A METHOD FOR DETECTING MEASUREMENT INVARIANCE IN THE LOG-LINEAR COGNITIVE DIAGNOSIS MODEL

by

SHERRILL AMANDA BRAMLETT

(Under the Direction of Allan S. Cohen)

ABSTRACT

The log-linear cognitive diagnosis model (LCDM) is a model comprised of categorical latent attributes said to represent specific constructs. It provides a method by which classification on a given attribute/trait can be statistically deduced from an individual's observed response pattern. Item bias or differential item function (DIF) represents the occasion when an item on an assessment produces disparate results for individuals possessing the same level of a particular trait or ability. In this dissertation, the LCDM is applied to a simulated dataset comprised of 12 items and measuring 3 attributes and an academic assessment dataset comprised of 3 items in which each item measures a single attribute. Using MPlus for analysis, an omnibus test for measurement invariance is implemented. The free baseline estimation approach is applied to the item and structural models. This baseline model is then compared to other, more constrained models. Model fit, item parameter estimates, structural model estimates, and impact of DIF or non-invariance are assessed. Results of the empirical study indicated that the estimation tool, MPlus, was challenged by the complexity of the simulated data. For this reason, estimation errors were common, and the omnibus approach was not entirely effective in identifying DIF

with these data. Overall, it was proven that greater instances of DIF in the items will produce differences or instability in the structural model. The omnibus testing method was most successful when applied to the simpler data structure of the academic assessment. These data were found to have invariant items but lacked invariance in the structural model.
INDEX WORDS: Measurement invariance, differential item function, log-linear cognitive diagnosis model, free baseline, MPlus

A METHOD FOR DETECTING MEASUREMENT INVARIANCE IN THE LOG-LINEAR COGNITIVE DIAGNOSIS MODEL

by

SHERRILL AMANDA BRAMLETT

B.S.F.C.S., University of Georgia, 2003

M.S. University of Virginia, 2005

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2018

© 2018

SHERRILL AMANDA BRAMLETT

All Rights Reserved

A METHOD FOR DETECTING MEASUREMENT INVARIANCE IN THE

LOG-LINEAR COGNITIVE DIAGNOSIS MODEL

by

SHERRILL AMANDA BRAMLETT

Approved:

Major Professor: Allan S. Cohen

Committee: Jonathan Templin
Seock-Ho Kim
Bonnie Cramond

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2018

DEDICATION

I dedicate this work to my mom, Sherry Bramlett, who has supported me every step of the way. She is my foundation, and I wouldn’t have made it this far without her love and encouragement.

ACKNOWLEDGEMENTS

While all dissertations take a healthy amount of time to complete, I think I managed to squeeze every minute out of the time allotted to me. There were times when I wondered what it would feel like to give up, but I knew I could never live with myself if that were the outcome. While it may have been a marathon to get here, it has most certainly been a sprint to the finish; a sprint that would not have been possible without the love and support of a few key individuals. First, Kerri Oransky, who made me face my fear of failure and confront the insecurities I had associated with re-learning my work after so much time away. I wouldn't have gotten this far without your encouragement to be brave and push through the discomfort of asking the "dumb" questions. Next, my dear friend Tonia Dousay who understood where my journey began and jumped in to help me cross this finish line. She was in my corner in a way that no one else could be. She is a rock for so many, and she gave me the encouragement I needed to press onward. I'm also incredibly thankful for my colleagues and mentors at Delta Air Lines. Without their support and the benefit of travel, I doubt I could have reached the finish line. And last, but certainly not least, I have to thank Jonathan Templin. We've probably set some kind of student/teacher record for longevity with this dissertation. Regardless, I'm glad you stuck with me. If we had it to do over again, I might pick a different topic, but I wouldn't pick a different mentor. Thanks for all you've done for me as a student and for being a friend as well.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Recent decades have seen a steady increase of psychometricians interested in statistical models containing latent variables intended to provide multidimensional classification of individuals (Rupp & Templin, 2008). Such interest has spurred the development of a body of psychometric models known as diagnostic classification models (or DCMs), which are comprised of categorical latent attributes said to represent specific constructs (e.g., knowledge, abilities, or psychological conditions). In diagnostic assessments, attributes are manifest through the individual's responses to specific items. Thus, an individual's response pattern (i.e., series of correct/incorrect responses) reflects whether that individual has mastered a given attribute. These models provide one method by which such classifications can be statistically deduced from an individual's observed response pattern (Rupp, Templin, & Henson, 2010). A great strength of DCMs lies in their ability to provide such detailed feedback regarding individuals' mastery or non-mastery of fine-grained attributes.

Measurement Invariance

The term *measurement invariance* refers to the consistency of measurement across groups. Invariance testing, like differential item functioning (DIF), is predicated on the assumption that the items of an assessment function in the same manner across relevant groups. Essentially, a test that is designed to measure a given attribute should reveal differences among individuals only if those individuals differ on the attribute(s) purported to be measured by an item and not on some other ancillary dimension(s). We should not find disparate results for people who are identical on the attribute but who might differ on other, less relevant, variables (Millsap, 2011).

Invariance testing looks to establish measurement invariance for both the item parameters and the characteristics of the distribution of examinees. Measurement invariance analyses utilized for multi-group confirmatory factor analysis (CFA) examine a series of testable hypotheses through a set of nested model comparisons. The stepwise imposition of additional model constraints results in a stringent invariance assessment protocol with each step requiring stronger evidence to support invariance. The CFA literature pertaining to measurement invariance tends to endorse the use of the following series of hypothesis tests for model fit to obtain evidence in support of measurement invariance:

1) Configural Invariance establishes that the same latent factor structure is present by fixing the same pattern of parameters (fixed or free) across groups. This model implies that similar, but not identical, latent variables are present in the groups,

2) Weak Factorial Invariance or Metric Invariance constrains the factor loadings to be equal across groups. Invariance in this model implies that the same latent variables are being measured across groups,

3) Strong Factorial Invariance or Scalar Invariance is assessed by specifying that both factor loadings and intercepts be invariant across groups. This constrained model still implies that the same latent variables are being measured across groups, but invariance in the intercepts and the mean structure allows us to evaluate mean differences in the latent variables, and

4) Strict Factorial Invariance or Residual Invariance extends the strong factorial invariance model by placing additional constraints that the unique error variances are invariant across groups. Constraining error variances to be invariant across groups implies that group differences in variances of the measured variables are a result of group differences in variances of the latent variables (MacCallum, 2012).

Cognitive Diagnostic Approach to Invariance Testing

The log-linear cognitive diagnosis model (LCDM) is a general case of DCM that provides a common framework through which all latent class-based DCMs can be expressed (e.g., Henson, Templin, & Willse, 2009). The foundations of classifications made with the LCDM, as with all DCMs, are the observed response data collected via diagnostic assessment. For comparisons to be made between groups of individuals, measurement invariance is essential. As the purpose of DCMs (or, more specifically, the LCDM) is to classify individuals as "masters" or "non-masters" or to render diagnoses, the invariance of items comprising diagnostic assessments is critical. Methods for assessing invariance in DCMs, more specifically the LCDM, have been based on those developed for use in CFA (Bozard, 2010). What has yet to be investigated, and is the focus of this dissertation, is the appropriateness of these methods for use with the LCDM.

To show how an investigation of measurement invariance would work under the LCDM, two other models, item response theory (IRT) and confirmatory factor analysis (CFA) are discussed in this dissertation. Methods of invariance analyses in IRT and CFA are briefly described to provide additional context leading into the discussion of invariance testing under the LCDM.

Overview of Dissertation

In this dissertation, I relate confirmatory factor analysis (CFA) and item response theory (IRT) to the LCDM to demonstrate the unique properties of the LCDM in terms of the best fit of measurement invariance procedures. To that end, I investigate invariance testing with the LCDM. Items that lack measurement invariance can yield biased estimates of examinee ability. Current research has yet to explore the appropriateness of measurement invariance testing methods as they are applied to the LCDM. This is cause for concern as all latent trait models are not created equal, and it is potentially detrimental to assume a one-size-fits-all approach to measurement invariance testing. For these reasons, this dissertation is important to the continued pursuit of invariance testing as it will provide evidence necessary for the continued application (or the reimaging) of invariance testing protocols within the LCDM framework.

The following chapter presents a review of existing research as well as a definition and a detailed discussion of measurement invariance as it pertains to confirmatory factor analysis and item response theory. Chapter 3 details how invariance methods apply to the LCDM and the design of the current study, including methods of estimation and analysis. Chapter 4 discusses the procedure and results of the simulation study. Finally, Chapter 5 details the application of invariance testing to a real data set and makes conclusions regarding the efficacy and applicability of this method to invariance testing for the LCDM.

CHAPTER 2

THEORETICAL BACKGROUND

In this chapter, I discuss the existing research related to measurement invariance analysis, addressing the theoretical aspects of invariance testing under multi-group confirmatory factor analysis (CFA) and item response theory (IRT) including likelihood ratio tests. Connections between these two methods are discussed to provide a theoretical background for the subsequent discussion of invariance testing under the log-linear cognitive diagnosis model (LCDM), which appears in Chapter 3.

Confirmatory Factor Analysis (CFA)

Confirmatory factor analysis (CFA) is a psychometric method in which the measurement model is determined a priori and explicitly specifies both the number of latent factors and the relationship of those factors to the specified indicators or items (Kline, 2005). CFA is characterized by continuous latent variables and is the term applied to latent variable modeling most frequently with continuous (or assumed continuous) data. Although CFA may be a multidimensional model, the basic CFA model involving the examination of a unidimensional trait or factor is discussed to demonstrate how invariance procedures are conducted.

Confirmatory factor analysis models are typically characterized by the following set of assumptions: 1) the data are continuous, each item is normally distributed and may be represented by a regression like function of the latent factor (λ) or factors in the multidimensional case, 2) the latent factors are continuous and normally distributed, 3) the indicated factor(s) are the only things measured by an item, and 4) residuals are assumed to be normally distributed.

Assumptions related to normality are prerequisite for the application of maximum likelihood estimation (MLE). Methods for testing measurement invariance in CFA with normal outcomes extend to CFA for categorical outcomes. In such cases, maximum likelihood estimation is still used; however, the assumed distribution of the outcome changes.

The unidimensional CFA model (2.1) posits that each item may be represented as a linear function of a particular latent variable with a random error term:

$$X_{ij} = \mu_i + \lambda_{ia}\xi_{ja} + \delta_{ij}, \qquad (2.1)$$

where X_{ij} symbolizes the continuous item response to item *i* by person *j*, μ_i represents the intercept term for the item *i*, λ_{ia} symbolizes the factor loading for item *i* on factor a, ξ_{ja} represents the latent variable value for person *j* on attribute a (for a test measuring *A* factors – subsequently called attributes for diagnostic classification models), and the residual or uniqueness of individuals for an item, δ_{ij} where *i* again indicates the item and *j* the individual.

The intercept μ_i is a constant and, conditional on the measured latent traits all having means of zero, is the mean for the item *i*. It can be interpreted as the item difficulty as it is a component of the larger mean structure for each item. The factor loading λ_{ia} can be interpreted as a measure of item discrimination, indicating how much or how little of a latent trait (attribute) is present for an item. Factor loadings are statistical estimates of direct effects, and in a regression context, they may be interpreted as the slope of the regression line indicating the expected change in the score of the item per a one-unit change in the factor. Items with larger factor loadings imply a stronger indication of the attribute over smaller factor loadings. A factor loading of zero would indicate an item does not measure the factor at all. Finally, the positive or negative direction of the residual or uniqueness δ_{ij} represents the difference between the actual predicted level of response to that which was observed directly.

In order for the CFA model to be identified, the scale of the factor (the mean and the variance) must be set. Commonly, the variance of the factor is identified by setting either an item's factor loading to one (one item per each factor measured) or the factor variances must be fixed to one (often referred to as the standardized factor variance identification method). When the factor loading for an item is fixed to equal one, it is then referred to a *marker* item, which allows for the factor variance to be estimated. Fixing the factor variances to one and factor means to zero *standardizes* the factor(s) and eliminates the need for item factor loadings to be constrained for identification. In the following section, procedures for determining item invariance for multi-group CFA are discussed.

Invariance Testing for Multi-group CFA Models

For any test that is intended to measure a specific trait (latent variable) and is used with multiple populations (i.e., genders, racial groups, etc.), a critical assumption is that the scale measures the same trait in all groups. If that assumption holds, then comparisons and analyses of those factor scores are acceptable and provide meaningful interpretations. However, if this assumption is false, then any comparisons and/or analyses will likely yield biased results (MacCallum, 2012). In other words, if an item or scale is shown to be non-invariant across groups, then the trait or attribute purported to be measured is somehow biased or potentially reflecting differences present between groups that are a result of something other than the attribute or factor in question. This is the fundamental case for measurement invariance testing.

Invariance Testing Via Likelihood-Ratio Tests

Likelihood ratio tests (2.2) offer a general framework for the investigation of measurement invariance or item bias. In this approach, the likelihood functions of two proposed models are compared, with one model said to be nested within the other. The first model, M_1 is considered the baseline model and typically allows all parameters except the referent or *marker* variable to be freely estimated. In the constrained model, M_2, the referent remains constrained and an item is studied by imposing additional equality constraints across groups one item at a time (Stark, Chernyshenko, & Drasgove, 2006). Assuming algorithm convergence for both models, one is able to obtain likelihood function values, L_1 and L_2, then the test statistic

$$X_{LR}^2 = -2\ln\left(\frac{L_1}{L_2}\right), \tag{2.2}$$

is chi-square distributed with degrees of freedom equal to the difference in degrees of freedom between the baseline and constrained models (Millsap, 2011). It should be noted that likelihood ratio tests are only appropriate for nested model comparisons.

