparison group as well.

4.3 Multiple-Group Measurement Invariance, Identification, And Standardization Issues In Mplus

Consider for simplicity the case of binary outcomes. Consider first the case with full measurement invariance, i.e. that the threshold τ_j and loading λ_j for outcome j are the same in all groups. A reference group is chosen with y^* variances standardized to unity, $\mathbf{\Delta}_g = 1$ for all variables, letting $\mathbf{\Delta}_g$ be estimated for the other groups. The Theta parameterization is analogous, standardizing residual variances to unity in the reference group and estimating them in other groups. The standardization of $\nu = 0$ for all variables in all groups and the factor mean $\alpha = 0$ in the reference group gives

$$E(y^*_{ref}) = 0, \tag{27}$$

$$E(y^*_{nonref}) = \lambda \, \alpha_{nonref}. \tag{28}$$

The univariate probability expression (12) shows that the standardizations identify the thresholds for all variables in the reference group. With the usual factor analysis standardization of one loading fixed at unity, the reference group identifies the loadings as well as the factor variance. Consider a non-reference group. Here, the correlations (14) identify $\mathbf{\Delta}_g$. This is because the loadings and factor variance are already known from the reference group, so that the $\mathbf{\Delta}_g$ play the role of loadings which are identified. Because (12) in the non-reference group identifies $E(y^*_{nonref})$, this means that α_{nonref} is identified in the non-reference group. This shows that all parameters are identified.

It may be noted that the choice of scale standardization $\mathbf{\Delta}_g = 1$ (or choice of θ standardization) is arbitrary (cf the earlier section LRV Model Identification). In the single-factor model considered here, a change from the standardization $\mathbf{\Delta}_g = 1$ to $\mathbf{\Delta}_g = c$ for all variables is absorbed into the parameters as (assuming the factor metric is set by one λ fixed at unity) $\tau_c = c^{-1}\ \tau$, $\alpha_c = c^{-1}\ \alpha_{nonref}$, $\psi_c = c^{-2}\ \psi$. The ratios of $\mathbf{\Delta}_g$ across the groups remain the same. The chi-square test of model fit and the ratios of estimates to standard errors remain the same. If instead the factor metric is set by fixing $\psi = 1$, the change is absorbed as $\lambda_c = c^{-1}\ \lambda$, with no change in α or ψ. In the Theta parameterization, changing from the $\theta = 1$ standardization for all variables in a reference group to $\theta = c$, results in the changes (assuming metric set as $\lambda_1 = 1$) $\mathbf{\Delta}_{gc} = \sqrt{c}\ \mathbf{\Delta}_g$ so that $\tau_c = \sqrt{c}\ \tau$, $\alpha_c = \sqrt{c}\ \alpha$, $\psi_c = c\ \psi$. The ratios of θ in the groups remain the same and the chi-square test of model fit and the ratios of estimates to standard errors remain the same.

Consider next binary outcomes where some outcomes do not have measurement invariance across groups. For such outcomes, it is not meaningul to compare y^* distributions across groups since the y^*'s are in different metric with measurement noninvariance. For such outcomes, the scaling factors $\mathbf{\Delta}_g$ can be fixed at unity since no across-group comparison is made and this is also necessary to avoid indeterminacies. Analogously, in the Theta parameterization, the residual variances are fixed at unity.

Millsap and Tein (2002) develops a set of minimal across-group invariance restrictions on thresholds and other parameters that provides sufficient conditions for identification and compares multiple-group model testing using LISREL and Mplus. In the discussion above it is assumed that a considerably higher degree of invariance is present. To make meaningful comparisons of factor distributions across groups and across time points, a majority of the variables should have both threshold and loading invariance so that the factors not only are in the same metric technically, but so that it is also plausible that the variables measure factors with the same meaning in the different groups or at the different time points. As a baseline model, all thresholds and loadings may be held invariant, using model modification indices to relax these assumptions for variables where this does not fit the data well.

