*Multivariate Behavioral Research*, *39* (3), 479-515


# Assessing Factorial Invariance in Ordered-Categorical Measures

Roger E. Millsap and Jenn Yun-Tein
Arizona State University

The factor analysis of ordered-categorical measures has been described in the literature on factor analysis, but the extension of the analysis to the multiple-population case is less well-known. For example, a comprehensive statement of identification conditions for the multiple-population case seems absent in the literature. We review this multiple-population extension here, with an emphasis on model specification and identification. The use of the method in the study of factorial invariance is described. New results on identification are given for a variety of factor structures and types of measures. Two widely-available software packages, LISREL 8.52 (Jöreskog & Sörbom, 1996) and Mplus 2.12 (Muthén & Muthén, 1998), are applied in simulated data to illustrate the method. The two programs are shown to have different model specifications for this method, leading to different fit results in some cases. The final section discusses some remaining problems facing researchers who wish to study factorial invariance in ordered-categorical data.

## *Introduction*

Common factor models for ordered-categorical measures are described in a number of sources (Bartholomew, 1980, 1984, 1987; Bock & Aitkin, 1981; Browne & Arminger, 1995; Christofferson, 1975; Jöreskog, 1990, 1993; Jöreskog & Moustaki, 2001; Mislevy, 1986; Muthén, 1978, 1984). An "ordered-categorical" measure is one whose values are both discrete and ordinal in scale. Examples would include dichotomously-scored test items, likert-scale questionnaire items, and partial-credit polytomous items. Most items used on tests and questionnaires in the social sciences could be classified as ordered-categorical, and so the factor model for such measures is potentially applicable if item-level analyses are to be conducted. In the factor analysis of such items, the use of the ordered-categorical factor model can avoid problems that result when the traditional continuous factor model is applied (Babakus, Ferguson, & Jöreskog, 1987; Bernstein & Teng, 1989;

We thank Bengt Muthén, Stephen du Toit, and four anonymous reviewers for helpful suggestions on earlier drafts. Correspondence concerning this article should be addressed to Roger E. Millsap, Department of Psychology, Box 871104, Arizona State University, Tempe, AZ 85283, or via email at millsap@asu.edu.

Olsson, 1979; Rigdon & Ferguson, 1991). If we wish to study the invariance of the factor structure for these items, we must consider the multiple population extension of the factor model. Unfortunately, in the ordered-categorical case, this extension has not received much attention in the factor analytic literature (Browne & Arminger, 1995; Lee, Poon, & Bentler, 1989; Muthén & Christofferson, 1981; Poon, Lee, Afifi, & Bentler, 1990). For example, a general statement of minimal conditions for identification in the multiple-population case seems lacking in the literature, although limited results are available (Browne & Arminger, 1995; Muthén & Asparouhov, 2002; Muthén & Christofferson, 1981).

The present article has two goals. The first goal is to describe the use of the factor model in studying factorial invariance for ordered-categorical measures, with an emphasis on model specification and identification. The description is intended to be general enough to include a variety of possible factor structures and data conditions. The second goal is to illustrate the foregoing ideas as they apply to two widely-available software packages that can handle the factor model for ordered-categorical measures in multiple populations: LISREL 8.52 (Jöreskog & Sörbom, 1996) and Mplus 2.12 (Muthén & Muthén, 1998). Here we will illustrate the different specifications used by these two programs, and the different results that accrue from these two specifications.

We begin the article with a description of the factor model for ordered-categorical measures in the multiple population case. The definition of factorial invariance in the context of the ordered-categorical model is discussed, and the unique nature of this case relative to the traditional continuous measure case is noted. New results on parameter identification in the ordered-categorical model are given in Appendix A, and are discussed in this section. The next section describes the model specifications used in the LISREL and Mplus software packages for factor models of ordered-categorical measure in multiple populations. We then illustrate the use of these two programs using several large simulated datasets whose factor structure is known. Three different model specifications are illustrated, two of which represent violations of factorial invariance. The performance of the two software packages in detecting the violations of invariance are displayed, emphasizing the fit evaluation in each case. The final section of the article discusses some further issues in the use of the factor model in ordered-categorical data.

## *The Multiple-Group Factor Model for Ordered-Categorical Measures*

Let $X_{ijk}$ be the score on the $j^{th}$ ordered-categorical measure for the $i^{th}$ person in the $k^{th}$ group or population. We will simplify the presentation by

assuming that all measured variables to be modeled have score ranges {0, 1, ..., $c$}, where $c$ is the largest possible score, common across measures. A more general description would permit $c$ to vary across variables, but this extension introduces needless complications for the purpose at hand. We will also confine discussion to the two population case ($k$ = 1, 2) because all of the relevant results can be illustrated with only two populations. The number of measured variables is $p$($j$ = 1, ..., $p$). When samples are drawn from the two populations, the number of persons in the $k^{\text{th}}$ group is $n_k$.

In the factor model for ordered-categorical data, the observed scores $X_{ijk}$ are assumed to be determined by unobserved scores on the latent response variates $X^*_{ijk}$. These latent response variates are continuous in scale, unlike the observed measures $X_{ijk}$. The observed measures can be viewed as discretized versions of the latent response variates, given that scores on the observed measures are determined through

$$X_{ijk} = m \quad \text{if} \quad \nu_{jkm} \leq X^*_{ijk} < \nu_{ij(m+1)}, \tag{1}$$

where $m$ = 0, 1, ..., $c$ and {$\nu_{jk0}$, $\nu_{jk1}$, ..., $\nu_{jk(c+1)}$} are latent threshold parameters for the $j^{\text{th}}$ variable as measured on persons from the $k^{\text{th}}$ group. Two of the thresholds are pre-defined: $\nu_{jk0} = -\infty$ and $\nu_{jk(c+1)} = +\infty$. The remaining $c$ threshold parameters may vary across variables and across groups. The probabilities associated with observed values for $X_{ijk}$ are determined by the probability distribution for $X^*_{ijk}$. Let $\mathbf{X}'_{ik} = \{X_{i1k}, X_{i2k}, ..., X_{ipk}\}$ be the $1 \times p$ vector of observed scores on the $p$ variables for the $i^{\text{th}}$ person in the $k^{\text{th}}$ group, with $\mathbf{X}'_{ik}$ the analogous vector of scores on the latent response variates. It is typically assumed that

$$\mathbf{X}_{ik} \sim MVN\left(\boldsymbol{\mu}^*_k, \boldsymbol{\Sigma}^*_k\right), \tag{2}$$

where $\boldsymbol{\mu}^*_k$ is a $p \times 1$ vector of means on the latent response variates, and $\boldsymbol{\Sigma}^*_k$ is a $p \times p$ covariance matrix for the latent response variates, each subscripted to permit differences in these parameters between groups. The probability of observing any value for $X_{ijk}$, or any joint set of values for $\mathbf{X}'_{ik}$ = {$X_{i1k}$, $X_{i2k}$, ..., $X_{ipk}$}, is found through integration of the multivariate density in Equation 2.

As described thus far, the observed values $\mathbf{X}'_{ik}$ = {$X_{i1k}$, $X_{i2k}$, ..., $X_{ipk}$} do not permit identification of both the threshold parameters {$\nu_{jk0}$, $\nu_{jk1}$, ..., $\nu_{jk(c+1)}$} and the response variate parameters ($\boldsymbol{\mu}^*_k, \boldsymbol{\Sigma}^*_k$) without further constraints. To illustrate, for any $p \times 1$ vector **a** and $p \times p$ diagonal matrix **B**, we can define new latent response variates

(3) $$\mathbf{Y}^*_{ik} = \mathbf{a} + \mathbf{B}\,\mathbf{X}^*_{ik},$$

with

(4) $$\mathbf{Y}^*_{ik} \sim MVN\left(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}^*_k, \mathbf{B}\boldsymbol{\Sigma}^*_k\mathbf{B}'\right),$$

and new thresholds $\omega_{jkm} = a_j + b_{jj}\nu_{jkm}$. The new latent response variates $\mathbf{Y}^*_{ik}$ in combination with the new thresholds $\{\omega_{jk0}, \omega_{jk1}, \omega_{jk(c+1)}\}$ yield the same probability structure for $\mathbf{X}'_{ik} = \{X_{i1k}, X_{i2k}, ..., X_{ipk}\}$ as do the original variates and thresholds. Hence we cannot use the observed values $\mathbf{X}_{ik}$ to estimate the thresholds or latent response parameters without placing restrictions on the thresholds, on the response variate distributions, or on both. For example, in the single-population case, typical restrictions are that $\boldsymbol{\mu}^*_k = \mathbf{0}$ and $diag(\boldsymbol{\Sigma}^*_k) = \mathbf{I}$, standardizing the latent response variates. Under this restriction, the observed response frequencies for $\mathbf{X}'_{ik} = \{X_{i1k}, X_{i2k}, ..., X_{ipk}\}$ lead to estimates of thresholds $\{\nu_{jk0}, \nu_{jk1}, ..., \nu_{jk(c+1)}\}$ as percentiles of the standard normal distribution. The off-diagonal elements of $\boldsymbol{\Sigma}^*_k$ are estimated as polychoric correlations. As will be discussed below, once multiple populations are considered, different identification constraints on the thresholds and on $(\boldsymbol{\mu}^*_k, \boldsymbol{\Sigma}^*_k)$ must be adopted.

Given the latent response variates $X^*_{ijk}$, the factor model is specified for these variates as

(5) $$X^*_{ijk} = \tau_{jk} + \boldsymbol{\lambda}'_{jk}\,\boldsymbol{\xi}_{ik} + u_{ijk},$$

where $\tau_{jk}$ is a latent intercept parameter, $\boldsymbol{\lambda}_{jk}$ is an $r \times 1$ vector of factor loadings for the $j^{th}$ variate on $r$ factors, $\boldsymbol{\xi}_{ik}$ is the $r \times 1$ vector of factor scores for the $i^{th}$ person in the $k^{th}$ group, and $u_{ijk}$ is the $j^{th}$ unique factor score for that person. Letting $\mathbf{u}'_{ik} = \{u_{i1k}, u_{i2k}, ..., u_{ipk}\}$ be the $1 \times p$ vector of unique factor scores, we assume that

(6) $$\boldsymbol{\xi}_{ik} \sim MVN(\boldsymbol{\kappa}_k, \boldsymbol{\Phi}_k), \quad \mathbf{u}_{ik} \sim MVN(\mathbf{0}, \boldsymbol{\Theta}_k),$$

with $\boldsymbol{\kappa}_k$ an $r \times 1$ vector of factor means, $\boldsymbol{\Phi}_k$ an $r \times r$ factor covariance matrix, and $\boldsymbol{\Theta}_k$ a $p \times p$ diagonal covariance matrix for the unique factors. We also assume that $\mathrm{Cov}(\boldsymbol{\xi}_{ik}, \mathbf{u}_{ik}) = \mathbf{0}$ for all $i$, $k$. These assumptions lead to the structure

(7) $$E(\mathbf{X}^*_{ik}) = \boldsymbol{\mu}^*_k = \boldsymbol{\tau}_k + \boldsymbol{\Lambda}_k\boldsymbol{\kappa}_k, \quad Cov(\mathbf{X}^*_{ik}) = \boldsymbol{\Sigma}^*_k = \boldsymbol{\Lambda}_k\boldsymbol{\Phi}_k\boldsymbol{\Lambda}'_k + \boldsymbol{\Theta}_k,$$

where $\boldsymbol{\tau}_k' = \{\tau_{1k}, \tau_{2k}, ..., \tau_{pk}\}$ and $\boldsymbol{\Lambda}_k$ is the $p \times r$ factor pattern matrix whose $j^{th}$ row is $\boldsymbol{\lambda}_{jk}'$. All factor model parameters are subscripted to permit group differences. Not all parameters are identified however, even in the case in which $(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$ are known for all $k$. In the typical case in which $(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$ are not themselves identified from the data as reviewed above, further identification constraints are needed. We return to this point.

