



Published in final edited form as:

*Psychometrika*. 2016 December ; 81(4): 1014–1045. doi:10.1007/s11336-016-9506-0.

## Identification of Confirmatory Factor Analysis Models of Different Levels of Invariance for Ordered Categorical Outcomes

Hao Wu and  
Boston College

Ryne Estabrook  
Northwestern University

### Abstract

This article considers the identification conditions of confirmatory factor analysis (CFA) models for ordered categorical outcomes with invariance of different types of parameters across groups. The current practice of invariance testing is to first identify a model with only configural invariance and then test the invariance of parameters based on this identified baseline model. This approach is not optimal because different identification conditions on this baseline model identify the scales of latent continuous responses in different ways. Once an invariance condition is imposed on a parameter, these identification conditions may become restrictions and define statistically non-equivalent models, leading to different conclusions. By analyzing the transformation that leaves the model-implied probabilities of response patterns unchanged, we give identification conditions for models with invariance of different types of parameters without referring to a specific parametrization of the baseline model. Tests based on this approach have the advantage that they do not depend on the specific identification condition chosen for the baseline model.

### Keywords

ordered categorical data; invariance testing; model identification

## 1. Introduction

Ordered categorical measures such as Likert scales are prevalent tools in psychology. Data collected with such measures had been analyzed as normal data until proper models for ordered categorical outcomes were introduced (Jöreskog and Moustaki, 2001; Mislevy, 1986; Muthén, 1984) and proved to be superior (Babakus, Ferguson and Jöreskog, 1987; Bernstein and Teng, 1989; Lubke and Muthén, 2004; Muthén and Kaplan, 1985). One of the most widely used models for ordered categorical outcomes is the threshold model (see, e.g. Muthén, 1984), where the ordered categories are considered a discretized version of a normally distributed latent continuous response through a number of threshold parameters. These thresholds indicate values of the latent continuous response where individuals cross

over from one ordinal category to the next. Item response theory (IRT) gives an alternative framework of modeling ordered categorical data (e.g., Baker, 1992) and uses a different set of parameters. Invariance issues under IRT (differential item functioning) is not in the scope of this paper.

In this paper we are particularly concerned with the use of different identification conditions in the threshold model. An identification condition is a set of constraints imposed on the parameters of a model so that they can be uniquely estimated. More details will be discussed in Section 3. Similar to the common factors in a factor analysis model, the latent continuous responses for ordered categorical items are not directly observed and therefore have unknown scales. Constraints are needed to properly define their scales so that relevant parameters can be uniquely estimated. The most widely used option is to set the intercept of each latent response to zero and its variances to one (e.g., Christofferson, 1975), which uses the  $\delta$ -parametrization in Mplus (Muthén and Muthén, 1998–2012). Instead of setting the variances of the latent responses, one may set the variances of residuals to one in the  $\theta$ -parametrization (e.g., Muthén and Christofferson, 1981) of Mplus. For both of these options, all thresholds are freely estimated. When the observed ordinal response contains three or more categories, one can set the scale of a latent response by constraining two thresholds to constants, typically through setting the first two thresholds to values of zero and one (Mehta, Neale and Flay, 2004). With this option, an intercept and a unique variance are estimated for each observed variable in addition to the remaining thresholds not constrained. Any of the three methods above can identify the model so that parameters can be estimated. Different identification conditions yield different parameter estimates but identical standardized parameter estimates and fit statistics. Changing identification condition simply rescales the latent continuous responses (or common factors), and has the same effect as any linear transformation of a variable.

Identification in multiple group analyses are considerably more complicated, as ordinal variables must be identified for each group in the analysis. When there is no constraint imposed across groups, each group is identified separately in the analysis and has its own scales, and parameters are in group-specific rather than shared units. Researchers interested in the invariance of parameters typically impose an identification condition across groups, with the goal of placing each group on the same scale. Subsequent tests of invariance (e.g., weak factorial invariance) are dependent on the chosen identification condition. If the constraints in the identification condition are not properly selected, invariance constraints and other tests of parameter equivalence may be inherently biased.

In this paper, we argue that constraints involved in common identification conditions are in and of themselves invariance constraints, restricting either the scale (mean and variance) or the allowed transformations applied to each group in an analysis regardless of unobserved group differences in scale. Because of this problem, we advocate a different approach to identifying multigroup models, where models with different invariance properties are identified on a case by case basis. We will give the identification conditions for models with different levels of invariance on a congeneric structure for categorical outcomes, and show that some of the invariance properties can not be tested alone without assuming other

invariance properties. We will also discuss their implementation in common software packages as well as some technical issues.

This paper is organized as follows. The next section will review the basic concepts in modeling ordered categorical data and factorial invariance. This will be followed by a discussion of identification condition in Section 3 where the sources of non-identifiability are represented by transformations of parameters that leave the model implied probabilities unchanged. Using this group of transformations, Section 4 shows the problems of identifying an invariant model by simply incorporating additional constraints. We give identification conditions for models of different levels of invariance in Section 5 by exploiting the group of transformations in Section 3 and summarize the relationships among them in Sections 6 and 7 supported by a numerical illustration and simulation studies in Section 8. Software usage is briefly discussed next before closing.

## 2. CFA model for ordered categories and factorial invariance

### 2.1. The statistical model

Let  $y_j$  ( $j = 1, 2, \dots, p$ ) be the random variable of observed categorical response to item  $j$ , which takes  $K + 1$  values  $0, 1, \dots, K$ . The threshold model (see, e.g. Muthén, 1984) assumes an underlying latent continuous response  $y_j^*$  for  $y_j$ . The observed categorical response variable  $y_j$  arises from discretizing its corresponding latent response  $y_j^*$  through thresholds  $\tau_{jk}$  ( $j = 1, 2, \dots, p; k = 0, 1, \dots, K + 1$ ), i.e.

$$y_j = k \quad \text{if} \quad \tau_{jk} < y_j^* < \tau_{j(k+1)}$$

where we assume  $\tau_{j0} = -\infty$  and  $\tau_{j(K+1)} = \infty$ . The latent responses for all items

$\mathbf{y}^* = (y_1^*, y_2^*, \dots, y_p^*)'$  follow a regular factor analysis model for continuous variables:

$$\mathbf{y}^* = \boldsymbol{\nu} + \boldsymbol{\Lambda} \boldsymbol{\xi} + \boldsymbol{\delta}$$

where  $\boldsymbol{\xi} \sim N(\boldsymbol{\kappa}, \boldsymbol{\Phi})$  and  $\boldsymbol{\delta} \sim N(\mathbf{0}, \boldsymbol{\Theta})$  are the vectors of  $m$  common factors and  $p$  unique factors,  $\boldsymbol{\nu}$  is the  $p \times 1$  vector of item intercepts and  $\boldsymbol{\Lambda}$  is the  $p \times m$  loading matrix. Note that in a factor analysis model,  $\boldsymbol{\xi}$  and  $\boldsymbol{\delta}$  are uncorrelated and  $\boldsymbol{\Theta}$  is diagonal. This model implies the following expression for the covariance matrix  $\boldsymbol{\Sigma}$  and mean vector  $\boldsymbol{\mu}$  of the latent response vector  $\mathbf{y}^*$ :

$$\boldsymbol{\mu} = \boldsymbol{\nu} + \boldsymbol{\Lambda} \boldsymbol{\kappa} \quad \text{and} \quad \boldsymbol{\Sigma} = \boldsymbol{\Lambda} \boldsymbol{\Phi} \boldsymbol{\Lambda}' + \boldsymbol{\Theta}. \quad (1)$$

In a typical CFA model, the loading matrix  $\boldsymbol{\Lambda}$  is assumed to be sparse and the locations and signs of the possible nonzero elements are known. In this paper, we only consider the case where  $\boldsymbol{\Lambda}$  has a clustered structure, i.e., where each item loads only on one common factor. Because items with negative loadings are usually reverse-coded, we assume all nonzero

loadings are positive. There are  $pK$  (finite) thresholds,  $p$  item intercepts,  $p$  loadings,  $p$  unique variances and  $m$  means,  $m$  variances and  $m(m-1)/2$  covariances of the common factors. The total number of parameters is  $p(K+3) + m(m+3)/2$ . When multiple groups are considered simultaneously with CFA models of the same configuration<sup>1</sup> but separate sets of parameters, we will use a superscript ( $g$ ) for  $g = 1, 2, \dots, G$ , where  $G$  is the total number of groups. In this case, the total number of parameters is  $G(p(K+3) + m(m+3)/2)$ . These parameters are counted in the first row of Table 3 and Table 4 for polytomous and binary data. Of course, these parameters are not identified.

## 2.2. Factorial invariance

Establishing some degree of metric invariance in factor analytic models is an important step in comparing relationships between latent variables across samples, groups and other types of selection. While methods and terminology for metric invariance can vary across disciplines, most applications of metric invariance to structural equation modeling derive from the hierarchy of invariance levels first established by Meredith (1964a,b, 1993) and subsequently reviewed by Widaman and Reise (1997), Vandenberg and Lance (2000), Millsap and Meredith (2007) and Estabrook (2012). These levels of invariance create a hierarchy of constraints on parameters across groups that increasingly attributes group differences to the common factors. The four levels of factorial invariance most frequently referred to are: configural, weak, strong, and strict factorial invariance, each with varying restrictions on confirmatory factor models for continuous data.

Configural invariance is either the least stringent form of factorial invariance or a precondition for the Meredith definitions of invariance described elsewhere in this section. The only requirements of configural invariance are that there are the same number of factors in each group, and the pattern of salient (non-zero) and non-salient (zero) loadings is the same across groups. No restriction is imposed on the values of nonzero loadings.

The three levels of factorial invariance described by Meredith (1993) place additional constraints on a configurally invariant structure to establish metric invariance. Weak factorial invariance constrains the whole of the factor loading matrix to equality across groups, making all group differences in observed covariances attributable to differences in the factor covariance structure. This does not fit the definition of metric invariance: the free estimation of item intercepts in each group prevent the estimation of factor means and the testing of group differences in factor means.

Strong factorial invariance extends weak factorial invariance, further constraining item intercepts to equality across groups. Under strong factorial invariance, differences in observed means and covariances are due to group differences in factor means and covariances. Strict factorial invariance further constrains residual variances to equality, making all observed differences in means, variances and covariances attributable to the common factors. Both levels satisfy Meredith (1993) definition of metric invariance.

---

<sup>1</sup>This includes the same number of factors, the same loading patterns, and that all nonzero loadings are positive. This assumption is generally satisfied for confirmatory analyses.

These definitions were all developed for continuous interval level responses. As such, all observed variables have an inherent scale and can be described as having means, variances and covariances with other variables. Ordinal and binary data lack these features, and thus alternative methods must be used. Millsap and Yun-Tein (2004) identified thresholds, factor loadings, intercepts and unique variances as parameters for studies of measurement invariance and obtained an identification condition for a model with configural invariance in both binary and polytomous data. Before reviewing these identification conditions in detail, we give a brief summary of the basic concepts in this area.

### 3. Identifying a Model

#### 3.1. Identifiability and Identification Conditions

A statistical model is a family of distributions indexed by a set of parameters. If different sets of parameters index different members in this family of distribution, the model is identified or identifiable. Two statistical models are equivalent if they contain the same family of distributions. They are nested if one model is only a subset of the other model. For example, an exploratory factor analysis (EFA) model for normal data is a family of multivariate normal distributions<sup>2</sup> with a mean vector and a covariance matrix that satisfy Equation 1. This class of distribution is indexed by all parameters mentioned in the Introduction except the thresholds. This model is not identified because two sets of parameters that are related to each other by

$$\begin{cases} \tilde{\Lambda} &= \Lambda \mathbf{U} \\ \tilde{\nu} &= \nu + \Lambda \beta \\ \tilde{\kappa} &= \mathbf{U}^{-1}(\kappa - \beta) \\ \tilde{\Phi} &= \mathbf{U}^{-1} \Phi \mathbf{U}^{-1'} \end{cases} \quad (2)$$

produce exactly the same mean and covariance matrix for the multivariate normal distribution (i.e., they index the same member in the family of distributions) and fit data equally well. In this transformation, the  $p \times p$  non-singular matrix  $\mathbf{U}$  represents the rotational indeterminacy of the common factors and the indeterminacy of their variances, while the vector  $\beta$  of length  $m$  represents the indeterminacy of factor means. For a CFA model with a clustered loading structure and only non-negative loadings,  $\mathbf{U}$  can only be a  $p \times p$  positive diagonal matrix in order to preserve the structure of  $\tilde{\Lambda}$ . In this case, it represents the indeterminacy in the standard deviations of the common factors. The system of equations above defines a transformation that transforms one set of parameter into another without changing the model implied mean vector and covariance matrix. It represents all sources of non-identifiability of a model and can be used to examine whether a given set of constraints is an identification condition. A similar transformation for ordered categorical data will serve as an important tool in this paper.

<sup>2</sup>Many estimation methods in factor analysis do not assume normality, but because they only fit the covariance and mean structures, the identification conditions mentioned here are still valid.

To identify a model, constraints on parameters must be introduced. A set of constraints may or may not identify the model, and may or may not restrict the model to a smaller class of distributions. A set of constraints is called an identification condition, if it makes the model identifiable<sup>3</sup> without restricting the class of distributions to a subset. By applying an identification condition, one obtains a statistically equivalent but identified model.

For any distribution in the model, one and only one set of parameters that satisfies a given identification condition can be found to index this distribution. As a result, an identification condition does not change the fit of the model. Instead, it merely chooses one of the many sets of parameters that fit the model equally well. For example,

$$\text{diag}(\Phi) = \mathbf{I} \quad \text{and} \quad \kappa = \mathbf{0} \quad (3)$$

are an identification condition for a CFA model for continuous outcomes with nonnegative loadings. This statement carries the following two-fold information: first, for any mean vector and covariance matrix that satisfy Equation 1, one may use Transformation 2 with  $\mathbf{U} = \text{diag}^{\frac{1}{2}}(\Phi)$  and  $\boldsymbol{\beta} = \boldsymbol{\kappa}$  to reparametrize this model to satisfy Identification Condition 3; second, for any mean vector and covariance matrix that satisfy both Equation 1 and Identification Condition 3, the parameter values are uniquely defined because the identification condition effectively fixes  $\mathbf{U}$  and  $\boldsymbol{\beta}$  in Transformation 2 to  $\mathbf{U} = \mathbf{I}$  and  $\boldsymbol{\beta} = \mathbf{0}$ . Condition 3, however, is not an identification condition for an EFA model: imposing these constraints on both the left- and right-hand-side parameters of Transformation 2 only leads to  $\boldsymbol{\beta} = \mathbf{0}$  and  $\text{diag}[(\mathbf{U}\mathbf{U}')^{-1}] = \mathbf{I}$ , leaving the possibility of an oblique rotation. The two-fold information of an identification condition discussed above will serve as key reasoning in the proofs of propositions in this paper.

Identification conditions may not be unique, but different ways to identify the same model should end up with the same number of parameters, which is also the intrinsic dimension of the class of distributions. For example, instead of setting the common factor variance to 1, one may choose to set one loading of each common factor to 1 to identify the scale of that factor, though this option may lead to numerical instability if the chosen loading is close to 0.