5 Multiple-Group Analysis of Antisocial Behavior

This section continues the analysis of the 8 property offense items from the ASB instrument in the NLSY. The shoplift item had been found noninvariant with respect to gender. The multiple-group output in the Mplus web note section considers several analyses, with and without item invariance across gender.

First, invariance is assumed for all items with respect to thresholds and loadings, allowing residual variances to be different across the groups (using the Theta parameterization), or letting the scale factors be different across groups (using the Delta parameterization). The factor means and variances are allowed to vary across gender. It should be noted that the gender difference in the factor mean is now insignificant. The difference relative to the earlier significance finding using factor analysis with covariates is presumably due to the group-varying residual variances (scale factors in the Delta parameterization), or due to allowing gender differences in the factor variance.

Next, threshold and loading noninvariance is allowed for the shoplift item.[3] It is seen that the model fit improves and that the shoplift thresholds and loadings are different across gender. For females the item has a higher loading and a lower threshold which implies that compared to males, the female conditional probability curve of (22) is steeper and shifted to the left. The probability curves can be computed[4] and plotted for different factor values η. The interpretation is that for given property offense factor value, females are more likely to admit to the shoplift item than males and the difference increases with increasing factor value. Note that the gender difference in the factor mean is now significant, whereas it was not in the invariance model. The misspecification of gender invariance for the shoplift item, for which females have a higher conditional probability, attenuated the higher male factor mean to the point of becoming insignificant.

The Theta parameterization can be used to estimate a model with gender invariance of the residual variances for the items. Here, the male (the second group) residuals are fixed at unity, the value for females (the first group). Note that chi-square difference testing cannot be done using the default WLSMV estimator (Muthén, du Toit, & Spisic, 1997), but that the WLS estimator will have to be used for such a purpose.

6 Growth Modeling

The longitudinal situation is analogous to the multiple-group situation, where the different time points correspond to the different groups. Muthén (1996) discusses details of growth modeling with binary outcomes, relates the Mplus approach to other approaches, and presents a Monte Carlo study using the weighted-least squares (WLS) estimator in Mplus. A linear probit growth model may be written as,

$$y^*_{ti} = \eta_{0i} + \eta_{1i}\, x_t + \epsilon_{ti}, \tag{29}$$

[3] Because the item is not invariant, the latent response variable is not in the same metric across the groups, so that its variance should not be compared across groups and cannot be identified. Due to this, the Delta parameterization fixes the scale factor for the item.

[4] In the Delta parameterization, the residual variance θ of (22) is obtained from the R^2 section of the output: 0.565 for females and 0.441 for males.

where for example with T time points, $x_t = 0, 1, \ldots, T-1$ so that η_{0i} is interpreted as an initial status (intercept) factor and η_{1i} is interpreted as a change (slope) factor. The model implies across-time differences in the individual values of y^*_{ti} due to the slope factor. In the CP formulation, this LRV formulation translates to individuals' probabilities changing over time. To make the growth model meaningful, the y^*_t values need to be in the same metric across time. This is achieved with threshold invariance across time. While the mean of the slope factor is a free parameter, the mean of the intercept factor can be either (1) fixed at zero with free, across-time invariant thresholds, or (2) free with one threshold fixed at all time points.

The mean and variance of y^*_{ti} change over time as,

$$E(y^*_t) = E(\eta_0) + E(\eta_1)\, x_t, \tag{30}$$

$$V(y^*_t) = V(\eta_0) + V(\eta_1)\, x_t^2 + 2\, x_t\, Cov(\eta_0, \eta_1) + V(\epsilon_t). \tag{31}$$

It is seen in (31) that the variance changes over time for three reasons: due to the slope variance, due to the intercept-slope covariance, and due to the time-specific variance for the residual. From experience with continuous outcomes, the variance of the time-specific residual ϵ_{ti} is likely to vary across time. By analogy with the multiple-group case, this implies that while the scaling factors $\mathbf{\Delta}$ (in the Delta parameterization), or the residual variances (in the Theta parameterization), can be fixed at unity for a reference time point such as the first time point, they should be let free for remaining time points in order to not distort the growth model structure.