*Factorial Invariance*

Unlike the traditional case involving continuous measured variables, the factor model in the ordered-categorical case only indirectly determines scores on the measured variables through the probabilities associated with the measured outcomes in Equation 1. The definition of factorial invariance in this case is built on the conditional probabilities for various observed outcomes given scores on the common factors. This approach is consistent with general definitions of measurement invariance that rely on conditional probability (Mellenbergh, 1989; Meredith, 1993; Meredith & Millsap, 1992). For example, if the foregoing factor model is an appropriate measurement model for the ordered-categorical measure $X_{ijk}$, we can state that measurement invariance for $X_{ijk}$ in relation to $\boldsymbol{\xi}_{ik}$ and in relation to the groups under study, holds if

$$P_k(X_{ijk} = x|\boldsymbol{\xi}_{ik}) = P(X_{ijk} = x|\boldsymbol{\xi}_{ik}), \tag{8}$$

for all $i$, $j$, $k$. Equation 8 states that the conditional probabilities of various outcomes for $X_{ijk}$, given the factor score $\boldsymbol{\xi}_{ik}$, do not depend on the group. The most obvious way in which measurement invariance in Equation 8 can hold is if the conditional distribution of $\mathbf{X}_{ik}^*$ given $\boldsymbol{\xi}_{ik}$ is itself invariant across groups. Under the multivariate normal model just described, invariance in the conditional distribution of $\mathbf{X}_{ik}^*$ given $\boldsymbol{\xi}_{ik}$ holds if

$$E(\mathbf{X}_{ik}^*|\boldsymbol{\xi}_{ik}) = \boldsymbol{\tau} + \boldsymbol{\Lambda}\boldsymbol{\xi}_{ik}, \tag{9}$$

$$\mathrm{Cov}(\mathbf{X}_{ik}^*|\boldsymbol{\xi}_{ik}) = \boldsymbol{\Theta}_k, \tag{10}$$

for all $k$. Equations 9 and 10 are not sufficient for measurement invariance in Equation 8 however, because measurement invariance refers to the observed measures $X_{ijk}$. Another condition is needed for measurement invariance:

$$\nu_{jkm} = \nu_{jm} \quad \text{for} \quad m = 0, 1, ..., c, \quad j = 1, 2, ..., p, \tag{11}$$

for all $k$. Taken together, Equations 9, 10, and 11 show that the relevant parameter set for studies of invariance are the thresholds $\{\nu_{jk0}, \nu_{jk1}, ..., \nu_{jk(c+1)}\}$ and the factor model parameters $(\boldsymbol{\tau}_k, \boldsymbol{\Lambda}_k, \boldsymbol{\Theta}_k)$. Measurement invariance does not require invariance in the factor means or factor covariance matrices $(\boldsymbol{\kappa}_k, \boldsymbol{\Phi}_k)$. Under measurement invariance, group differences in $(\boldsymbol{\kappa}_k, \boldsymbol{\Phi}_k)$ would still lead to group differences in the unconditional latent response variate parameters $(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$, and subsequent differences in the moment structure for the observed measures.

*Model Identification*

The ordinary confirmatory factor analysis model for continuous measures requires some constraints on the model parameters if the model is to be identified (Anderson & Rubin, 1956). The extension of the factor model to multiple populations and ordered-categorical measures raises additional identification problems. In this extension, we can approach the identification problem by dividing the constraints needed into two categories. The first category contains constraints needed to identify $(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$ within each group. The second category contains constraints that will identify the factor model parameters $\{\boldsymbol{\tau}_k, \boldsymbol{\Lambda}_k, \boldsymbol{\Theta}_k, \boldsymbol{\Phi}_k, \boldsymbol{\kappa}_k\}$. If $(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$ are identified and estimable, the second constraint category includes the constraints ordinarily needed in any confirmatory factor analysis using latent means. A general approach to identification would proceed in two stages. Stage one would impose constraints on the thresholds $\{\nu_{jk0}, \nu_{jk1}, ..., \nu_{jk(c+1)}\}$ and/or on $(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$ to yield identification for $(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$. Stage two would then impose constraints on the factor model $\{\boldsymbol{\tau}_k, \boldsymbol{\Lambda}_k, \boldsymbol{\Theta}_k, \boldsymbol{\Phi}_k, \boldsymbol{\kappa}_k\}$ to identify these parameters. This two-stage approach is sufficient to resolve the identification problem generally. One difficulty however is that depending on the structure of $\boldsymbol{\Lambda}_k$ and the number of response categories $c$, we may be able to employ fewer constraints in stage one by taking advantage of the model structure. In this case the division into two stages is less clear, and the identification problem is attacked simultaneously for the thresholds, for $(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$, and for the factor model parameters $\{\boldsymbol{\tau}_k, \boldsymbol{\Lambda}_k, \boldsymbol{\Theta}_k, \boldsymbol{\Phi}_k, \boldsymbol{\kappa}_k\}$. This point is illustrated next.

*Congeneric Structure*. Suppose that in the factor model in Equation 7, each row of $\boldsymbol{\Lambda}_k$ has only one nonzero element. This structure would include any single-factor model, and any multiple factor model in which each latent response variate loads on only one factor. The general factor analytic identification problem is greatly simplified in this case (Bollen, 1989). We will denote this factor structure as "congeneric", to distinguish it from the more complex case in which variables load on multiple factors. In this congeneric

case, the following constraints are sufficient to achieve model identification when $c > 1$ (i.e., each measured variable has at least three categories):

1. In one group, fix $\boldsymbol{\mu}_k^* = \mathbf{0}$ and $Diag(\boldsymbol{\Sigma}_k^*) = \mathbf{I}$. These constraints are sufficient to identify all threshold parameters in this group.

2. Fix $\boldsymbol{\kappa}_k = \mathbf{0}$ in the above group.

3. In all groups, fix $\boldsymbol{\tau}_k = \mathbf{0}$, and also pick one nonzero loading for each factor to fix to one. The variate chosen for a given factor will be denoted the "reference variate."

4. Require that for a chosen value of $m$, $\nu_{jkm} = \nu_{jm}$ for all $k$, with $j = 1, ..., p$. This constrains one threshold per latent response variate to be invariant. In addition, for each of the $r$ reference variates, require a second threshold to be invariant. The number of such additional invariance constraints is equal to the number of factors. Hence we require a total of $p + r$ thresholds to be invariant.

Appendix A shows that the above conditions are sufficient to identify the factor model. Constraints 1-3 have already appeared in the literature (Muthén & Christofferson, 1981). A notable feature of constraint 4 is that complete invariance of all threshold parameters is not required. All variates have one invariant threshold parameter. A subset of the variates have a second invariant threshold, with the size of the subset being determined by the number of factors. This choice of number corresponds to the number of factor means, and leads to the identification of those means as shown in Appendix A. No invariance constraints are imposed on the factor means themselves, on the factor covariance matrices, or on the unique factor variances. The only invariance constraints imposed on the nonzero factor loadings are those used for the reference variates. The identification constraints described here are not unique; alternative sets of conditions exist that are sufficient for identification. Different software programs may adopt different identification conditions, leading to differences between programs in parameter estimates and even model fit, as shown below.

One special case that requires further comment arises when the measured variables are all dichotomous ($c = 1$). It is not possible in this case to implement the threshold constraints as described in 4. Using "d" to denote the dichotomous case, we can replace constraint 4 with:

4d. Require $\nu_{jkm} = \nu_{jm}$ for all $k$, with $j = 1, ..., p$.

5d. Pick $r$ diagonal elements of $\boldsymbol{\Sigma}_k^*$ to fix to unit values for all $k$. The chosen elements should correspond to the reference variates chosen for constraints in $\boldsymbol{\Lambda}_k$.

The implication of the first constraint set in 4d is that for dichotomous measured variables, all thresholds will be invariant. As in the general ordered-categorical case, the thresholds become identified by the requirement that $\boldsymbol{\mu}_k^* = \mathbf{0}$ and $Diag(\boldsymbol{\Sigma}_k^*) = \mathbf{I}$ in the first group. Full invariance

of the thresholds alone is not a sufficient condition for identification in the dichotomous case. The constraint set in 5d above supplies additional constraints that resolve this problem. Appendix A describes why these further constraints lead to identification, and also presents some alternative identification constraints that are also sufficient. Taken together, the constraints 1-3, 4d, and 5d permit identification of any congeneric factor model for dichotomous variables.

*General Factor Structure*. We now consider factor models for ordered-categorical variables in which $r > 1$ and no restrictions are placed on the structure in $\mathbf{\Lambda}_k$ apart from the requirement that $\mathbf{\Lambda}_k$ have full rank for all $k$. Minimal identification conditions that cover all possible structures for $\mathbf{\Lambda}_k$, for all values $r > 1$, are unknown even in the case of continuous measured variables (Anderson & Rubin, 1956; Shapiro, 1985). The ordered-categorical case introduces further problems. We will not solve this general identification problem here. Instead, we will give constraints that are sufficient to render the identification problem for ordered-categorical variables to be equivalent to the continuous variable case.

Consider first the case in which each measured variable assumes values in at least three categories ($c > 1$). Using "nc" to denote the general (non-congeneric case), the following set of model constraints can be used:

1. In one group, fix $\mathbf{\mu}_k^* = \mathbf{0}$ and $Diag(\mathbf{\Sigma}_k^*) = \mathbf{I}$. These constraints identify all threshold parameters in this group.

2. Fix $\mathbf{\kappa}_k = \mathbf{0}$ in the above group.

3nc. Fix $\mathbf{\tau}_k = \mathbf{0}$ in all groups, and impose constraints on $\mathbf{\Lambda}_k$ that will render $\mathbf{\Lambda}_k$ to be rotationally unique within each group. These constraints can take several forms (Bollen & Jöreskog, 1985; Jöreskog, 1979; Millsap, 2001). A common choice is to pick $r$ rows of $\mathbf{\Lambda}_k$ to fix as rows of an $r \times r$ identity matrix.

4nc. Pick two values of $m$, and require that for each chosen value of $m$, $\nu_{jkm} = \nu_{jm}$ for all $k$, with $j = 1, ..., p$. These constraints force two thresholds per measured variable to be invariant.

As shown in Appendix A, constraints 1 and 4nc are sufficient to identify $(\mathbf{\mu}_k^*, \mathbf{\Sigma}_k^*)$ for all $k$, and to identify all remaining thresholds. The constraints in 3nc are standard constraints needed to achieve uniqueness for $\mathbf{\Lambda}_k$.

The previous threshold constraints will not work when the measured variables are dichotomous ($c = 1$). In this case, the constraints to be used will include 1, 2 and 3nc, on the previous page, but will replace 4nc with:

4d. Require $\nu_{jkm} = \nu_{jm}$ for all $k$, with $j = 1, ..., p$.

5dnc. Require $Diag(\mathbf{\Sigma}_k^*) = \mathbf{I}$ for all $k$.

The set (4d) above renders all thresholds to be invariant, as in the congeneric case. The set (5dnc) standardizes the latent response variates to unit variance across all groups. Note that these variates are not required to have invariant means $\boldsymbol{\mu}_k^*$ however.

From the standpoint of invariance investigations, identification constraints that require $Diag(\boldsymbol{\Sigma}_k^*) = \mathbf{I}$ have one disadvantage: it becomes difficult to evaluate the invariance of the unique factor covariance matrices $\boldsymbol{\Theta}_k$. For example, suppose that the loadings $\boldsymbol{\Lambda}_k$ are invariant and the constraint $Diag(\boldsymbol{\Sigma}_k^*) = \mathbf{I}$ is in place across all groups. Then group differences in the common factor covariance matrices $\boldsymbol{\Phi}_k$ will generally imply that the matrices $\boldsymbol{\Theta}_k$ also vary across groups. To avoid this problem, an alternative set of identification constraints based on $\boldsymbol{\Theta}_k = \mathbf{I}$ are described in Appendix A. These constraints are denoted the "theta parameterization" in Mplus (Muthén & Muthén, 1998).