### 3.2. Identifying a model for a single group

A CFA model for ordered categorical data has two types of latent variables: the  $m$  common factors and the  $p$  latent continuous responses for items. Neither has a fixed scale and both introduce indeterminacy of parameters into the model. To identify this model in a single group, we shall use the following transformation:

---

<sup>3</sup>It is usually good enough to identify a subset of the parameter space as long as its complement has lower dimensions.

$$\begin{cases} \tilde{\mathbf{T}} &= \boldsymbol{\gamma}\mathbf{1}' + \boldsymbol{\Delta}^{-1}\mathbf{T} \\ \tilde{\boldsymbol{\Lambda}} &= \boldsymbol{\Delta}^{-1}\boldsymbol{\Lambda}\mathbf{D} \\ \tilde{\boldsymbol{\nu}} &= \boldsymbol{\Delta}^{-1}\boldsymbol{\nu} + \boldsymbol{\Delta}^{-1}\boldsymbol{\Lambda}\boldsymbol{\beta} + \boldsymbol{\gamma} \\ \tilde{\boldsymbol{\Theta}} &= \boldsymbol{\Delta}^{-1}\boldsymbol{\Theta}\boldsymbol{\Delta}^{-1} \\ \tilde{\boldsymbol{\kappa}} &= \mathbf{D}^{-1}(\boldsymbol{\kappa} - \boldsymbol{\beta}) \\ \tilde{\boldsymbol{\Phi}} &= \mathbf{D}^{-1}\boldsymbol{\Phi}\mathbf{D}^{-1} \end{cases} \quad (4)$$

where  $\mathbf{T} = \{\tau_{jk}\}$  is a  $p \times K$  matrix that includes all finite thresholds. The  $m \times m$  positive diagonal matrix  $\mathbf{D}$  represents the indeterminacy of the common factor variances; the  $m \times 1$  vector  $\boldsymbol{\beta}$  represents the indeterminacy of the common factor means; the  $p \times 1$  vector  $\boldsymbol{\gamma}$  and the  $p \times p$  positive diagonal matrix represent the indeterminacy of locations and scales of the latent continuous responses. The two positive diagonal matrices and the two vectors above can be freely chosen to transform one set of parameters into another set without changing the probabilities of response patterns. The total number of freely chosen transformation constants is  $2(p + m)$ , which is also the number of redundant parameters in the model. We have the following proposition:

**Proposition 1**—Two sets of parameters  $(\mathbf{T}, \boldsymbol{\Lambda}, \boldsymbol{\nu}, \boldsymbol{\Theta}, \boldsymbol{\kappa}, \boldsymbol{\Phi})$  and  $(\tilde{\mathbf{T}}, \tilde{\boldsymbol{\Lambda}}, \tilde{\boldsymbol{\nu}}, \tilde{\boldsymbol{\Theta}}, \tilde{\boldsymbol{\kappa}}, \tilde{\boldsymbol{\Phi}})$  are related to each other through Transformation 4 iff they produce the same probabilities of response patterns for the model in Section 2.1 with a clustered loading pattern and non-negative loadings.

**Proof:** We first prove the “only if” part. Suppose the two sets of parameters are related to each other through Transformation 4. It is straightforward to verify that the model-implied mean vector and covariance matrix of the latent continuous responses by the tilde parameters are (from Equation 1)

$$\tilde{\boldsymbol{\mu}} = \boldsymbol{\Delta}^{-1}\boldsymbol{\mu} + \boldsymbol{\gamma} \quad \text{and} \quad \tilde{\boldsymbol{\Sigma}} = \boldsymbol{\Delta}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Delta}^{-1}.$$

If we standardize the latent continuous responses, the thresholds become

$$\text{diag}^{-\frac{1}{2}}\left(\sum\right)(\tilde{\mathbf{T}} - \tilde{\boldsymbol{\mu}}\mathbf{1}') = \text{diag}^{-\frac{1}{2}}\left(\sum\right)\boldsymbol{\Delta}(\boldsymbol{\Delta}^{-1}\mathbf{T} - \boldsymbol{\Delta}^{-1}\boldsymbol{\mu}\mathbf{1}') = \text{diag}^{-\frac{1}{2}}\left(\sum\right)(\mathbf{T} - \boldsymbol{\mu}\mathbf{1}')$$

which are the same to the standardized thresholds in the original parameters. Because the two sets of parameters also imply the same correlations among the latent continuous responses (i.e. polychoric correlations), they produce the same probabilities of the response patterns.

Now we prove the “if” part. If two sets of parameters produce the same probabilities for the response patterns, they must produce the same multinomial probabilities for each item and therefore the same standardized thresholds. As a result, there must exist diagonal matrix and vector  $\boldsymbol{\gamma}$  such that

$$\tilde{\mathbf{T}} = \gamma \mathbf{1}' + \Delta^{-1} \mathbf{T}, \quad \tilde{\boldsymbol{\mu}} = \Delta^{-1} \boldsymbol{\mu} + \gamma \quad \text{and} \quad \text{diag}(\tilde{\boldsymbol{\Sigma}}) = \Delta^{-1} \text{diag}(\boldsymbol{\Sigma}) \Delta^{-1}.$$

Because the two sets of parameters must produce the same response probabilities for every pair of items, they must produce the same polychoric correlations, which strengthens the last equation above to  $\tilde{\boldsymbol{\Sigma}} = \Delta^{-1} \boldsymbol{\Sigma} \Delta^{-1}$ . Using the expressions of mean and covariance structures in Equation 1, we have

$$\tilde{\boldsymbol{\nu}} + \tilde{\boldsymbol{\Lambda}} \tilde{\boldsymbol{\kappa}} = \Delta^{-1} \boldsymbol{\nu} + \Delta^{-1} \boldsymbol{\Lambda} \boldsymbol{\kappa} + \gamma \quad \text{and} \quad (5)$$

$$\tilde{\boldsymbol{\Lambda}} \tilde{\boldsymbol{\Phi}} \tilde{\boldsymbol{\Lambda}}' + \tilde{\boldsymbol{\Theta}} = \Delta^{-1} \boldsymbol{\Lambda} \boldsymbol{\Phi} \boldsymbol{\Lambda}' \Delta^{-1} + \Delta^{-1} \boldsymbol{\Theta} \Delta^{-1}. \quad (6)$$

The two sides of Equation 6 gives two decompositions of the same covariance matrix into a diagonal matrix and a matrix of a given lower rank  $m < p$ . Because such decomposition is unique, we have  $\tilde{\boldsymbol{\Theta}} = \Delta^{-1} \boldsymbol{\Theta} \Delta^{-1}$  and  $\tilde{\boldsymbol{\Lambda}} \tilde{\boldsymbol{\Phi}} \tilde{\boldsymbol{\Lambda}}' = \Delta^{-1} \boldsymbol{\Lambda} \boldsymbol{\Phi} \boldsymbol{\Lambda}' \Delta^{-1}$ . The second equation implies that there must be a non-singular matrix  $\mathbf{D}$  such that  $\tilde{\boldsymbol{\Lambda}} = \Delta^{-1} \boldsymbol{\Lambda} \mathbf{D}$  and  $\tilde{\boldsymbol{\Phi}} = \mathbf{D}^{-1} \boldsymbol{\Phi} \mathbf{D}^{-1}$ . Because both  $\tilde{\boldsymbol{\Lambda}}$  and  $\boldsymbol{\Lambda}$  has the same clustered structure of non-negative loadings,  $\mathbf{D}$  must be a positive diagonal matrix. Now Equation 5 becomes

$$\tilde{\boldsymbol{\nu}} = \Delta^{-1} \boldsymbol{\nu} + \Delta^{-1} \boldsymbol{\Lambda} (\boldsymbol{\kappa} - \mathbf{D} \tilde{\boldsymbol{\kappa}}) + \gamma$$

We denote the difference  $\boldsymbol{\kappa} - \mathbf{D} \tilde{\boldsymbol{\kappa}}$  by  $\boldsymbol{\beta}$ , and all equations in Transformation 4 have been proved.

For a model for categorical outcomes, non-identifiability arises when two sets of parameters imply the same set of probabilities of response patterns. Transformation 4 links such two sets of parameters with two diagonal matrices  $\mathbf{D}$  and  $\boldsymbol{\Phi}$  and two vectors  $\boldsymbol{\beta}$  and  $\boldsymbol{\gamma}$ , which explain the mechanism of such non-identifiability. To identify a model for a single group,  $2(p + m)$  constraints must be introduced to remove the indeterminacy of parameters without restricting the model.

For example, one widely used identification condition (e.g., Christofferson, 1975) is

$$\text{diag}(\boldsymbol{\Phi}) = \mathbf{I}, \quad \boldsymbol{\kappa} = \mathbf{0}, \quad \boldsymbol{\nu} = \mathbf{0}, \quad \text{and} \quad \text{diag}(\boldsymbol{\Sigma}) = \mathbf{I}. \quad (7)$$

The first equation sets  $\mathbf{D} = \mathbf{I}$  in Transformation 4 (because  $\text{diag}(\tilde{\boldsymbol{\Phi}}) = \text{diag}(\boldsymbol{\Phi}) = \mathbf{I}$  implies  $\mathbf{D} = \mathbf{I}$ ); the second equation sets  $\boldsymbol{\beta} = \mathbf{0}$ ; the third equation sets  $\boldsymbol{\gamma} = \mathbf{0}$ . Note the last equation defines  $p$  constraints on the variances of the latent responses, which are functions of the unique variances, factor loadings and factor variances. This sets  $\boldsymbol{\Sigma} = \mathbf{I}$  in the Transformation



4. It can also be verified that any set of parameters can be transformed to satisfy Condition 7 using Transformation 4 with properly chosen  $\mathbf{D}$ ,  $\boldsymbol{\beta}$  and  $\boldsymbol{\gamma}$ . An alternative identification condition constrains the variances of the residuals rather than the latent responses:

$$\text{diag}(\Phi)=\mathbf{I}, \quad \boldsymbol{\kappa}=\mathbf{0}, \quad \boldsymbol{\nu}=\mathbf{0}, \quad \text{and} \quad \Theta=\mathbf{I}. \quad (8)$$

Both of these options identify the model, but the  $\theta$ -parametrization of Condition 8 may give numerical instability if one of the unique variances is close to 0. Similar to models for continuous outcomes, in both Conditions 7 and 8 one may fix a reference loading of a factor rather than the factor variance.

### 3.3. Identifying a model with configural invariance

A model with configural invariance is the starting point of invariance study. In this paper we call it a baseline model. To identify this model, one may identify each group separately with Condition 7 or Condition 8. Alternatively, one may first identify one group and then identify the remaining groups through equality constraints across groups. The identification condition proposed by Millsap and Yun-Tein (2004, p. 485) for polytomous data is given below<sup>4</sup>.

1. For all groups set  $\boldsymbol{\nu}^{(g)} = \mathbf{0}$  and a reference loading to 1 for each factor;
2. In the first group set  $\boldsymbol{\kappa}^{(1)} = \mathbf{0}$  and  $\text{diag}(\Sigma^{(1)}) = \mathbf{I}$ ;
3. Set invariant one threshold per item and a second threshold in the reference item of each factor.

The total number of parameters in this model is  $G(pK + p + m(m-1)/2)$ , the same number as we use Condition 7 or 8 for each group. We refer to this parametrization (identification condition) of the baseline model as MYT parametrization (identification condition).

When data are dichotomous, imposing invariance on a second threshold is impossible. In this case, the latent response variance is set to 1 for the reference items. Parameter counts of the three different ways to identify this model are listed in the second block of Tables 3 and 4 for polytomous and binary outcomes.

## 4. Identifying models of different invariance levels

To test invariance, models with different invariant parameters can be compared. The issue of model comparison will not be discussed in this paper; instead, we focus on the identification conditions for models with different invariance properties. We limit our attentions to the invariance of the four measurement parameters: thresholds, factor loadings, intercepts and unique variances (Millsap and Yun-Tein, 2004, pp.483–484).

The intuitive way of building a model with parameter invariance is to impose invariance constraints on an identified baseline model. For example, a model with factor loading

<sup>4</sup>Millsap and Yun-Tein (2004) used slightly different notations: the thresholds are  $\nu$  (we use  $\tau$ ), the intercepts are  $\tau$  (we use  $\nu$ ), groups are indexed in subscript by  $k$  (we use superscript  $(g)$ ), thresholds indexed as  $m = 0, 1, \dots, c+1$  (we use  $k = 0, 1, \dots, K+1$ ).

invariance can be built by constraining the loadings to be invariant across groups in the baseline model identified by MYT parametrization (Millsap and Yun-Tein, 2004, p.494). Unfortunately, this approach has its drawbacks, as explained below.

An identification condition of the baseline model defines the scale of each latent continuous response and provides each with a mean and a variance for subsequent analyses. Because the estimates of the measurement parameters depend on these scales, constraints on these parameters across groups can have wildly different results depending on which identification condition is used in the baseline model.

This problem can be explained with a brief illustration whose parameters and results are presented in Tables 1 and 2. Consider a test of invariance involving one common factor each measured on four variables in two groups. The true values as listed in the first column of Table 1 have invariant thresholds, factor loadings and intercepts, but residual variances vary across groups. In the first group the latent continuous responses have unit variance but in the second group this is true only for the second item. We obtain a sample of size  $10^4$  in each group with these true values for use in this illustration.

Table 1 shows parameter estimates in the baseline model identified in different ways: estimates in Column 2 are produced with each group identified by Condition 7, but instead of fixing the factor variance, the first loading is set to 0.8; estimates in Columns 3 and 4 are produced with MYT parametrization with the first or second loading set to 0.8. These estimates were estimated with maximum likelihood using OpenMx (Boker, Neale, Maes et al., 2011; Neale, Hunter, Pritkin et al., 2015) in R (R Development Core Team, 2013). These identification conditions constrain the scale of each latent variable, and thus the residual variances, in different ways. As the true parameters do not imply unit variances for the latent responses in the second group (except for the second item), the first two identification methods effectively rescale the underlying continuous responses, which in turn transforms the factor loadings or thresholds, leading to pronounced differences in their estimates despite their invariance in the true values. The third identification condition is satisfied in the true parameters and therefore its use only leads to minor discrepancies in the loadings between the two groups during to sampling error. Despite the different parameter estimates, all three identification methods produce the same likelihood value as shown in Table 2.

Now we construct models with loadings invariance by imposing constraints on loadings across groups in the three versions of the baseline model. Because of the different sizes of group differences in the loading estimates of the baseline models, models with loading invariance find different levels of loss of fit as shown in the second half of Table 2. The first two methods find larger deficits in fit, leading to the rejection of loading invariance. The third method has smaller loss of fit and fails to reject loading invariance. Although the three models with loading invariance are intended as different parameterizations of the same model, they resulted in different maximum likelihood value, suggesting that they are not equivalent to each other. As we will see later, none of these three models is equivalent to a model with only loading invariance because such model should be equivalent to the baseline model.

This scaling problem can be equivalently stated using Transformation 4. An identification condition transforms true parameters that may not satisfy it into a new set of parameters that satisfies it. The loadings ( $\tilde{\Lambda}$ ) under a specific identification condition can be expressed as the product of the true loadings ( $\Lambda$ ) and two scaling matrices ( $\mathbf{C}$  and  $\mathbf{D}$ ) that depend on the identification condition. Even though the true loadings are invariant across groups, different scaling matrices may arise for different groups and result in non-invariant loadings in the identified model.

As we have seen above, imposing invariance constraints on loadings in a baseline model identified in different ways may lead to statistically non-equivalent models. Mathematically, this problem arises because the identification condition used in the baseline model may become restrictive when additional invariance conditions are introduced. They restrict the model with loading invariance in different ways, so the resultant models differ. Such restrictions change the model and the model is no longer “just” loading invariant.

To resolve this problem, we advocate a different approach that does not start with an identified baseline model. Instead, we identify models with different levels of invariance on a case by case basis. The model with only the desired invariance properties is first considered, and then appropriate constraints are added to “just” identify this model without imposing restrictions. Below we give the identification conditions for these models. Parameter counts of the models with the identification conditions are listed in Tables 3 and 4 for polytomous and binary data respectively. Surprisingly, we found that for some parameters such as factor loadings, its invariance cannot be tested alone: a model with only invariant loadings is equivalent to the baseline model.