Muthén (1996) points out that incorporating correlations among the ϵ residuals is straightforward when using the Mplus weighted least squares estimators, while harder with maximum-likelihood estimation. The y^* covariance between time points t and t' is

$$V(\eta_0) + V(\eta_1)\, x_t\, x_{t'} + (x_t + x_{t'})\, Cov(\eta_0, \eta_1) + Cov(\epsilon_t, \epsilon_{t'}). \tag{32}$$

Mplus also allows growth modeling where the dependent variable in (29) is a factor measured with multiple categorical indicators at each time point (see also the Mplus User's Guide, Technical Appendix 7, page 366). This type of growth modeling with a factor analytic measurement model was proposed in Muthén (1983). Assuming for simplicity a model with a single factor η_f, the measurement model for indicator j and the structural (growth) model at time point t are expressed as

$$y^*_{jti} = \nu_{jt} + \lambda_{jt}\, \eta_{fti} + \epsilon_{jti}, \tag{33}$$

$$\eta_{fti} = \eta_{0i} + \eta_{1i}\, x_t + \zeta_{ti}. \tag{34}$$

Typically, $\nu_{jt} = 0$ since location parameters are captured by the thresholds. As a baseline invariance assumption, $\lambda_{jt} = \lambda_j$ for all t with corresponding threshold invariance. Here, the standardization of $E(\eta_0) = 0$ is used unless one threshold is fixed for one of the indicators. With multiple indicators it is possible to identify the growth model even with only partial measurement invariance of thresholds and loadings, including across-time changes in scaling factors or residual variances. The multiple-indicator case also makes it possible to identify time-specific factor residual variances $V(\zeta_t)$ in addition to the indicator- and time-specific residual variances $V(\epsilon_{jt})$.

7 The LISREL Approach

7.1 Multiple-Group Analysis

The LISREL approach ultimately amounts to an analysis of mean vectors and covariance matrices for each group in line with a conventional continuous-outcome analysis. This gives an easy to understand approach, but it comes with the tradeoff of assuming across-group invariance of thresholds.

A technical description of the multiple-group analysis in LISREL is hard to find and the following is based on our understanding of notes by Jöreskog on the topic of analyzing ordinal variables, which have been posted on the SSI web site (see Jöreskog, 2002).[5] Using PRELIS and LISREL, a 3-stage approach can be taken based on the LRV formulation (an alternative two-stage approach is also given below). In stage 1, the thresholds are estimated using the $\mu^* = 0$, $\sigma^* = 1$ standardization in a single-group analysis of data from all groups. In stage 2, the thresholds are held fixed at their stage 1 values and the estimation concerns the y^* μ^*, σ^*, and σ^*_{jk} elements for all groups together with their asymptotic covariance matrices. Stage 3 then amounts to a regular multiple-group analysis for continuous outcomes based on these means, variances, and covariances. Because the stage 3 analysis is in a continuous-outcome framework, the $\mu^* = 0$, $\sigma^* = 1$ standardization is not used in stage 3. Intercepts ν and residual variances θ can be identified. The conventional standardization of $\alpha = 0$ in a reference group is used when imposing intercept (ν) invariance across groups.

The LISREL stage 2 identification issues are important for understanding the procedure. Stage 2 does not use the standardization $\mu^* = 0$, $\sigma^* = 1$ because these quantities are identified from univariate probabilities due to the thresholds being fixed, known. This can be seen as follows. More than two outcome categories are required (the binary case is discussed below). Consider a 3-category y variable ($y = 0, 1, 2$), where

$$P(y = 2) = \int_{\mathbf{\Delta}(\tau_2 - \mu^*)}^{\infty} \phi_1(y^*)\, dy^*, \tag{35}$$