## *Multiple-Group CFA of Ordered-Categorical Variables: LISREL and Mplus*

Effective use of the foregoing model to evaluate factorial invariance requires software that can (a) specify models for independent groups, and (b) specify models for ordered-categorical measures. Two widely-available software programs that fulfill these requirements are LISREL 8.52 (Jöreskog & Sörbom, 1996) and Mplus 2.12 (Muthén & Muthén, 1998). In this section, we describe the model specifications used by each of these programs in conducting confirmatory factor analysis (CFA) in multiple-groups with ordered-categorical data. We illustrate how the model specifications used by the two programs differ, leading to different fit results. The results given in the previous section are used to help understand both the program specifications and the different fit results. We also show how each program's model specification performs under violations of factorial invariance. Throughout this section, our emphasis lies in model specification and the evaluation of fit. Estimation issues will not be the focus of interest.

### *LISREL*

The specification of the LISREL 8.52 (Jöreskog & Sörbom, 1996) CFA model for ordered-categorical variables in multiple groups differs from the general description given earlier in several respects. First, the LISREL model imposes invariance on all thresholds regardless of the number of categories required by each measured variable. Thresholds are then assigned values as described below, and are treated as fixed parameters for

purposes of estimation. Second, latent intercepts $\boldsymbol{\tau}_k$ are not fixed to zero in all groups, and the factor means $\boldsymbol{\kappa}_k$ are not fixed to zero in one group. The user is free to constrain these parameters as the need arises. Third, in the case of dichotomous indicators, the condition $Diag(\boldsymbol{\Sigma}_k^*) = \mathbf{I}$ is imposed in all groups regardless of the factor structure or number of factors. In the polytomous case, this condition is not imposed, and both $\boldsymbol{\mu}_k^*$ and $\boldsymbol{\Sigma}_k^*$ are estimable given the fixed values of the thresholds.

The LISREL analysis proceeds in the following three stages, the first two of which use PRELIS 2.3 (Jöreskog & Sörbom, 1996):

1. Pooling data from all groups, PRELIS is used to estimate pooled threshold estimates in the combined group. For polytomous items, the first two thresholds are fixed to 0 and 1 respectively. For dichotomous items, thresholds are fixed to zero. All threshold estimates are saved to a file for later input in stage two.

2. Using the threshold estimates from stage one as fixed values, PRELIS then estimates $\boldsymbol{\mu}_k^*$ and $\boldsymbol{\Sigma}_k^*$ within each group separately, using the same threshold values in all groups. This analysis is a multiple-group analysis. Estimates of $(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$ are saved, along with appropriate estimates of their asymptotic covariance matrices. Note that in general, $\boldsymbol{\mu}_k^* \neq \mathbf{0}$ and $Diag(\boldsymbol{\Sigma}_k^*) \neq \mathbf{I}$ for any $k$. An exception occurs when $c = 1$. In this dichotomous case, $Diag(\boldsymbol{\Sigma}_k^*) = \mathbf{I}$ for all groups.

3. The estimates of $(\boldsymbol{\mu}_k^*, \boldsymbol{\Sigma}_k^*)$ are input to LISREL, along with the asymptotic covariance matrices. A multiple-group CFA model is fit to the data. The model is the same as in Equation 7, with $\boldsymbol{\tau}_k \neq \mathbf{0}$ generally. Several estimation methods are available. We focus on weighted-least squares (WLS) on the next page.

In stage three, the user may evaluate invariance in $(\boldsymbol{\tau}_k, \boldsymbol{\Lambda}_k, \boldsymbol{\Theta}_k)$ using a series of nested hypothesis tests, or conventional goodness-of-fit criteria. Invariance in the thresholds is assumed, rather than evaluated. For some situations (e.g., dichotomous indicators), full threshold invariance does not exceed what is required for identification. For other situations (e.g., polytomous indicators with $c > 2$), the full threshold invariance exceeds what is required. We will explore some consequences of full threshold invariance below.

### *Mplus*

The Mplus 2.12 (Muthén & Muthén, 1998) model also differs in several respects from the general description given earlier. First, $p \times p$ diagonal scaling matrices $\boldsymbol{\Delta}_k$ are introduced such that

$$\mathbf{R}_k^* = \boldsymbol{\Delta}_k \boldsymbol{\Sigma}_k^* \boldsymbol{\Delta}_k, \tag{12}$$

where $\mathbf{R}_k^*$ contains the polychoric correlations among the latent response variates in the $k^{\text{th}}$ group. The scaling matrices $\boldsymbol{\Delta}_k$ have diagonal elements equal to the reciprocal square roots of the diagonal elements of $\boldsymbol{\Sigma}_k^*$. The elements of $\boldsymbol{\Delta}_k$ may or may not be free parameters in the model, as explained below. Second, the unique factor covariance matrices $\boldsymbol{\Theta}_k$ may or may not be independent parameters in the model, again depending on the choice of the user. Third, although the thresholds are fully invariant by default in Mplus, the user may override this default and employ any pattern of threshold invariance constraints desired. Several estimation methods are available. We focus on WLS estimation here.

Mplus offers two parameterizations in relation to $\boldsymbol{\Delta}_k$ and $\boldsymbol{\Theta}_k$ that are relevant to studies of invariance. In the "theta parameterization", the $\boldsymbol{\Delta}_k$ are not independent parameters in the model, and are found as

$$\boldsymbol{\Delta}_k^{-2} = diag(\boldsymbol{\Delta}_k \boldsymbol{\Phi}_k \boldsymbol{\Lambda}_k' + \boldsymbol{\Theta}_k), \tag{13}$$

and so group differences in $\boldsymbol{\Delta}_k$ are created by group differences in $(\boldsymbol{\Delta}_k \boldsymbol{\Phi}_k \boldsymbol{\Theta}_k)$. In the reference group, the unique factor covariance matrix $\boldsymbol{\Theta}_k$ is fixed as $\boldsymbol{\Theta}_k = \mathbf{I}$, with $\boldsymbol{\Theta}_k$ varying across the remaining groups. Given that the $\boldsymbol{\Delta}_k$ are not of direct interest in tests of invariance, this theta parameterization is suitable for invariance research. The second "delta parameterization" fixes $\boldsymbol{\Delta}_k = \mathbf{I}$ in the reference group, letting the remaining $\boldsymbol{\Delta}_k$ be free. Now the unique factor covariance matrices $\boldsymbol{\Theta}_k$ are no longer free parameters:

$$\boldsymbol{\Theta}_k = \boldsymbol{\Delta}_k^{-2} - diag(\boldsymbol{\Delta}_k \boldsymbol{\Phi}_k \boldsymbol{\Lambda}_k'). \tag{14}$$

Group differences in the $\boldsymbol{\Theta}_k$ are due to group differences in $(\boldsymbol{\Delta}_k \boldsymbol{\Phi}_k)$. It is not possible to evaluate the unique factor covariance matrices $\boldsymbol{\Theta}_k$ for invariance under the delta parameterization. These matrices are calculated as residual matrices in Equation 14 under this parameterization, and are not available for invariance constraints. In what follows, we will focus primarily on the theta parameterization, with mention of the alternative delta parameterization as the need arises.

*Baseline Models*

The next section will illustrate the use of LISREL and Mplus in the investigation of factorial invariance in ordered-categorical data. All of the examples reported below use the same baseline model when conducting the

analysis using LISREL, and the same baseline model when conducting the analysis using Mplus. These baseline models provide a plausible starting point for the study of factorial invariance, given that the number of factors is the same in all groups. The LISREL and Mplus baseline models are different in structure due to the differing specifications used by these two programs. In the LISREL case, the baseline model has the following characteristics:

1. All thresholds $\{\nu_{jk0}, \nu_{jk1}, ..., \nu_{jk(c+1)}\}$ are fixed and invariant.
2. All common factor means $\boldsymbol{\kappa}_k$ and covariance matrices $\boldsymbol{\Phi}_k$ are free.
3. All unique factor covariance matrices $\boldsymbol{\Theta}_k$ are free.
4. Factor loadings $\boldsymbol{\Lambda}_k$ are free, except for $r^2$ fixed elements used to achieve rotational uniqueness in each group. Further constraints on the loadings would be introduced if $r > 1$ and the congeneric factor structure is specified.
5. Latent intercepts $\boldsymbol{\tau}_k$ are free, except for $r$ elements in each group that are fixed (usually to zero). The placement of the fixed elements in 5 is coordinated with 4: the $r$ elements selected in 5 correspond to the variables whose loadings are fixed in 4.

In the case of Mplus, the baseline model would have the following characteristics:

1. Thresholds $\{\nu_{jk0}, \nu_{jk1}, ..., \nu_{jk(c+1)}\}$ are free except for a subset that are constrained to invariance across groups. The size of this subset depends on the structure imposed on the loadings, and on whether $c > 1$.
2. All common factor covariance matrices $\boldsymbol{\Phi}_k$ are free. Common factor means $\boldsymbol{\kappa}_k$ in one group are fixed to zero, but means are free in the other groups.
3. The unique factor covariance matrix $\boldsymbol{\Theta}_k$ is fixed as $\boldsymbol{\Theta}_k = \mathbf{I}$ in one group, but is free in all other groups unless $c = 1$. When $c = 1$, additional elements of $\boldsymbol{\Theta}_k$ are fixed to unit values in other groups, depending on whether the factor structure is congeneric.
4. Factor loadings $\boldsymbol{\Lambda}_k$ are free, except for $r^2$ fixed elements used to achieve rotational uniqueness in each group. Further constraints on the loadings would be introduced if $r > 1$ and the congeneric factor structure is specified.
5. All latent intercepts $\boldsymbol{\tau}_k$ are fixed to zero. This constraint can be relaxed (see Muthén & Asparouhov, 2002), but this option will not be illustrated here.
6. Scaling factors $\boldsymbol{\Delta}_k$ are free in all groups, but are not independent parameters under the theta parameterization.

The exact configuration of constraints in Mplus depends on the number of thresholds needed per observed measure, and on the structure in the loadings. This dependence is explained in Appendix A.

To illustrate the calculation of degrees of freedom (*df*) for the two baseline models, we first count the number of independent parameters to be estimated. Table 1 gives the count for the case of two groups, $p$ measured

variables, $r = 1$, and $c > 1$. This case corresponds to the examples presented next. As revealed by Table 1, the number of parameters need not be equal for the LISREL and Mplus models. The next step in the *df* calculation requires a count of the number of sample covariance matrix elements, plus the number of sample means (LISREL) or the number of sample response proportions (Mplus). For Mplus, we have $2pc$ independent response proportions across variables and groups. We also have $p(p - 1)/2$ unique polychoric correlations per group, making a total of $2pc + p(p - 1)$ sample statistics. For LISREL, we have $2p$ sample means across variables and groups. We also have $p(p + 1)/2$ unique polychoric covariance matrix elements in each group, making a total of $3p + p^2$ sample statistics. The baseline model df calculation for Mplus is then

$$df = 2pc + p(p - 1) - 2p(c + 1) = p^2 - 3p.$$

The baseline model *df* calculation for LISREL is then

$$df = 3p + p^2 - 6p = p^2 - 3p.$$

In spite of the different models used for the two programs, their *df* values for the baseline model are identical. Note also that the *df* value does not depend on *c*, the number of thresholds, in this case.