In CFA, measurement invariance analyses are performed on one set of parameters at a time. At each step in the sequence, we are interested in (1) the fit of that model and (2) the degree of decrease in fit compared to the previous model. At any step, if the model fits well and the decrease in quality of fit is not statistically significant, then the constraints imposed at that step are deemed plausible for the population (MacCallum, 2012). For the purposes of this dissertation, invariance analyses are pursued by first establishing configural invariance as a "prerequisite" for metric (loading), scalar (intercept), and residual invariance. The following sections further detail these invariance procedures.

Configural Invariance

In order to have configural invariance, the overall factor structure must be the same in each group. This means that the groups have the same number of factors and the same pattern of fixed and free factor loadings (and other parameters). Figure 2.1 displays the factor structure for a two group, two factor model with eight items. Each item has a unique relationship with each factor, and this relationship is mirrored in both groups. No constraints are enforced at this stage. Groups are merely investigated for the presence of the same factor structure. If there is a lack of configural invariance, the pattern of factor loadings is not the same for both groups and further comparisons should not be made across groups because the observed variables are indicators of different factors or attributes.

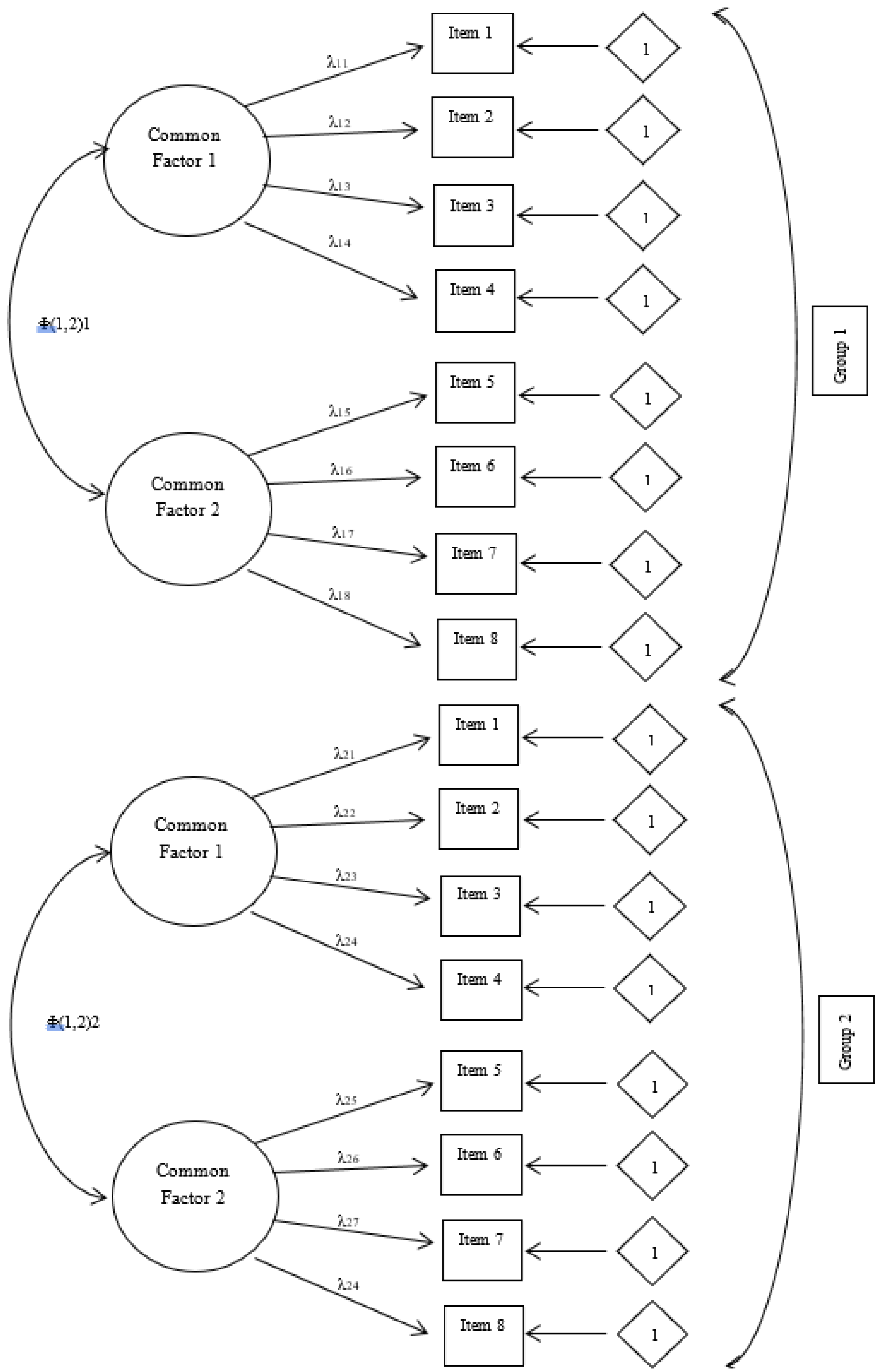

Figure 2.1 *Two group measurement model*

This model implies that similar, but not identical, latent variables are present for the groups in question. Configural invariance allows all of the model parameters to be free and is considered the baseline model within which subsequent models are nested. This baseline model is used for model fit comparison in the next step in the invariance sequence, *metric invariance*. If configural invariance is not supported, then the argument for similar factor structures/patterns across groups does not hold and further analysis is not pursued.

Metric Invariance

If configural invariance is established, the next step in an invariance analysis is to confirm that the loadings for items on each factor are the same across groups. These so-called tests of "metric" or factorial invariance for multi-group CFA are essentially the test of whether the factor loading matrix, $\widehat{\boldsymbol{\Lambda}}_{\boldsymbol{g}}$, is invariant across groups (Reise et al. 1993). The test of the null hypothesis of full measurement invariance for two groups can be expressed formally as

$$H_o: \boldsymbol{\Lambda}_1 = \boldsymbol{\Lambda}_2. \tag{2.3}$$

No additional restrictions are placed on the variances as groups are likely to differ on the latent factors and unique (error) factors (Reise et. al., 1993). For a single item, X_{ig} where *i* indicates the item and *g* indicates the group, the factor loading is in bold (2.4):

$$X_{ig} = \mu_{ig} + \boldsymbol{\lambda}_{ig}\xi + \delta_{ig}. \tag{2.4}$$

At this stage, the factor loadings are constrained to be equal across groups but no other constraints are imposed. Metric invariance testing provides evidence for or against the presence of equivalent common latent factors having identical impact across groups. Differences in factor loadings indicate that the regressions of the measured variables on the factor scores are not parallel across groups. If the regression slopes are varying, group differences in the measurement intercepts are also likely because two regression lines with different slopes (i.e., factor loadings) will ordinarily also have different intercepts (Hoyle, 2012). Item factor loadings shown to be invariant across groups are subsequently tested for scalar invariance while those items that are found to be non-invariant are considered as such for all subsequent analyses and remain freely estimated. The final metric invariant model (i.e., the model for which there is not statistically significant improvement in model fit when additional factor loadings are freely estimated) becomes the *baseline* model for comparisons in the next invariance testing step, *scalar* invariance. Figure 2.2 graphically represents four different invariance scenarios for an item. The upper row displays two graphs in which metric invariance is present. The upper left graph indicates both metric and scalar invariance (discussed in the next section) while the upper right graph represents an item that is metric invariant but NOT scalar invariant across groups. The two graphs on the lower row reflect instances in which an item is metric non-invariant.

Scalar Invariance

Scalar invariance examines item intercepts for equality across groups, implying that the population differences in the means of the measured variables must be due to the influence of the common factors. It should be noted that only the items that have been shown to be metric invariant by the previous analyses are tested for scalar invariance. Scalar invariance testing is done by specifying both item factor loadings and item intercepts to be invariant across groups ($\lambda_{i1} = \lambda_{i2} = \lambda_i$, $\mu_{i1} = \mu_{i2} = \mu_i$). The presence of scalar invariance implies that differences in the means of the observed items are due to differences in the means of the underlying latent factor or attribute. This is because the item mean is partially due to the item and partially due to the factor. However, when item intercepts are non-invariant or allowed to vary across populations, the differences in observable variables may be due to differences in the true values of the associated latent variables or the item intercept. Heterogeneous item intercepts can represent systematic bias on all or part of a scale. These differences across populations are also called *additive response bias*, a situation in which influences not related to the common factor (i.e., social desirability or rater leniency) may result in biased item responses (positive bias or negative bias) in one population as compared to another (Gregorich, 2006).

Using a single item X_{ig}, the part of the model tested for scalar invariance is indicated in bold (2.5):

$$X_{ig} = \boldsymbol{\mu}_{ig} + \lambda_{ig}\xi + \delta_{ig}. \quad (2.5)$$

The two graphs on the left in Figure 2.2 reflect instances in which the item is scalar invariant. The item indicated in the upper left graph is both metric and scalar invariant while the item displayed in the lower left graph is only scalar invariant. In practice, this item would have been eliminated from further invariance analysis during metric invariance testing. The two graphs on the right reflect items that are scalar non-invariant (i.e., for each group, the item has a unique intercept).

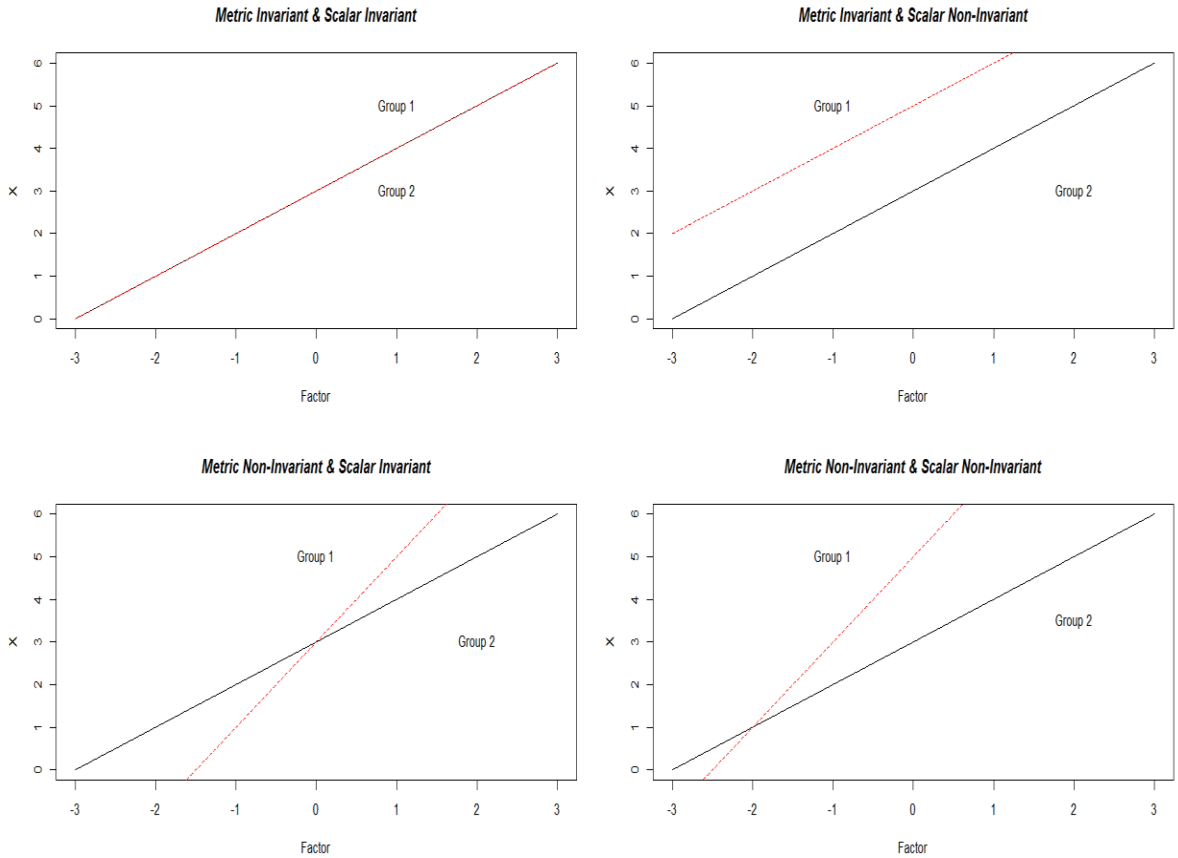

Figure 2.2 *Display of different types of item invariance for two groups under CFA*

Residual Invariance

Residual invariance testing is the final step in the investigation of measurement invariance. It is the strictest form of factorial invariance investigating only those items that are metric and scalar invariant. Testing for residual invariance involves restraining the factor loadings, item intercepts, and residual variances to be equal across groups ($\lambda_{i1} = \lambda_{i2} = \lambda_i, \mu_{i1} = \mu_{i2} = \mu_i, \delta_{i1} = \delta_{i2} = \delta_i$). It answers the questions of whether group differences on the items are solely due to group differences on the latent factors, since the error variances are held invariant across groups (MacCallum, 2012). Using a single item X_{ig}, the part of the model tested for residual invariance is related to the residual noted in bold (2.6):

$$X_{ig} = \mu_{ig} + \lambda_{ig}\xi + \boldsymbol{\delta}_{ig}, \tag{2.6}$$

where the parameter tested is the residual variance. Under the basic CFA model, δ_{ig} is assumed to be normally distributed with a (fixed) zero mean and a unique variance Ψ_{ig}. The unique variance parameters are the terms that are tested for invariance across groups.