## 5. Identification conditions of models of different invariance levels

It is well known that a model with full measurement invariance can be identified with the following condition:

$$\nu=0, \quad \text{diag}(\sum^{(1)})=\mathbf{I}, \quad \kappa^{(1)}=0 \quad \text{and} \quad \text{diag}(\Phi^{(1)})=\mathbf{I} \quad (9)$$

In Sections 5.1 through 5.7, we consider the various situations where only some but not all of these four types of parameters are invariant and provide proofs of their identification conditions. Section 5.1 covers threshold invariance. Section 5.2 describes which types of invariance are testable or untestable by themselves. Sections 5.3 through 5.5 cover invariance of thresholds and one other parameter type. Section 5.6 covers invariance of loadings and unique variances. Section 5.7 describes remaining cases where three of the four types of parameters are invariant. Section 5.8 briefly discusses the possibility of testing factor means and variances under models with different levels of invariance.

Please note that the proofs of all propositions below follow the same logic: to establish a set of constraints as an identification condition, we need to prove: (a) this condition does not restrict the model, and (b) with this condition the model is identified. For (a), we prove that any set of parameters of this model can be transformed to satisfy this condition without

changing the probabilities of response patterns, which can be done by properly choosing the matrices  $\mathbf{D}$  and  $\mathbf{\Lambda}$  and the vectors  $\boldsymbol{\beta}$  and  $\boldsymbol{\gamma}$  for each group in Transformation 4. For (b), we prove that if this condition is satisfied by both the tilde and regular parameters in Transformation 4, we can reduce  $\mathbf{D}$  and  $\mathbf{\Lambda}$  to an identity matrix and  $\boldsymbol{\beta}$  and  $\boldsymbol{\gamma}$  to a null vector. This general logical framework will not be repeated in each proof.

### 5.1. Invariance of thresholds

We first consider the case of threshold invariance. For binary ( $K = 1$ ) or ternary ( $K = 2$ ) outcomes, a linear transformation on the latent response variable  $y_j^*$  in one group can match its thresholds to those of another group. As a result, any set of parameters with configural invariance can be transformed to satisfy threshold invariance through Transformation 4. This is summarized in the following propositions. Note that in all propositions below “model” refers to a CFA model with a clustered factor structure of nonnegative loadings.

**Proposition 2**—*The following constraints are an identification condition to a baseline model for ordered ternary data:*

$$\boldsymbol{\nu}^{(1)} = \mathbf{0}, \quad \text{diag}(\boldsymbol{\Sigma}^{(1)}) = \mathbf{I}, \quad \text{and for all } g: \quad \mathbf{T}^{(g)} = \mathbf{T}^{(1)}, \quad \boldsymbol{\kappa}^{(g)} = \mathbf{0} \quad \text{and} \quad \text{diag}(\boldsymbol{\Phi}^{(g)}) = \mathbf{I} \quad (10)$$

**Proof:** For ternary items in multiple groups, Transformation 4 becomes

$$\begin{cases} \tilde{\mathbf{T}}^{(g)} &= \boldsymbol{\gamma}^{(g)} \mathbf{1}' + \boldsymbol{\Delta}^{(g)-1} \mathbf{T}^{(g)} \\ \tilde{\boldsymbol{\Lambda}}^{(g)} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda}^{(g)} \mathbf{D}^{(g)} \\ \tilde{\boldsymbol{\nu}}^{(g)} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\nu}^{(g)} + \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda}^{(g)} \boldsymbol{\beta}^{(g)} + \boldsymbol{\gamma}^{(g)} \\ \tilde{\boldsymbol{\Theta}}^{(g)} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Theta}^{(g)} \boldsymbol{\Delta}^{(g)-1} \\ \tilde{\boldsymbol{\kappa}}^{(g)} &= \mathbf{D}^{(g)-1} (\boldsymbol{\kappa}^{(g)} - \boldsymbol{\beta}^{(g)}) \\ \tilde{\boldsymbol{\Phi}}^{(g)} &= \mathbf{D}^{(g)-1} \boldsymbol{\Phi}^{(g)} \mathbf{D}^{(g)-1} \end{cases} \quad (11)$$

We first prove that Condition 10 does not restrict the model. We first choose  $\boldsymbol{\beta}^{(g)} = \boldsymbol{\kappa}^{(g)}$ ,  $\mathbf{D}^{(g)} = \text{diag}^{\frac{1}{2}}(\boldsymbol{\Phi}^{(g)})$ ,  $\boldsymbol{\Delta}^{(1)} = \text{diag}^{\frac{1}{2}}(\boldsymbol{\Sigma}^{(1)})$  and  $\boldsymbol{\gamma}^{(1)} = -\boldsymbol{\Delta}^{(1)-1} \boldsymbol{\nu}^{(1)} - \boldsymbol{\Delta}^{(1)-1} \boldsymbol{\Lambda}^{(1)} \boldsymbol{\beta}^{(1)}$  and the transformation above can transform any set of parameters to satisfy all equations in Condition 10 except  $\mathbf{T}^{(g)} = \mathbf{T}^{(1)}$ . Now we need to choose  $\boldsymbol{\gamma}^{(g)}$  and  $\boldsymbol{\beta}^{(g)}$  for  $g = 2, \dots, G$  so that for all the  $G - 1$  groups

$$\boldsymbol{\gamma}^{(g)} \mathbf{1}' + \boldsymbol{\Delta}^{(g)-1} \mathbf{T}^{(g)} = \tilde{\mathbf{T}}^{(1)} \quad (12)$$

holds, where  $\tilde{\mathbf{T}}^{(1)} = \boldsymbol{\gamma}^{(1)}\mathbf{1}' + \Delta^{(1)-1}\mathbf{T}^{(1)}$  is already determined by our choice of  $\boldsymbol{\gamma}^{(1)}$  and  $\Delta^{(1)}$  above. For each  $g$ , Equation 12 is a system of  $p \times 2$  linear equations of  $2p$  unknowns. In particular, its  $i$ th row comprises of the following system of equations:

$$\gamma_i^{(g)} + \delta_{ii}^{(g)-1} \tau_{i1}^{(g)} = \tilde{\tau}_{i1}^{(1)} \quad \text{and} \quad \gamma_i^{(g)} + \delta_{ii}^{(g)-1} \tau_{i2}^{(g)} = \tilde{\tau}_{i2}^{(1)}$$

which has a unique solution of  $\gamma_i^{(g)}$  and  $\delta_{ii}^{(g)}$  because  $\tau_{i1}^{(g)} < \tau_{i2}^{(g)}$ .

Now we prove that Condition 10 identifies the model. If this condition is satisfied by on both sides of Transformation 11: the last equation of Transformation 11 along with  $\text{diag}(\Phi^{(g)}) = \text{diag}(\tilde{\Phi}^{(g)}) = \mathbf{I}$  implies  $\mathbf{D}^{(g)} = \mathbf{I}$ ; the fifth equation of Transformation 11 along with  $\boldsymbol{\kappa}^{(g)} = \tilde{\boldsymbol{\kappa}}^{(g)} = \mathbf{0}$  implies  $\boldsymbol{\beta}^{(g)} = \mathbf{0}$ ; the constraint  $\text{diag}(\Sigma^{(1)}) = \text{diag}(\tilde{\Sigma}^{(1)}) = \mathbf{I}$  implies  $\Delta^{(1)} = \mathbf{I}$ ; the third equation of Transformation 11, which now becomes  $\tilde{\boldsymbol{\nu}}^{(g)} = \Delta^{(g)-1}\boldsymbol{\nu}^{(g)} + \boldsymbol{\gamma}^{(g)}$  along with  $\boldsymbol{\nu}^{(1)} = \tilde{\boldsymbol{\nu}}^{(1)} = \mathbf{0}$  implies  $\boldsymbol{\gamma}^{(1)} = \mathbf{0}$ . Now the first equation of Transformation 11 for  $g = 1$  becomes  $\tilde{\mathbf{T}}^{(1)} = \mathbf{T}^{(1)}$ . If the equations  $\mathbf{T}^{(g)} = \mathbf{T}^{(1)}$  (Condition 10) and  $\tilde{\mathbf{T}}^{(g)} = \tilde{\mathbf{T}}^{(1)}$  hold for  $g = 2, 3, \dots, G$ , we must have  $\tilde{\mathbf{T}}^{(g)} = \mathbf{T}^{(g)}$  for  $g = 2, 3, \dots, G$ , which leads to the system of Equations

$$\gamma_i^{(g)} + \delta_{ii}^{(g)-1} \tau_{i1}^{(g)} = \tau_{i1}^{(g)} \quad \text{and} \quad \gamma_i^{(g)} + \delta_{ii}^{(g)-1} \tau_{i2}^{(g)} = \tau_{i2}^{(g)}$$

for  $g = 2, 3, \dots, G$  and  $i = 1, 2, \dots, p$ . This implies  $\boldsymbol{\gamma}^{(g)} = \mathbf{0}$  and  $\Delta^{(g)} = \mathbf{I}$ . Now all four sources of non-identifiability in Transformation 11 are fixed.

**Proposition 3**—*The following constraints are an identification condition for a baseline model for binary data:*

$$\text{For all } g: \quad \boldsymbol{\tau}^{(g)} = \mathbf{0}, \quad \boldsymbol{\kappa}^{(g)} = \mathbf{0}, \quad \text{diag}(\Phi^{(g)}) = \mathbf{I}, \quad \text{and} \quad \text{diag}(\sum^{(g)}) = \mathbf{I}. \quad (13)$$

**Proof:** For binary items, Transformation 11 becomes

$$\begin{cases} \tilde{\boldsymbol{\tau}}^{(g)} = & \boldsymbol{\gamma}^{(g)} + \Delta^{(g)-1} \boldsymbol{\tau}^{(g)} \\ \tilde{\boldsymbol{\Lambda}}^{(g)} = & \Delta^{(g)-1} \boldsymbol{\Lambda}^{(g)} \mathbf{D}^{(g)} \\ \tilde{\boldsymbol{\nu}}^{(g)} = & \Delta^{(g)-1} \boldsymbol{\nu}^{(g)} + \Delta^{(g)-1} \boldsymbol{\Lambda}^{(g)} \boldsymbol{\beta}^{(g)} + \boldsymbol{\gamma}^{(g)} \\ \tilde{\boldsymbol{\Theta}}^{(g)} = & \Delta^{(g)-1} \boldsymbol{\Theta}^{(g)} \Delta^{(g)-1} \\ \tilde{\boldsymbol{\kappa}}^{(g)} = & \mathbf{D}^{(g)-1} (\boldsymbol{\kappa}^{(g)} - \boldsymbol{\beta}^{(g)}) \\ \tilde{\Phi}^{(g)} = & \mathbf{D}^{(g)-1} \Phi^{(g)} \mathbf{D}^{(g)-1} \end{cases} \quad (14)$$

Note that  $\tilde{\boldsymbol{\tau}}^{(g)}$  and  $\boldsymbol{\tau}^{(g)}$  are now vectors.

First, if we choose  $\beta^{(g)} = \kappa^{(g)}$ ,  $D^{(g)} = \text{diag}^{\frac{1}{2}}(\Phi^{(g)})$ ,  $\Delta^{(g)} = \text{diag}^{\frac{1}{2}}(\Sigma^{(g)})$  and  $\gamma^{(g)} = -\Delta^{(g)-1}\tau^{(g)}$  in the transformation then any set of parameters can be transformed to satisfy Constraints 13.

Second, we prove that if both sides of Transformation 14 satisfy Constraints 13, then  $\beta^{(g)} = \mathbf{0}$ ,  $D^{(g)} = \mathbf{I}$ ,  $\Sigma^{(g)} = \mathbf{I}$  and  $\gamma^{(g)} = \mathbf{0}$ . In fact, setting  $\text{diag}(\tilde{\Phi}^{(g)}) = \text{diag}(\Phi^{(g)}) = \mathbf{I}$  in the last equation of Transformation 14 gives  $D^{(g)} = \mathbf{I}$ , setting  $\tilde{\kappa}^{(g)} = \kappa^{(g)} = \mathbf{0}$  in the penultimate equation of Transformation 14 gives  $\beta^{(g)} = \mathbf{0}$ , setting  $\text{diag}(\tilde{\Sigma}^{(g)}) = \text{diag}(\Sigma^{(g)}) = \mathbf{I}$  gives  $\Sigma^{(g)} = \mathbf{I}$  and setting  $\tilde{\tau}^{(g)} = \tau^{(g)} = \mathbf{0}$  in the first equation of Transformation 14 gives  $\gamma^{(g)} = \mathbf{0}$ .

The propositions above indicate that for binary and ternary items, threshold invariance cannot be tested alone because they can be automatically satisfied. In other words, a model with invariant thresholds is statistically equivalent to a configural invariant model. When there are three or more thresholds for each item, a linear transformation that matches those thresholds across groups generally does not exist, so threshold invariance becomes testable. To identify a model with threshold invariance, the following identification condition can be used:

$$\nu^{(1)} = \mathbf{0}, \text{diag}(\Sigma^{(1)}) = \mathbf{I}, \text{ and for all } g: \kappa^{(g)} = \mathbf{0} \text{ and } \text{diag}(\Phi^{(g)}) = \mathbf{I} \quad (15)$$

Intuitively, the first group is identified because it uses parameterization specified by Condition 7. For the remaining groups the scale of their latent continuous responses are identified through invariant thresholds; the scale of common factors are identified by fixing the factor variances and means as in a model with observed continuous responses. Note that Condition 15 also identifies a model with threshold invariance for ordered ternary items ( $K = 2$ ). The total number of parameters is  $p(K-2) + G(3p + m(m-1)/2)$ , which coincide with that of the baseline model for  $K = 2$ . Formally we have the following proposition.

**Proposition 4**—Condition 15 is an identification condition for a model with more than one thresholds and threshold invariance.

**Proof:** For a multi-group model with threshold invariance, Transformation 4 becomes

$$\begin{cases} \tilde{\mathbf{T}} & = & \gamma^{(g)}\mathbf{1}' + \Delta^{(g)-1}\mathbf{T} \\ \tilde{\Lambda}^{(g)} & = & \Delta^{(g)-1}\Lambda^{(g)}D^{(g)} \\ \tilde{\nu}^{(g)} & = & \Delta^{(g)-1}\nu^{(g)} + \Delta^{(g)-1}\Lambda^{(g)}\beta^{(g)} + \gamma^{(g)} \\ \tilde{\Theta}^{(g)} & = & \Delta^{(g)-1}\Theta^{(g)}\Delta^{(g)-1} \\ \tilde{\kappa}^{(g)} & = & D^{(g)-1}(\kappa^{(g)} - \beta^{(g)}) \\ \tilde{\Phi}^{(g)} & = & D^{(g)-1}\Phi^{(g)}D^{(g)-1} \end{cases} \quad (16)$$

Note that the matrices of thresholds  $\mathbf{T}$  and  $\tilde{\mathbf{T}}$  do not bear a superscript.