Table 1
Parameter Count for the Baseline LISREL and Mplus Models[a]

| Parameters | Mplus | LISREL |
|---|---|---|
| Thresholds | $2pc - p - 1$ | 0 |
| Intercepts | 0 | $2(p - 1)$ |
| Loadings | $2(p - 1)$ | $2(p - 1)$ |
| Factor means | 1 | 2 |
| Factor Cov Matrix | 2 | 2 |
| Unique variances | $p$ | $2p$ |
| Scaling factors | 0 | 0 |
| Total Count | $2p(c + 1)$ | $6p$ |

[a] Assumes two groups, one common factor, $c > 1$, $p$ = number of measured variables, $c$ = number of threshold parameters per variable. The Mplus count uses the theta parameterization.

*Illustrations*

To illustrate the use of both LISREL and Mplus in testing factorial invariance in ordered-categorical data, simulated data were created so that the model underlying the data could be known. The generation of the data is described in Appendix B. In all cases, two independent groups were created with $n_k = 5000$ observations per group. Single-factor models for six measured variables were used throughout. All measured variables were simulated to have four possible values (0, 1, 2, 3), requiring $c = 3$ threshold parameters per variable.

Three models were used to create three separate sets of data for purposes of illustration. All parameter values are given in Appendix B. The three models, denoted as "true" models in the sense that they generated the data, are:

*True Model One*. Full factorial invariance holds in this model, with $\nu_{jkm} = \nu_{jkm}$ for $j = 1, ..., p$, $\boldsymbol{\Lambda}_k = \boldsymbol{\Lambda}$, $\boldsymbol{\tau}_k = \boldsymbol{\tau}$, , and $\boldsymbol{\Theta}_k = \boldsymbol{\Theta}$. Group differences in factor means $\kappa_k$ and factor variances $\varphi_k$ are present.

*True Model Two*. Thresholds are *not* invariant in this model. The invariance status of all other parameters is identical to that of True Model One.

*True Model Three*. Unique variances are *not* invariant. The invariance status of all other parameters is identical to that of True Model One.

The first model above represents the fully invariant case. It is anticipated that both programs will indicate full invariance under this model. The second model above permits group differences in thresholds. LISREL should have some difficulty with this model because group differences in thresholds are not permitted within the LISREL specification. The third model permits group differences in unique factor variances. Both programs should detect these group differences. The results obtained using LISREL and Mplus in data generated by each of the above three models are described below in turn.

*True Model One*

Table 2 gives the fit results provided by LISREL for a sequence of four models fit to the data generated under True Model One. The first model is the baseline model described earlier in which all thresholds are fixed at values that are invariant over groups. No other invariance constraints are imposed, except that the loading of the first variable is fixed to one, and its

Table 2
LISREL Fit Results for True Model One

| Model | Chi-square | *df* | *p*-value | RMSEA | CFI |
|---|---|---|---|---|---|
| Baseline | 8.82 | 18 | .96 | 0 | 1.00 |
| Invariant $\boldsymbol{\Lambda}$ | 13.34 | 23 | .94 | 0 | 1.00 |
| Invariant $\boldsymbol{\Lambda}$, $\boldsymbol{\tau}$ | 14.08 | 28 | .99 | 0 | 1.00 |
| Invariant $\boldsymbol{\Lambda}$, $\boldsymbol{\tau}$, $\boldsymbol{\Theta}$ | 26.88 | 34 | .80 | 0 | 1.00 |

*Note*. $\boldsymbol{\Lambda}$ = factor loadings, $\boldsymbol{\tau}$ = latent intercepts, $\boldsymbol{\Theta}$ = unique factor covariance matrix.

corresponding latent intercept is fixed to zero. Using earlier results, the df for this baseline model should be $p^2 - 3p = 6^2 - 3p(6) = 18$. The second model adds invariance of the loadings to the baseline model. No loadings are fixed in this model, but the intercept that was fixed in the first model is fixed here as well. The factor variance in the first group is fixed to one to set the scale. The third model adds invariance of the latent intercepts to the second model. No intercepts are fixed in this model. The factor mean is fixed to zero in the first group. The fourth model adds invariance of the unique factor variances to the third model. As expected given the invariant nature of True Model One, all fit statistics in Table 2 for each of the four specified LISREL models indicate retention of the null hypothesis of perfect fit.

Table 3 gives the fit results for the same data, as provided by Mplus. Four models were fit to the data using Mplus. The first model is the baseline model for Mplus as described earlier. This baseline model includes partial invariance on the thresholds, a structure for the loadings that is identical to that used for the baseline model in LISREL, fixed zero values for the latent

Table 3
Mplus Fit Results for True Model One

| Model | Chi-square | *df* | *p*-value | RMSEA | CFI |
|---|---|---|---|---|---|
| Baseline | 8.81 | 18 | .96 | 0 | 1.00 |
| Invariant $\boldsymbol{\Lambda}$ | 12.39 | 23 | .96 | 0 | 1.00 |
| Invariant $\boldsymbol{\Lambda}$, $\boldsymbol{\nu}$ | 19.20 | 34 | .98 | 0 | 1.00 |
| Invariant $\boldsymbol{\Lambda}$, $\boldsymbol{\nu}$, $\boldsymbol{\Theta}$ | 26.17 | 40 | .96 | 0 | 1.00 |

*Note*. $\boldsymbol{\Lambda}$ = factor loadings, $\boldsymbol{\nu}$ = thresholds, $\boldsymbol{\Theta}$ = unique factor covariance matrix.

intercepts in all groups, and a fixed zero value for the factor mean in group one only. The partial threshold constraints require the first finite threshold for each variate to be invariant, and the second threshold for the first variable to also be invariant. The second model adds invariance constraints on the loadings to the baseline model, as done in the second model under LISREL. No loadings are fixed in this second model. The factor variance in group one is fixed to one. The third model adds complete invariance in the thresholds to the second model. The fourth model adds invariance constraints on the unique factor variances.

The fit results in Table 3 demonstrate that the four models cannot be rejected, as expected. The chi-square values for the first two models are close to those given by LISREL. The third model in Mplus has *df* equal to that of the fourth model in LISREL, but different chi-square values are obtained. Models three in LISREL and Mplus both imply that for a given factor score, we expect the same latent response variate scores across groups. These models have different *df* values however because Mplus fixes all latent intercepts to zero. LISREL imposes invariance on the intercepts but estimates the common values of the six intercepts.

*True Model Two*

Tables 4 and 5 give the LISREL and Mplus results respectively in data generated by True Model Two. True Model Two permits some thresholds to have different values in the two groups (see Appendix B). The LISREL baseline model in Table 4 constrains all thresholds to invariance, but the fit results show that LISREL does not detect group differences in thresholds under the baseline model. Models two and three in LISREL are rejected statistically by the chi-square test of exact fit. The RMSEA values for these models are within the bound of .05 used to denote a good approximation

Table 4
LISREL Fit Results for True Model Two

| Model | Chi-square | *df* | *p*-value | RMSEA | CFI |
|---|---|---|---|---|---|
| Baseline | 18.66 | 18 | .41 | .003 | 1.00 |
| Invariant $\boldsymbol{\Lambda}$ | 130.25 | 23 | <.001 | .031 | .99 |
| Invariant $\boldsymbol{\Lambda}$, $\boldsymbol{\tau}$ | 270.50 | 28 | <.001 | .042 | .98 |
| Invariant $\boldsymbol{\Lambda}$, $\boldsymbol{\tau}$, $\boldsymbol{\Theta}$ | 860.17 | 34 | <.001 | .070 | .93 |

*Note*. $\boldsymbol{\Lambda}$ = factor loadings, $\boldsymbol{\tau}$ = latent intercepts, $\boldsymbol{\Theta}$ = unique factor covariance matrix.

Table 5
Mplus Fit Results for True Model Two

| Model | Chi-square | *df* | *p*-value | RMSEA | CFI |
|---|---|---|---|---|---|
| Baseline | 19.60 | 18 | .36 | .004 | 1.00 |
| Invariant $\Lambda$ | 26.79 | 23 | .27 | .006 | 1.00 |
| Invariant $\Lambda$, $\nu$ | 1244.64 | 34 | <.001 | .084 | .89 |
| Invariant $\Lambda$, $\nu$, $\Theta$ | 1387.53 | 40 | <.001 | .082 | .88 |

*Note*. $\Lambda$ = factor loadings, $\nu$ = thresholds, $\Theta$ = unique factor covariance matrix.

(Browne & Cudeck, 1993). Only model four yields both a significant chi-square value and an RMSEA value that would reject a close fit. Given the increment at model four, researchers might erroneously attribute the poor fit to group differences in unique factor variances, rather than thresholds.

The Mplus results in Table 5 reveal a different pattern. The baseline model fits well here, given partial invariance of the thresholds. When the loadings are constrained to invariance in model two, the model is again retained, implying no group differences in loadings. Model three adds full invariance of the thresholds, and is rejected in both tests of exact and close fit. Mplus can properly distinguish between group differences in loadings and group differences in thresholds. Model four is also rejected, given that this model retains the threshold constraints.

*True Model Three*

Tables 6 and 7 present the fit results for the LISREL and Mplus models respectively in data generated by True Model Three. True Model Three is identical to True Model One, except that group differences in unique factor variances are present (see Appendix B). The LISREL fit results in Table 6 show that the baseline model fits well by both tests of exact and close fit. Models two and three also fit well, as neither of these models place constraints on the unique factor variances. Model four is rejected by both tests of exact and close fit. This model requires invariance of the unique factor variances. LISREL properly locates the source of the group difference in model parameters.

The Mplus results in Table 7 reveal a similar pattern. The baseline model and model two are each retained by the exact fit test. Model three includes full invariance constraints on both thresholds and loadings, leading to an

Table 6
LISREL Fit Results for True Model Three

| Model | Chi-square | *df* | *p*-value | RMSEA | CFI |
|---|---|---|---|---|---|
| Baseline | 19.00 | 18 | .39 | .003 | 1.00 |
| Invariant $\boldsymbol{\Lambda}$ | 27.02 | 23 | .26 | .006 | 1.00 |
| Invariant $\boldsymbol{\Lambda}$, $\boldsymbol{\tau}$ | 31.61 | 28 | .29 | .005 | 1.00 |
| Invariant $\boldsymbol{\Lambda}$, $\boldsymbol{\tau}$, $\boldsymbol{\Theta}$ | 434.42 | 34 | <.001 | .049 | .95 |

*Note*. $\boldsymbol{\Lambda}$ = factor loadings, $\boldsymbol{\tau}$ = latent intercepts, $\boldsymbol{\Theta}$ = unique factor covariance matrix.

increase in the chi-square statistic that approaches statistical significance. The model is not rejected by the close-fit indices. Model four adds invariance constraints on the unique factor variances, and is rejected by the exact-fit test. Mplus is able to detect the violation of invariance that is located in the unique factor variances.

*Discussion*

The multiple-population extension of the factor model for ordered-categorical measures is a neglected topic in the literature on factor analysis. The first goal of this article was to fully describe this model, especially with respect to model specification and identification. It was shown that the definition of factorial invariance in this model must incorporate general notions of measurement invariance. New results on model identification were also given. The second goal of the article was to illustrate the use of the model within two major software programs, LISREL 8.52 and Mplus 2.12. The

Table 7
Mplus Fit Results for True Model Three

| Model | Chi-square | *df* | *p*-value | RMSEA | CFI |
|---|---|---|---|---|---|
| Baseline | 19.27 | 18 | .38 | .004 | 1.00 |
| Invariant $\boldsymbol{\Lambda}$ | 20.77 | 23 | .60 | 0 | 1.00 |
| Invariant $\boldsymbol{\Lambda}$, $\boldsymbol{\nu}$ | 45.92 | 34 | .08 | .008 | .999 |
| Invariant $\boldsymbol{\Lambda}$, $\boldsymbol{\nu}$, $\boldsymbol{\Theta}$ | 639.71 | 40 | <.001 | .055 | .929 |

*Note*. $\boldsymbol{\Lambda}$ = factor loadings, $\boldsymbol{\nu}$ = thresholds, $\boldsymbol{\Theta}$ = unique factor covariance matrix.

model specifications used by these two programs were described and contrasted. Large simulated datasets were used to show how the two programs handle different violations of invariance.