The presence of all four types of invariance is considered “strict” measurement invariance and indicates that group differences in the covariances, variances, and means for the indicator variables are solely due to group differences on the latent factors and not due to differences in factor structure (Millsap & Meredith, 2007).

Two additional sets of constraints could be of interest as they pertain to the structural portion of the factor analysis model: 1) Invariance of variances and covariances of the latent variables, and 2) invariance of latent variable means. Both are considered to be extensions of the already “strict” state of residual invariance. In the first case, strict measurement invariance (having loadings, intercepts and residuals constrained) would be extended by constraining the variances and covariances of the latent variables to be equal across groups. Then the difference between the fit of this model and the strict factorial invariance model are compared via a likelihood ratio test. If latent variance/covariance invariance holds, then the implication is that the population covariance matrix is invariant across groups (MacCallum, 2012). In the second case, strict measurement invariance would be extended by constraining the latent factor means to be equal across groups. If latent mean invariance holds, then the implication is that the means of the latent factors are invariant across groups. However, if there is a significant difference between this constrained model and the strict invariance model, then this is evidence that the population means on the latent factors are significantly different across groups suggesting one group may, on average, have more of a trait than another (MacCallum, 2012).

It is important to note that pursuit of these additional forms of invariance may not be worth pursuing if only partial measurement invariance has been obtained (i.e., some but not all item factor loadings have been found to be invariant, Byrne, Shavelson, & Muthen, 1989). If only partial measurement invariance exists for a scale, then any comparison of latent factor means may not be meaningful (MacCallum, 2012). Byrne et al. (1989) gave the requirement that at least two parameters need to be invariant to assure significance of group comparisons.

Item Response Theory (IRT)

The previous section detailed procedures for assessing measurement invariance in CFA via a series of nested model comparisons using likelihood ratio tests. The nested model comparison approach may also be implemented for the detection of *differential item functioning* (DIF) in item response theory (IRT). In this section, the 2-paramter logistic model for IRT is introduced, procedures for detection of DIF via likelihood ratio tests are discussed, and the relationship between CFA and IRT methods is highlighted.

2PL IRT Model

The expansion and development of IRT models over recent decades has resulted in the creation of a wide range of models and measurement tools equipped to yield a plausible account of the relationship between the item and underlying trait (Reise, Widaman, & Pugh, 1993). IRT and CFA models quantify the probability that an individual will provide the correct response to a particular item. Because individuals will certainly possess differing levels of ability, and item difficulty will vary across items, applying an IRT model that best fits the data is essential.

One objective of this chapter is to highlight the shared elements of CFA and IRT. A few basic model assumptions are among these areas of overlap. Unidimensional IRT assumes 1) item responses are unidimensional (only measuring one latent trait) and are locally independent (given ability, a person's item responses are uncorrelated). Like CFA, IRT assumes that the only thing measured by the item is the factor, 2) while CFA assumes that each item is continuous and normally distributed, this is not the case under IRT because items are not continuous. The 2PL model discussed here relies on dichotomous data, 3) unlike CFA which requires that the relationship between items and latent factors be linearly modeled via a regression line, IRT models each item's relationship to the latent factor via a logistic regression line (item characteristic curve), and 4) both CFA and IRT assume a continuous and normally distributed latent factor or factors in the multidimensional case.

For dichotomous items, the probability of a correct response, $(P_{ij}(X_{ij} = 1|\theta_j)$, is a function of the individual's ability, or θ_j. Three unidimensional IRT models are often used to model discrete items having one, two, or three parameters that characterize the relationship between continuous, normally distributed ability and the probability of a correct response to an item X_{ij}, where *i* indicates the item and *j* the individual. For the purposes of this dissertation, the discussion of IRT based analyses will be focused on the two-parameter logistic model because parameterization of the 2PL model is more readily aligned with CFA (and the LCDM) and consequently provides a clear link between the three models and ultimately the invariance methods used for each.

In order to show how CFA and the LCDM may be linked to the 2-PL IRT model, some rewriting of the two-parameter model must be performed. By multiplying the item parameter a_i through, the traditional two-parameter model can be redefined as:

$$P_{ij}\left(X_i = 1 \middle| \theta_j\right) = \frac{\exp\left(a_i(\theta_j - b_i)\right)}{1+\exp\left(a_i(\theta_j - b_i)\right)} = \frac{\exp(-a_i b_i + a_i\theta_j)}{1+\exp(-a_i b_i + a_i\theta_j)}, \tag{2.7}$$

for an item *i*, and person *j* with ability parameter θ_j. The item difficultly parameter, b_i also referred to as the location parameter, indicates the point of inflection or the point on the ability scale for which the probability of correct response is 50% for the examinee. The item discrimination parameter represents the slope of the item characteristic function at the point of inflection. The higher the a_i value, the better the item discriminates between examinees of low and high ability levels. A notational link to CFA and the LCDM may be provided if $-a_i b_i$ is interpreted as the intercept (μ_i in CFA, and, as shown next chapter, $\lambda_{i,0}$ in the LCDM) and a_i is interpreted as the slope (λ_i in CFA and $\lambda_{i,1,(1)}$ in the LCDM). Because this IRT model is implemented with dichotomous data, the link function is necessary to transform these non-normal outcomes into continuous values. In this parameterization, the link function expresses the discrimination parameter as a component of both the intercept and slope which marks an additional distinction between IRT and CFA. In CFA, the mean or intercept is distinct from the loading or slope. The residual or error term is also not present in this model. This is because errors are not estimated but assumed to follow a logistic distribution with a known residual variance. (Errors are still considered independent in IRT.)

In order to extend the previous discussion of the relationship between CFA and IRT to include the LCDM, it is necessary to first address multidimensional item response theory (MIRT). In many cases, a test item measures several abilities rather than a single latent trait. For example, a student answering a mathematics word problem may rely on both reading and mathematics abilities. When this is the case, a multidimensional model should be used in order to adequately estimate item parameters and respondent abilities. In MIRT, the probability of a correct item response is a function of the vector of abilities, $\boldsymbol{\theta}$, rather than a single measure of ability, θ. When success on an item truly depends upon multiple traits, MIRT models may be considered more appropriate than unidimensional models. Additionally, MIRT models can be used to estimate an examinee's ability along several dimensions simultaneously (e.g., along both algebra and geometry scales in a mathematics test) and thus offer the potential to provide enhanced diagnostic information (Finkelman, 2013). The two-parameter multidimensional IRT model (M2PL) yielding the probability that an examinee with ability $\boldsymbol{\theta}$ correctly responds to item *i*. The M2PL may be defined as:

$$P(X_{ij} = |\boldsymbol{\theta_j}) = \frac{exp\,[\boldsymbol{a_i'}\boldsymbol{\theta_j} - a_i b_i]}{1 + exp\,[\boldsymbol{a_i'}\boldsymbol{\theta_j} - a_i b_i]}, \quad (2.8)$$

where $\boldsymbol{a_i'}$ is a vector of discrimination parameters and $a_i b_i$ is the difficulty parameter.

The M2PL model may be re-parameterized to more closely resemble the LCDM:

$$P(X_{ij} = 1|\boldsymbol{\theta_j}) = \frac{\mathrm{e}^{\lambda i0 + \sum \lambda i1a\, \theta_{ja}}}{1 + \mathrm{e}^{\lambda i0 + \sum \lambda i1a\, \theta_{ja}}}. \quad (2.9)$$

Similar to the rewriting of the unidimensional 2-parameter model from equation (2.7), now for the multidimensional case, the parameter $\boldsymbol{a'_i\theta_j}$ may represent a vector of slopes ($\lambda_{i1a}\theta_{ja}$) and the single difficulty parameter b_i may represent the intercept or threshold (λ_{i0}) for item *i*. One additional distinction for this re-parameterized model is the inclusion of the subscripted indicator *a*, signifying which latent trait is being measured.

MIRT models have the potential to provide more detailed information than unidimensional models and offer greater diagnostic information. The following section offers a discussion of invariance testing for MIRT as it relates to that of CFA and ultimately the LCDM.

Measurement Invariance (Differential Item Functioning) in IRT

In IRT, a lack of measurement invariance is referred to as differential functioning. When differential functioning occurs at the item level, it is called *differential item functioning* (DIF). "DIF is defined as a difference in the probability of endorsing an item across comparison groups when the scores are on a common metric" (Stark, Chernyshenko, & Drasgove, 2006, p. 1293). Essentially, an item shows DIF if individuals having the same ability, but from different groups, do not have the same probability of getting the item right (Hambleton et.al., 1991). An assumption inherent in IRT models is that examinees with the same value on one or more latent trait(s), $\boldsymbol{\theta}$, will have the same probability of correct response on any item that is purported to measure $\boldsymbol{\theta}$. When an item exhibits DIF, it suggests that some other variable (either latent or observed) influences the probability of correct response on that item. DIF is of utmost interest in test development and analysis because it relates directly to the accuracy/fairness and validity of an assessment.

There exist a number of methods of DIF detection both parametric and non-parametric some IRT-based and others non-IRT based. One commonly used non-parametric, non-IRT method is the Mantel-Haenszel procedure; the popularity of this method is most likely due to the relative simplicity of its execution (Li, 2008). Parametric methods for IRT include Raju's area measures (Raju, 1990), likelihood-ratio tests (Thissen, Steinberg, & Wainer, 1988), and Lord's chi-square method (Lord, 1980). This discussion emphasizes the application of the likelihood-ratio test method, as a parametric, IRT model-based procedure for measurement invariance analysis because this method can be used for both CFA and IRT and is subsequently proposed for invariance testing procedures in the LCDM.

Under the IRT framework, DIF can be defined as occurring when the item characteristic curves (ICCs) differ between the reference and focal groups. Since the ICC is defined by its item parameters, it follows that one effective method of detecting DIF is to compare the parameters involved in generating the ICCs for these groups. Assuming that the IRT model fits, one should be able to find a simple (often linear) transformation that will put the item parameter estimates on a common scale. This process is known as parameter linkage and is needed when direct comparisons are to be made between item parameter estimates from different groups (Millsap, 2011).

Configural Invariance

Applying likelihood ratio tests in IRT, like in CFA, involves a series of nested model comparisons in which item parameters are calculated and compared across examinee groups. Under CFA, the first step in the investigation of measurement invariance involves the establishment of configural invariance or the same factor structure for both groups (i.e., same items loading onto the same factors across groups). The execution of this step in IRT can be done in two different ways, the *constrained baseline approach* or *the free-baseline approach*. The *constrained baseline approach* involves fixing the group means, item loadings and item thresholds to be equal across groups. The *free-baseline approach* involves the free estimation of all item thresholds and factor loadings less one item identified as the "marker" item. This item loading is constrained to equal 1 and should therefore be reasonably considered to be DIF free (Stark et al., 2006). The *free-baseline* method is considered here as it more closely aligns to procedures implemented in CFA. This MIRT model,

$$P(X_{ij} = 1|\theta_j) = \frac{e^{\lambda i0 + \sum \lambda i1a\, \theta_{ja}}}{1+e^{\lambda i0 + \sum \lambda i1a\, \theta_{ja}}}, \quad (2.10)$$

may serve as the *free-baseline* model discussed above.

Metric Invariance & Scalar Invariance

Under CFA, the second step in measurement invariance testing pertains to the item factor loadings or metric invariance. The third step is that of scalar invariance or invariance of the item intercepts. In IRT the discrimination and location parameters are analogous to the factor loadings and intercepts in CFA analyses. However, in IRT, the metric (factor loading) and scalar (intercept) invariance are examined at the same time. (Stark et al., 2006). The reason simultaneous invariance testing of both intercept and slope for a binary response item is possible becomes clear when we consider the parameterization from Equation (2.8). Both the M2PL threshold and loading contain the discrimination parameter, a_i, as a component, thereby making simultaneous analysis preferable.

The constrained model is proposed for each item, one-by-one, with both the item thresholds (intercepts) and factor loadings simultaneously constrained or held equal. In the following multidimensional IRT model, the intercept and loadings are highlighted in bold:

$$P(X_{ij} = 1|\theta_j) = \frac{e^{\boldsymbol{\lambda i0} + \sum \boldsymbol{\lambda i1a}\, \theta_{ja}}}{1+e^{\boldsymbol{\lambda i0} + \sum \boldsymbol{\lambda i1a}\, \theta_{ja}}}. \quad (2.11)$$

The estimated log-likelihood for a model with *one item* constrained or held equal across groups is then compared to the baseline model in which all loadings for all items are freely estimated. The degree to which model fit is impacted determines whether the item demonstrates non-invariance or DIF. This procedure is performed for all items, and the item for which unique estimation across groups results in the “largest” (over statistical significance) decrease in model fit is identified as non-invariant and remains freely estimated.

If the constraining of items does not result in significant decrease in model fit, then all items are invariant. The upper left graph in Figure 2.3 displays the ICCs for an invariant item administered to two groups. However, if an item has been identified as non-invariant, then the item-by-item comparisons resume with the baseline model now referring to the estimated log-likelihood for the model in which the non-invariant item has been freed. This process continues until the constraining of items no longer significantly decreases model fit.