First, any set of parameters with invariant thresholds can be transformed to satisfy Condition 15 using the transformation above with  $\mathbf{D}^{(g)} = \text{diag}^{\frac{1}{2}}(\Phi^{(g)})$ ,  $\boldsymbol{\beta}^{(g)} = \boldsymbol{\kappa}^{(g)}$ ,  $\boldsymbol{\Delta}^{(g)} = \text{diag}^{\frac{1}{2}}(\sum^{(1)})$  and  $\boldsymbol{\gamma}^{(g)} = -^{(1)-1}(\boldsymbol{\nu}^{(1)} + \boldsymbol{\Lambda}^{(1)}\boldsymbol{\kappa}^{(1)})$ . Note that  $\boldsymbol{\Delta}^{(g)}$  and  $\boldsymbol{\gamma}^{(g)}$  do not differ across groups, so the transformed threshold matrix  $\tilde{\mathbf{T}}$  is still invariant across groups. This shows that Condition 15 does not restrict the model.

Second, by applying Condition 15 to parameters on both sides of Transformation 16, we can easily obtain  $\mathbf{D}^{(g)} = \mathbf{I}$ ,  $\boldsymbol{\beta}^{(g)} = \mathbf{0}$ ,  $\boldsymbol{\gamma}^{(1)} = \mathbf{0}$  and  $\boldsymbol{\Delta}^{(1)} = \mathbf{I}$ . Now we only need to extend the last two equalities to the remaining  $G-1$  groups. Note that the first equation in Transformation 16 defines, for each item  $j$ ,  $G$  linear functions with coefficients  $\gamma_j^{(g)}$  and  $\delta_{jj}^{(g)}$  that map the  $K$  thresholds on the  $j$ th row of  $\mathbf{T}$  onto the  $K$  thresholds on the  $j$ th row of  $\tilde{\mathbf{T}}$ . Because both  $\mathbf{T}$  and  $\tilde{\mathbf{T}}$  do not vary across groups, the  $G$  linear functions share  $K > 1$  common points, so the  $G$  linear functions must be identical, with  $\boldsymbol{\gamma}^{(g)} = \boldsymbol{\gamma}^{(1)} = \mathbf{0}$  and  $\boldsymbol{\Delta}^{(g)} = \boldsymbol{\Delta}^{(1)} = \mathbf{I}$ . Now all four sources of indeterminacy are removed, so Condition 15 is able to identify the model.

### 5.2. Types of invariance that are not testable

A baseline model with Identification Condition 8 for each group automatically satisfies the invariance of unique variances and that of intercepts. As a result, invariance of either or both intercepts and unique variances cannot be tested alone.

We further have the following proposition.

**Proposition 5**—*The following constraints are an identification condition for a baseline model:*

*For all  $g$ : the nonzero loadings of  $\boldsymbol{\Lambda}^{(g)}$  be 1,  $\boldsymbol{\kappa}^{(g)} = \mathbf{0}$ ,  $\boldsymbol{\nu}^{(g)} = \mathbf{0}$  and  $\text{diag}(\Phi^{(g)}) = \mathbf{I}$*

$$(17)$$

**Proof:** First, Condition 17 does not restrict the model. In fact, in Transformation 11 we can choose  $\boldsymbol{\Delta}^{(g)}$  to be a diagonal matrix of factor loadings of the items, and choose

$\mathbf{D}^{(g)} = \text{diag}^{\frac{1}{2}}(\Phi^{(g)})$ ,  $\boldsymbol{\beta}^{(g)} = \boldsymbol{\kappa}^{(g)}$  and  $\boldsymbol{\gamma}^{(g)} = -^{(g)-1}(\boldsymbol{\nu}^{(g)} + \boldsymbol{\Lambda}^{(g)}\boldsymbol{\kappa}^{(g)})$ . The transformed parameters satisfy Condition 17.

Second, if this condition is satisfied by parameters on both sides of Transformation 11, we immediately have  $\boldsymbol{\beta}^{(g)} = \mathbf{0}$ ,  $\mathbf{D}^{(g)} = \mathbf{I}$ ,  $\boldsymbol{\gamma}^{(g)} = -^{(g)-1}\boldsymbol{\Lambda}^{(g)}\boldsymbol{\beta}^{(g)} = \mathbf{0}$ . Now the second equation of Transformation 11 becomes  $\tilde{\mathbf{A}}^{(g)} = ^{(g)-1}\boldsymbol{\Lambda}^{(g)}$ . Because the loading matrices on both sides are the same binary matrix, we have  $\boldsymbol{\Delta}^{(g)} = \mathbf{I}$ .

Because a baseline model with the identification condition above automatically satisfies the invariance of loadings and intercepts, the invariance of either or both loadings and intercepts cannot be tested alone.

For four or more categories, there are a total of five models equivalent to the baseline model as discussed above. For three categories, the model with threshold invariance is also equivalent to the baseline model as discussed in the last subsection. For binary data, three more models (invariance of thresholds and loadings, of thresholds and intercepts, and of thresholds and unique variances) are equivalent to the baseline model. They are discussed in their respective subsections below.

### 5.3. Invariance of thresholds and loadings

For dichotomous data (i.e.  $K = 1$ ), a model with both invariant thresholds and invariant loadings is statistically equivalent to the baseline model. In fact, we have the following proposition.

**Proposition 6**—*The constraints below are an identification condition for a baseline model for binary data:*

For all  $g$ : the nonzero loadings of  $\Lambda^{(g)}$  be 1,  $\kappa^{(g)} = \mathbf{0}$ ,  $\tau^{(g)} = \mathbf{0}$  and  $\text{diag}(\Phi^{(g)}) = \mathbf{I}$

(18)

**Proof:** First, Condition 18 is not restrictions. In fact, in Transformation 14 if we choose  $\mathbf{D}^{(g)}$  to be a diagonal matrix of factor loadings of the items and choose  $\mathbf{D}^{(g)} = \text{diag}(\frac{1}{2} \Phi^{(g)})$ ,  $\beta^{(g)} = \kappa^{(g)}$  and  $\gamma^{(g)} = -\mathbf{D}^{(g)-1} \tau^{(g)}$ , the transformed set of parameters satisfy Condition 18.

Second, Condition 18 identifies the model. In fact, if it is satisfied by parameters on both sides of Transformation 14, it is straightforward to obtain  $\beta^{(g)} = \mathbf{0}$ ,  $\mathbf{D}^{(g)} = \mathbf{I}$ ,  $\gamma^{(g)} = \mathbf{0}$  and  $\text{diag}(\Phi^{(g)}) = \mathbf{I}$ .

The proposition above shows that the baseline model for binary data is equivalent to a model with threshold and loading invariance, so such invariance cannot be tested alone.

If there are more than two categories, a model with threshold and loading invariance is more restrictive than the baseline model and can therefore be tested.

**Proposition 7**—*The following is an identification condition for a model with threshold and loading invariance for three or more ordered categories:*

$$\text{diag}(\sum^{(1)}) = \mathbf{I}, \quad \nu^{(1)} = \mathbf{0}, \quad \text{diag}(\Phi^{(1)}) = \mathbf{I}, \quad \text{and for all } g: \kappa^{(g)} = \mathbf{0}. \quad (19)$$

**Proof:** Transformation 11 now becomes



$$\begin{cases} \tilde{\mathbf{T}} &= \boldsymbol{\gamma}^{(g)} \mathbf{1}' + \boldsymbol{\Delta}^{(g)-1} \mathbf{T} \\ \tilde{\boldsymbol{\Lambda}} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda} \mathbf{D}^{(g)} \\ \tilde{\boldsymbol{\nu}}^{(g)} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\nu}^{(g)} + \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda} \boldsymbol{\beta}^{(g)} + \boldsymbol{\gamma}^{(g)} \\ \tilde{\boldsymbol{\Theta}}^{(g)} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Theta}^{(g)} \boldsymbol{\Delta}^{(g)-1} \\ \tilde{\boldsymbol{\kappa}}^{(g)} &= \mathbf{D}^{(g)-1} (\boldsymbol{\kappa}^{(g)} - \boldsymbol{\beta}^{(g)}) \\ \tilde{\boldsymbol{\Phi}}^{(g)} &= \mathbf{D}^{(g)-1} \boldsymbol{\Phi}^{(g)} \mathbf{D}^{(g)-1} \end{cases} \quad (20)$$

Note that both thresholds and loadings do not bear a superscript.

First, any set of parameters with invariant thresholds and loadings can be transformed to satisfy Condition 19 using this Transformation with  $\boldsymbol{\Delta}^{(g)} = \text{diag}^{\frac{1}{2}}(\boldsymbol{\Sigma}^{(1)})$ ,  $\boldsymbol{\gamma}^{(g)} = -\boldsymbol{\Delta}^{(1)-1}(\boldsymbol{\nu}^{(1)} + \boldsymbol{\Lambda} \boldsymbol{\kappa}^{(1)})$ ,  $\boldsymbol{\beta}^{(g)} = \boldsymbol{\kappa}^{(g)}$  and  $\mathbf{D}^{(g)} = \text{diag}^{\frac{1}{2}}(\boldsymbol{\Phi}^{(1)})$ . Note that  $\boldsymbol{\Delta}^{(g)}$ ,  $\mathbf{D}^{(g)}$  and  $\boldsymbol{\gamma}^{(g)}$  are only related to parameters in the first group and are therefore invariant across groups, so that the resultant transformed loadings and thresholds can be invariant cross groups.

Second, if Condition 19 is satisfied on both sides of Transformation 20, we immediately have  $\boldsymbol{\beta}^{(g)} = \mathbf{0}$ ,  $\mathbf{D}^{(1)} = \mathbf{I}$ ,  $\boldsymbol{\Delta}^{(1)} = \mathbf{I}$  and  $\boldsymbol{\gamma}^{(1)} = \mathbf{0}$ . Now we need to extend the last three equalities to the remaining  $G - 1$  groups. Similar to the reasoning in the proof of Proposition 4, the first equation of Transformation 20 implies  $\boldsymbol{\gamma}^{(g)} = \boldsymbol{\gamma}^{(1)} = \mathbf{0}$  and  $\boldsymbol{\Delta}^{(g)} = \boldsymbol{\Delta}^{(1)} = \mathbf{I}$ . The second equation of Transformation 20 now implies  $\mathbf{D}^{(g)} = \mathbf{D}^{(1)} = \mathbf{I}$ .

#### 5.4. Invariance of thresholds and intercepts

Now we consider threshold and intercept invariance. We have the following proposition

**Proposition 8**—An identification condition for a model with invariant thresholds and intercepts for three or more ordered categories is given by

$$\boldsymbol{\nu} = \mathbf{0}, \quad \text{diag}(\boldsymbol{\Sigma}^{(1)}) = \mathbf{I}, \quad \text{and for all } g: \quad \text{diag}(\boldsymbol{\Phi}^{(g)}) = \mathbf{I}. \quad (21)$$

**Proof:** For a multi-group model with threshold and intercept invariance, Transformation 11 becomes

$$\begin{cases} \tilde{\mathbf{T}} &= \boldsymbol{\gamma}^{(g)} \mathbf{1}' + \boldsymbol{\Delta}^{(g)-1} \mathbf{T} \\ \tilde{\boldsymbol{\Lambda}}^{(g)} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda}^{(g)} \mathbf{D}^{(g)} \\ \tilde{\boldsymbol{\nu}} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\nu} + \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda}^{(g)} \boldsymbol{\beta}^{(g)} + \boldsymbol{\gamma}^{(g)} \\ \tilde{\boldsymbol{\Theta}}^{(g)} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Theta}^{(g)} \boldsymbol{\Delta}^{(g)-1} \\ \tilde{\boldsymbol{\kappa}}^{(g)} &= \mathbf{D}^{(g)-1} (\boldsymbol{\kappa}^{(g)} - \boldsymbol{\beta}^{(g)}) \\ \tilde{\boldsymbol{\Phi}}^{(g)} &= \mathbf{D}^{(g)-1} \boldsymbol{\Phi}^{(g)} \mathbf{D}^{(g)-1} \end{cases} \quad (22)$$

Again, invariant parameters do not have a superscript. First, any set of parameters with invariant thresholds and intercepts can be transformed to satisfy Condition 21 using

$$\text{Transformation 22 with } \mathbf{D}^{(g)} = \text{diag}^{\frac{1}{2}}(\Phi^{(g)}), \boldsymbol{\beta}^{(g)} = \mathbf{0}, \Delta^{(g)} = \text{diag}^{\frac{1}{2}}(\sum^{(1)}) \text{ and } \boldsymbol{\gamma}^{(g)} = -\text{diag}^{-\frac{1}{2}}(\sum^{(1)})\boldsymbol{\nu}.$$

We then show that Condition 21 identifies the parameters. If this condition is satisfied on both sides of Transformation 22, we immediately have  $\mathbf{D}^{(g)} = \mathbf{I}$  and  $\sum^{(1)} = \mathbf{I}$ . Similar to the reasoning in the proof of Proposition 4, the first equation of Transformation 20 implies  $\boldsymbol{\gamma}^{(g)} = \boldsymbol{\gamma}^{(1)}$  (denoted by  $\boldsymbol{\gamma}$  below) and  $\Delta^{(g)} = \Delta^{(1)} = \mathbf{I}$ . Because  $\boldsymbol{\nu} = \tilde{\boldsymbol{\nu}} = \mathbf{0}$ , we must have  $\boldsymbol{\gamma} = -\Lambda^{(g)}\boldsymbol{\beta}^{(g)}$ . For the general case where loadings on the same factor are not proportional across groups, we must have  $\boldsymbol{\beta} = \mathbf{0}$  and  $\boldsymbol{\gamma} = \mathbf{0}$ . Now all four sources of indeterminacy in the transformation have been fixed.

It should be noted that technically Identification Condition 21 cannot identify the model if, after imposing this condition, loadings on the same factor are proportional across groups. Because this singularity occurs on a subset of the parameter space of lower dimensions, it does not affect the practical use of Condition 21.

For dichotomous items ( $K = 1$ ), a model with threshold and intercept invariance is equivalent to the baseline model. In fact, the MYT parametrization of a baseline model automatically satisfies these invariance relations.

### 5.5. Invariance of thresholds and unique variances

**Proposition 9**—*The following constraints are an identification condition for a model with invariant thresholds and unique variances:*

$$\boldsymbol{\nu}^{(1)} = \mathbf{0}, \quad \text{diag}(\sum^{(1)}) = \mathbf{I} \quad \text{and for all } g: \quad \boldsymbol{\kappa}^{(g)} = \mathbf{0} \quad \text{and} \quad \text{diag}(\Phi^{(g)}) = \mathbf{I} \quad (23)$$

**Proof:** For the current situation Transformation 11 becomes

$$\begin{cases} \tilde{\mathbf{T}} &= \boldsymbol{\gamma}^{(g)}\mathbf{1}' + \Delta^{(g)-1}\mathbf{T} \\ \tilde{\Lambda}^{(g)} &= \Delta^{(g)-1}\Lambda^{(g)}\mathbf{D}^{(g)} \\ \tilde{\boldsymbol{\nu}}^{(g)} &= \Delta^{(g)-1}\boldsymbol{\nu}^{(g)} + \Delta^{(g)-1}\Lambda^{(g)}\boldsymbol{\beta}^{(g)} + \boldsymbol{\gamma}^{(g)} \\ \tilde{\Theta} &= \Delta^{(g)-1}\Theta\Delta^{(g)-1} \\ \tilde{\boldsymbol{\kappa}}^{(g)} &= \mathbf{D}^{(g)-1}(\boldsymbol{\kappa}^{(g)} - \boldsymbol{\beta}^{(g)}) \\ \tilde{\Phi}^{(g)} &= \mathbf{D}^{(g)-1}\Phi^{(g)}\mathbf{D}^{(g)-1} \end{cases} \quad (24)$$

First, any set of parameters with invariant thresholds and unique variances can be transformed to satisfy Condition 23 using Transformation 24 with  $\mathbf{D}^{(g)} = \text{diag}^{\frac{1}{2}}(\Phi^{(g)}), \boldsymbol{\beta}^{(g)} = \boldsymbol{\kappa}^{(g)}, \Delta^{(g)} = \text{diag}^{\frac{1}{2}}(\sum^{(1)})$  and  $\boldsymbol{\gamma}^{(g)} = -\text{diag}^{-\frac{1}{2}}(\sum^{(1)})\boldsymbol{\nu}^{(1)} + \Lambda^{(1)}\boldsymbol{\beta}^{(1)}$ . Second, if Condition 23 is satisfied on both sides of Transformation 24, we immediately have  $\mathbf{D}^{(g)} = \mathbf{I}, \boldsymbol{\beta}^{(g)} = \mathbf{0}$ ,

$\mathbf{1}^{(1)} = \mathbf{I}$  and  $\boldsymbol{\gamma}^{(1)} = \mathbf{0}$ . Now the first and fourth equation of Transformation 24 implies  $\mathbf{1}^{(g)} = \mathbf{1}^{(1)} = \mathbf{I}$  and  $\boldsymbol{\gamma}^{(g)} = \boldsymbol{\gamma}^{(1)} = \mathbf{0}$ .