The definition of factorial invariance in the ordered-categorical case differs from the definition in the continuous case in several respects. First, the threshold parameters are a new potential source of violations of invariance in the ordered-categorical case. Second, the factor model itself is only indirectly connected to the measured variables in the ordered-categorical case. In the continuous case, one can argue that the requirement of invariance in the factor model has intrinsic value because the model has a direct connection to the observed means and covariance structure. In the ordered-categorical case however, no direct connection exists, and distributional assumptions play a major role. If the latent response variates are not multivariate normal, invariance in the thresholds and in the factor model parameters will not guarantee measurement invariance for the measured variables. Higher-order moment structure in the latent response variates must be considered in such cases, making the study of invariance more difficult.

Appendix A described sets of identification constraints that are sufficient to identify the factor model under congeneric factor structures. The identification constraints for the non-congeneric case will usually be sufficient for identification, but may fail to be so in some cases, as explained in Appendix A. The identification constraints presented in Appendix A are also minimal in the sense that identification will ordinarily be impossible with fewer constraints. Should investigators routinely begin invariance studies with a model that includes minimal constraints, or should a more constrained model be taken as the starting point? The advantage of the minimal constraint approach is that if the baseline model fails to fit, the number of alternative explanations for the misfit is reduced. On the other hand, prior experience with the measured variables may suggest which parameters are responsible for violations of invariance. In this case, a more constrained model that focuses on the parameters in question may be more useful. The choice of baseline model in practice will be determined in large part by the investigator's prior knowledge of the measures under study. When prior knowledge does not point to a specific source for the violation of invariance, the minimal constraints in Appendix A are useful.

Are the constraints given in Appendix A unique in any sense? There will nearly always exist alternative sets of constraints that will identify the model. For example, many choices of $p + r$ thresholds in the polytomous, congeneric case could be constrained to invariance. This situation is not unusual. In the traditional factor model, many sets of $r^2$ loadings can be used for model

identification, and the ordered-categorical case is no different. Given that alternative sets of constraints exist, do the alternative sets yield equivalent fits for the resulting models? We expect to find equivalent fits generally, but there is no guarantee. Experience with the traditional factor model has shown that alternative sets of identification constraints may yield different fits in applications, expecially when the specified model is incorrect (Millsap, 2001).

Model specification and fit evaluation within two software packages, LISREL 8.51 and Mplus 2.12, were illustrated using simulated data. As shown, the two programs use different model specifications at baseline. These different model specifications have consequences for each program's sensitivity to violations of invariance. LISREL's baseline model includes full invariance constraints on the thresholds. As a result, LISREL is unable to evaluate whether thresholds differ across groups, and so group differences in thresholds may be mistaken for group differences in other model parameters, as shown in the examples. This limitation in LISREL is potentially serious for invariance studies because ordinarily we have no reason to expect that all thresholds are invariant. LISREL's inclusion of nonzero intercepts, with the potential for varying intercepts across groups, could be used in an attempt to model group differences in thresholds that are confined to shifts by a constant amount across groups. In practice, this attempt is likely to be only partially successful whenever threshold differences across groups vary depending on which threshold is considered. Conversely, group differences in intercepts could be absorbed in Mplus by permitting group differences in thresholds, but invariance in some thresholds prevents full absorption. To model group differences in intercepts in Mplus, Muthen and Asparouhov (2002) indicate that nonzero intercepts can be introduced into the Mplus model with ordered-categorical indicators. On balance, Mplus appears at present to offer a more flexible system for invariance modeling in ordered-categorical data.

Estimation issues in the ordered-categorical factor model have not been addressed in this article. While WLS estimation was used in the examples, this method is known to require very large samples for adequate performance (Boomsma & Hoogland, 2001; Muthén & Kaplan, 1985). Large simulated datasets were used in the examples for this reason. In the modest samples typically available in practice, which estimation method should be used? Mplus offers several alternative methods for ordered-categorical data that appear to work well in modest samples (Muthén, du Toit, & Spisac, in press; Satorra, 1992; Satorra & Bentler, 1994). Diagonally-weighted least squares (DWLS) estimation is available in both Mplus and LISREL. This method is more practical than WLS when the number of variables is modest or large. A different approach to the problem is to rely on normal-theory maximum likelihood, with adjustments to the

standard errors and to the chi-square fit statistic for nonnormality (Satorra, 1993a, 1993b).

The examples illustrated the use of a particular sequence of model fit evaluations that began with tests of invariance for loadings, followed by either tests of invariance for thresholds (Mplus) or intercepts (LISREL). The last model evaluated included invariance constraints on unique factor variances. In practice, researchers may wish to modify this sequence by adding further model fit evaluations. For example, preliminary tests to assess whether the factor loadings fit the congeneric pattern in all groups may be conducted prior to the test of loading invariance. Preliminary tests of this sort are likely to be useful in multiple-factor models. When a set of invariance constraints is rejected, the researcher may attempt to locate the source of the violation through a series of invariance tests on individual parameters. If the invariance of the factor loading matrix is rejected for example, the next step is to determine which loadings vary across groups. One approach to this search is to begin with invariance constraints on all loadings, relaxing the constraints sequentially to improve fit until no further meaningful improvement is found. A second approach is to begin with no invariance constraints apart from those needed for identification, and add constraints sequentially until the fit begins to deteriorate significantly. It is not clear whether these two strategies will arrive at the same solution in practice. Furthermore, when the number of variables is large, sequential explorations of this type may produce models that fail to cross-validate. These problems are not unique to the ordered-categorical factor model, but the presence of many threshold parameters in this model in the polytomous case may worsen the problem. Here the researcher will benefit greatly from any psychological theory that can be applied to suggest the location of potential violations of invariance.

Finally, it is worth noting that item response theory (IRT) offers an alternative modeling approach for ordered-categorical measures in multiple populations. No continuous latent response variates are invoked in the IRT approach. Instead, an IRT model directly expresses the conditional probability of achieving a score on the observed measure, given the person's standing on a (generally) unidimensional latent trait. Relationships between the parameters of some IRT models and those of the single-factor model for ordered-categorical measures are known. Similarly, relationships between the parameters in the multiple-factor model and some multidimensional IRT models are known, although multidimensional IRT models are not yet widely used in practice. For discussions of the relationship between the IRT and common-factor models, see Lord and Novick (1968), Takane and de Leeuw (1987), Muthén and Lehman (1985), or McDonald (1999).

## References


Anderson, T. W. & Rubin, H. (1956). Statistical inference in factor analysis. In J. Neyman (Ed.), *Proceedings of the third Berkeley symposium on mathematical statistics and probability*, (pp. 111-150). Berkeley, CA: University of California Press.

Babakus, E., Ferguson, C. E., & Joreskog, K. G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research*, *37*, 72-141.

Bartholomew, D. J. (1980). Factor analysis for categorical data. *Journal of the Royal Statistical Society, Series B*, *42*, 293-321.

Bartholomew, D. J. (1984). Scaling binary data using a factor model. *Journal of the Royal Statistical Society, Series B*, *46*, 120-123.

Bartholomew, D. J. (1987). *Latent variable models and factor analysis.* London: Charles Griffin & Company.

Bernstein, I. H. & Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin*, *105*, 467-477.

Bock, R. D. & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, *46*, 443-459.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Bollen, K. A. & Joreskog, K. G. (1985). Uniqueness does not imply identification. *Sociological Methods and Research*, *14*, 155-163.

Boomsma, A. & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. Du Toit & D. Sorbom (Eds.), *Structural equation modeling: Present and future*. (pp. 139-168). Lincolnwood, IL: Scientific Software Inc.

Browne, M. W. & Arminger, G. (1995). Specification and estimation of mean and covariance structure models. In G. Arminger, C. C. Clogg, & M. E. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 185-249). New York: Plenum.

Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Christoffersson, A. (1975). Factor analysis of dichotomized variables. *Psychometrika*, *40*, 5-32.

Jöreskog, K. G. (1979). A general approach to confirmatory factor analysis with addendum. In K. G. Jöreskog, D. Sörbom, & J. Magidson (Eds.), *Advances in factor analysis and structural equation models* (pp. 21-43). New York: Abt Associates.

Jöreskog, K. G. (1990). New developments in LISREL: Analysis of ordinal variables using polychoric correlations and weighted least squares. *Quality and Quantity*, *24*, 387-404.

Jöreskog, K. G. (1993). Latent variable modeling with ordinal variables. In K. Haagen, D. J. Bartholomew, & M.Deistler (Eds.), *Statistical modeling and latent variables* (pp. 163-171). Amsterdam: North-Holland.

Jöreskog, K. G. & Moustaki, I. (2001). Factor analysis of ordinal variables: A comparison of three approaches. *Multivariate Behavioral Research*, *36*, 347-387.

Jöreskog, K. G. & Sörbom, D. (1996). *LISREL 8 User's Reference Guide*. Chicago: Scientific Software International.

Lee, S. Y., Poon, W. Y., & Bentler, P. M. (1989). Simultaneous analysis of multivariate polytomous variates in several groups. *Psychometrika*, *54*, 63-73.

Lord, F. M. & Novick, M. E. (1968). *Statistical theories of mental test scores*. Reading: Addison-Wesley.

Mathsoft (1999). *S-Plus 2000 user's guide*. Seattle: Mathsoft Inc.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mellenbergh, G. J. (1989). Item bias and item response theory. *International Journal of Educational Research*, *13*, 127-143.

Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika*, *58*, 525-543.

Meredith, W. & Millsap, R. E. (1992). On the misuse of manifest variables in the detection of measurement bias. *Psychometrika*, *57*, 289-311.

Millsap, R. E. (2001). When trivial constraints are not trivial: The choice of uniqueness constraints in confirmatory factor analysis. *Structural Equation Modeling*, *8*, 1-17.

Mislevy, R. J. (1986). Recent developments in the factor analysis of categorical variables. *Journal of Educational Statistics*, *11*, 3-31.

Muthén, B. O. (1978). Contributions to factor analysis of dichotomized variables. *Psychometrika*, *43*, 551-560.

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

Muthén, B. O. & Asparouhov, T. (2002). *Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus*. Mplus Web Note No. 4, at http://www.statmodel.com/mplus/examples/webnote.html.

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

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

Muthén, B. O. & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of nonnormal Likert variables. *British Journal of Mathematical and Statistical Psychology*, *38*, 171-189.

Muthén, B. O. & Lehman, J. (1985). Multiple group IRT modeling: Applications to item bias analysis. *Journal of Educational Statistics*, *10*, 133-142.

Muthén, L. K. & Muthén, B. O. (1998). *Mplus user's guide*. Los Angeles: Author.

Olsson, U. H. (1979). On the robustness of factor analysis against crude classification of the observations. *Multivariate Behavioral Research*, *14*, 485-500.

Poon, W. Y., Lee, S. Y., Afifi, A. A., & Bentler, P. M. (1990). Analysis of multivariate polytomous variates in several groups via the partition maximum likelihood approach. *Computational Statistics and Data Analysis*, *10*, 17-27.

Rigdon, E. E. & Ferguson, C. E. (1991). The performance of the polychoric correlation coefficient and selected fitting functions in confirmatory factor analysis with ordinal data. *Journal of Marketing Research*, *28*, 491-497.

Satorra, A. (1992). Asymptotic robust inferences in the analysis of mean and covariance structures. In P. V. Marsden (Ed.), *Sociological methodology 1992* (pp. 249-278). Oxford: Blackwell Publishers.

Satorra, A. (1993a). Asymptotic robust inferences in multiple-group analysis of augmented-moment structures. In C. M. Cuadras & C. R. Rao (Eds.), *Multivariate analysis: Future directions 2* (pp. 211-229). Amsterdam: Elsevier.