Simultaneous detection of loading and threshold invariance is the method by which to identify non-invariant items, and Figure 2.3 provides further evidence in support of this process as a reasonable means by which to determine item parameter invariance. Although in CFA, each type of non-invariance plays a unique role with regard to the fit of the linear item function, in IRT, the item characteristic curve (ICC) is influenced in both difficulty and discrimination regardless of which parameter lacks invariance. The upper right graph and two lower graphs reflect non-uniform DIF. The upper right graph presents the ICCs for an item having non-invariant loadings but invariant thresholds (difficulty) across groups. The lower left graph presents the ICCs for an item having invariant loadings but non-invariant thresholds across groups, and the lower right graph represents the ICCs for an item in which both thresholds and loadings are non-invariant. Regardless of the type of DIF, each graph reflecting an item with a non-invariant component displays item ICCs having different intercepts and slopes.

Residual Invariance

The final step for measurement invariance testing in CFA involves the investigation of each item's error variances. As previously discussed, because IRT is implemented with dichotomous data, this step is unnecessary. Residuals or error terms are not estimated because they are assumed to have a logistic distribution with known variance, $\pi^2/3$. Therefore, once items are found invariant in the previous analysis, testing for invariance is concluded.

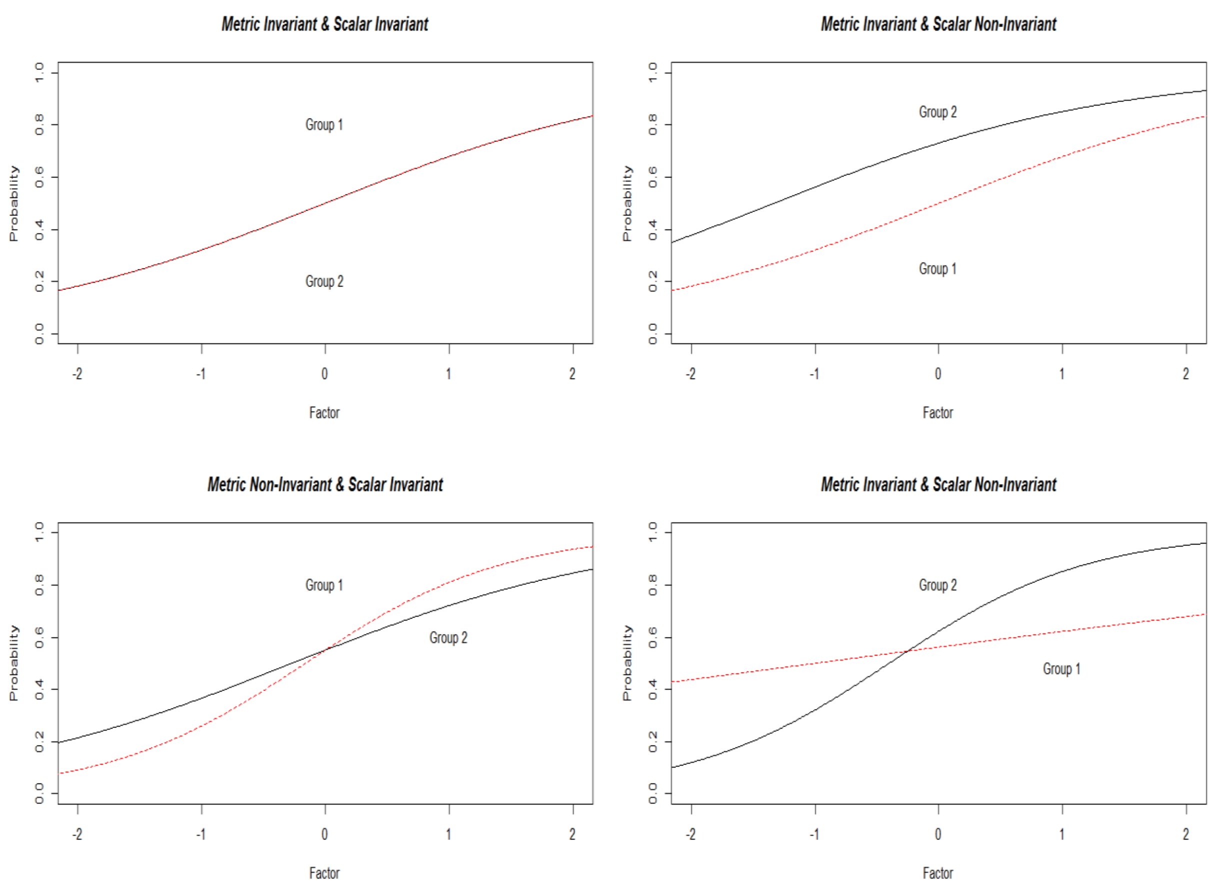

Figure 2.3 *Display of different degrees of item invariance for two groups under IRT*

Summary

The purpose of this chapter has been to provide a theoretical as well as mathematical linkage between item response theory (IRT) and confirmatory factor analysis (CFA). Assuming the presence of a single latent trait or variable, the CFA model supposes that the latent trait is linearly related to each of the items or indicators. That is, individual differences on the latent trait, ξ_{ij}, are linearly related to the individual differences on a given item, $X_{ij},$ and the factor loading, λ_{ij}, is the raw regression coefficient representing this linear relationship. Error variances or residuals for items that are linearly unrelated to the latent variable are represented by the residual indicator, δ_{ij} (Reise et. al., 1993). Additionally, the latent variable, θ, in IRT models represents individual differences in item response, but differences on θ are related monotonically to each item. In IRT models, *a* coefficients relate the latent trait (θ) to the item responses much like the λ_{ij} estimates from CFA. As with λ_{ij} estimates, the larger the *a* coefficient, the more closely the item is related to the latent variable (Reise et. al., 1993). Although there is not a direct connection between item difficulty parameters (b_i), in IRT and CFA, the item intercept μ_i offers the nearest approximation as it contributes to the item mean structure, or average response to an item (MacCallum, 2012).

Table 1 provides a summary of each of the steps described for the measurement invariance procedures as they apply to both item response theory and confirmatory factor analysis. The models are parameterized for each stage of analysis with the parameter of interest at each level of analysis reported in bold.

Table 2.1

Summary of Measurement Invariance Procedures for IRT and CFA

Invariance Steps	**Theoretical Description**	**Item Response Theory (MIRT)**	**Confirmatory Factor Analysis**
Measurement Mode			
Configural Invariance	Baseline fit in which all factors are freely estimated	$P(X_{ij} = 1 \vert \theta_j) = \frac{e^{\lambda i0 + \sum \lambda i1a\, \theta_{ja}}}{1+e^{\lambda i0 + \sum \lambda i1a\, \theta_{ja}}}$	$X_{ig} = \mu_{ig} + \lambda_{ig}\xi + \delta_{ig}$
Metric Invariance	Item factor loadings are fixed to be invariant. This describes intercepts as well as slopes for IRT	$P(X_{ij} = 1 \vert \theta_j) = \frac{e^{\boldsymbol{\lambda i0} + \sum \boldsymbol{\lambda i1a}\, \theta_{ja}}}{1+e^{\boldsymbol{\lambda i0} + \sum \boldsymbol{\lambda i1a}\, \theta_{ja}}}$	$X_{ig} = \mu_{ig} + \boldsymbol{\lambda}_{ig}\xi + \delta_{ig}$
Scalar Invariance	Item intercepts are fixed to be invariant	NA	$X_{ig} = \boldsymbol{\mu}_{ig} + \lambda_{ig}\xi + \delta_{ig}$
Residual Invariance	Item residuals are fixed to be invariant	NA	$X_{ig} = \mu_{ig} + \lambda_{ig}\xi + \boldsymbol{\delta}_{ig}$

CHAPTER 3

MEASUREMENT INVARIANCE & THE LCDM

The previous chapter specified similarities present in the relationship between CFA and IRT and offered a detailed discussion of invariance testing and DIF detection procedures for both models. In this chapter, the purpose of the dissertation is further explicated by first drawing a connection from IRT and CFA to the LCDM. Then, methods of measurement invariance analysis as they pertain to the LCDM are discussed.

Log-linear Cognitive Diagnosis Model

The log-linear cognitive diagnosis model (LCDM), a general Diagnostic Classification Model (DCM), shares several essential characteristics with other latent variable models discussed previously. These include, but are not limited to its multidimensionality, confirmatory nature, and ability to provide diagnostic information concerning individuals' abilities/performances (Rupp & Templin, 2008). Much like multidimensional IRT and multidimensional CFA, the LCDM can be used to assess multiple latent variables, each one representing a unique attribute of an assessment. Also, like IRT, the LCDM assumes each item can be modeled via a monotonically increasing function of the latent attributes.

Fundamental to the LCDM is the Q-matrix, a unique loading structure in which the confirmatory nature of the LCDM is characterized through the representation of the relationship(s) between items and latent variables through a pattern matrix of zeros and ones. The Q-matrix shows which items measure which categorical latent variables (Rupp et al., 2010). The

Q-matrix further reflects the capability of the LCDM to extend beyond the modeling capabilities of other latent variable models with its ability to support complex loading structures.

Under DCMs (and the LCDM), an additional assumption is of the presence of a latent variable or attribute that is split into (typically) two categories. This latent attribute dichotomy is meant to represent “mastery” or “non-mastery” of each measured latent attribute (mastery indicating an individual has a greater amount of the attribute than does an individual in the non-mastery status). Attributes in DCMs typically represent constructs such as content knowledge or psychological conditions, and the classification of individuals is made based on their “mastery” or “non-mastery” of these attributes. It is the work of the LCDM to statistically deduce the relative “mastery’ or “non-mastery” by an individual of an attribute via observed response data (Rupp et al., 2010).

The dichotomous latent variables in the LCDM distinguish it from IRT and CFA which both model continuous latent variables. The LCDM models the relationship of observed categorical response data to categorical latent variables and provides the conditional probability that a given individual’s attribute profile yields an accurate response to an item (Rupp et al., 2010). Person estimates in the form of probabilities exceeding .50 are considered evidence in favor of “mastery” while probabilities lower than .50 are perceived as evidence of “non-mastery” for the latent variable or attribute in question. Probabilities that are near .50 would indicate that the item does not provide sufficient information to make a diagnosis. Item responses are assumed to be independent conditional on the multiple latent variables that are included in the model.

Assuming item responses are conditionally independent given an examinee’s class membership, the latent class model indicates the probability of observing a vector of person *j’s* scored item responses to all items (indicated by $\boldsymbol{x}_j$) as a function of the class membership c of examinee j ($\boldsymbol{\alpha}_j$) as

$$P(\boldsymbol{X}_j = \boldsymbol{x}_j) = \sum_{c=1}^{2^A} \upsilon_c \prod_{i=1}^{I} \pi_{i|\boldsymbol{\alpha}_j}^{x_{ij}} \left(1 - \pi_{i|\boldsymbol{\alpha}_j}\right)^{1-x_{ij}}. \quad (3.1)$$

The parameter υ_c reflects the proportion of examinees belonging to class c (i.e., having the attribute pattern corresponding with class c). The υ_c values are probabilities and sum to one. These parameters also explain the relationship(s) between the attributes (i.e., the correlations) and make up the structural components of the LDCM. The item parameter $\pi_{i|\boldsymbol{\alpha}_j}$ represents the conditional probability that person j provides a correct response for item i given his or her attribute pattern ($\boldsymbol{\alpha}_j$) and x_{ij} indicates the dichotomous item response ($x_{ij} = 0$ or $x_{ij} = 1$) for examinee j to item I (Bradshaw, 2011).

The measurement component of the model is represented by the product term. It specifies the relationship between the observed response data and the latent variable and expresses the joint probability of the observed responses as a product of the conditional probabilities of each item response (Bradshaw, 2011). This is similar to the way items are modeled in IRT (i.e., conditional independence given ability) except that in unidimensional IRT the item response probability is conditional on a continuous latent ability (θ) and in DCMs it is a latent attribute pattern or class.

In sum, monotonicity is a constraint or assumption of both the LCDM and IRT models but is not for CFA (linearity is assumed under CFA). Finally, latent variables or attributes and Q-matrix data are defined as 0/1 (dichotomous) under the LCDM. Although data assumptions in IRT state that items are dichotomous, IRT models continuous latent variables. CFA models both continuous data and continuous attributes or latent factors. Each of the assumptions or constraints of the LCDM mentioned here are employed to ensure model identification (Henson et al., 2009). Next, I introduce the parameterization of the LCDM, highlight this relationship to both IRT and CFA, and present a two-attribute LCDM for the discussion of measurement invariance testing.