$$P(y = 1) = \int_{\mathbf{\Delta}(\tau_1 - \mu^*)}^{\mathbf{\Delta}(\tau_2 - \mu^*)} \phi_1(y^*)\, dy^*, \tag{36}$$

$$P(y = 0) = \int_{-\infty}^{\mathbf{\Delta}(\tau_2 - \mu^*)} \phi_1(y^*)\, dy^*, \tag{37}$$

where $\mathbf{\Delta} = 1/\sqrt{\sigma^*}$. Let the probits of the two integration limits be denoted p_2 and p_1,

$$p_j = \mathbf{\Delta}(\tau_j - \mu^*). \tag{38}$$

The three probabilities contain two independent pieces of information, which may be viewed in terms of the two probits. The ratio p_2/p_1 eliminates $\mathbf{\Delta}$ which, given known

[5]We also acknowledge helpful communication with Roger Millsap.

thresholds, identifies μ^*. Given the thresholds and μ^*, any of the probits can be used to identify $\mathbf{\Delta}$ and therefore σ^*.

In principle, stage 1 and stage 2 can also be carried out in a single-step, multiple-group analysis. Here, the thresholds are estimated, holding them equal across groups. The model has two indeterminacies per y^* variable. The standardization $\mu^* = 0$, $\sigma^* = 1$ can be imposed in a reference group. The probit expression (38) for the reference group identifies the thresholds. In the other groups, the p_2/p_1 ratio eliminates $\mathbf{\Delta}$ and identifies μ^* for those groups. Given thresholds and μ^*, the σ^* variance is then identified in those other groups using any of the probits. Instead of the standardization $\mu^* = 0$, $\sigma^* = 1$ imposed in a reference group, the restrictions of average means over groups being zero and average variances being one may be used to eliminate the indeterminacies.

7.1.1 Latent Response Variable Intercepts

From the point of view of the CP formulation, the ν intercept parameters in (4) can be used to impose a less restrictive form of across-group invariance for the thresholds. This is clear from (7), where the $\tau_c - \nu$ term can have across-group invariance of the τ_c's while across-group differences in the ν's. Viewing $\kappa_g = \tau_c - \nu_g$ as the effective thresholds in the CP formulation, the κ_g thresholds are not group invariant but allow a rigid ν_g shift (maintaining the distance between them for a given variable) of all the thresholds for a given variable. This is another generalization of an ordered probit model and is used in the LISREL approach. From an LRV point of view, where the y^* variables are thought of as substantively meaningful, this is straightforward to interpret. The interpretation from a CP point of view is less compelling because it is hard to motivate why threshold non-invariance would take place in the form of a rigid shift. It seems more likely that threshold non-invariance occurs quite differently with respect to different thresholds for a variable.

7.1.2 Binary Case

The case where all outcomes are binary requires special treatment in the LISREL approach. In this case, there is only information on a single probit $\mathbf{\Delta}(\tau - \mu^*)$ per group. Even when holding thresholds invariant across groups, this does not identify both μ^* and σ^* in all groups. In the LISREL approach this is handled by fixing $\sigma^* = 1$ in all groups, only allowing μ^* to vary across groups.

7.2 Longitudinal Analysis

In line with the Mplus longitudinal section, no special consideration arise beyond those of multiple-group analysis.

8 Comparing The Mplus And LISREL Approaches

This section compares the parameterizations used in the Mplus and LISREL approaches and their consequences. While the focus is on multiple-group modeling, analogous conclusions hold for growth modeling. The last stage of the LISREL approach is analogous to analysis of continuous outcomes.[6] There is, however, a drawback to this simplicity. The threshold invariance assumption that the approach is based on can be questioned.