As a special case, the Identification Condition 23 also applies to dichotomous items, but in this case the model is statistical equivalent to the baseline model. This can be easily verified by calculating the effective number of parameters.

### 5.6. Invariance of loadings and unique variances

**Proposition 10**—*The following constraints are an identification condition for a model with invariant loadings and unique variances:*

$$\text{diag}(\boldsymbol{\Sigma}^{(1)}) = \mathbf{I}, \quad \text{diag}(\boldsymbol{\Phi}^{(1)}) = \mathbf{I}, \quad \text{and for all } g: \quad \boldsymbol{\kappa}^{(g)} = \mathbf{0} \quad \text{and} \quad \boldsymbol{\nu}^{(g)} = \mathbf{0}. \quad (25)$$

**Proof:** Transformation 11 now becomes

$$\begin{cases} \tilde{\mathbf{T}}^{(g)} &= \boldsymbol{\gamma}^{(g)} \mathbf{1}' + \boldsymbol{\Delta}^{(g)-1} \mathbf{T}^{(g)} \\ \tilde{\boldsymbol{\Lambda}} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda} \mathbf{D}^{(g)} \\ \tilde{\boldsymbol{\nu}}^{(g)} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\nu}^{(g)} + \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda} \boldsymbol{\beta}^{(g)} + \boldsymbol{\gamma}^{(g)} \\ \tilde{\boldsymbol{\Theta}} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Theta} \boldsymbol{\Delta}^{(g)-1} \\ \tilde{\boldsymbol{\kappa}}^{(g)} &= \mathbf{D}^{(g)-1} (\boldsymbol{\kappa}^{(g)} - \boldsymbol{\beta}^{(g)}) \\ \tilde{\boldsymbol{\Phi}}^{(g)} &= \mathbf{D}^{(g)-1} \boldsymbol{\Phi}^{(g)} \mathbf{D}^{(g)-1} \end{cases} \quad (26)$$

First, any set of parameters with invariant loadings and unique variances can be transformed to satisfy Condition 25 using Transformation 26 with  $\mathbf{D}^{(g)} = \text{diag}^{\frac{1}{2}}(\boldsymbol{\Phi}^{(1)})$ ,  $\boldsymbol{\beta}^{(g)} = \boldsymbol{\kappa}^{(g)}$ ,

$\boldsymbol{\Delta}^{(g)} = \text{diag}^{\frac{1}{2}}(\boldsymbol{\Sigma}^{(1)})$  and  $\boldsymbol{\gamma}^{(g)} = -\mathbf{1}^{(1)-1}(\boldsymbol{\nu}^{(g)} + \boldsymbol{\Lambda} \boldsymbol{\kappa}^{(g)})$ . Second, if Condition 25 is satisfied on both sides of Transformation 26, we immediately have  $\mathbf{D}^{(1)} = \mathbf{I}$ ,  $\boldsymbol{\beta}^{(g)} = \mathbf{0}$ ,  $\mathbf{1}^{(1)} = \mathbf{I}$  and  $\boldsymbol{\gamma}^{(g)} = \mathbf{0}$ . The second and fourth equations of Transformation 24 further implies  $\mathbf{1}^{(g)} = \mathbf{1}^{(1)} = \mathbf{I}$  and  $\mathbf{D}^{(g)} = \mathbf{D}^{(1)} = \mathbf{I}$ .

### 5.7. Invariance of three types of parameters

Because Identification Condition 25 for a model with invariant loadings and unique variances sets the intercepts to 0, this model is statistically equivalent to a model with invariant loadings, intercepts and unique variances. In the propositions below we discuss the remaining cases where three types of parameters are invariant.

**Proposition 11**—*The following is an identification condition for a model with invariant thresholds, loadings and intercepts:*

$$\boldsymbol{\nu} = \mathbf{0}, \quad \text{diag}(\boldsymbol{\Sigma}^{(1)}) = \mathbf{I}, \quad \text{diag}(\boldsymbol{\Phi}^{(1)}) = \mathbf{I} \quad \text{and} \quad \boldsymbol{\kappa}^{(1)} = \mathbf{0} \quad (27)$$

**Proof:** Transformation 11 now becomes

$$\begin{cases} \tilde{\mathbf{T}} &= \boldsymbol{\gamma}^{(g)} \mathbf{1}' + \boldsymbol{\Delta}^{(g)-1} \mathbf{T} \\ \tilde{\boldsymbol{\Lambda}} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda} \mathbf{D}^{(g)} \\ \tilde{\boldsymbol{\nu}} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\nu} + \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda} \boldsymbol{\beta}^{(g)} + \boldsymbol{\gamma}^{(g)} \\ \tilde{\boldsymbol{\Theta}}^{(g)} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Theta}^{(g)} \boldsymbol{\Delta}^{(g)-1} \\ \tilde{\boldsymbol{\kappa}}^{(g)} &= \mathbf{D}^{(g)-1} (\boldsymbol{\kappa}^{(g)} - \boldsymbol{\beta}^{(g)}) \\ \tilde{\boldsymbol{\Phi}}^{(g)} &= \mathbf{D}^{(g)-1} \boldsymbol{\Phi}^{(g)} \mathbf{D}^{(g)-1} \end{cases} \quad (28)$$

First, any set of parameters with invariant thresholds, loadings and intercepts can be transformed to satisfy Condition 27 using Transformation 28 with  $\mathbf{D}^{(g)} = \text{diag}^{\frac{1}{2}}(\boldsymbol{\Phi}^{(1)})$ ,  $\boldsymbol{\beta}^{(g)} = \boldsymbol{\kappa}^{(1)}$ ,  $\boldsymbol{\Delta}^{(g)} = \text{diag}^{\frac{1}{2}}(\sum^{(1)})$  and  $\boldsymbol{\gamma}^{(g)} = -\text{diag}^{-\frac{1}{2}}(\sum^{(1)})(\boldsymbol{\nu} + \boldsymbol{\Lambda} \boldsymbol{\kappa}^{(1)})$ .

Second, if condition holds on both sides of Transformation 28, we have  $\mathbf{D}^{(1)} = \mathbf{I}$ ,  $\boldsymbol{\Theta}^{(1)} = \mathbf{I}$ ,  $\boldsymbol{\beta}^{(1)} = \mathbf{0}$  and  $\boldsymbol{\gamma}^{(1)} = \mathbf{0}$ . Now we need to extend these equalities to the remaining  $G - 1$  groups, for which we need to consider the cases of binary and polytomous data separately below.

If there are more than one thresholds, similar to the reasoning in the proof of Proposition 4, the first equation of Transformation 28 implies  $\boldsymbol{\Theta}^{(g)} = \boldsymbol{\Theta}^{(1)} = \mathbf{I}$  and  $\boldsymbol{\gamma}^{(g)} = \boldsymbol{\gamma}^{(1)} = \mathbf{0}$ . Now the second and third equations in Transformation 28 implies  $\mathbf{D}^{(g)} = \mathbf{D}^{(1)} = \mathbf{I}$  and  $\boldsymbol{\beta}^{(g)} = \mathbf{0}$ .

For binary data, the second equation in Transformation 28 implies that  $\boldsymbol{\Theta}^{(g)}$  must take the following structure:  $\delta_{jj}^{(g)} = c_j d_{l(j)l(j)}^{(g)}$ , where  $l(j)$  refers to the index of the factor on which the  $j$ th item loads. Because  $\boldsymbol{\Theta}^{(1)} = \mathbf{I}$  and  $\mathbf{D}^{(1)} = \mathbf{I}$ , we must have  $c_j = 1$ . Because  $\tilde{\boldsymbol{\nu}} = \boldsymbol{\nu} = \mathbf{0}$ , we have  $\boldsymbol{\gamma}^{(g)} = -\boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda} \boldsymbol{\beta}^{(g)}$ . Because  $\boldsymbol{\gamma}^{(1)} = \mathbf{0}$  and  $\boldsymbol{\Theta}^{(1)} = \mathbf{I}$ , we have  $\tilde{\boldsymbol{\tau}} = \boldsymbol{\tau}$ . (Remember for binary items  $\mathbf{T}$  becomes a vector  $\boldsymbol{\tau}$ .) Now the first equation of Equation 28 becomes  $\boldsymbol{\tau} = -\boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda} \boldsymbol{\beta}^{(g)} + \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\tau}$ , or equivalently  $(\mathbf{I} - \boldsymbol{\Delta}^{(g)}) \boldsymbol{\tau} = \boldsymbol{\Lambda} \boldsymbol{\beta}^{(g)}$ . The  $j$ th row of this equation is  $(1 - \delta_{jj}^{(g)}) \tau_j = \lambda_{jl(j)} \beta_{l(j)}^{(g)}$ . Because we have proved  $\delta_{jj}^{(g)} = d_{l(j)l(j)}^{(g)}$ , this equation becomes  $(1 - d_{l(j)l(j)}^{(g)}) \tau_j / \lambda_{jl(j)} = \beta_{l(j)}^{(g)}$ . Note that in general the ratio  $\tau_j / \lambda_{jl(j)}$  varies with  $j$  but  $1 - d_{l(j)l(j)}^{(g)}$  and  $\beta_{l(j)}^{(g)}$  stay the same for items on the same factor, so we must have  $\mathbf{D}^{(g)} = \mathbf{I}$  and  $\boldsymbol{\beta}^{(g)} = \mathbf{0}$ , which further implies  $\boldsymbol{\Theta}^{(g)} = \mathbf{I}$  and  $\boldsymbol{\gamma}^{(g)} = \mathbf{0}$ . Now all four sources of indeterminacy of parameters are fixed, so the model is identified.

It should be noted that for binary data, Identification Condition 27 is not able to identify the model if, after imposing this condition, the ratios of correspondent thresholds and loadings are the same for all items of a given factor. Because this singularity occurs on a lower dimensional subset of the parameter space, it does not affect the practical use of this identification condition.

**Proposition 12**—*Below is an identification condition for a model with invariant thresholds, loadings and unique variances:*

$$\nu^{(1)}=0, \quad \text{diag}(\Sigma^{(1)})=\mathbf{I}, \quad \text{diag}(\Phi^{(1)})=\mathbf{I}, \quad \text{and for all } g: \quad \kappa^{(g)}=0 \quad (29)$$

**Proof:** For our current situation, Transformation 11 becomes

$$\begin{cases} \tilde{\mathbf{T}} &= \boldsymbol{\gamma}^{(g)} \mathbf{1}' + \Delta^{(g)-1} \mathbf{T} \\ \tilde{\boldsymbol{\Lambda}} &= \Delta^{(g)-1} \boldsymbol{\Lambda} \mathbf{D}^{(g)} \\ \tilde{\boldsymbol{\nu}}^{(g)} &= \Delta^{(g)-1} \boldsymbol{\nu}^{(g)} + \Delta^{(g)-1} \boldsymbol{\Lambda} \boldsymbol{\beta}^{(g)} + \boldsymbol{\gamma}^{(g)} \\ \tilde{\boldsymbol{\Theta}} &= \Delta^{(g)-1} \boldsymbol{\Theta} \Delta^{(g)-1} \\ \tilde{\boldsymbol{\kappa}}^{(g)} &= \mathbf{D}^{(g)-1} (\boldsymbol{\kappa}^{(g)} - \boldsymbol{\beta}^{(g)}) \\ \tilde{\boldsymbol{\Phi}}^{(g)} &= \mathbf{D}^{(g)-1} \boldsymbol{\Phi}^{(g)} \mathbf{D}^{(g)-1} \end{cases} \quad (30)$$

This identification condition does not restrict the model because any set of parameters with invariant thresholds, loadings and unique variances can be transformed to satisfy Condition 29 using Transformation 30 with  $\mathbf{D}^{(g)} = \text{diag}^{\frac{1}{2}}(\Phi^{(1)})$ ,  $\boldsymbol{\beta}^{(g)} = \boldsymbol{\kappa}^{(g)}$ ,  $\Delta^{(g)} = \text{diag}^{\frac{1}{2}}(\Sigma^{(1)})$  and  $\boldsymbol{\gamma}^{(g)} = -\text{diag}^{-\frac{1}{2}}(\Sigma^{(1)})(\boldsymbol{\nu}^{(1)} + \boldsymbol{\Lambda} \boldsymbol{\kappa}^{(1)})$ . Condition 29 removes all indeterminacy of parameters. In fact, if this condition holds in both sides of Transformation 30, we immediately have  $\boldsymbol{\beta}^{(g)} = \mathbf{0}$ ,  $\boldsymbol{\nu}^{(1)} = \mathbf{I}$ ,  $\mathbf{D}^{(1)} = \mathbf{I}$  and  $\boldsymbol{\gamma}^{(1)} = \mathbf{0}$ . The fourth equation of Transformation 30 implies  $\boldsymbol{\Theta}^{(g)} = \boldsymbol{\Theta}^{(1)} = \mathbf{I}$ ; the first and second equations of Transformation 30 now imply  $\boldsymbol{\gamma}^{(g)} = \boldsymbol{\gamma}^{(1)} = \mathbf{0}$  and  $\mathbf{D}^{(g)} = \mathbf{D}^{(1)} = \mathbf{I}$ .

Identification Condition 29 applies to both binary and ordered polytomous data, but for binary data, the model is statistically equivalent to one with only invariant loadings and unique variances. This can be verified by checking the effective number of parameters of both models.

**Proposition 13**—*Below is an identification condition for a model with invariant thresholds, intercepts and unique variances:*

$$\boldsymbol{\nu}=\mathbf{0}, \quad \text{diag}(\Sigma^{(1)})=\mathbf{I}, \quad \text{and for all } g \quad \text{diag}(\Phi^{(g)})=\mathbf{I} \quad (31)$$

**Proof:** For our current situation, Transformation 11 becomes

$$\begin{cases} \tilde{\mathbf{T}} &= \boldsymbol{\gamma}^{(g)} \mathbf{1}' + \boldsymbol{\Delta}^{(g)-1} \mathbf{T} \\ \tilde{\boldsymbol{\Lambda}}^{(g)} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda}^{(g)} \mathbf{D}^{(g)} \\ \tilde{\boldsymbol{\nu}} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\nu} + \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Lambda}^{(g)} \boldsymbol{\beta}^{(g)} + \boldsymbol{\gamma}^{(g)} \\ \tilde{\boldsymbol{\Theta}} &= \boldsymbol{\Delta}^{(g)-1} \boldsymbol{\Theta} \boldsymbol{\Delta}^{(g)-1} \\ \tilde{\boldsymbol{\kappa}}^{(g)} &= \mathbf{D}^{(g)-1} (\boldsymbol{\kappa}^{(g)} - \boldsymbol{\beta}^{(g)}) \\ \tilde{\boldsymbol{\Phi}}^{(g)} &= \mathbf{D}^{(g)-1} \boldsymbol{\Phi}^{(g)} \mathbf{D}^{(g)-1} \end{cases} \quad (32)$$

Condition 31 does not restrict the model because any set of parameters with invariant thresholds, intercept and unique variances can be transformed to satisfy this condition using

Transformation 32 with  $\mathbf{D}^{(g)} = \text{diag}^{\frac{1}{2}}(\boldsymbol{\Phi}^{(g)})$ ,  $\boldsymbol{\beta}^{(g)} = \mathbf{0}$ ,  $\boldsymbol{\Delta}^{(g)} = \text{diag}^{\frac{1}{2}}(\boldsymbol{\Sigma}^{(1)})$  and  $\boldsymbol{\gamma}^{(g)} = -\text{diag}^{-\frac{1}{2}}(\boldsymbol{\Sigma}^{(1)})\boldsymbol{\nu}$ .