Specifying the LCDM

The LCDM can be written as an equation using a logit link function. Representing the LCDM in this manner serves to parallel the structure for the multidimensional IRT model discussed in chapter 2. Like IRT, the link function is used to make model-predicted probabilities between zero and one (Rupp et al., 2010). As in IRT, the LCDM models the log-odds of a correct response conditional on a respondent's attribute pattern $\boldsymbol{\alpha}_j$. The log-odds or logit is:

$$\text{Logit}\left(X_{ij} = 1 \middle| \boldsymbol{\alpha_j}\right) = \ln\left(\frac{P(X_{ij}=1|\boldsymbol{\alpha}_j)}{1-P(X_{ij}=1|\boldsymbol{\alpha}_j)}\right). \quad (3.2)$$

The inverse logit function converts logit values to probabilities and is also the form IRT takes:

$$P\left(X_{ij} = 1 \middle| \boldsymbol{\alpha_j}\right) = \frac{\exp\left(\text{Logit}(X_{ij}=1|\boldsymbol{\alpha}_j)\right)}{1+\exp\left(\text{Logit}(X_{ij}=1|\boldsymbol{\alpha}_j)\right)}. \quad (3.3)$$

However, as also discussed, a major difference between IRT and the LCDM is in the binary instead of continuous latent predictors. Binary indicators specify the presence or absence of the latent predictors or attributes. Effects of individual attributes (main effects) and effects of combinations of attributes (interaction effects) are modeled in the item response (Rupp et al., 2010). The LCDM is specified as

$$\pi_{i|\boldsymbol{\alpha}_j} = P\left(X_{ij} = 1 \middle| \boldsymbol{\alpha_j}\right) = \frac{\exp\left(\lambda_{i\,0}+\boldsymbol{\lambda}_i^T \boldsymbol{h}(\boldsymbol{\alpha_j},\boldsymbol{q_i})\right)}{1+\exp\left(\lambda_{i\,0}+\boldsymbol{\lambda}_i^T \boldsymbol{h}(\boldsymbol{\alpha_j},\boldsymbol{q_i})\right)}, \quad (3.4)$$

where the term $\lambda_{i,0}$ is the intercept that quantifies the log-odds (logit) of a correct response if examinee *j* has not mastered any of the attributes measured by item *i*. This value also indicates the reference group. The term $\boldsymbol{\lambda}_i^T \boldsymbol{h}(\boldsymbol{\alpha_j}, \boldsymbol{q_i})$ is a linear combination of main and interaction effects in the model. The main effects and interactions are given in the row vector, $\boldsymbol{\lambda}_i^T$, where superscript *T* represents the transpose. The term $\boldsymbol{h}(\boldsymbol{\alpha_j}, \boldsymbol{q_i})$ is a column vector of indicators used to specify whether the main effects and interactions are present for the examinee and item. The term $\boldsymbol{q_i} = [q_{i1}, q_{i2}, \dots, q_{iA}]^{T}$ and symbolizes the Q-matrix entries for item *i* and $\boldsymbol{\alpha_j} = [\alpha_{j1}, \alpha_{j2}, \dots, \alpha_{jA}]$ the attribute pattern for examinee *j*. It follows that an element of $\boldsymbol{h}(\boldsymbol{\alpha_j}, \boldsymbol{q_i})$ equals one if and only if (1) the item measures the attribute(s) that correspond to the effect (q_{ia} =1), and (2) the examinee possesses the attribute(s) that correspond to the effect (a_{ja} = 1). Otherwise, the element will equal zero, which will eliminate any main effect or interaction effect parameter that is associated with the unmeasured attribute(s) for the item or for un-mastered attributes in the respondents attribute profile. Via the LCDM, the linear combination $\boldsymbol{\lambda_i^T h}(\boldsymbol{\alpha_j}, \boldsymbol{q_i})$ can be further explicated as:

$$\boldsymbol{\lambda_i^T h}(\boldsymbol{\alpha_j}, \boldsymbol{q_i}) = \sum_{a=1}^{A} \lambda_{i,1(a)}(\alpha_{ja} q_{ia}) + \sum_{a=1}^{A-1} \sum_{b=a+1}^{A} \lambda_{i,2(ab)}(\alpha_{ja} \alpha_{ja} q_{ia} q_{ib}) + \cdots, \qquad (3.5)$$

where $\lambda_{i,1(a)}$ is the main effect for attribute *a* for item *i*, $\lambda_{i,2(ab)}$ is the two-way interaction effect between attributes *a* and *b* for item *i*, and the ellipses suggest the third through A^{th} higher-order interactions where $\lambda_{i,A}$ represents the A-way interaction effect between all possible attributes. As in IRT, this model contains a difficulty parameter, represented by the intercept $\lambda_{i,0}$ and a discrimination parameter in the form of the attribute specific loadings, which are also known as main effects. The intercept for an item provides the probability of correct response for the reference group (i.e., the lowest ability group or the group having no attribute mastery) when the latent trait is zero, and the main effect for an attribute can be interpreted as providing a measure of discrimination between attribute patterns that do or do not have that attribute (Bradshaw, 2011).

Although the connection between IRT and the LCDM has been touched upon, yet to be discussed is the way CFA relates to the LCDM. Simply stated, the parameterization of LCDM attribute effects are like that of the factor loadings found in CFA. The following parameterization of the LCDM provides the probability of correct response for an item that requires two attributes:

$$P\left(X_{ji} = 1 \middle| \boldsymbol{\alpha}\right) = \frac{\exp(\lambda_{i0}+\lambda_{i1}\alpha_1+\lambda_{i2}\alpha_2+\lambda_{i12}\alpha_1\alpha_2)}{1+\exp(\lambda_{i0}+\lambda_{i1}\alpha_1+\lambda_{i2}\alpha_2+\lambda_{i12}\alpha_1\alpha_2)}. \quad (3.6)$$

This model illustrates the similarities between the parameterization of CFA and the LCDM. However, it is important to note that one critical difference lies in the interpretation of these parameters; a point that will be discussed further in the following invariance section. The discussion of measurement invariance testing for the LCDM will reference this model and its parallels to CFA and IRT. There also exists a unique relationship between the LCDM and multi-factor ANOVA. The LCDM can be considered a special case of a generalized linear model, one that models the item response similarly to the multi-factor ANOVA. Taken in this context, the attributes or latent variables of the LCDM can be considered “equivalent” to the factors in ANOVA with the state of “mastery” or “non-mastery” indicating the levels of the ANOVA factor (Rupp et al., 2010).

Assessment of Measurement Invariance with the LCDM

Once mastery probabilities have been assigned to individuals, a major objective of the LCDM, as with any latent variable model, is to accurately make inferences about latent variable means across populations. When multiple groups are involved, we may wish to make comparisons between those groups. Measurement invariance, as it has been discussed for CFA and IRT, allows us to infer that the differences we see in observed variables are in fact due to the differences present for the latent factors or attributes. Often, measurement model parameters are either assumed invariant or their invariance is verified through statistical tests. Invariance tests via nested model comparisons using likelihood ratio tests have already been described for IRT and CFA respectively. What remains to be determined is if this method/procedure is the best approach for use with invariance testing of the LCDM.

The unique set of circumstances provided by the LCDM (e.g., multidimensionality, complex structure) suggest that procedures developed and implemented with CFA may be appropriate tools; however, as previously mentioned, the LCDM data structure is a pattern matrix of binary responses. The binary nature of these data relates more closely to data often investigated via IRT. It is the unique combination of this multidimensionality with binary data that supports the case for further investigation of a best approach for invariance analysis under the LCDM.

Configural Invariance

Invariance testing as it is performed within a multi-group CFA framework, addresses configural invariance first, to establish that the overall factor structure fits across groups. This procedure as it is applied to the LCDM is very similar. A baseline model is established at this point and all parameter estimates are freely estimated while factor means and variances are held constant. This model merely implies that similar, but not identical, latent factors (attributes) are present for the groups.

To determine if this procedure is appropriate for use with the LCDM, a preliminary simulation was performed. Data were generated to represent two groups of 5,000 individuals with scores on a 2-attribute, 10-item measure. The items were simulated to have five items measuring each attribute with three of these items measuring both attributes. As data were simulated to have a known factor/attribute structure, the tests for configural invariance ran as expected and a baseline model having five items loading onto each attribute was established for further investigation.

Initially, data were generated to be entirely invariant across groups to verify that the method for detecting non-invariance was reliable. This means that item intercepts and all item loadings were generated to be equal for both groups. A graph of the probability of correct response to an item for two groups when all item parameters are invariant is shown in Figure 3.1.

Subsequent data simulations introduced varying degrees of measurement non-invariance across parameter types. First, data were simulated to contain non-invariant thresholds or intercepts. A graph of the probability of correct response to an item for two groups when intercepts are non-invariant is displayed in Figure 3.1. Another dataset was generated to contain non-invariant main effects or loadings. A graph displaying the probability of correct response to an item for two groups when the item main effect loading are non-invariant is included in figure 3.1. A fourth dataset was generated to contain non-invariant intercepts and non-invariant main effects simultaneously. Figure 3.1 displays a graph of the probabilities of correct response to an item for two groups both intercepts and an attribute main effect are non-invariant. Non-invariant interaction terms were avoided at this stage.

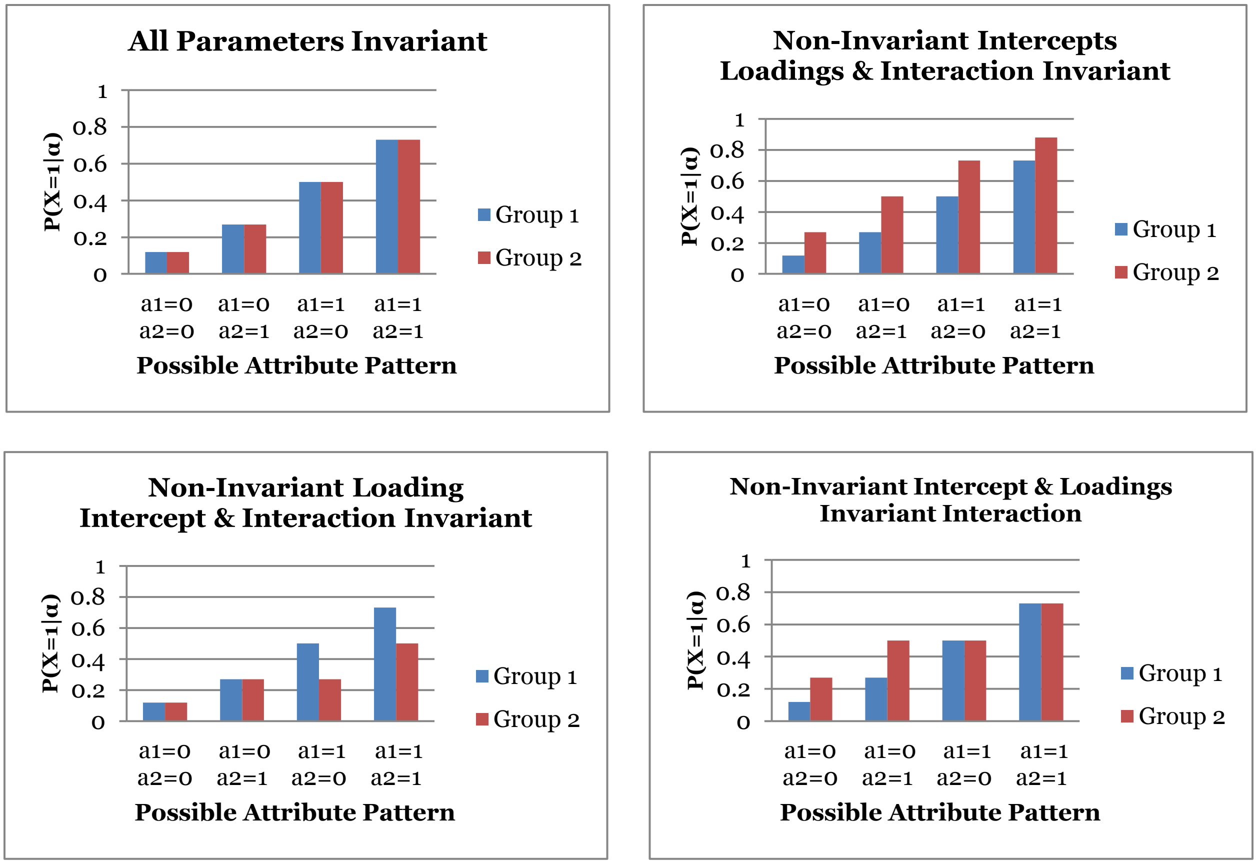

Figure 3.1 *Display of different degrees of invariance for two groups under LCDM*

Metric Invariance

Once configural invariance is established, subsequent invariance tests can be performed to verify item invariance for main effects. This process is analogous to the factorial or metric invariance employed for CFA. The analysis is performed by first holding all main effects for an item to be constant across groups and comparing the resulting model fit. Because these models are nested, fit values may be tested via likelihood ratio tests. If this constrained model yields a poorer fit (which is likely to occur should any parameters be non-invariant), then the item's main effect loadings are freed. At this stage, we expect all items with main effect parameters simulated to be non-invariant, to statistically significantly decrease the overall model fit once they are constrained to be equal. When constraining main effects no longer significantly decreases model fit, this reduced model (with some invariant item loadings held equal) is then used as the new baseline for subsequent scalar or intercept invariance tests.

At this point it is also important to note that this order of approach may be appropriate when only main effects are included in the model but oftentimes interaction terms are present for LCDM data. In these instances, what has yet to be determined is what if any priority should be given to the potential non-invariant interaction terms. It may be pertinent to first investigate the factor loading invariance of interaction terms as a precursor to analyses of "main effects" invariance. The following two-attribute LCDM indicates the parameter of interest for the metric invariance analysis in bold:

$$P\left(X_{ji} = 1 \middle| \boldsymbol{\alpha}\right) = \frac{\exp(\lambda_{i0} + \boldsymbol{\lambda_{i1}}\alpha_1 + \boldsymbol{\lambda_{i2}}\alpha_2 + \lambda_{i12}\alpha_1\alpha_2)}{1 + \exp(\lambda_{i0} + \boldsymbol{\lambda_{i1}}\alpha_1 + \boldsymbol{\lambda_{i2}}\alpha_2 + \lambda_{i12}\alpha_1\alpha_2)}. \tag{3.8}$$

Scalar Invariance

Once factorial invariance has been established, CFA invariance analyses move to the investigation of item intercepts. As previously mentioned in the discussion of CFA invariance testing, this portion of the analysis addressing model intercepts recognizes that these values represent the average for the item whenever the latent variable is equal to zero. However, for the LDCM, intercepts indicate the probability of correct response for the reference group when the latent trait is zero. What this means for invariance testing is that not an item average score, but a score tied to a reference class is being tested. If this is the case, when we hold item intercepts constant, we are asserting that the item has the same value in the reference class across groups. The following two-attribute LCDM model indicates the parameter of interest for this analysis in bold:

$$P\left(X_{ji} = 1 \middle| \boldsymbol{\alpha}\right) = \frac{\exp(\boldsymbol{\lambda_{i0}} + \lambda_{i1}\alpha_1 + \lambda_{i2}\alpha_2 + \lambda_{i12}\alpha_1\alpha_2)}{1 + \exp(\boldsymbol{\lambda_{i0}} + \lambda_{i1}\alpha_1 + \lambda_{i2}\alpha_2 + \lambda_{i12}\alpha_1\alpha_2)}. \tag{3.9}$$

For the pilot simulated data, it was determined that freeing item intercepts or thresholds was a means of successfully recovering items that were simulated to be non-invariant regardless of the type of non-invariance (intercept or main effect). Based on this outcome, it may be worth further investigation to determine which parameters, main effects or intercepts, should be investigated for non-invariance first.