Satorra, A. (1993b). Multi-sample analysis of moment structures: Asymptotic validity of inferences based on second-order moments. In K. Haagen, D. J. Bartholomew, & M. Deistler (Eds.), *Statistical modeling and latent variables* (pp. 283-298). Amsterdam: North-Holland.

Satorra, A. & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variable analysis: Applications to developmental research* (pp. 399-419). Newbury Park: Sage.

Shapiro, A. (1985). Identifiability of factor analysis: Some results and open problems. *Linear Algebra and its Applications*, *70*, 1-7.

Takane, Y. & de Leeuw, J. (1987). On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika*, *52*, 393-408.

*Accepted October, 2003.*

## *Appendix A*

Identification conditions are here described for multiple-group confirmatory factor analytic models in ordered-categorical data. The conditions to be described are sufficient for identification, with one exception that is explained below. In describing the identification conditions, we consider four cases formed by the cross-classification of two binary classifications. The first classification concerns the number of scale points in the variables being modeled: dichotomous ($c = 1$) or polytomous ($c > 1$). The second classification concerns the structure of the factor pattern matrix: congeneric or non-congeneric. Congeneric patterns require each variable to have only one nonzero loading. Any single-factor model is congeneric, as are multiple factor models whose pattern matrices have one nonzero element per row. All other factor structures are non-congeneric.

### *The Polytomous Congeneric Case*

We assume here that each factor is defined by at least three variables or indicators. Cases in which a single variable defines the factor can still be handled, but the two-variable case would require some additional constraints beyond those considered here. Using notation defined earlier, sufficient conditions for identification of the factor model are:

1. Fix $\boldsymbol{\mu}_k^* = \mathbf{0}$ and $diag(\boldsymbol{\Sigma}_k^*) = \mathbf{I}$ in one group.
2. Fix $\boldsymbol{\kappa}_k = \mathbf{0}$ in the above group.
3. Fix $\boldsymbol{\tau}_k = \mathbf{0}$ in all groups, and pick one variate per factor to use as a reference variate, fixing that variate's loading on that factor to one. All other loadings for the reference variate are fixed to zero. These loading constraints are imposed in all groups.
4. For some chosen value of $m$, require that $\nu_{jkm} = \nu_{jm}$ for all $k$, with $j = 1, ..., p$. In addition, require a second threshold to be invariant for each reference variate. Together, these constraints involve $p + r$ thresholds in each group.

To show that the above four sets of constraints are sufficient to identify the factor structure, first note that condition 1 implies identification of all thresholds in the group referenced in condition 1. The response variates are standard normal variates in this group, and estimates for all thresholds are

formed as percentiles from the standard normal distribution based on the observed frequencies for $\mathbf{X}_{ik}$, the measured variables. A second consequence of condition 1 is that all thresholds that are constrained to invariance in condition 4 are now identified, using values from the group referenced in 1. The constraints in conditions 2 and 3 are sufficient to identify ($\mathbf{\Lambda}_k$, $\mathbf{\Phi}_k$, $\mathbf{\Theta}_k$, $\boldsymbol{\kappa}_k$, $\boldsymbol{\tau}_k$) in the group referenced in 1 (e.g., Bollen, 1989). All model parameters in this group are now identified.

Next, consider the group *not* referenced in condition 1. For any latent response variate in which two thresholds are identified and estimated, we can find $\mu^*_{jk}$ and the corresponding diagonal element of $\mathbf{\Sigma}^*_k$, denoted $\sigma^{2*}_{jk}$. To see this, let $\nu_{jkm}$ and $\nu_{jks}$ be the two identified thresholds, and define new "standardized" thresholds

$$z_{jkm} = \frac{\nu_{jkm} - \mu^*_{jk}}{\sigma^*_{jk}}, \quad z_{jks} = \frac{\nu_{jks} - \mu^*_{jk}}{\sigma^*_{jk}}. \tag{A1}$$

These standardized thresholds would apply if the $j^{\text{th}}$ variate is transformed to standard score metric. We can estimate $z_{jkm}$ and $z_{jks}$ directly using the observed frequencies on $X_{ijk}$, as percentiles of the standard normal distribution. Hence Equation A1 represents two linear equations in the two unknowns $\mu^*_{jk}$ and $\sigma^*_{jk}$. We can solve for $\mu^*_{jk}$ and $\sigma^*_{jk}$. The solutions for $\mu^*_{jk}$ for each of the $r$ reference variates immediately lead to a solution for $\boldsymbol{\kappa}_k$ because

$$\boldsymbol{\mu}^*_k = \mathbf{\Lambda}_k \boldsymbol{\kappa}_k, \tag{A2}$$

and the $r$ elements of $\boldsymbol{\mu}^*_k$ whose solutions were found correspond to fixed unit elements in $\mathbf{\Lambda}_k$ due to condition 3. Hence $\boldsymbol{\kappa}_k$ is equated to the vector of means for the latent reference variates.

Next, consider any pair of latent response variates that share a common factor, and for whom only one threshold per variate is known, given the invariance constraints in condition 4. Let these variates be variates $h$ and $g$. For these two variates, the identified (i.e., invariant) thresholds can be expressed in standardized form as

$$z_{gk1} = \frac{\nu_{g1} - \mu^*_{gk}}{\sigma^*_{gk}}, \quad z_{hk1} = \frac{\nu_{h1} - \mu^*_{hk}}{\sigma^*_{hk}}, \tag{A3}$$

where the first threshold has been designated as the invariant threshold for purposes of argument, and the group subscript has been dropped from the

raw thresholds to denote invariance. The standardized thresholds in Equation A3 are directly estimable, as argued earlier. Also, we know from the factor model that $\mu^*_{gk} = \lambda_{gk}\kappa_k$ and $\mu^*_{hk} = \lambda_{hk}\kappa_k$, where $\kappa_k$ is the factor mean for the factor containing variates *g* and *h*. All factor means are identified, and after substituting these expressions in Equation A3, and rearranging, we have

(A4) $$\sigma^*_{gk} = \frac{\nu_{g1}}{z_{gk1}} - \frac{\kappa_k}{z_{gk1}}\lambda_{gk} = d_g + f_g\lambda_{gk},$$

(A5) $$\sigma^*_{hk} = \frac{\nu_{h1}}{z_{hk1}} - \frac{\kappa_k}{z_{hk1}}\lambda_{hk} = d_h + f_h\lambda_{hk}.$$

Note that $d_g$, $f_g$, $d_h$, $f_h$ are identified in the above by previous arguments; the group subscripts on these quantities have been dropped for convenience in what follows. The factor model also implies that the polychoric correlations $r_{ghk}$ and $r_{jhk}$ can be written

(A6) $$r_{ghk} = \frac{\lambda_{gk}\lambda_{hk}\varphi_k}{\left(d_g + f_g\lambda_{gk}\right)\left(d_h + f_h\lambda_{hk}\right)},$$

(A7) $$r_{jhk} = \frac{\lambda_{hk}\varphi_k}{\sigma^*_{jk}\left(d_h + f_h\lambda_{hk}\right)},$$

where $\varphi_k$ is the variance of the factor common to variates *g*, *h*, and *j*, with variate *j* being the reference variate for this factor. The polychoric correlations in Equations A6 and A7 are directly estimable. Also, the variance $\sigma^{2*}_{jk}$ is identified as noted earlier. We can rearrange the above expressions

(A8) $$\frac{r_{ghk}}{\sigma^*_{jk} r_{jhk}} = \frac{\lambda_{gk}}{d_g + f_g\lambda_{gk}} = q_{ghk},$$

with $q_{ghk}$ being estimable. Equation A8 leads to an expression for the factor loading $\lambda_{gk}$ in terms of identified quantities as

(A9) $$\lambda_{gk} = \frac{q_{ghk} d_g}{1 - q_{ghk} f_g}.$$

This factor loading is identified. Proceeding analogously, we can extend the above series of operations to derive expressions for all nonzero loadings that do not correspond to reference variates. All factor loadings in $\boldsymbol{\Lambda}_k$ are therefore identified.

Once the loadings are all identified, we can give expressions for all remaining elements of $\boldsymbol{\mu}_k^*$ using the identity in Equation A2. This expression establishes identification for $\boldsymbol{\mu}_k^*$. Identification of the factor loadings also leads to identification for all diagonal elements of $\boldsymbol{\Sigma}_k^*$ using Equations A4 and A5. All off-diagonal elements in $\boldsymbol{\Sigma}_k^*$ are then identified using expressions involving the polychoric correlations. For an arbitrary pair of variates *s* and *t*, we have

(A10) $$\sigma_{stk}^* = \sigma_{sk}^* \sigma_{tk}^* r_{stk}.$$

We also know from the factor model that

(A11) $$\sigma_{stk}^* = \lambda_{sk} \lambda_{tk} \varphi_{stk},$$

where $\varphi_{stk}$ is the factor covariance between the factors corresponding to variates *s* and *t* if these variates load on separate factors. If variates *s* and *t* load on the same factor, $\varphi_{stk}$ is the corresponding diagonal element of $\boldsymbol{\Phi}_k$. Given that all loadings are identified, we can identify all elements of $\boldsymbol{\Phi}_k$ through Equations A10 and A11. Next, we can identify all unique factor variances through the relation $\boldsymbol{\Theta}_k = \boldsymbol{\Sigma}_k^* - \boldsymbol{\Lambda}_k \boldsymbol{\Phi}_k \boldsymbol{\Lambda}_k'$. Finally, all remaining thresholds are identified given knowledge of $\boldsymbol{\mu}_k^*$ and $\boldsymbol{\Sigma}_k^*$.

Mplus 2.12 introduced an alternative set of identification constraints to those given in constraint 1. In constraint 1, the "theta parameterization" would replace $diag(\boldsymbol{\Sigma}_k) = \mathbf{I}$ with $\boldsymbol{\Theta}_k = \mathbf{I}$, fixing the unique factor variances in the reference group to unit values. The unique variances in other groups are unconstrained. This alternative parameterization, together with constraints 2-4, is also sufficient for identification.

To understand this point, consider the group referenced in 1. An expression for the $m^{\text{th}}$ standardized threshold of a reference variable in this group using the theta parameterization is

(A12) $$z_{jkm} = \frac{\nu_{jkm}}{\sigma_{jk}^*} = \frac{\nu_{jkm}}{\sqrt{\varphi_{sk} + 1}},$$

where $\varphi_{sk}$ is the variance of the common factor on which the $j^{\text{th}}$ variate loads in the $k^{\text{th}}$ group. Note that the factor loading is fixed to one for the $j^{\text{th}}$ variate in constraint 3. For the $g^{\text{th}}$ latent response variate that shares a common factor with the $j^{\text{th}}$ variate, the standardized threshold is

(A13) $$z_{gkm} = \frac{\nu_{gkm}}{\sigma^{*}_{gk}} = \frac{\nu_{gkm}}{\sqrt{\lambda^2_{gk}\varphi_{sk} + 1}}.$$

Considering a third latent response variate, the $h^{\text{th}}$ variate that also shares a common factor with the $j^{\text{th}}$ and $g^{\text{th}}$ variates, we can write expressions for three polychoric correlations as

(A14) $$r_{gjk} = \frac{\varphi_{sk}\lambda_{gk}}{\sqrt{\varphi_{sk} + 1}\sqrt{\lambda^2_{gk}\varphi_{sk} + 1}},$$

(A15) $$r_{hjk} = \frac{\varphi_{sk}\lambda_{hk}}{\sqrt{\varphi_{sk} + 1}\sqrt{\lambda^2_{hk}\varphi_{sk} + 1}},$$

(A16) $$r_{ghk} = \frac{\lambda_{hk}\lambda_{gk}\varphi_{sk}}{\sqrt{\lambda^2_{gk}\varphi_{sk} + 1}\sqrt{\lambda^2_{hk}\varphi_{sk} + 1}}.$$

These polychoric correlations are directly estimable, and so the ratio

(A17) $$\frac{r_{ghk}}{r_{gjk}r_{hjk}} = \frac{\varphi_{sk} + 1}{\varphi_{sk}},$$

provides an estimate for $\varphi_{sk}$. All other diagonal elements of $\mathbf{\Phi}_k$ are found analogously. The correlation between any pair of reference variates, say the $j^{\text{th}}$ and $i^{\text{th}}$ variates, is

(A18) $$r_{ijk} = \frac{\varphi_{stk}}{\sqrt{\varphi_{sk} + 1}\sqrt{\varphi_{tk} + 1}}.$$

This correlation is directly estimable, and coupled with the estimated values for $\varphi_{sk}$ and $\varphi_{tk}$, we are led to an estimate for $\varphi_{stk}$. In this manner, the full

factor covariance matrix $\mathbf{\Phi}_k$ is identified. Using appropriate elements of $\mathbf{\Phi}_k$, all raw thresholds for the reference variates are identified using Equation A12. Furthermore, all factor loadings in $\mathbf{\Lambda}_k$ now become identified using the polychoric correlations in Equations A14 and A15. Having identified all loadings, the covariance matrix $\mathbf{\Sigma}^*_k$ is now identified, as are all remaining thresholds using Equation A13. We have now identified all factor models parameters and thresholds in the group referenced in constraint 1. The identification proof for the remaining groups then proceeds as in the foregoing case that used the $diag(\mathbf{\Sigma}^*_k) = \mathbf{I}$ constraints.