Condition 31 identifies this model. In fact, if it holds on both sides of Transformation 32, we have  $\boldsymbol{\Sigma}^{(1)} = \mathbf{I}$  and  $\mathbf{D}^{(g)} = \mathbf{I}$ . The fourth equation in Transformation 32 implies  $\boldsymbol{\Sigma}^{(g)} = \boldsymbol{\Sigma}^{(1)} = \mathbf{I}$ . The first equation in Transformation 32 implies that  $\boldsymbol{\gamma}^{(g)}$  does not depend on  $g$ , so we can drop its superscript. The condition  $\tilde{\boldsymbol{\nu}} = \boldsymbol{\nu} = \mathbf{0}$  implies  $\boldsymbol{\gamma} = -\boldsymbol{\Lambda}^{(g)}\boldsymbol{\beta}^{(g)}$ , which, similar to the reasoning in the proof of Proposition 8, implies  $\boldsymbol{\beta}^{(g)} = \mathbf{0}$  and  $\boldsymbol{\gamma} = \mathbf{0}$  for the general case where loadings on the same factor are not proportional across groups.

### 5.8. The comparison of factor means and variances

A related issue in the invariance of measurement parameters is whether each of the invariance levels discussed in this paper allows the comparison of factor means and factor variances. One way to address this question is to check whether the identification condition of this model involves setting all factor means to zero or all factor variances to one. If such constraints are present in the identification condition, factor means or factor variances cannot be compared because the null and alternative models in this comparison would be equivalent. For example, the invariance of loadings and unique variances does not allow the comparison of factor means, because Identification Condition 25 contains the constraint  $\boldsymbol{\kappa}^{(g)} = \mathbf{0}$  for all groups.

However, to confirm that a model allows the comparison of factor means (or variances) one need to make sure that all possible identification conditions of this model do not involve such constraints, which would not be discussed in this paper. Fortunately, because threshold invariance (for polytomous data) effectively equates the scales of the latent responses, the general rules for continuous outcomes apply. For example, since for continuous outcomes the invariance of loadings and intercepts guarantees the comparison of both factor means and variances, for ordered polytomous data one can do the same comparison with invariant thresholds, loadings and intercepts.

## 6. Nesting and Equivalence Relations

The number of different parameters in models with different levels of invariance is summarized in Tables 3 and 4 for ordered polytomous and binary data, respectively. Note

that the symbols in the first column denote invariant parameters in a model. The 16 models with different levels of invariance form a hierarchy where the full measurement invariant model is the most constrained and the baseline model is the least constrained. This hierarchy is depicted in Figures 1 and 2 for polytomous and binary data. The nesting relationships among them are determined by their levels of invariance. For example, the model with invariant thresholds and unique variances is nested in the model with invariant thresholds only. In this case, the two models have the same identification conditions (Conditions 15 and 23) and the more restricted model can be obtained by applying additional constraints (in this case the invariance of unique variances) to the less restricted model. However, this is not always the case. Most nesting relationships among the 16 models are not obvious from the constraints they bear on their parameters. In most cases, when an additional invariance constraint is introduced, a part of the identification condition of the original model need to be removed because the newly introduced invariance provides information to identify additional parameters. For example, to obtain the model with invariant thresholds, loadings and unique variances from the model with only invariant loadings and unique variances, one need to *remove* the constraints on intercepts for  $G-1$  groups in addition to adding the constraint of threshold invariance. This does not change their nesting relationship because the nesting relationship between two models is a relationship between the two sets of statistical distributions, not a relationship between their free parameters or the constraints on their parameters. This is related to the major point of this paper that an invariant model should not be built by simply applying invariance constraints on an identified baseline model. For example, in all three parameterizations of the baseline model listed in Table 3, item intercepts are fixed to zero, which need to be freed when a model with invariant thresholds is built.

It is possible that adding one or multiple invariance conditions leads to an equivalent model. In this case, those invariance conditions cannot be tested because they can be satisfied trivially by a reparametrization. For example, not all invariance conditions imposes restrictions on the baseline model. Models with invariant loadings, intercepts or unique variances alone are equivalent to the baseline model. Each of those invariance conditions cannot be tested alone. Once thresholds are assumed invariant, invariance of loadings, intercepts and unique variances can be tested either separately or jointly. In light of this, we recommend investigating threshold invariance before considering other types of invariance. Once threshold invariance is established, the invariance of the remaining three types of parameters can be tested. For dichotomous items, more models are equivalent to the baseline model, but one is still able to compare models that are not statistically equivalent as depicted in Figure 2.

## 7. Singularities

It is not uncommon that an (otherwise) identified model is not identified on a subset of its parameter space. For example, a CFA model (for continuous items) with two standardized factors and two items per factor is identified, but when the factor correlation is zero, the remaining parameters are not identified; an otherwise identified mixture model is not identified when a mixture component (that involves unique parameters) has zero probability; a model with a covariance matrix of random effects parametrized in Cholesky

decomposition is not identified when the true value of this covariance matrix is singular; a fixed effect 3PL IRT model can be identified by fixing the slope and difficulty of one item (Wu, accepted), but that is not enough for a fixed effect Rasch model with guessing (van der Linden and Barrett, 2016). These singularities occur at a lower dimensional subset of the parameter space, so they do not affect the identification of the model for *almost* the whole parameter space. However, if parameter estimates are close to this subset, the model is near un-identified and parameter estimation may suffer numerical instability.

Several models in this paper also have such singularities. The model with threshold and intercept invariance has singularities on the subset of its parameter space where loadings on the same factor are proportional across groups. The same situation arises for the model with invariant thresholds, intercepts and unique variances. The model for binary data with invariant thresholds, loadings and intercepts has singularities on the subset of its parameter space where the threshold to loading ratio is constant across items on the same factor.

Although these singularities do not affect the use of the model, the likelihood ratio test (LRT) statistic will generally not follow a  $\chi^2$  distribution if the true parameter values are in a singularity set of at least one of the two compared models (e.g. Drton, 2009). For example, it is well known that in mixture modeling the LRT for the number of mixture components is a nonstandard test (Davies, 1977, 1987; Jeffries, 2003) and that tests on singular multivariate variance components suffer from the same problem (Carey, 2005; Wu and Neale, 2013). For the CFA models discussed above, the LRT comparing a full measurement invariance model and a model with only threshold and intercept invariance is a nonstandard test because under the null hypothesis, the true model has equal loadings across groups, falling in the singularities of the alternative model. In this test, the asymptotic distribution of the LRT statistic may not be  $\chi^2$  distributed.

## 8. Numerical Illustration and Simulation Studies

### 8.1. A numerical illustration of equivalence relations

We introduce an numerical illustration to demonstrate the equivalence relationships in the first box in Figure 1. Consider a model for four ternary items of one standardized latent factor in two groups and let parameters take true values in the row “Set I” of Table 5. These parameters satisfies Identification Condition 7 in each group and also invariance of intercepts. They imply the following correlations for the latent responses for the two groups:

$$\mathbf{P}^{(1)} = \begin{pmatrix} 1 & 1/4 & \sqrt{2}/4 & \sqrt{2}/4 \\ 1/4 & 1 & \sqrt{2}/4 & \sqrt{2}/4 \\ \sqrt{2}/4 & \sqrt{2}/4 & 1 & 1/2 \\ \sqrt{2}/4 & \sqrt{2}/4 & 1/2 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{P}^{(2)} = \begin{pmatrix} 1 & 3/4 & 1/2 & 1/2 \\ 3/4 & 1 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1 & 1/3 \\ 1/2 & 1/2 & 1/3 & 1 \end{pmatrix}.$$

The standardized thresholds are the same as the thresholds because the latent responses are standardized under this parametrization. Below we will transform this set of true parameter values into three different sets of parameter values that satisfy different invariance conditions without changing the model implied standardized thresholds and polychoric correlations.



Because a configural invariant model is equivalent to a model with invariant loadings (see Proposition 5), we can transform the parameters to satisfy Identification Condition 17 using Transformation 11. Details of this transformation are given in the proof of Proposition 5. An easier account of this transformation is that both thresholds of an item in Set I are divided by the loading of that item to produce the transformed thresholds and its residual variance is divided by the squared loading to produce the transformed residual variance. The resultant parameters are listed in the row “Set II” of Table 5, where all factor loadings become one.

The row “Set III” of this table shows another parameterization of the same true values. Now the parameter values satisfy Identification Condition 8 in each group and the invariance of both intercepts and unique variances. This time the loading and thresholds of an item in Set I are divided by the square root of its residual variance to produce the transformed values.

For ternary data, the model with configural invariance is equivalent to the model with threshold invariance (see proposition 2). We can transform the parameters to satisfy threshold invariance. The resultant parameters are in row “Set IV” of Table 5.

It is straightforward to find that all four sets of parameters imply the same correlations for the latent responses as shown above and the same standardized thresholds. As a result, they also imply the same probabilities for response patterns. Since the common factors stay standardized, the only difference across the four sets of parameters is the scale of the four latent responses. This illustration shows that for ordered ternary data, any set of parameters with configural invariance can be transformed to satisfy the invariance of thresholds, of loadings and intercepts, or of intercepts and unique variances. As a result, there is no degree of freedom to test these invariance conditions.

## 8.2. Simulation studies

We now introduce two simulation studies to demonstrate the regular and singular nesting relations among the models. For regular nesting relations we expect a difference test with ML estimation to produce an asymptotic  $\chi^2$  distribution; for singular nesting relations the test statistic may not have a  $\chi^2$  distribution. This simulation study also verifies the dfs marked in Figures 2 and 1.

In the first simulation, binary data were generated with a single factor model with four indicators in two groups. The true parameter values satisfy full measurement invariance. The four loadings are 0.7, 0.7, 0.5 and 0.5. The unique variances are 0.51, 0.51, 0.75 and 0.75. The intercepts are all zero. The four thresholds are equally space between  $-0.5$  and  $1$ . The factor variances are 1 and 2.25 for the two groups. The factor means are 0 and 1. The per group sample size is 1000. 2000 replications are used. For each replication, all five statistically non-equivalent models in Figure 2 were fitted using OpenMx with ML estimation<sup>5</sup> and all models converged.

Each arrow in Figure 2 corresponds to a nesting relationship and a difference test. The 1st-99th percentiles of the LRT statistic of the six difference tests are plotted in the six

<sup>5</sup>We choose ML because it produces a  $\chi^2$  distributed statistic if regularity conditions are met.

panels of Figure 3 against those of a  $\chi^2$  distribution with the desired dfs marked on the arrows in Figure 2. In the upper left corner of each panel the null and alternative models of this LRT are marked. Greek letters represent the invariant parameters. “Full” refers to the model invariant in all four types of parameters. Among the six panels, four show a good alignment between the empirical and desired distribution of the LRT statistic. The remaining two panels are tests that involve the model with invariant thresholds, intercepts and unique variances. As discussed in Sections 7, this model involves singularities and the true parameter values of this simulation is in the singularities of this model, so the resultant LRT statistic does not have the desired  $\chi^2$  distribution.

The second simulation study is the same to the first study except that the outcomes are in four categories with the following true threshold matrix (four columns for four items)

$$\mathbf{T} = \begin{pmatrix} -0.5 & -0.25 & 0 & 0.25 \\ -0.25 & 0 & 0.25 & 0.5 \\ 0 & 0.25 & 0.5 & 0.75 \end{pmatrix}$$

The per-group sample size and number of replications are both 2000. For each replication, all 10 statistically non-equivalent models in Figure 1 were fitted using OpenMx with ML estimation. Except one replication which did not converge when fitted to the model with threshold and intercept invariance, all other models and replications converged.

Each arrow in Figure 1 corresponds to a nesting relationship and a difference test. The 1st-99th percentiles of the LRT statistic of the 15 tests are plotted in the 15 panels of Figure 4 against those of a  $\chi^2$  distribution with dfs marked on the arrows in Figure 1. The two nested models are marked on the upper left corner of each panel in the same way as in Figure 3. In four of the 15 panels the LRT does not produce a statistic with the desired  $\chi^2$  distribution. These four tests involve either the model with threshold and intercept invariance or the model with invariant thresholds, intercepts and unique variances. Again, this happens because the true parameters of this simulation lies on the singularities of both models. Interestingly, the distribution of the LRT statistic comparing these two models seems to be close to the desired  $\chi^2$  distribution even though the regularity condition is still violated. All remaining panels of this figure show a good alignment between the theoretical and simulated distributions of the statistic.

## 9. Model Specifications in Mplus and OpenMx

For ordered categorical data from multiple groups, Mplus (Muthén and Muthén, 1998–2012) imposes several default constraints on the parameters. Some of these constraints can be overridden by carefully specifying all parameters explicitly in the `MODEL` command. However, some default cannot be overridden easily. For example, the intercepts are fixed at 0 for ordered categorical data. This is hard to override because intercept parameters are not explicitly specified in the `MODEL` command. Below we give an alternative way to specify models with which all the 16 models discussed above can be specified easily.

To specify a model with ordered categorical data with explicit intercept parameters, one may explicitly specify the unobserved response variables  $y^*$  as latent indicator variables of the factors and then specify the observed categorical variables as single indicators of the  $y^*$ s. The loadings of these single indicators must be fixed to 1 and the error variances of the  $y^*$  variables must be fixed at 0. The intercepts of these  $y^*$  variables can then be explicitly specified as the intercepts of the ordered categorical variables. In the Appendix, a model with a single common factor, four indicators and two groups is specified in this way. Threshold and loading invariance is imposed.

In OpenMx (Boker, Neale, Maes et. al., 2011; Neale, Hunter, Pritkin et. al., 2015), a model can be specified by defining the threshold, mean and covariance structures as matrix valued functions of parameters. Equality constraints on model parameters are specified by using the same label to corresponding elements in the matrices. Although a functional constraint such as  $\text{diag}(\Sigma) = \mathbf{I}$  can be specified using `mxConstraint`, it adds to the computational complexity and should be avoided if possible. We recommend that users specify the diagonal elements of  $\text{diag}(\Sigma)$  as explicit parameters if they are set to one in the model and express the unique variances as a function of parameters if needed. The OpenMx code of a model for binary data with invariant loadings and unique variances is illustrated in the Appendix. It has one common factor and four indicators.