As previously mentioned during the discussion of invariance testing for IRT, the analysis of item residuals does not apply to the LCDM because the distribution of error is assumed. Therefore, the investigation of measurement invariance under this protocol would conclude at this point.

Table 3.1 at the end of this chapter reports the differences in invariance testing of the LCDM model as compared to IRT and CFA models. Each model is parameterized with the parameter of interest at each level indicated in bold.

Figure 3.2 displays graphs of the logit response functions for four different invariance/non-invariance cases for two groups on an item having two attributes. These graphs are like those typically used in ANOVA to investigate interaction terms. Parallel lines mean that a significant interaction term is not present. Conversely, a lack of parallel lines in these graphs would indicate a significant interaction term for the LCDM. The *All Parameters Invariant* graph demonstrates the same logit values for both groups when all parameters are equal, and no interaction is present. For every non-invariant case, the manipulated value was in *Group 2*. The non-invariant intercept in the upper right graph resulted in changes for both attribute 1 and attribute 2 logit values in Group 2 (shifted up one unit). The non-invariant main effect for attribute 1 shown in the lower left graph, increased the logit values by one unit when attribute 1 is present. Finally, the graph of the non-invariant interaction term is reflected in the lower right graph and results in non-parallel logit functions for Group 2.

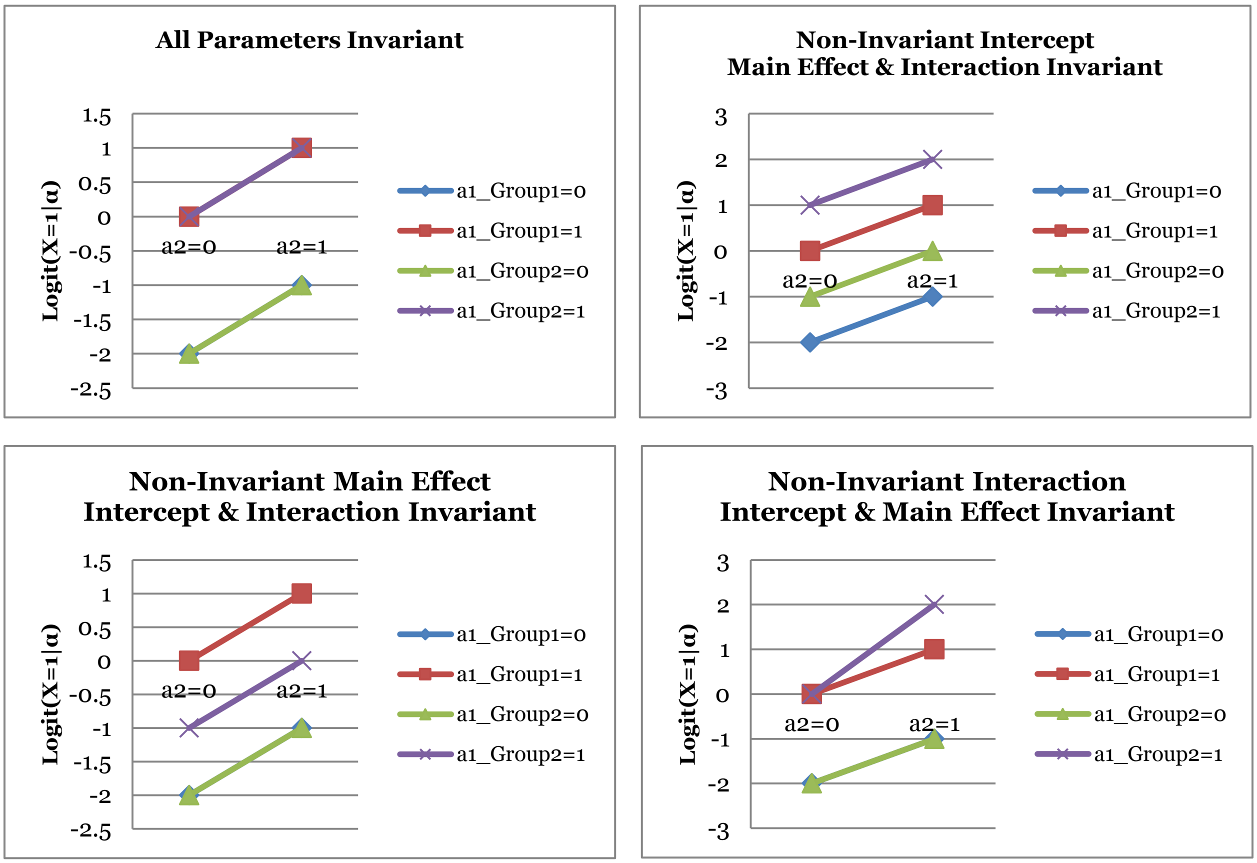

Figure 3.2 *Logit response functions for different degrees of invariance under the LCDM*

The graphs reflect the impact that various sources of non-invariance have on logit values (and subsequent probabilities) in the LCDM.

Study Design

To verify the best approach for establishing measurement invariance for the LCDM, a series of larger scale simulations must be performed. Because an initial simulation or pilot study of non-invariant LCDM data indicated discrepancies between what is considered best practice for CFA invariance testing and what may be necessary for best practice for measurement invariance testing of the LCDM, the proposed simulation is designed to assess the efficacy of measurement invariance procedures using likelihood ratio tests to determine the fit of item parameters via a series nested model comparisons.

Data Generation

The proposed simulation involves generation of data to represent a twelve-item test taken by 2,500 respondents. The test will assess the mastery/non-mastery of three attributes where $\boldsymbol{\alpha}_e = [\alpha_{e1}, \alpha_{e2}, \alpha_{e3}]^T$ with no item measuring more than two attributes simultaneously. The resulting q-matrix will represent six unique item types. A test length of twelve items is adequate for the purposes of this study as it allows for the desired item-type diversity while remaining manageable for the desired number of replications. The balanced Q-matrix shown in Table 3.2 is a pattern matrix containing two of each of the six unique item types resulting in the 12-item test. The eight unique membership classes resulting from this q-matrix are shown in Table 4.

Sample sizes of at least 1,000 subjects are considered large enough for reliable application of likelihood ratio tests for nested model comparisons (e.g., Li, 2008); therefore, a total sample size of 2,500 respondents is used for this study. Two hypothetical groups, each containing 1,250 responses, will be simulated where $G_e = 1$ for group 1 and $G_e = 0$ for group 2.

Two items were simulated to have DIF, one item measuring attribute 1 and the other item measuring both attributes 2 & 3.

The model for an item measuring attribute 1 may be written as follows,

$$logit\big(P(X_{ei} = 1|\alpha_{e1})\big) = \big(\lambda_{i,0} + \lambda_{i,0,g}G_e\big) + \big(\lambda_{i,1,(1)}\alpha_{e1} + \lambda_{i,1,(1),g}\alpha_{e1}G_e\big). \quad (3.10)$$

The model for an item measuring attributes 2 & 3 simultaneously may be written as follows,

$$logit\big(P(X_{ei} = 1|\alpha_{e2}, \alpha_{e3})\big) = \big(\lambda_{i,0} + \lambda_{i,0,g}G_e\big) + \big(\lambda_{i,1,(2)}\alpha_{e2} + \lambda_{i,1,(2)}\alpha_{e2}G_e\big) + \big(\lambda_{i,1,(3),g}\alpha_{e3} + \lambda_{i,1,(3),g}\alpha_{e3}G_e\big) + \big(\lambda_{i,2,(2,3)}\alpha_{e2}\alpha_{e3} + \lambda_{i,2,(2,3),g}\alpha_{e2}\alpha_{e3}G_e\big). \quad (3.11)$$

Item parameters were simulated for each group in the study using a uniform distribution. The distribution for item intercepts in Group 2 will take the following range of values,

$$\lambda_{i,0} \sim U(-2, .5). \quad (3.12)$$

Because data were simulated to represent two groups, and the potential and magnitude for DIF will be allowed to vary, the item intercept differences between Group 2 and Group 1 may take any of the following range of values in a Uniform cumulative distribution function,

$$\lambda_{i,0,g} \sim \begin{cases} 0 & with\ probability \quad .5 \\ U(-1,1) & with\ probability \quad .5 \end{cases}. \quad (3.13)$$

Item main effects for Group 2 were simulated using a uniform distribution taking the following range of values,

$$\lambda_{i,1,(a)} \sim U(.5,2). \quad (3.14)$$

The differences for item main effects between Group2 and Group 1 was also be allowed to vary in magnitude and may take the following range of values under the cumulative distribution function:

$$\lambda_{i,1,(a)} \sim \begin{cases} 0 & with\ probability \quad .5 \\ U(-1,1) & with\ probability \quad .5 \end{cases}. \quad (3.15)$$

Item interactions for Group 2 were simulated using the following uniform distribution,

$$\lambda_{i,2,(a,a')} \sim U(-.75, .75). \quad (3.16)$$

Item interactions differences between Group2 and Group 1 were simulated using the following Uniform cumulative distribution function,

$$\lambda_{i,2,(a,a')} \sim \begin{cases} 0 & with\, probability \quad .5 \\ U(-.5,.5) & with\, probability \quad .5 \end{cases} . \tag{3.17}$$

The structural model below indicates the probability function for data having three attributes. The simulated structural model ensured that attributes had a bivariate tetrachoric correlation around .5, which is within the range typically used in DCM-related studies. Group differences are modeled for attribute three and the interaction between attributes one and two,

$$\log\left(P(\boldsymbol{\alpha}_e)\right) = \gamma_{1,(1)}\alpha_{e1} + \gamma_{1,(2)}\alpha_{e2} + \left(\gamma_{1,(3)}\alpha_{e3} + \gamma_{1,(3),g}\alpha_{e3}G_e\right) + \left(\gamma_{2,(1,2)}\alpha_{e1}\alpha_{e2} + \gamma_{2,(1,2),g}\alpha_{e1}\alpha_{e2}G_e\right) + \gamma_{2,(1,3)}\alpha_{e1}\alpha_{e3} + \gamma_{2,(2,3)}\alpha_{e2}\alpha_{e3}. \tag{3.18}$$

For each replication in the study, the following uniform distribution was used for structural main effects in Group 2,

$$\gamma_{1,(a)} \sim U(-1.1,-0.9). \tag{3.19}$$

The differences between Group 2 and Group 1 for the structural main effects were simulated using the following uniform distribution function,

$$\gamma_{1,(a),g} \sim \begin{cases} 0 & with\, probability \quad .5 \\ U(-.25,.25) & with\, probability \quad .5 \end{cases} . \tag{3.20}$$

The structural interaction for Group 2 was simulated using the following distribution,

$$\gamma_{2,(a,a')} \sim U(0.9,1.1) \tag{3.21}$$

The differences between Group 2 and Group 1 for the structural interaction terms were simulated according to the following distribution function,

$$\gamma_{2,(a,a'),g} \sim \begin{cases} 0 & with\, probability \quad .5 \\ U(-.25,.25) & with\, probability \quad .5 \end{cases} . \tag{3.22}$$

Data were generated using a base R program according to the above specifications and were assessed under four unique comparison conditions via Mplus. These comparisons are designed to identify different instances of DIF as they may occur in the LCDM (e.g. structural and item-level) as well as the efficacy of log-likelihood ratio tests for the identification of DIF.

First, data were analyzed using the *free baseline* approach in which all parameters are allowed to vary across groups. This method should yield the best fit for simulated data and will be the comparison or control set for each of the subsequent models. Data were also analyzed under a fully constrained condition in which both structural and item level parameters are held constant across groups. A third condition for analysis was an item-level constrained case in which the structural parameters were allowed to vary across groups, but item level parameters were held constant. Finally, the fourth analysis condition was a “best case scenario” in which the data were analyzed in accordance with their simulated structure in a partial-item invariant model; items that were simulated to have DIF were allowed to vary across groups while other items were constrained.

The first comparison set was between the fit of the *free baseline* model and the fully constrained model that does not allow structural or item-level parameters to vary across groups. Success is quantified as the proportion of times the likelihood ratio test statistics correctly indicates a poorer fit for the constrained model due to the presence of DIF. The second condition for comparison is the assessment of the change in model fit between the *free baseline* model and a partially constrained model in which the structural parameters are freely estimated, but the item-level parameters are constrained or held equal across groups. Again, success is defined as a statistically significantly poorer likelihood ratio test statistic for instances in which a parameter was simulated to have DIF. The third model comparison is made between the fully estimated model and the partially item invariant model in which only the DIF items are allowed to vary across groups. It is expected that the fully constrained model will always have a poorer fit than the partially constrained model when DIF is present in some form. The question is whether the degradation in fit reaches a statistically significant level. Finally, the impact of DIF at the structural level is analyzed by comparing the fit statistics of the fully constrained model versus a freely estimated structural model.