*The Dichotomous Congeneric Case*

The dichotomous case implies that only one threshold parameter is needed for each latent response variate. The constraints in 4 under the polytomous case cannot be fully implemented in this case. An alternative set of constraints that replace 4 from the polytomous case are

4d. Fix $\nu_{jkm} = \nu_{jm}$ for all $k$, with $j = 1, ..., p$.

5d. In all groups besides the group used in constraint 1, fix $r$ diagonal elements of $\mathbf{\Sigma}^*_k$ to unit values. The chosen $r$ elements should correspond to the reference variates in constraint 3.

The combination of constraints 1, 2, and 3 from the polytomous congeneric case, and 4d, 5d above, is sufficient to identify all model parameters. To understand this point, first note that 1 and 4d result in complete identification of all threshold parameters in all groups. Consider the standardized thresholds corresponding to the reference variates in 3

$$z_{jkm} = \nu_{jm} - \mu^*_{jk}, \tag{A19}$$

since $\sigma^*_{jk} = 1$ in these cases by constraint 5d. These standardized thresholds are directly estimable, and because the raw thresholds are themselves identified, we can identify $\mu^*_{jk}$. Then given constraint 3 and Equation A2, we can identify $\boldsymbol{\kappa}_k$.

From this point, the proof of identification for the remaining parameters proceeds in the same fashion as in the polytomous case. We again pick variates $g$ and $h$ that share a common factor but are not reference variates. Using expressions essentially identical to Equations A3 to A9 but involving tetrachoric correlations, we can demonstrate that all factor loadings are identified. Identification of the loadings leads to identification of the latent variate means in Equation A2. The remaining diagonal elements of $\mathbf{\Sigma}^*_k$ are then found using relations as in Equations A4 and A5. The off-diagonal elements of $\mathbf{\Sigma}^*_k$ are found using expressions as in Equation A10. Equation

A11 then yields identification for the elements in $\mathbf{\Phi}_k$, and we can then identify the unique factor variances in $\mathbf{\Theta}_k$.

The theta parameterization in Mplus 2.12 provides an alternative route to identification that parallels the polytomous congeneric case. As in the polytomous case, we replace the condition $diag(\mathbf{\Sigma}_k^*) = \mathbf{I}$ in constraint 1 with $\mathbf{\Theta}_k = \mathbf{I}$ for the group defined in 1. The dichotomous case requires a further change however: we replace 5d with the requirement that $r$ diagonal elements of $\mathbf{\Theta}_k$ be fixed to unit values in all groups. The chosen $r$ elements should correspond to the reference variates in 3.

Proof of identification in this dichotomous case proceeds as in the polytomous case under the theta parameterization. Equations A12 and A13 again display the standardized threshold parameters of a reference variable and a non-reference variable respectively. The ratio in Equation A17 is again formed using the (now) tetrachoric correlations in Equations A14-A16. The ratio leads to identification for $diag(\mathbf{\Phi}_k)$, with the off-diagonal elements identified by the tetrachoric correlations in Equation A18. Identification of $\mathbf{\Phi}_k$ leads in turn to identification of the raw thresholds in Equation A12 for the reference variates. Ultimately, both the factor loadings $\mathbf{\Lambda}_k$ and the elements of $\mathbf{\Sigma}_k^*$ are identified for the group referenced in constraint 1.

For the remaining groups, all threshold parameters are identified given 4d. A general expression for the standardized threshold of the $j^{th}$ latent response variate is

$$z_{jk} = \frac{\nu_j - \mu_{jk}^*}{\sigma_{jk}^*}, \tag{A20}$$

with $\mu_{jk}^* = \lambda_{jk}\kappa_k$ and $\sigma_{jk}^{2*} = \lambda_{jk}^2\varphi_k + \theta_{jk}$, and with $\kappa_k$ being the factor mean and $\varphi_k$ being the factor variance for the factor on which the $j^{th}$ variate loads. If the $j^{th}$ variate is a reference variate, Equation A20 reduces to

$$z_{jk} = \frac{\nu_j - \kappa_k}{\sqrt{\varphi_k + 1}}, \tag{A21}$$

and if the $j^{th}$ variate is not a reference variate, we have

$$z_{jk} = \frac{\nu_j - \lambda_{jk}\kappa_k}{\sqrt{\lambda_{jk}^2\varphi_k + \theta_{jk}}}. \tag{A22}$$

Consider a set of three measures that share a common factor. The tetrachoric correlation between the $g^{th}$ and $j^{th}$ measures is

$$r_{gjk} = \frac{\lambda_{gk}\lambda_{jk}\varphi_k}{\sigma^*_{gk}\sigma^*_{jk}}. \tag{A23}$$

Then it can be shown that if the $j^{th}$ variate is a reference variate,

$$\frac{r_{ghk}}{r_{gjk}r_{hjk}} = \frac{\varphi_k + 1}{\varphi_k}, \tag{A24}$$

leading to identification for $diag(\boldsymbol{\Phi}_k)$. Once these diagonal elements are identified, the factor mean vector $\boldsymbol{\kappa}_k$ is identified using the reference variable threshold expressions in Equation A21. The remaining elements of $\boldsymbol{\Phi}_k$ are identified using the expression for the tetrachoric correlation between any pair of reference variates

$$r_{ijk} = \frac{\varphi_{srk}}{\sqrt{\varphi_{sk}+1}\sqrt{\varphi_{rk}+1}}, \tag{A25}$$

where the $i^{th}$ and $j^{th}$ response variates are reference variates.

Next, we can identify the loadings $\boldsymbol{\Lambda}_k$ by considering the correlation between any reference variate and another variate that shares a common factor with the reference variate. Letting the $j^{th}$ variate be the reference variate, we have

$$r_{jgk} = \frac{\lambda_{gk}\varphi_k}{\sigma^*_{jk}\sigma^*_{gk}}, \tag{A26}$$

where $\varphi_k$ and $\sigma^*_{jk}$ are already identified. From the threshold expression for the $g^{th}$ variate, we have

$$\sigma^*_{gk} = \frac{\nu_{gk}}{z_{gk}} - \frac{\kappa_k}{z_{gk}}\lambda_{gk}. \tag{A27}$$

Substituting Equation A27 into the tetrachoric formula in Equation A26, only one unknown remains, leading to identification for $\lambda_{gk}$. This process can be repeated for all non-reference variates, leading to full identification of $\mathbf{\Lambda}_k$. Finally, the free elements of $\mathbf{\Theta}_k$ can be identified through the threshold expressions in Equation A22. Identification of $\mathbf{\Sigma}_k^*$ is also achieved.

*The Polytomous Non-congeneric Case*

Here we permit arbitrary structure in $\mathbf{\Lambda}_k$, with no requirement that each variate loads on only one factor. This general structure presents identification problems even when the measured variables are continuous in scale. While sufficient conditions for rotational uniqueness are known for this case, full identification conditions that are minimal in some sense are unknown (Anderson & Rubin, 1956; Shapiro, 1985). The identification problem has been solved for some special cases, but no solution is known for arbitrary $r$ and $p$. We will not solve this problem here. Instead, we offer conditions that are sufficient to identify all threshold parameters and all parameters ($\boldsymbol{\mu}_k^*, \mathbf{\Sigma}_k^*$). The conditions are sufficient for rotational uniqueness in the parameters ($\mathbf{\Lambda}_k$, $\mathbf{\Phi}_k$), and will usually be sufficient for identification of all model parameters, especially when $r$ is small relative to $p$ (Shapiro, 1985). The conditions are

1. Fix $\boldsymbol{\mu}_k^* = \mathbf{0}$ and $diag(\mathbf{\Sigma}_k^*) = \mathbf{I}$ in one group.
2. Fix $\boldsymbol{\kappa}_k = \mathbf{0}$ in the above group.

3nc. Fix $\boldsymbol{\tau}_k = \mathbf{0}$ in all groups, and impose constraints on $\mathbf{\Lambda}_k$ that will render $\mathbf{\Lambda}_k$ to be rotationally unique within each group. These constraints can take several forms (Bollen & Jöreskog, 1985; Jöreskog, 1979; Millsap, 2001). A common choice is to pick $r$ rows of $\mathbf{\Lambda}_k$ to fix as rows of an $r \times r$ identity matrix.

4nc. Pick two thresholds per response variate to constrain to invariance: $\nu_{jkm} = \nu_{jm}$ for all $k$, with $j = 1, ..., p$ and for two values of $m$.

Given conditions 1 and 4nc, we can identify two thresholds per response variate, for all variates. As a result, the parameters $\boldsymbol{\mu}_k^*$ and the diagonal elements of $\mathbf{\Sigma}_k^*$ are all identifiable using the two-threshold argument made earlier. The off-diagonal elements of $\mathbf{\Sigma}_k^*$ are then estimable by rescaling the polychoric correlations as in Equation A10. All remaining thresholds are then estimable and identified.

Given that $\mathbf{\Sigma}_k^*$ is estimated in each group, we can use the constraints in condition 3nc to yield estimates for ($\mathbf{\Lambda}_k$, $\mathbf{\Phi}_k$, $\mathbf{\Theta}_k$) in most cases. Cases in which this cannot be done are cases in which the conditions for rotational uniqueness do not fully identify the factor model. Such cases would require individual study, and further constraints. Once estimates for $\mathbf{\Lambda}_k$ are available, we can identify $\boldsymbol{\kappa}_k$ from Equation A2.