## 10. Summary and Conclusion

Studies of measurement invariance for ordered categorical data usually proceed by first identifying the model with only configural invariance as a baseline and then introducing additional invariance constraints to build a hierarchy of models. In this approach, the scale of the latent responses is set by the identification condition used in the baseline model and all subsequent additional invariance constraints are built on this particular scale. This may lead to the incorrect rejection of parameter invariance if the original parameters are invariant under some different scale of the latent responses. This would also lead to different conclusions for different choices of identification condition for the baseline model. Mathematically, these occur because the identification condition used in the baseline model may become restrictions when additional invariance conditions are introduced. In this article we advocate an approach where each model in the invariance hierarchy is identified separately based on Transformation 11 that reveals the four sources of non-identifiability of the CFA model for ordered categorical data. In this approach models with different levels of invariance are not defined through modification of the baseline model and are therefore not affected by its identification condition. Using this approach we give the identification condition for models with different levels of invariance. We also find that the invariance of some parameters cannot be tested alone. For example, the invariance of loadings can not be tested when there is no threshold invariance. For data with more than two categories, we recommend that invariance be tested first for thresholds, and then proceed to other parameters.

This article only discusses the invariance of all parameters of the same type (e.g. the invariance of all loading parameters), but not partial invariance (see, e.g., Millsap and Meredith, 2007, pp.145–149) where parameters are invariant only for some of the items.

Tests of partial invariance for continuous data are sensitive to the how the model is standardized and the choice of the reference items (Cheung and Rensvold, 1998, 1999; Rensvold and Cheung, 2001), a problem of the same essence as discussed here, and has the additional issue of specification search (Oort, 1998; Yoon and Millsap, 2007). For ordered categorical data, the issue of identification will be more complex due to the presence of threshold parameters. Identification conditions for partially invariant models for ordered categorical data are beyond the scope of this article.

Testing measurement invariance in categorical data is inherently complex due to the sometimes competing needs of identification and invariance constraints. Given the importance of measurement invariance to substantive work, we urge researchers to exercise great care in selecting identification conditions to yield testable and interpretable measurement structures.

## Acknowledgments

Work on this research by the second author was partially supported by the National Institute of Drug Abuse research education program R25DA026-119 (Director: Michael C. Neale) and by grant R01 AG18436 (20112016, Director: Daniel K. Mroczek) from National Institute on Aging, National Institute on Mental Health.

Roger Millsap, whose work inspired this paper, unexpectedly passed away when we were preparing this manuscript. We would like to honor him for his pioneering work in measurement invariance.

## References

- Baker, FB. Item response theory parameter estimation techniques. New York, NY: Marcel Dekker, Inc; 1992.
- Babakus E, Ferguson CE, Jöreskog KG. The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research*. 1987; 24(2):222–228.
- Bernstein IH, Teng G. Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin*. 1989; 105:467–477.
- Boker S, Neale M, Maes H, Wilde M, Spiegel M, Brick T, et al. OpenMx: an open source extended structural equation modeling framework. *Psychometrika*. 2011; 76:306–317. [PubMed: 23258944]
- Carey G. Cholesky problems. *Behavioral Genetics*. 2005; 35:653–665.
- Cheung GW, Lau RS. A direct comparison approach for testing measurement invariance. *Organizational Research Methods*. 2012; 15(2):167–198.
- Cheung GW, Rensvold R. Cross cultural comparisons using non-invariant measurement items. *Applied Behavioral Science Review*. 1998; 6:93–110.
- Cheung GW, Rensvold R. Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management*. 1999; 25:1–27.
- Christofferson A. Factor analysis of dichotomized variables. *Psychometrika*. 1975; 40:5–32.
- Davies RB. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*. 1977; 64:247–254.
- Davies RB. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*. 1987; 74:33–43.
- Drton M. Likelihood ratio tests and singularities. *The Annals of Statistics*. 2009; 37(2):979–1012.
- Estabrook, R. Factorial Invariance: Tools and Concepts for Strengthening Research. In: Tenenbaum, G.Eklund, R., Kamata, A., editors. *Measurement in Sport and Exercise Psychology*. Champaign, IL: Human Kinetics; 2012.
- Jeffries NO. A note on ‘Testing the number of components in a normal mixture’. *Biometrika*. 2003; 90(4):991–994.

- Jöreskog KG, Moustaki I. Factor analysis of ordinal variables: A comparison of three approaches. *Multivariate Behavioral Research*. 2001; 36:347–387.
- Lubke GH, Muthén BO. Applying multiple group confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons. *Structural Equation Modeling*. 2004; 11(4):514–534.
- Mehta PD, Neale MC, Flay BR. Squeezing Interval Change From Ordinal Panel Data: Latent Growth Curves With Ordinal Outcomes. *Psychological Methods*. 2004; 9(3):301–333. [PubMed: 15355151]
- Meredith W. Notes on factorial invariance. *Psychometrika*. 1964a; 29:177–185.
- Meredith W. Rotation to achieve factorial invariance. *Psychometrika*. 1964b; 29:186–206.
- Meredith W. Measurement invariance, factor analysis and factor invariance. *Psychometrika*. 1993; 58:525–543.
- Millsap, RE., Meredith, W. Factorial invariance: Historical perspectives and new problems. In: Cudeck, R., MacCallum, RC., editors. *Factor analysis at 100: Historical developments and future directions*. Mahwah, New Jersey: Lawrence Erlbaum Associates; 2007. p. 131-152.
- Millsap RE, Yun-Tein J. Assessing factorial invariance in ordered categorical measures. *Multivariate Behavioral Research*. 2004; 39(3):479–515.
- Mislevy RJ. Recent developments in the factor analysis of categorical variables. *Journal of Educational Statistics*. 1986; 11:3–31.
- Muthén BO. A general structural equation model for dichotomous, ordered categorical and continuous latent variable indicators. *Psychometrika*. 1984; 49:115–132.
- Muthén B, Christofferson. Simultaneous factor analysis of dichotomous variables in several groups. *Psychometrika*. 1981; 46:407–419.
- Muthén BO, Kaplan D. A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology*. 1985; 38:171–189.
- Muthén, LK., Muthén, BO. *Mplus Users Guide*. 7. Los Angeles, CA: Muthén & Muthén; 1998–2012.
- Neale MC, Hunter MD, Pritkin J, Zahery M, Brick TR, Kirkpatrick RM, et al. OpenMx 2.0: Extended Structural Equation and Statistical Modeling. *Psychometrika*. 2015; :80.doi: 10.1007/s11336-014-9435-8
- Oort FJ. Simulation study of item bias detection with restricted factor analysis. *Structural Equation Modeling*. 1998; 5:107–124.
- R Development Core Team. R: A Language and Environment for Statistical Computing. 2013. <http://www.R-project.org>
- Rensvold RB, Cheung GW. Testing for metric invariance using structural equation models, solving the standardization problem. *Research in Management*. 2001; 1:25–50.
- Strom DO, Lee JW. Variance components testing in the longitudinal mixed effects model. *Biometrics*. 1994; 50:1171–1177. Corrected in *Biometrics*, 51, 1196. [PubMed: 7786999]
- van der Linden WJ, Barrett MD. Linking item response model parameters. *Psychometrika*. 2016
- Vandenberg RJ, Lance CE. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*. 2000; 3(1):4–69.
- Widaman, KF., Reise, SP. Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In: Bryant, KJ., Windle, M., West, SG., editors. *The science of prevention: Methodological advances from alcohol and substance abuse research*. American Psychological Association; 1997. p. 281-324.
- Wu H. A note on the identifiability of fixed effect 3PL models. *Psychometrika*. (accepted).
- Wu H, Neale MC. On the likelihood ratio tests in bivariate ACDE models. *Psychometrika*. 2013; 78(3): 441–463. [PubMed: 25106394]
- Yoon M, Millsap RE. Detecting violations of factorial invariance using data-based specification searches: A Monte-Carlo study. *Structural Equation Modeling*. 2007; 14(3):435–463.

## A. Mplus and OpenMx Codes

### A.1. Model specification in Mplus

```

TITLE: A model with invariant thresholds and loadings of two groups,
      one common factor, four indicators and five categories per
indicator
DATA: FILE IS (file name)
VARIABLE: NAMES ARE x1 x2 x3 x4 group;
          USEVARIABLES ARE x1 x2 x3 x4 group;
          GROUPING IS group (0=male 1=female);
          CATEGORICAL ARE x1 x2 x3 x4
ANALYSIS: ESTIMATOR IS WLSMV;
MODEL:
      y1 By x1@1;
      y2 By x2@1;
      y3 By x3@1;
      y4 By x4@1;
      factor BY y1* (L1) y2 (L2) y3 (L3) y4 (L4);
      factor@1;
      [factor@0];
      {x1-x4@1};
      [y1-y4@0];
      y1-y4@0;
      [x1$1-x1$4*](T1-T4);
      [x2$1-x2$4*](T5-T8);
      [x3$1-x3$4*](T9-T12);
      [x4$1-x4$4*](T13-T16);
MODEL female:
      factor BY y1* (L1) y2 (L2) y3 (L3) y4 (L4);
      factor*;
      [factor@0];
      {x1-x4*};
      [y1-y4*];
      y1-y4@0;
      [x1$1-x1$4*](T1-T4);
      [x2$1-x2$4*](T5-T8);
      [x3$1-x3$4*](T9-T12);
      [x4$1-x4$4*](T13-T16);

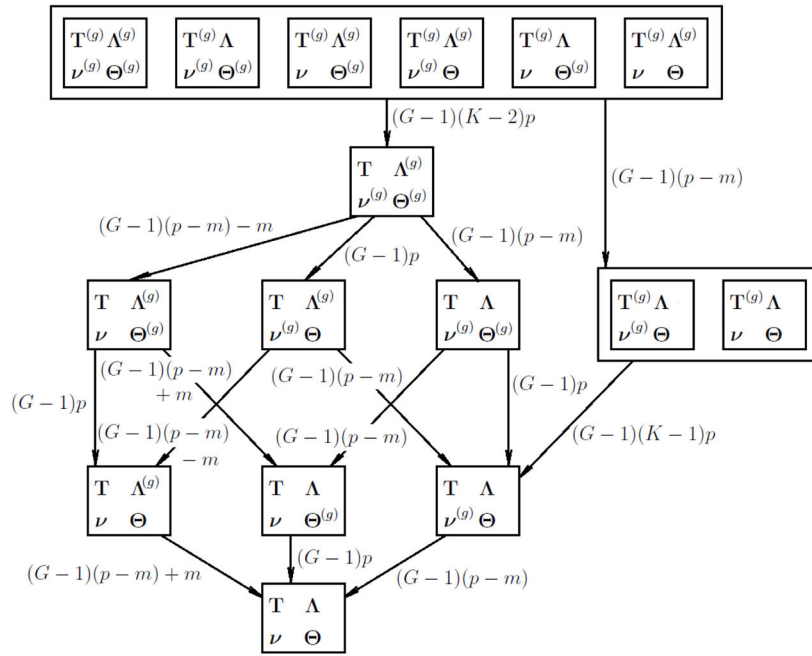
```

### A.2. Model specification in OpenMx

```

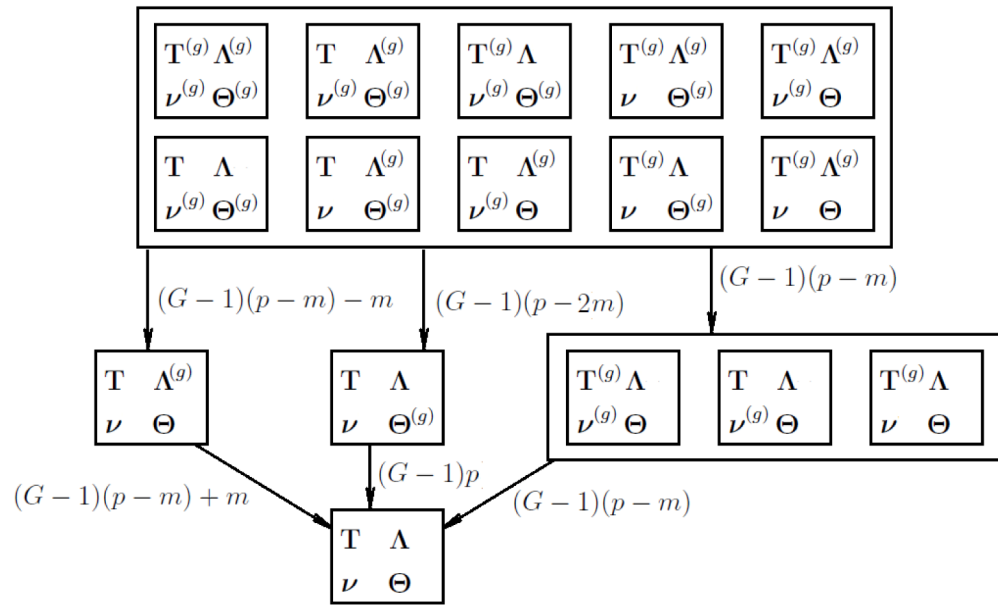
# A model with invariant loadings and unique variances of two groups,
# one common factor, four indicators and two categories per indicator
library(OpenMx)
group1data <- read.table("myGroup1Data.txt");
group2data <- read.table("myGroup2Data.txt");
items <- 4;
label.tau<-paste("t",1:items,sep="");
# specify components
tau<-mxMatrix("Full",1,items,TRUE,c(-.5,0,.5,1),label.tau, name="tau");
L<-mxMatrix("Full",1,items,TRUE,0.8,paste("l",1:items,sep=""),
lbound=0.001, name="L");
nu<-mxMatrix("Full", 1,items,FALSE, 0,paste("n",1:items,sep=""),name="nu");
D<-mxMatrix("Diag",items,items,FALSE,1,paste("d",1:items,sep=""),name="D");
Phi<-mxMatrix("Symm",1,1,FALSE,1,"facVar",name="Phi");
alpha<-mxMatrix("Full",1,1,FALSE,0,"a",name="alpha");
Theta<-mxAlgebra(D-vec2diag(diag2vec(t(L)**Phi**L)),name="Theta");
Sigma<-mxAlgebra(t(L)**Phi**L+Theta,name="Sigma");
M<-mxAlgebra(alpha**L+nu,name="M");
# specify group 1 and then group 2 by altering group 1.
group1 <- mxModel(L, Phi, D, Theta, Sigma, nu, alpha, M, tau,
  mxExpectationNormal("Sigma", "M", paste("X", 1:items,
sep=""),"tau"),mxFitFunctionML())
group2 <- omxSetParameters(group1, labels=label.tau,
newlabels=paste("t", 1:items, "2", sep=""),free=TRUE, name="G2");
group2<-omxSetParameters(group2, labels="facVar",newlabels="facVar2",
free=TRUE);
Theta2<-mxAlgebra(G1.Theta, name="Theta");
group2$Theta<-Theta2;
group2$D<-NULL;
# add data, combine two groups and run model
G1<-mxModel(group1,mxData(group1data, "raw"), name="G1");
G2<-mxModel(group2,mxData(group2data, "raw"), name="G2");
multiGroup <- mxModel("multiGroupModel", G1, G2,
  mxAlgebra(G1.fitfunction + G2.fitfunction,
name="obj"),mxFitFunctionAlgebra("obj"));
mgRes <- mxRun(multiGroup)
summary(mgRes)

```

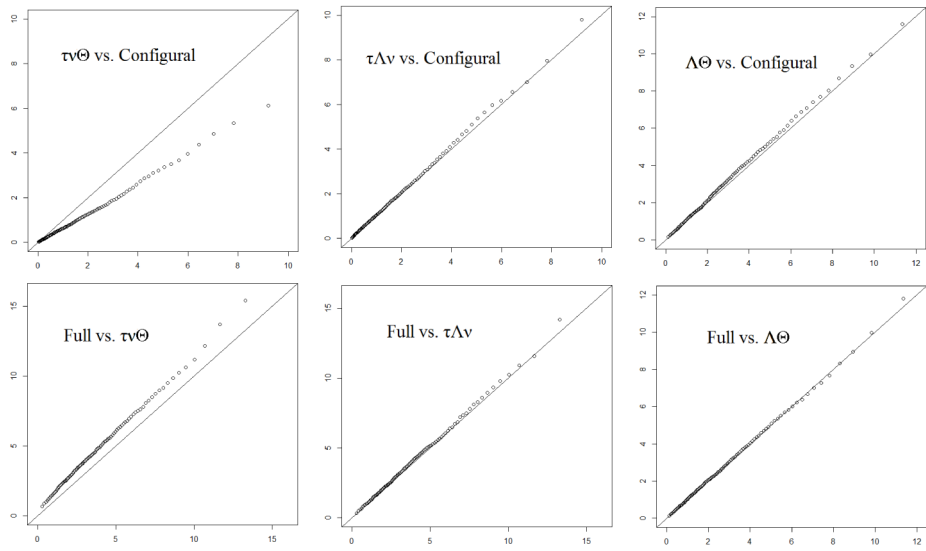


**Figure 1.** Nesting relations among models for polytomous items. Models are labeled by its parameters, where a superscript  $(g)$  denotes different parameters for different groups. Models in the same box are statistically equivalent. An arrow points from a less restricted model to a more restricted model in a nesting relationship. Expressions on the arrows show the difference in model dimensions.