The LISREL estimation approach is appropriate only if threshold invariance holds for all thresholds and all variables, or if the rigid shift of thresholds across groups, discussed above, holds. This seems like a strong assumption.[7] In IRT contexts, non-invariance of the threshold parameters is typically of central interest. This is certainly true in Rasch modeling, where thresholds are the only measurement parameters. Therefore, it is limiting to base the approach on the assumption that all thresholds are invariant. In this connection it may be mentioned that the authors' experience with continuous outcomes is that invariance is more often found with respect to loadings than intercepts (intercepts being analogous to thresholds for categorical outcomes). Furthermore, with continuous outcomes, typically only partial measurement invariance is found, so that a subset of variables have non-invariant intercepts (and loadings). Partial invariance still allows comparisons across groups and time, and so is probably the most useful approach in practice.

In the binary case, the LISREL approach fixes the variances at unity for all variables and groups. This not only precludes a study of across-group differences in variances but also distorts a meaningful multiple-group analysis. For example, a model with full measurement invariance of thresholds and loadings, and also invariant residual variances, would be distorted due to across-group differences in factor variances because they cause latent respose variable variance differences. Binary growth modeling is not possible using the LISREL approach because the restriction of unit variances for all time points precludes representation of the variance structure of (31).

In the Mplus approach, the process of studying group invariance avoids the multiple stages of the LISREL approach and is accomplished in a single analysis. A key value of the Mplus approach lies in its ability to jointly study invariance of threshold, loading, and error variance parameters. This same flexibility is critical also in growth modeling, where the multiple time points play the role of multiple groups. Threshold invariance is

[6] The LISREL approach can be done in Mplus in two steps. Step 1 is the same as LISREL's stage 1. Step 2 is a combination of LISREL's stages 2 and 3, where the unrestricted mean vectors and covariance matrices are not estimated in an intermediate step, but the model parameters are directly estimated. Although Mplus' default setup for categorical outcomes does not include the ν intercept parameters in (4), a perfectly measured factor may be introduced behind each y^* variable, representing the ν parameters as α intercept parameters. It should be noted that the degrees of freedom will be inflated as a function of the fixed thresholds.

[7] Jöreskog (2002), pp. 28-29 shows how to use PRELIS to test the threshold equality for a pair of variables in a longitudinal setting, but this is not in the context of the latent variable model and no allowance for partial violations of equality is made in the final analysis.

not presupposed. Threshold non-invariance is possible with respect to any groups, any subset of variables, and with respect to any threshold of a variable.

9 Monte Carlo Simulation Examples

Five examples are given using the Mplus Monte Carlo option. Two studies concern multiple-group modeling and three studies concern growth modeling. The particular Monte Carlo facility for categorical outcomes in Mplus 2.12 requires a population y^* mean vector and covariance matrix for each group. Using these population values, observations on multivariate normal y^* variables are randomly drawn, followed by a categorization. The Monte Carlo studies below use 500 replications. The Mplus default WLSMV estimator is used throughout (Muthén, du Toit, & Spisic, 1997). Summarizing the 500 replications, the study focuses on quality of parameter estimation in terms of parameter estimate bias, the agreement between the standard deviation of estimates and average standard error, and the 95% coverage. Chi-square model testing is also considered, focusing on the agreement between the rejection proportion at the 5% level and the nominal value 0.05. Because all analysis models agree with how the data are generated, this reflects the Type I error. Power is not considered here. For general aspects of model testing with categorical outcomes, see Muthén (1993).

The Mplus Web Note section (see www.statmodel.com) gives the Mplus input, population y^* mean vector and covariance matrix data, and output for each of the five studies. Readers can easily modify the input to study other situations.