Power & Type I Error

Power is defined as the proportion of times where the model comparisons correctly identify the presence of DIF (i.e. hit rate) between item parameters in Group 1 and Group 2 when DIF is truly present. For these data, if no DIF is present, but the likelihood ratio tests indicate DIF (i.e. false positive), then Type I error has occurred. However, if there was truly DIF present, but the likelihood ratio tests did not show it, then that counts against power (Type II error).

Because the true model parameters can vary, the magnitude of DIF will vary across items and replications. Therefore, in addition to the identification of Power and Type I error, these data allow for the exploration of the impact effect size or the relative magnitude of DIF has on accurate identification of item parameter differences.

Providing evidence in support of an amended protocol for investigating measurement invariance with the LCDM via simulation study is the primary task addressed by this research. Once complete, the same procedures were applied using an existing dataset. Specifically, the procedure that is identified as most appropriate for the LCDM via the simulation was then implemented with this existing data.

Real World Application: Kansas MTSS DCM Pilot

In early 2018, the Kansas Multi-Tier Systems of Support group (MTSS) conducted a pilot study of a novel assessment system geared toward providing real-time feedback to students regarding their mastery status on a set of state standards. The pilot study was officially commissioned in January of 2018, with assessments being delivered to students during the last week of February and the first week of March. Assessments were delivered by MasteryConnect LLC, an educational technology company.

The purpose of the pilot was to provide a proof-of-concept test for a series of assessments designed to aid MTSS in progress monitoring at-risk students throughout an academic year. Current curriculum-based measures provide information that is either too far removed from the formative process (i.e., a math score that does not provide insight into where students need help) or provide sub-scores that are so highly correlated they likely come from an assessment calibrated using a single unidimensional psychometric model (where any differences in scores likely represent measurement error).

The Sample

Two Kansas school districts volunteered to participate in the pilot: Wichita (Unified School District 259) and Olathe (Unified School District 233). Districts picked the schools where the pilot was to be administered. Olathe had 12 participating schools (eight elementary, two middle, and two high schools). Wichita had nine participating schools (six elementary, one middle, and two high schools).

The assessments of the pilot were engineered to provide multidimensional information for each student rather than a single summary score on a broad content domain. As such, four sets of standards from the 2017 Kansas College and Career Ready Standards were chosen for assessment, two from English Language Arts (Strands ELA1 and ELA2) and two from Mathematics (Strands Math1 and Math2).

To demonstrate DIF methods using the LCDM, this study used Mathematics standard 3.OA.3, which is described as:

3. Use multiplication and division within 100 to solve word problems in situations involving equal groups, arrays, and measurement quantities, *(e.g. by using drawings and equations with a symbol for the unknown number to represent the problem.)*

Items used in the assessments came from either MTSS staff or the Navigate Item Bank, as provided by Certica Solutions. The Navigate Item Bank was initially developed by the Educational Testing Service. It was purchased by NWEA (Northwest Evaluation Association) and used in formative assessments. Items have undergone extensive alignment studies and content checks throughout the existence of the item bank.

All items were delivered by MasteryConnect using either interview, pencil-and-paper, or electronic delivery methods. Assessments at grades Pre-K, K, 1, and 2 were given via interview where proctors read questions aloud and recorded student responses. All items were multiple choice format and administered in the Wichita School District using the electronic MasteryConnect online platform. Assessments administered in the Olathe School District were delivered using pencil-and-paper bubble sheets and scored by MasteryConnect's GradeCam optical character recognition software. Olathe's assessments were administered offline for two reasons: (1) to provide a fail-safe method for obtaining data should any technological issues occur in Wichita, and, (2) due to Olathe's formative assessment schedule having students take assessments from other vendors online at the time of the pilot, thereby limiting bandwidth for online access. Responses were scored correct/incorrect, and all missing responses were scored as incorrect.

In each grade level, two forms per subject area (ELA and Math) were created, called Form A and Form B. Each form was comprised of between three and five standards from each of the two content area strands, for a total of between six to ten standards assessed simultaneously. At the core of each form were the two on-grade standards from each strand, each assessed with five items. Then, depending on grade, up to four off-grade standards, with standards from two other grade levels, above or below being assessed. Pre-Kindergarten forms had standards from K and 1st grade for each strand (6 total), Kindergarten forms had standards from Pre-Kindergarten, 1st, and 2nd grades (8 standards in total), forms for 1st through 10th grade had two standards above/below grade level (10 total), 11th grade forms had standards from 9th, 10th, and 12th grades (8 standards in total), and 12th grade forms had standards from 10th and 11th grade (6 standards in total).

Sample Assessment

Each form contained five items per on-grade standard and two items per off-grade standard (for a total of 22, 24, or 26 items per form, depending on grade). A common item design was employed to link Forms A and B. For each on-grade standard, one item was chosen to appear on both forms. For each off-grade standard a separate item was chosen to appear on both forms. This design ensured standards were linked not just across forms, but also across grade levels through a spiral linking design.

Conclusion

This dissertation discusses a method of invariance testing via likelihood ratio tests of nested model comparisons and specifies how this method is applied to both CFA invariance testing procedures and IRT DIF detection procedures. Likelihood ratio tests have also been proposed as a means of investigating non-invariance in the LCDM, and an outline for how this may be implemented has been described.

A definitive invariance testing protocol for the LCDM has yet to be defined. Therefore, it is a primary purpose of this study to provide evidence in support of specific invariance testing procedures for the LCDM. This study aims to determine the most effective method for identifying non-invariance in the LCDM by exploring different conditions or types of non-invariance as well as differing degrees or effect sizes of non-invariance. Non-invariance may occur in any item parameter type (intercept, main effect, or interaction) or structural parameter (main effect and interaction).

The results of this study will make a novel and important contribution to the LCDM and invariance testing literature. Not only will this investigation result in the provision of a practical guide for invariance testing procedures with the LCDM but it will yield information regarding the power and therefore the utility of this method for detecting differing types and magnitudes of non-invariance.

Table 3.1

Summary of Invariance Procedures for IRT, CFA, and the LCDM

Invariance Steps	**Theoretical Description**	**Item Response Theory (MIRT)**	**Confirmatory Factor Analysis**	**Log-linear Cognitive Diagnostic Model**
Measurement Model				
Configural Invariance	Baseline fit in which all factors are freely estimated	$P\left(X_{ij} = 1 \middle| \theta_j\right) = \frac{\mathrm{e}^{\lambda i0 + \sum \lambda i1a\, \theta_{ja}}}{1 + \mathrm{e}^{\lambda i0 + \sum \lambda i1a\, \theta_{ja}}}$	$X_{ig} = \mu_{ig} + \lambda_{ig}\xi + \delta_{ig}$	$P\left(X_{ji} = 1 \middle| \boldsymbol{\alpha}\right) = \frac{\exp(\lambda_{i0}+\lambda_{i1}\alpha_1+\lambda_{i2}\alpha_2+\lambda_{i12}\alpha_1\alpha_2)}{1+\exp(\lambda_{i0}+\lambda_{i1}\alpha_1+\lambda_{i2}\alpha_2+\lambda_{i12}\alpha_1\alpha_2)}$
Metric Invariance	Item factor loadings are fixed to be invariant. This describes intercepts as well as slopes for IRT	$P\left(X_{ij} = 1 \middle| \theta_j\right) = \frac{\mathrm{e}^{\boldsymbol{\lambda i0} + \sum \boldsymbol{\lambda i1a}\, \theta_{ja}}}{1 + \mathrm{e}^{\boldsymbol{\lambda i0} + \sum \boldsymbol{\lambda i1a}\, \theta_{ja}}}$	$X_{ig} = \mu_{ig} + \boldsymbol{\lambda}_{ig}\xi + \delta_{ig}$	$P\left(X_{ji} = 1 \middle| \boldsymbol{\alpha}\right) = \frac{\exp(\lambda_{i0}+\boldsymbol{\lambda_{i1}}\alpha_1+\boldsymbol{\lambda_{i2}}\alpha_2+\lambda_{i12}\alpha_1\alpha_2)}{1+\exp(\lambda_{i0}+\boldsymbol{\lambda_{i1}}\alpha_1+\boldsymbol{\lambda_{i2}}\alpha_2+\lambda_{i12}\alpha_1\alpha_2)}$
Scalar Invariance	Item intercepts are fixed to be invariant	NA	$X_{ig} = \boldsymbol{\mu}_{ig} + \lambda_{ig}\xi + \delta_{ig}$	$P\left(X_{ji} = 1 \middle| \boldsymbol{\alpha}\right) = \frac{\exp(\boldsymbol{\lambda_{i0}}+\lambda_{i1}\alpha_1+\lambda_{i2}\alpha_2+\lambda_{i12}\alpha_1\alpha_2)}{1+\exp(\boldsymbol{\lambda_{i0}}+\lambda_{i1}\alpha_1+\lambda_{i2}\alpha_2+\lambda_{i12}\alpha_1\alpha_2)}$
Residual Invariance	Item residuals are fixed to be invariant	NA	$X_{ig} = \mu_{ig} + \lambda_{ig}\xi + \boldsymbol{\delta}_{ig}$	NA

Table 3.2
Q-matrix for 3 Attribute, 12 Item Test

Item	Attribute 1	Attribute 2	Attribute 3
1	1	0	0
2	0	1	0
3	0	0	1
4	1	1	0
5	0	1	1
6	1	0	1
7	1	0	0
8	0	1	0
9	0	0	1
10	1	1	0
11	0	1	1
12	1	0	1

Table 3.3
Attribute Mastery by Class Membership for 3 Attribute, 30 Item Test

Class	Attribute 1	Attribute 2	Attribute 3
1	0	0	0
2	0	0	1
3	0	1	0
4	0	1	1
5	1	0	0
6	1	0	1
7	1	1	0
8	1	1	1

CHAPTER 4

SIMULATION STUDY

Chapter 2 provided a review of item response theory, confirmatory factor analysis, and the process by which each of these assesses differential item function (DIF). In Chapter 3, the log-linear cognitive diagnosis model was detailed and the need for a standard DIF detection methodology was highlighted. Chapter 3 also proposed a simulation design and testing protocol for DIF detection under the LCDM.

The results of this analysis are presented in two sections: In the first portion, the results for the simulation study are explored; in the second portion, a real data application is discussed. The simulation study includes a detailed discussion of the Differential Item Functioning (DIF) detection across four scenarios. Then, the impact of DIF effect size as it pertains to Type I error rates and power is discussed. As proposed in Chapter 3, power, Type I error control, and the effect size or magnitude of DIF were evaluated across four separate testing conditions in which each were allowed to vary randomly.

Data were simulated to represent two groups of 1250 examinees. Each observation consisted of a 12-item test measuring three attributes. The four proposed model conditions were successfully simulated and represent the conditions of *free baseline* estimation, fully constrained item and structural parameter estimation, constrained item-level but freely estimated structural parameter estimation, and partially constrained item-level parameter estimation where items simulated to have DIF could vary across groups but the other items and the structural parameters were held equal.

Table 5 contains the specifications under which each test condition was estimated. These constraints were selected so as to limit each parameter to take a realistic set of values for a given condition.

Table 4.1
Simulation Estimation Constraints

	Estimation Constraints					
Level	Intercept		Main Effect		Interaction	
Structural	N/A	N/A	-1.10	-0.90	0.90	1.10
Item	-2.00	-0.50	0.50	2.00	-0.75	0.75

The four proposed estimation methods were implemented and assessed to determine the most effective means of identifying DIF at the item and structural levels. Model 1, or the *free baseline* approach, allowed all item parameters to be freely estimated and served as the control condition. Each of the other models constrained the estimation in some way. Model 2 constrained all item parameters and structural parameters to be equal across groups. Model 3 constrained all item parameters to be equal across groups but allowed structural parameters to be freely estimated. Model 4 constrained all parameters excepting those known to possess simulated DIF. The log-likelihood of the *free baseline* model was used as the benchmark for comparison against each of the subsequent constrained models. Additionally, the log-likelihood of model 3 was compared to the log-likelihood of model 2 to determine if fit was significantly impacted by freely estimating the structural model.

Rather than assessing each item parameter for DIF in a stepwise fashion as is performed in CFA invariance analyses, each parameter was estimated for DIF simultaneously in an omnibus test for model fit.

Table 4.2
Simulation Study Model Comparison Description

Comparison	Explanation
Model1 vs Model2	Fully Estimated model vs Fully Constrained model
Model1 vs Model3	Fully Estimated model vs Structural Model estimated/Items Constrained
Model1 vs Model4	Fully Estimated model vs only DIF elements estimated
Model2 vs Model3	Fully Constrained model vs Structural Model estimated/Items Constrained

Data in model 1 were estimated via the *free-baseline* method which allows all item parameters to be freely estimated. The resulting fit statistic is considered the best fit because no elements have been constrained to equality across groups or conditions. For this reason, the fit of this model is used for comparison against other more constrained models. Data in model 2 were fully constrained in both the item and structural parameters and should represent a significantly poorer fit for models containing simulated DIF. In model 3, the structural model is freely estimated and represents the represents the theoretical condition in which groups differ on the construct being measured but not necessarily on the items that purport to measure the construct. Model 4 represents the ideal case in which we “know” where the DIF exists and allow these item parameters alone to vary. Table 4.3 contains the count of valid observations for each model.