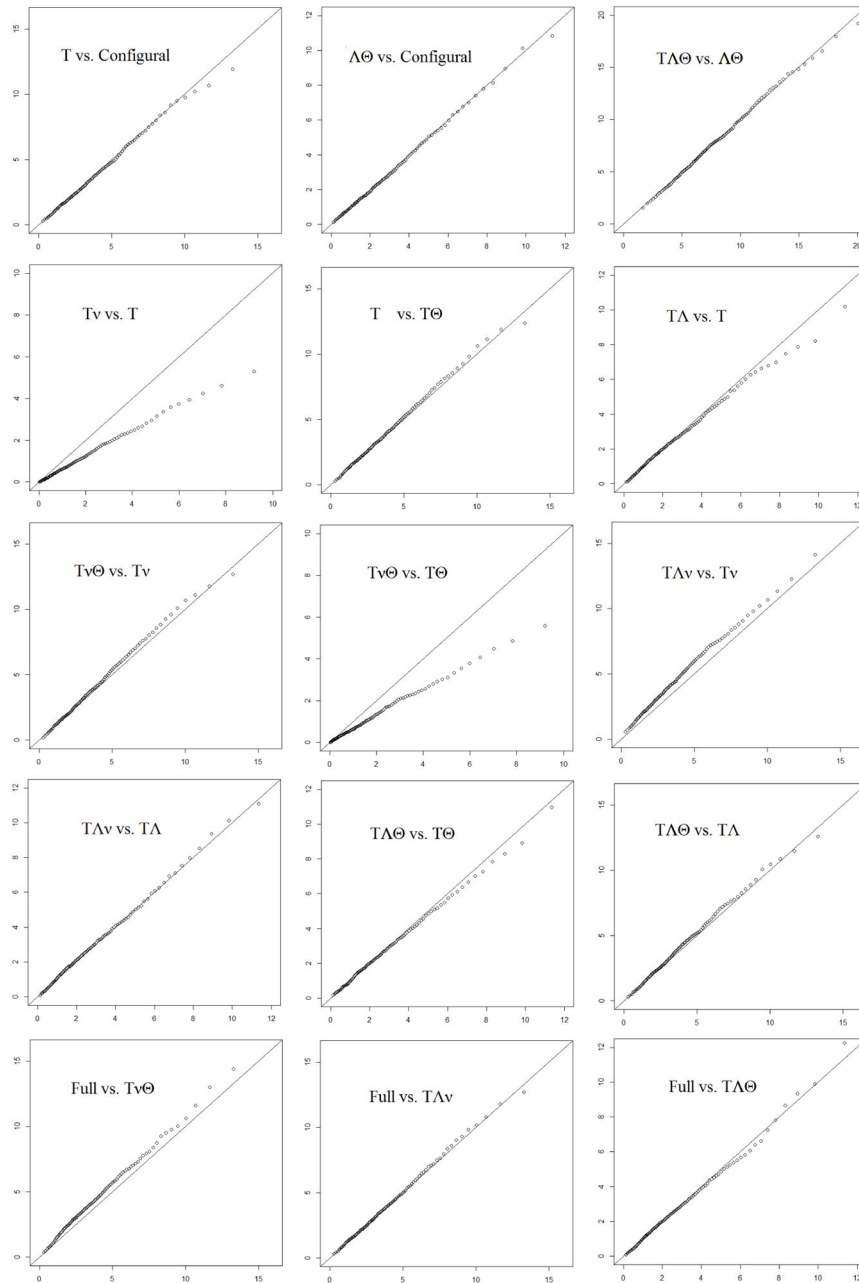




**Figure 2.** Nesting relations among models for dichotomous items. The structure is similar to Figure 1.



**Figure 3.** Plots of the 1st-99th percentiles of the simulated distribution of the LRT statistic against those of a  $\chi^2$  distribution with appropriate dfs. Data are dichotomous. Each panel corresponds to a difference test represented by an arrow in Figure 2. Greek letters denote invariant parameters in a model.



**Figure 4.** Plots of the 1st-99th percentiles of the simulated distribution of the LRT statistic against those of a  $\chi^2$  distribution with appropriate dfs. Data are in four categories. Each panel corresponds to a difference test represented by an arrow in Figure 1. In each panel, Greek letters denote invariant parameters in a model.

Parameter estimates of a two-group model for four binary items of one factor identified in different ways.

**Table 1**

	True Value		$\delta$ -1		MYT		1 MYT 2	
	G 1	G 2	G 1	G 2	G 1	G 2	G 1	G 2
Loadings	0.8	0.8	0.8*	0.8*	0.8*	0.8*	0.80	0.83
	0.8	0.80	0.72	0.80	2.47	0.8*		
	0.8	0.78	0.68	0.78	0.33	0.78	0.83	
	0.8	0.78	0.64	0.78	0.48	0.79	0.85	
Intercepts	0	0*	0*	0*	0*	0*	0*	0*
	0.36	0.18			***		0.19	
Unique	0.36	0.36	***	***	6.41	***	***	***
Variances	0.36	0.54			0.16		0.51	
	0.36	0.72			0.44		0.70	
Thresholds	-0.75	-0.76	-0.83	-0.76	-0.76	-0.76	-0.76	-0.76
	-0.25	-0.25	-0.25	-0.25	-0.25	-0.25	-0.25	-0.25
	0.25	0.25	0.24	0.25	0.25	0.25	0.25	0.25
	0.75	0.75	0.64	0.75	0.75	0.75	0.75	0.75
Factor M.	0	0*	0*	0*	0.32	0*	0*	-0.01
Factor Var.	1	1.03	1.21	1.03	1.80	1.02	1.02	0.93

\* fixed value;

\*\*\* fixed at one less the communality of that item.

Fit statistics for the models in Table 1. The configural invariance model corresponds to the models and parameter estimates presented in Table 1. Models with weak factorial invariance further constrain factor loadings to invariance across groups, yielding 3 more degrees of freedom.

**Table 2**

Identification		Model Fit			
Invariance	Method	$-2lnL$	$\chi^2$	df	p
Configural	All three	84592.24			
	$\delta$ -1	84646.36	54.11	3	<.001
Weak	MYT-1	84627.39	35.15	3	<.001
	MYT-2	84593.66	1.41	3	.70

**Table 3**

Model parameter counts ( $K > 1$ ). The 1st row counts all parameters. The 2nd through 4th rows list three parameterizations of a baseline model. The remaining rows list models whose invariant parameters are noted in Column 1. For all rows,  $\frac{1}{2}m(m-1)G$  factor covariances are not listed but counted. Negative numbers are the number of constraints on  $\text{diag}(\Sigma)$ . The entry with \*\*\* is  $pKG - (G-1)(p+m)$ .

	T	$\Lambda$	$\nu$	$\text{diag}(\Theta)$	$\kappa$	$\text{diag}(\Phi)$	$\text{diag}(\Sigma)$	Total
Total	$pKG$	$pG$	$pG$	$pG$	$mG$	$mG$	0	$G(p(K+3) + m(m+3)/2)$
Condition 7	$pKG$	$pG$	0	$pG$	0	0	$-pG$	
Condition 8	$pKG$	$pG$	0	0	0	0	0	$G(p(K+1) + m(m-1)/2)$
MYT	***	$(p-m)G$	0	$pG$	$m(G-1)$	$mG$	$-p$	
T	$pK$	$pG$	$p(G-1)$	$pG$	0	0	$-p$	$p(K-2) + G(3p + m(m-1)/2)$
$\Lambda$	$pKG$	0	0	$pG$	0	0	0	same as baseline
$\nu$ or $\Theta$				same as Condition 8				same as baseline
T & $\Lambda$	$pK$	$p$	$p(G-1)$	$pG$	0	$m(G-1)$	$-p$	$G(2p + m(m+1)/2) + p(K-1) - m$
T & $\nu$	$pK$	$pG$	0	$pG$	$mG$	0	$-p$	$G(2p + m(m+1)/2) + p(K-1)$
T & $\Theta$	$pK$	$pG$	$p(G-1)$	$p$	0	0	$-p$	$G(2p + m(m-1)/2) + p(K-1)$
$\Lambda$ & $\nu$			same as loading invariance					same as baseline
$\Lambda$ & $\Theta$	$pKG$	$p$	0	$p$	0	$m(G-1)$	$-p$	$G(pK + m(m+1)/2) + p - m$
$\nu$ & $\Theta$			same as Condition 8					same as baseline
T, $\Lambda$ & $\nu$	$pK$	$p$	0	$pG$	$m(G-1)$	$m(G-1)$	$-p$	$G(p + m(m+3)/2) + pK - 2m$
T, $\Lambda$ & $\Theta$	$pK$	$p$	$p(G-1)$	$p$	0	$m(G-1)$	$-p$	$G(p + m(m+1)/2) + pK - m$
T, $\nu$ & $\Theta$	$pK$	$pG$	0	$p$	$mG$	0	$-p$	$G(p + m(m+1)/2) + pK$
$\Lambda$ , $\nu$ & $\Theta$			same as loading and unique variance invariance					
all four	$pK$	$p$	0	$p$	$m(G-1)$	$m(G-1)$	$-p$	$G(m(m+3)/2 + p(K+1) - 2m)$

**Table 4**

Parameter count for models with different identification conditions ( $K = 1$ ). The structure is the same to Table 3.

	<b>T</b>	<b><math>\Lambda</math></b>	<b><math>\nu</math></b>	<b>diag(<math>\Theta</math>)</b>	<b><math>\kappa</math></b>	<b>diag(<math>\Phi</math>)</b>	<b>diag(<math>\Sigma</math>)</b>	<b>total</b>
Total	$pG$	$pG$	$pG$	$pG$	$mG$	$mG$	0	$G(4p + m(m+3)/2)$
Condition 7	$pG$	$pG$	0	$pG$	0	0	$-pG$	
Condition 8	$pG$	$pG$	0	0	0	0	0	$G(2p + m(m-1)/2)$
MYT	$p$	$(p-m)G$	0	$pG$	$m(G-1)$	$mG$	$-p - m(G-1)$	
<b>T</b>	0	$pG$	$pG$	$pG$	0	0	$-pG$	same as baseline
<b><math>\Lambda</math></b>	$pG$	0	0	$pG$	0	0	0	same as baseline
<b><math>\nu</math> or <math>\Theta</math></b>				same as Condition 8				same as baseline
<b>T &amp; <math>\Lambda</math></b>	0	0	$pG$	$pG$	0	0	0	same as baseline
<b>T &amp; <math>\nu</math></b>				same as MYT				same as baseline
<b>T &amp; <math>\Theta</math></b>	$p$	$pG$	$p(G-1)$	$p$	0	0	$-p$	same as baseline
<b><math>\Lambda</math> &amp; <math>\nu</math></b>				same as loading invariance				same as baseline
<b><math>\Lambda</math> &amp; <math>\Theta</math></b>	$pG$	$p$	0	$p$	0	$m(G-1)$	$-p$	$G(p + m(m+1)/2) + p - m$
<b><math>\nu</math> &amp; <math>\Theta</math></b>				same as Condition 8				same as baseline
<b>T, <math>\Lambda</math> &amp; <math>\nu</math></b>	$p$	$p$	0	$pG$	$m(G-1)$	$m(G-1)$	$-p$	$G(p + m(m+3)/2) + p - 2m$
<b>T, <math>\Lambda</math> &amp; <math>\Theta</math></b>	$p$	$p$	$p(G-1)$	$p$	0	$m(G-1)$	$-p$	$G(p + m(m+1)/2) + p - m$
<b>T, <math>\nu</math> &amp; <math>\Theta</math></b>	$p$	$pG$	0	$p$	$mG$	0	$-p$	$G(p + m(m+1)/2) + p$
<b><math>\Lambda, \nu</math> &amp; <math>\Theta</math></b>				same as loading and unique variance invariance				
all four	$p$	$p$	0	$p$	$m(G-1)$	$m(G-1)$	$-p$	$Gm(m+3)/2 + 2p - 2m$

**Table 5**

Four sets of parameters that produce the same model implied probabilities for response patterns of four ternary items in two groups. For all four cases, the latent factor is standardized. The first set gives standardized parameters. The second set has loading and intercept invariance. The third set has invariant unique variances and intercepts. The fourth set satisfies threshold invariance.

		Thresholds				Loadings				Intercepts				Residual var.				diag( $\Sigma$ )			
Set I	G1	1	0	1/2	-1/2	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	0	0	0	0	$\frac{3}{4}$	$\frac{3}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	1	1	1	1
		0	1	1	3/2																
	G2	-3/2	-1	0	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$	0	0	0	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{2}{3}$	$\frac{2}{3}$	1	1	1	1
		1/2	-1/2	1	2																
Set II	G1	-2	0	$\sqrt{2}/2$	$-\sqrt{2}/2$	1	1	1	1	0	0	0	0	3	3	1	1	4	4	2	2
		0	2	$\sqrt{2}$	$3\sqrt{2}/2$																
	G2	$-3/\sqrt{3}$	$-2/\sqrt{3}$	0	$\sqrt{3}$	1	1	1	1	0	0	0	0	$\frac{1}{3}$	$\frac{1}{3}$	2	2	$\frac{4}{3}$	$\frac{4}{3}$	3	3
		$1/\sqrt{3}$	$-1/\sqrt{3}$	$\sqrt{3}$	$2\sqrt{3}$																
Set III	G1	$-2/\sqrt{3}$	0	$\sqrt{2}/2$	$-\sqrt{2}/2$	$\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$	1	1	0	0	0	0	1	1	1	1	$\frac{4}{3}$	$\frac{4}{3}$	2	2
		0	$2/\sqrt{3}$	$\sqrt{2}$	$3\sqrt{2}/2$																
	G2	-3	-2	0	$\sqrt{3/2}$	$\sqrt{3}$	$\sqrt{3}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	0	0	0	0	1	1	1	1	4	4	$\frac{3}{2}$	$\frac{3}{2}$
		1	-1	$\sqrt{3/2}$	$\sqrt{6}$																
Set IV	G1	0	0	0	0	$\frac{1}{2}$	$\frac{1}{2}$	$\sqrt{2}$	$\frac{\sqrt{2}}{4}$	1	0	-1	$\frac{1}{4}$	$\frac{3}{4}$	$\frac{3}{4}$	2	$\frac{1}{8}$	1	1	4	$\frac{1}{4}$
		1	1	1	1																