9.1 Multiple-Group Examples

The multiple-group examples consider a one-factor model for six 4-category variables in two groups. Two versions are considered: all thresholds invariant across groups, and a small set of thresholds invariant across groups. [8] In the first version, thresholds are $-.7$, 0, and .7 for variables one, two, four and five, and $-.8$, 0, and .8 for variables three and six. The loadings are all invariant with values .4, .5, .6, .4, .5, and .6. The residual variances are constant across variables but vary across groups with the value .30 in the first group and .49 in the second group. The factor mean is 0 in the first group and .25 in the second group. The factor variance is 1 in the first group and 1.2 in the second group. This implies the following y^* mean and variance values for each of the six variables and each group:

$$E(y^*_{group\ 1}) = (\ 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0\ ), \tag{39}$$

$$E(y^*_{group\ 2}) = (\ 0.10 \quad 0.125 \quad 0.15 \quad 0.10 \quad 0.125 \quad 0.15\ ), \tag{40}$$

[8] The parameter choices are similar to those chosen in Millsap and Tein (2002), except absorbing the latent response variable intercepts in the thresholds.

$$V(y^*_{group\ 1}) = (\ 0.46 \quad 0.55 \quad 0.66 \quad 0.46 \quad 0.55 \quad 0.66\ )\,, \tag{41}$$
$$V(y^*_{group\ 2}) = (\ 0.682 \quad 0.79 \quad 0.922 \quad 0.682 \quad 0.79 \quad 0.922\ )\,. \tag{42}$$

9.1.1 Study A: Full Measurement Invariance

Study A uses full threshold invariance. The Theta parameterization is used.[9] The results are good already at a sample size of 100 in each of the two groups.

The output in the Mplus Web Note section shows that the WLSMV chi-square test works well, with rejection proportion .054 at the 5% level for the 500 replications. The parameter estimate bias is low, the agreement between the standard deviation of estimates and average standard error is good, and the 95% coverage is good. The coverage is somewhat low for the group 2 residual variances, but an increased sample size avoids this. [10]

9.1.2 Study B: Partial Threshold Invariance

Study B uses invariance across groups for only a small set of the thresholds. This analysis is chosen to show the performance of the analysis in a more complex case. For the first variable the first two thresholds are held invariant, for the second variable the first threshold is held invariant, for the third variable the first and third thresholds are held invariant, and for the fourth, fifth and sixth variable the first threshold is held invariant. The true values of the thresholds are as given above, i.e. they are all invariant across the groups. The sample size is still 100 in each of the two groups and the Theta parameterization is used.

The output in the Mplus Web Note section shows that the WLSMV chi-square test works well and that the parameter estimation works well. The coverage for the group 2 residual variances is somewhat less good than in Study A. Again, increasing the sample size improves on this.

It should be noted that larger sample sizes than used here are often required for good estimation performance. This is because the examples shown here do not have strongly skewed distributions for the categorical outcomes, i.e. do not include categories with few individuals.

[9] It may be noted that the residual variances θ are not unity in this example. To match the population values, the analyses therefore do not use the typical group 1 $\theta = 1$ standardization but instead the group 1 population values. As discussed in the section Multiple-Group Measurement Invariance, Identification, And Standardization Issues In Mplus, such a change of standardization is inconsequential.

[10] From the standard errors it is seen that the sample sizes are not quite large enough to be able to reject group differences in factor means and variances.

9.2 Growth Examples

The growth modeling examples uses the model of (29) for four time points. The means of the intercept and slope growth factors are 0.5 and -0.5, respectively. The variances of the intercept and slope factors are 0.5 and 0.10, respectively, with covariance 0 (free in the estimation). Three versions of the growth model were considered: binary outcomes with thresholds 0 at all time points; 4-category outcomes with thresholds -0.7, 0, and 0.7; and multiple-indicator, 4-category outcomes. The y^* (or factor in the multiple-indicator case) means and variances, the residual variances, the scale factors ($\mathbf{\Delta}$), and the probability of $y = 1$ are as follows for the four time points:

$$E(y_t^*) = (\ 0.5 \quad 0 \quad -0.5 \quad -1.0\ )\,, \tag{43}$$

$$V(y_t^*) = (\ 1.0 \quad 1.2 \quad 1.8 \quad 2.8\ )\,, \tag{44}$$

$$V(\epsilon_t) = (\ 0.5 \quad 0.6 \quad 0.9 \quad 1.4\ )\,, \tag{45}$$

$$\mathbf{\Delta}_t = (\ 1.0 \quad 0.913 \quad 0.745 \quad 0.598\ )\,, \tag{46}$$

$$P(y_t = 1) = (\ 0.69 \quad 0.50 \quad 0.36 \quad 0.28\ )\,. \tag{47}$$

The 1/5 ratio of slope to intercept variance is often seen in real data. Residual variances are increasing over time in line with what is often seen in real data. The chosen parameter values give R^2 values, i.e. the proportion of y^* variance explained by the growth factors, that are constant across time at the value 0.5. Such a low value is realistic with categorical outcomes. The Mplus WLSMV estimator is used.[11]