The theta parameterization in Mplus 2.12 replaces $diag(\boldsymbol{\Sigma}_k^*) = \mathbf{I}$ in condition 1 with $\boldsymbol{\Theta}_k = \mathbf{I}$. All other constraints remain the same. The proof of identification under the theta parameterization begins with the group referenced in 1. Within this group, the correlation matrix for the latent response variates is written as

$$\mathbf{R}_k^* = \boldsymbol{\Delta}_k(\boldsymbol{\Lambda}_k\boldsymbol{\Phi}_k\boldsymbol{\Lambda}_k' + \mathbf{I})\boldsymbol{\Delta}_k = \boldsymbol{\Delta}_k\boldsymbol{\Lambda}_k\boldsymbol{\Phi}_k\boldsymbol{\Lambda}_k'\boldsymbol{\Delta}_k + \boldsymbol{\Delta}_k^2, \tag{A28}$$

with $\boldsymbol{\Delta}_k = \mathbf{D}_{\boldsymbol{\Sigma}_k^*}^{-1/2}$ being the diagonal "scaling matrix" in Mplus. We can generate a factoring of $\mathbf{R}_k^*$ for $r$ factors as

$$\mathbf{R}_k^* = \mathbf{P}_k\boldsymbol{\Omega}_k\mathbf{P}_k' + \boldsymbol{\Psi}_k, \tag{A29}$$

with $\boldsymbol{\Psi}_k$ a $p \times p$ diagonal matrix, $\boldsymbol{\Omega}_k$ an $r \times r$ factor covariance matrix, and $\mathbf{P}_k$ a $p \times r$ loading matrix. We assume that this factoring is unique given sufficient constraints on ($\mathbf{P}_k$, $\boldsymbol{\Omega}_k$) to yield rotational uniqueness. This assumption is violated when rotational uniqueness does not lead to general identification of the factor model as discussed earlier. Relating the factoring in Equation A29 to the model in Equation A28, we can set $\boldsymbol{\Delta}_k^2 = \boldsymbol{\Psi}_k$. This identifies $diag(\boldsymbol{\Sigma}_k^*)$, and we can identify the full matrix $\boldsymbol{\Sigma}_k^*$ through a rescaling of $\mathbf{R}_k^*$. Given that $\boldsymbol{\Sigma}_k^*$ is identified, we can identify all thresholds using Equation A12. The matrix $\boldsymbol{\Phi}_k$ is then identified using the submatrix of $\boldsymbol{\Sigma}_k^*$ corresponding to the covariance structure for the $r$ reference variates. This step is possible if the uniqueness constraints in condition 3 fix an $r \times r$ submatrix of $\boldsymbol{\Lambda}_k$ to an indentity matrix. Finally, consider the partitioned structure of $\boldsymbol{\Sigma}_k^*$ under these uniqueness constraints:

$$\boldsymbol{\Sigma}_k^* = \begin{vmatrix} \boldsymbol{\Phi}_k + \mathbf{I} & \boldsymbol{\Phi}_k\boldsymbol{\Lambda}_{sk}' \\ \boldsymbol{\Lambda}_{sk}\boldsymbol{\Phi}_k & \boldsymbol{\Lambda}_{sk}\boldsymbol{\Phi}_k\boldsymbol{\Lambda}_{sk}' + \mathbf{I} \end{vmatrix}. \tag{A30}$$

Let $\boldsymbol{\Sigma}_{srk}^* = \boldsymbol{\Lambda}_{sk}\boldsymbol{\Phi}_k$ be the $(p - r) \times r$ submatrix of covariances between the $r$ reference variates and the $p - r$ other variates. We can then identify $\boldsymbol{\Lambda}_{sk}$using $\boldsymbol{\Sigma}_{srk}^*\boldsymbol{\Phi}_k^{-1}$. This identifies all model parameters in the group referenced in condition 1.

Turning to the other groups, we have two invariant and identified raw thresholds per latent response variate from the foregoing arguments in the group referenced in condition 1. These two thresholds permit identification of $\boldsymbol{\mu}_k^*$ and $diag(\boldsymbol{\Sigma}_k^*)$. Having the latter, we can identify $\boldsymbol{\Sigma}_k^*$ through the rescaling of $\mathbf{R}_k^*$. Also, we can identify $\boldsymbol{\kappa}_k$ from $\boldsymbol{\mu}_k^*$ using the $r$ reference variables, whose loadings are fixed to unit values. All remaining thresholds

are identified given knowledge of $\boldsymbol{\mu}_k^*$ and $\boldsymbol{\Sigma}_k^*$. Finally, assuming that the constraints for rotational uniqueness are sufficient for global identification of the factor model, we can factor $\boldsymbol{\Sigma}_k^*$ to obtain

$$\boldsymbol{\Sigma}_k^* = \boldsymbol{\Lambda}_k \boldsymbol{\Phi}_k \boldsymbol{\Lambda}_k' + \boldsymbol{\Theta}_k, \tag{A31}$$

with an $r \times r$ submatrix of $\boldsymbol{\Lambda}_k$ being fixed to an identity matrix. We have then identified $\boldsymbol{\Lambda}_k$, $\boldsymbol{\Phi}_k$, and $\boldsymbol{\Theta}_k$. This step completes the identification.

*The Dichotomous Non-congeneric Case*

Here we have only one threshold per variate, and the general structure for $L_k$ noted in the previous polytomous case. We cannot implement the constraints in 4nc because only one threshold per variate is available. We therefore replace 4nc with

4d. Fix $\nu_{jkm} = \nu_{jm}$ for all $k$, with $j = 1, ..., p$.

5dnc. In all groups besides the group used in 1, fix all diagonal elements of $\boldsymbol{\Sigma}_k^*$ to unit values.

Combining conditions 1, 2, and 3nc defined earlier with 4d and 5dnc given here, all model parameters can be shown to be identified apart from the problem of rotational uniqueness versus full identification explained earlier for the polytomous non-congeneric case. Under condition 5dnc, all elements of $\boldsymbol{\Sigma}_k^*$ are directly estimable as tetrachoric correlations. Constraints 1 and 4d render all threshold parameters to be identified. The means $\boldsymbol{\mu}_k^*$ are then identified using expressions for the standardized thresholds, which are themselves directly estimable as noted earlier. Given estimates for ($\boldsymbol{\mu}_k^*$, $\boldsymbol{\Sigma}_k^*$), we can ordinarily obtain estimates for ($\boldsymbol{\Lambda}_k$, $\boldsymbol{\Phi}_k$, $\boldsymbol{\Theta}_k$) using the constraints in 3nc. Estimates cannot be obtained if the conditions for rotational uniqueness do not lead to full identification of the model. Such cases would require individual study and further constraints. Once estimates of $\boldsymbol{\Lambda}_k$ are available, we can identify $\boldsymbol{\kappa}_k$ using Equation A2.

Under the theta parameterization in Mplus, we replace $diag(\boldsymbol{\Sigma}_k^*) = \mathbf{I}$ with $\boldsymbol{\Theta}_k = \mathbf{I}$. In addition, we replace 5dnc with the requirement that in all groups, $\boldsymbol{\Theta}_k = \mathbf{I}$. In other words, the unique factor variances are fixed to unit values in all groups, and for all variates. Taken together with the other constraints 2, 3, 4d, these constraints will identify the model parameters apart from the limitations on identification for the factor model generally. The proof parallels the polytomous case. We begin in the group referenced in 1 with the factorization in Equation A29, leading to identification of $diag(\boldsymbol{\Sigma}_k^*)$. The rescaling of $\mathbf{R}_k^*$ in turn leads to the full covariance matrix $\boldsymbol{\Sigma}_k^*$. We also identify all thresholds using $\boldsymbol{\Sigma}_k^*$ in Equation A12. We can identify $\boldsymbol{\Phi}_k$ using

the submatrix of $\boldsymbol{\Sigma}_k^*$ corresponding to the $r$ reference variables, as shown in Equation A30. The loadings $\boldsymbol{\Lambda}_k$ are then identified via the same reasoning described earlier for the polytomous case. All model parameters are now identified in the group referenced in 1.

For the other groups, all thresholds are identified using 4d. We can identify $diag(\boldsymbol{\Sigma}_k^*)$ using the factoring argument that was used above. Knowledge of $diag(\boldsymbol{\Sigma}_k^*)$ leads to a rescaling of $\mathbf{R}_k^*$ that identifies the whole matrix $\boldsymbol{\Sigma}_k^*$. The factor mean vector $\boldsymbol{\kappa}_k$ is identified given $\boldsymbol{\Sigma}_k^*$ and the threshold equations for the reference variates. The covariance matrix $\boldsymbol{\Phi}_k$ is identified using the reference variate submatrix in $\boldsymbol{\Sigma}_k^*$, as shown in Equation A30. Finally, the loadings $\boldsymbol{\Lambda}_k$ are identified using the same arguments used in the polytomous case.

## *Appendix B*

All simulated data were created using Splus 2000 (Mathsoft, 1999). Data were simulated under a series of three single-factor models representing varying levels of factorial invariance. In all cases, data were generated for two independent groups, with $n_1 = n_2 = 5000$. This large sample size was chosen to minimize sampling variability and to achieve adequate performance using weighted least squares (WLS) estimation. Data were simulated to create $p = 6$ ordered-categorical measured variables in all models, using the latent response variate and threshold formulation described earlier. Within a given model, the steps in the simulation were as follows.

1. For group 1, generate a random common factor score $\xi_{i1}$ from a normal $(\kappa_1, \varphi_1)$ density.

2. For group 1, generate a $p \times 1$ random vector $\mathbf{u}_{i1}$ of unique factor scores from a MVN($\mathbf{0}$, $\mathbf{I}$) density. These unique factor scores are generated independently of step 1.

3. Calculate $\mathbf{X}_{i1}^* = \boldsymbol{\tau}_1 + \boldsymbol{\Lambda}_1\xi_{i1} + \boldsymbol{\Theta}_1^{1/2}\mathbf{u}_{i1}$.

4. Create for $j = 1, ..., 6$, $X_{ij1} = m$ if $\nu_{j1m} \leq X_{ij1} < \nu_{j1(m+1)}$, where $m = 0, 1, 2, 3$ with $\nu_{j10} = -\infty$ and $\nu_{j14} = +\infty$.

5. Repeat steps 1-4 $n = 5000$ times, creating a $5000 \times 6$ data matrix $\mathbf{X}_1$ for group 1.

6. Repeat steps 1-4 $n = 5000$ times using parameter values for group 2, creating another $5000 \times 6$ data matrix $\mathbf{X}_2$ for group 2.

The data matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ were used as input to PRELIS and to Mplus. Steps 1-6 were executed three times, once under each of three models. In all models, $\kappa_1 = 0$, $\varphi_1 = 1$ in the first group, and $\kappa_2 = .25$, $\varphi_2 = 1.2$

in the second group. Factor loadings and intercepts are invariant in all models, with

$$\text{(B1)} \qquad \boldsymbol{\Lambda} = \begin{vmatrix} .4 \\ .5 \\ .6 \\ .4 \\ .5 \\ .5 \end{vmatrix} \qquad \tau = \begin{vmatrix} .25 \\ .25 \\ .50 \\ .50 \\ .70 \\ .70 \end{vmatrix}.$$

All other model parameters varied, depending on the model. The parameter values used in each of the three models are given below. Note that in the baseline models that were fit to the simulated data, one loading was fixed to one for identification. This constraint was dropped once the loadings were constrained to invariance because the factor variance in group one was fixed to one in any model with invariant loadings. Also, in all models fit to the simulated data, invariance constraints used for identification purposes were applied only to parameters that were truly invariant.

*True Model One*

This model represents the fully invariant factor model, with invariant thresholds. Let $\boldsymbol{\nu}_k$ be the 6 × 3 matrix of threshold parameters in the $k^{\text{th}}$ group. In this model,

$$\text{(B2)} \qquad \boldsymbol{\nu}_1 = \boldsymbol{\nu}_2 = \begin{vmatrix} -.45 & .25 & .95 \\ -.45 & .25 & .95 \\ -.30 & .50 & 1.30 \\ -.20 & .50 & 1.20 \\ 0 & .70 & 1.40 \\ -.10 & .70 & 1.50 \end{vmatrix}.$$

All unique factor variances were set to .30 in both groups.

*True Model Two*

This model was identical to True Model One, except that thresholds were not invariant. The thresholds used in group one were as in Equation B2, but the thresholds in group two were

(B3)
$$\boldsymbol{\nu}_2 = \begin{vmatrix} -.45 & .25 & .75 \\ -.45 & .55 & 1.10 \\ -.30 & .70 & 1.30 \\ -.20 & .50 & 1.10 \\ 0 & .40 & 1.30 \\ -.10 & .50 & 1.60 \end{vmatrix}.$$

*True Model Three*

This model was identical to True Model One, except that the unique factor variances were not invariant. The variances in group one were identical to those used in True Model One. The variances in group two were all equal to .49.