9.2.1 Study C: Binary Outcomes

Study C uses the Delta approach for the binary outcome case at a sample size of 100. The output in the Mplus Web Note section shows that already at this low sample size, the results are good both in terms of chi-square and parameter estimation. [12] The coverage is good also for the scaling factors estimated in group 2. The alternative Theta approach[13] works poorly in this setting. Although reasonable coverage is obtained, the Theta approach gives poor average estimates and across-replication agreement between estimate variability and standard errors. At the sample size of 100 very poor results are obtained and poor results are also obtained at 500 (good results are obtained at 1000). This appears to be due to a few replications where if the Delta approach was used, the residual variance would show a very low value for the first time point. The Theta approach of fixing the residual variance then causes the residual variances for remaining time points to blow up by attempting to keep the variance ratios correct.

[11] Version 2.12 of Mplus does not allow for missing data with categorical outcomes; such features will be added in future versions.

[12] 498 out of the 500 replications converge.

[13] It may be noted that the θ values are not unity in this example. To match the population values, the analyses therefore do not use the typical standardization $\theta = 1$ at the first time point but instead the population value. As discussed in the section Multiple-Group Measurement Invariance, Identification, And Standardization Issues In Mplus, such a change of standardization is inconsequential.

The results for the Delta and Theta approaches for this example suggest an important approach in real-data analyses. Unless the sample size is very large, the Delta approach should be used first. This gives the residual variances as remainders as described earlier. If the residual variances are not very small, a Theta approach can be taken if this is of interest. In the Theta approach it is then recommended that the residual variance is fixed for the time point with the largest residual variance value.

It may be noted that if the across-time variation in the variance of y^* is ignored, distorted results are obtained. This can be illustrated by using the Delta approach with the scaling factors $\mathbf{\Delta}$ fixed at unity at all time points. Using a large sample of $10,000$, there is a clear bias toward zero for the growth factor means and a strong underestimation of the slope variance. A similar picture is obtained by using the Theta approach with residual variances fixed at the value of the first time point.

9.2.2 Study D: Polytomous Outcomes

Study D uses the Theta approach for the 4-category outcome case at a sample size of 250. Here, full across-time invariance of all thresholds is imposed.[14] The output in the Mplus Web Note section shows that good results are obtained. The large improvement over the binary case is presumably due to having more information available with 4-category outcomes, reducing the sampling variability in the estimates.

As with the multiple-group case, it should be noted that larger sample sizes than used here are often required for good estimation performance. This is because the examples shown here do not have strongly skewed distributions for the categorical outcomes, i.e. do not include categories with few individuals.

9.2.3 Study E: Multiple-Indicator, Polytomous Outcomes

Study E uses the Theta approach with the same 4-category outcomes as in Study D, but with three such indicators at each time point in line with (33) and (34). The example allows changes over time in some of the thresholds and loadings. While a majority of thresholds and loadings are invariant across time, the first of the three indicators is not invariant at the first time point and the third of the three indicators is not invariant at the last time point. This is intended to illustrate a situation where an item is not age appropriate; the first item is not age appropriate at the beginning of the study and the third item is not age appropriate at the end of the study. Instead of deleting such items at those time points, the model can take the item non-invariance into account. By keeping the item, the precision with which the factor is measured is not deteriorated. The particular threshold noninvariance chosen in this example is as follows. For the

[14]For convenience in matching the population values, the Mplus setup has the intercept factor mean free to be estimated, while fixing the middle threshold at 0. A real-data analysis may instead have the intercept factor mean fixed at zero and all thresholds estimated with equality across time.

first item at the first time point, the last threshold is 0.2 instead of 0.7, illustrating an over-reporting of the highest category. For the third item at the last time point, the first threshold is -0.2 instead of -0.7, illustrating an over-reporting of the lowest category. The remaining thresholds are invariant and at the values of Study D, -0.7, 0.0, and 0.7. The invariant loadings are 0.6, 1.0, and 0.8. The loading noninvariance appears for the same items as the threshold noninvariance so that the first item at the first time point has loading 0.3 instead of 0.6 and the third item at the last time point has loading 0.3 instead of 0.8. The WLSMV estimator with the Theta parameterization is used at a sample size of 250.

The output in the Mplus Web Note section shows that good results are obtained. The 5% chi-square rejection percentage is 6.8%. The coverage is good in all cases, although a bit low for residual variances. The performance at a sample size of 500 is excellent.

References

Agresti, A. (1990). Categorical data analysis. New York: John Wiley & Sons.

Gallo, J.J., Anthony, J. & Muthén, B. (1994). Age differences in the symptoms of depression: A latent trait analysis. Journals of Gerontology: Psychological Sciences, 49, 251-264.

Jöreskog, K.G. (2002). Structural equation modeling with ordinal variables using LISREL. SSI note, June 5, 2002 (http://ssicentral.com/lisrel/ - see "ordinal.pdf").

MacIntosh. R & Hashim, S. (2000). Converting MIMIC model parameters to IRT parameters: a comparison of variance estimation models. Forthcoming in Applied Psychological Measurement.

Millsap, R. & Tein, J.Y. (2002). Model specification and identification in multiple-group factor analysis of ordered-categorical measures. Under review.

Muthén, B. (1979). A structural probit model with latent variables. Journal of the American Statistical Association, 74, 807-811.

Muthén, B. (1983). Latent variable structural equation modeling with categorical data. Journal of Econometrics, 22, 48-65.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. Psychometrika, 49, 115-132.

Muthén, B. (1988). Some uses of structural equation modeling in validity studies: Extending IRT to external variables. In H. Wainer, & H. Braun (Eds.), Test Validity (pp. 213-238). Hillsdale, NJ: Erlbaum Associates.

Muthén, B. (1989). Latent variable modeling in heterogeneous populations. Psychometrika, 54, 557-585.

Muthén, B. (1993). Goodness of fit with categorical and other non-normal variables. In K. A. Bollen, & J. S. Long (Eds.), Testing Structural Equation Models (pp. 205-243). Newbury Park, CA: Sage.

Muthén, B. (1996). Growth modeling with binary responses. In A. V. Eye, & C. Clogg (Eds.), Categorical Variables in Developmental Research: Methods of Analysis (pp. 37-54). San Diego, CA: Academic Press.

Muthén, B. & Christofferson (1981). Simultaneous factor analysis of dichotomous variables in several groups. Psychometrika, 46, 407-419.

Muthén, B., du Toit, S.H.C., & Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. Accepted for publication in Psychometrika.

Muthén, B., Kao, Chih-Fen, & Burstein, L. (1991). Instructional sensitivity in mathematics achievement test items: Applications of a new IRT-based detection technique. Journal of Educational Measurement, 28, 1-22.

Muthén, L. & Muthén, B. (1998-2002). Mplus User's Guide. Los Angeles, CA: Muthén & Muthén.

Muthén, B., Tam, T., Muthn, L., Stolzenberg, R. M., & Hollis, M. (1993). Latent variable modeling in the LISCOMP framework: Measurement of attitudes toward career choice. In D. Krebs, & P. Schmidt (Eds.), New Directions in Attitude Measurement, Festschrift for Karl Schuessler (pp. 277-290). Berlin: Walter de Gruyter.