Table 4.3
Parameter Estimate N

Model	N Valid Analyses
Model 1	495
Model 2	898
Model 3	641
Model 4	766

Table 4.4
Simulation Study Model Comparison Results

Comparisons	N Comparisons Type 1	Type 1 p05	Type 1 p01	N Comparisons Type 2	Type 2 p05	Type 2 p01
Model 1 vs. Model 2	5	.400	.200	457	0.123	0.204
Model 1 vs. Model 3	87	.000	.000	275	1.000	1.000
Model 1 vs. Model 4	5	.200	.200	408	0.902	0.980
Model 2 vs. Model 3	141	.879	.766	454	0.101	0.214

Table 4.4 presents the overall model comparison results. Type I comparisons were made for cases in which no DIF was present and no estimation errors occurred. In line one, Model 1 vs Model 2, only five cases met the conditions for Type I error estimation resulting in poor Type I error control ($\alpha > .05$). Type 2 Error/Power analyses included models that contained DIF items but were without estimation errors. For Model 1 vs Model 2, this resulted in a much larger sample size and better Type 2 Error/Power ($1-\beta > .80$). Ideally, the comparison of Model 1 with Model 2 should act as a baseline model in omnibus tests for DIF; however, in this simulation, Model 1 estimation was fraught with errors and failed to converge much of the time. Further, Wald statistics for item parameters with DIF were unobtainable for Model 1.

Model 1 and Model 3 both allowed for the structural parameters to be freely estimated. This model comparison produced very strong Type I error control ($\alpha < .05$). Conversely, Power was extremely low despite sufficient sample size and may reflect the impact of DIF in the structural model on the model fit statistics for Model 3 ($1-\beta = 0$).

Table 4.4, line 3 shows the results for Model 1 vs Model 4 comparisons. These also experienced problems with Type I error due to small sample size ($\alpha > .05$); however, Type II error/Power rates were also poor for these comparisons (*1-β* <.20). It is unclear if this is due to the manner in which the DIF was simulated or the estimation difficulties in MPlus.

In the last line of Table 4.4, For Model 2 vs Model 3 comparisons are evaluated. The Type I error was extremely poor for these tests (*α*>.50) which highlights the impact of a freely estimated structural model. Conversely, Power/Type II error rates were much better for these comparisons (*1-β* >.80).

Table 4.5
Model 1 vs Model 2 Type II Error by Number of DIF Parameters

N DIF Parameters	N Comparisons Type 2	Type 2 p05	Type 2 p01
1	19	.368	.684
2	54	.148	.278
3	106	.142	.189
4	113	.088	.177
5	97	.113	.155
6	52	.058	.154
7	15	.133	.133
8	1	.000	.000

Table 4.5 contains the Type II error rates by the number of DIF parameters in each replication. For Model 1 vs Model 2, as the number of DIF parameters increases, the probability of Type II error decreases. The DIF break-out captures one aspect of effect size for DIF indicating that the greater the DIF, the more powerful the omnibus test to accurately detect it.

The simulated conditions in model 3 allowed for the structural parameters to be freely estimated across groups. Because DIF was simulated to randomly occur for the main effect for attribute 3 as well as the interaction between main effect 1 and 2 at the structural level, the fit of this model was compared to that of the *free baseline* model to determine if DIF at the structural level would negatively impact the overall model fit. This simulation also allowed magnitude of

DIF to vary across observations. Log likelihood comparison tests between model 3 and the *free-baseline* model were performed in order to determine if a statistically significant decline in overall model fit occurred for model 3 cases in which DIF was present at the structural level.

Table 4.6
Model 1 vs Model 3 Error Rates by Number of DIF Item Parameters

N DIF Item Parameters	N Comparisons Type 1	Type 1 p05	Type 1 p01	N Comparisons Type 2	Type 2 p05	Type 2 p01
0	5	.000	.000	7	1.000	1.000
1	10	.000	.000	23	1.000	1.000
2	27	.000	.000	64	1.000	1.000
3	22	.000	.000	94	1.000	1.000
4	19	.000	.000	60	1.000	1.000
5	4	.000	.000	24	1.000	1.000
6	1	.000	.000	3	1.000	1.000

In Table 4.6, the Type I and Type II error rates are reported for the Model 1 vs Model 3 comparison. Type I error represents the case in which no DIF was present for structural parameters, but DIF may have been present in the item parameters. Type II error represents that case in which DIF is present in the structural parameters and also may be present at the item level.

Regardless of the number of item parameters containing DIF, the fully estimated model was consistently preferred to the constrained model (Type I error). Conversely, Type II error rate/Power was poor for this comparison. Because the items in Model 3 were all constrained, when DIF was present, it manifested in the freely estimated structural model and produced significant differences between Model 1 and Model 3 fit statistics. Model 1 was favored in every comparison indicating the impact of constraining DIF items on overall model fit. The lack of power in this case is likely due to small DIF effect size in the structural parameters.

Table 4.7
Model 1 vs Model 4 Type II Error by Number of DIF Parameters

N DIF Parameters	N Comparisons Type 2	Type 2 p05	Type 2 p01
1	17	0.941	1.000
2	46	0.913	0.978
3	97	0.887	0.979
4	100	0.910	0.960
5	89	0.876	0.989
6	47	0.936	1.000
7	11	0.909	1.000
8	1	1.000	1.000

Table 4.7 highlights the Type II error rates by number of DIF parameters for the Model 1 vs Model 4 comparisons. This comparison contended with small sample size issues due to estimation errors in Model 1. Ultimately, these results indicate that the omnibus test does not work well for this type of model comparison, even when the DIF parameters are known and freely estimated as in Model 4. A better test for DIF in Model 4 turned out to be the Wald test for item parameter differences. Difference parameters were calculated in Model 4, and these results were then evaluated using the Wald statistic. Table 4.8 shows the results for the individual DIF parameters in the item and structural models.

Table 4.8
Error Rates: Parameter Estimates with DIF

Model Part	Parameter Type	Item	Affected Attributes	Type1 p05	Type1 p01	Type2 p05	Type2 p01
Structural	Main Effect	N/A	3	.08	.03	.33	.44
Structural	Interaction	N/A	1 & 2	.04	.01	.44	.48
Measurement	Intercept	5	N/A	.02	.00	.32	.41
Measurement	Main Effect	5	2	.02	.01	.42	.45
Measurement	Main Effect	5	3	.03	.01	.41	.45
Measurement	Interaction	5	2 & 3	.04	.02	.49	.51
Measurement	Intercept	7	N/A	.01	.00	.23	.29
Measurement	Main Effect	7	1	.03	.01	.32	.40

When the individual parameters are evaluated for DIF, the Type I and Type II error rates are much better. These results further support the finding that the omnibus test for DIF may not be the best method for DIF detection via MPlus. It is worth noting that DIF effect sizes for interaction terms were the smallest and this appears to manifest in the Type II error rates for these parameters.

Table 4.9
Model 2 vs Model 3 Error Rates by Number of DIF Item Parameters

N DIF Item Parameters	N Comparisons Type 1	Type 1 p05	Type 1 p01	N Comparisons Type 2	Type 2 p05	Type 2 p01
0	8	0.875	0.750	8	.250	.500
1	16	0.813	0.750	39	.051	.282
2	38	0.868	0.737	114	.096	.228
3	34	0.853	0.706	154	.091	.195
4	32	0.938	0.844	91	.143	.231
5	10	1.000	1.000	42	.071	.095
6	3	0.667	0.333	6	.167	.167

Table 4.9 contains the Type I and Type II error rates by DIF parameter for the Model 2 vs Model 3 comparisons. Despite small samples sizes for Type I error, this table captures the impact of freeing the structural model across all levels of DIF. As the number of DIF items increased, the model with more estimated parameters was favored. Model 3 is favored when there is more DIF present in the items which suggests that not controlling for item level DIF results in larger differences in the structural model.

Impact of DIF on Classification Rates

Classification rates and percent reduction in error estimates (Cohen's Kappa) for each of the attributes are shown in Tables 4.10 – 4.15. The classification rates are consistently strong across models and for increasing numbers of DIF parameters. This suggests that the magnitude of simulated DIF was not large because it did not yield poorer classification. Cohen's Kappa statistics show moderate to strong agreement for attribute classification. The only exception to this may be for Model 2. Across attributes, the proportion of correct attribute classification was slightly smaller than for the other models. This is likely due to the high degree of misspecification (e.g. fully constrained model estimation) whenever DIF was present.

Pattern classification rates and Cohen's Kappa are shown in Tables 4.16 and 4.17. While pattern classification was consistent across all models and for each level of DIF, it was lower for Model 2. Cohen's Kappa statistic showed fair to moderate agreement for pattern classifications and showed some decline as instances of DIF increased.

Table 4.10

Attribute 1 Classification Rates by DIF Count

Model	All	All SE	0	SE 0	1	SE 1	2	SE 2	3	SE 3
Model 1	0.792	0.036	0.776	0.032	0.802	0.033	0.800	0.036	0.791	0.034
Model 2	0.785	0.042	0.785	0.036	0.795	0.043	0.794	0.034	0.786	0.040
Model 3	0.789	0.043	0.791	0.031	0.805	0.036	0.794	0.037	0.792	0.039
Model 4	0.793	0.035	0.776	0.008	0.796	0.036	0.801	0.031	0.795	0.033

Model	4	SE 4	5	SE 5	6	SE 6	7	SE 7	8	SE 8
Model 1	0.788	0.039	0.793	0.035	0.788	0.038	0.791	0.037	0.823	NA
Model 2	0.782	0.041	0.788	0.043	0.779	0.049	0.775	0.050	0.810	0.008
Model 3	0.789	0.043	0.792	0.038	0.778	0.055	0.773	0.063	0.754	0.078
Model 4	0.788	0.038	0.797	0.033	0.790	0.039	0.788	0.038	0.772	0.035

Table 4.11
Attribute 1 Kappa Statistic

Model	All	All SE	0	SE 0	1	SE 1	2	SE 2	3	SE 3
Model 1	0.582	0.073	0.551	0.063	0.603	0.068	0.598	0.071	0.581	0.069
Model 2	0.571	0.083	0.570	0.070	0.589	0.085	0.587	0.068	0.572	0.079
Model 3	0.577	0.086	0.581	0.062	0.610	0.073	0.587	0.074	0.582	0.078
Model 4	0.585	0.071	0.553	0.014	0.591	0.073	0.602	0.061	0.588	0.066

Model	4	SE 4	5	SE 5	6	SE 6	7	SE 7	8	SE 8
Model 1	0.574	0.078	0.585	0.070	0.576	0.077	0.583	0.074	0.645	NA
Model 2	0.563	0.081	0.576	0.085	0.557	0.100	0.549	0.101	0.619	0.016
Model 3	0.576	0.087	0.582	0.075	0.555	0.110	0.545	0.125	0.507	0.155
Model 4	0.574	0.078	0.592	0.067	0.579	0.077	0.576	0.076	0.544	0.072

Table 4.12
Attribute 2 Classification Rates by DIF Count

Model	All	All SE	0	SE 0	1	SE 1	2	SE 2	3	SE 3
Model 1	0.794	0.033	0.803	0.029	0.797	0.042	0.792	0.031	0.794	0.033
Model 2	0.789	0.037	0.801	0.020	0.793	0.040	0.788	0.037	0.789	0.037
Model 3	0.794	0.035	0.810	0.023	0.801	0.032	0.793	0.036	0.799	0.031
Model 4	0.796	0.034	0.803	0.027	0.803	0.030	0.793	0.034	0.796	0.037

Model	4	SE 4	5	SE 5	6	SE 6	7	SE 7	8	SE 8
Model 1	0.794	0.033	0.797	0.032	0.787	0.039	0.795	0.031	0.790	NA
Model 2	0.789	0.037	0.788	0.038	0.786	0.036	0.792	0.038	0.792	0.018
Model 3	0.794	0.034	0.792	0.037	0.789	0.036	0.786	0.040	0.801	0.001
Model 4	0.795	0.033	0.796	0.033	0.794	0.033	0.800	0.035	0.788	0.003

Table 4.13
Attribute 2 Kappa Statistic

Model	All	All SE	0	SE 0	1	SE 1	2	SE 2	3	SE 3
Model 1	0.586	0.067	0.606	0.059	0.593	0.084	0.584	0.062	0.587	0.066
Model 2	0.577	0.073	0.602	0.040	0.584	0.081	0.576	0.073	0.578	0.073
Model 3	0.587	0.069	0.619	0.046	0.601	0.065	0.585	0.071	0.598	0.062
Model 4	0.590	0.067	0.604	0.054	0.605	0.061	0.585	0.067	0.591	0.072

Model	4	SE 4	5	SE 5	6	SE 6	7	SE 7	8	SE 8
Model 1	0.586	0.066	0.593	0.064	0.573	0.077	0.590	0.062	0.579	NA
Model 2	0.577	0.074	0.576	0.075	0.573	0.071	0.584	0.074	0.585	0.034
Model 3	0.586	0.067	0.583	0.075	0.577	0.074	0.572	0.080	0.597	0.004
Model 4	0.589	0.066	0.591	0.066	0.586	0.067	0.597	0.071	0.573	0.004

Table 4.14
Attribute 3 Classification Rates by DIF Count

Model	All	All SE	0	SE 0	1	SE 1	2	SE 2	3	SE 3
Model 1	0.792	0.035	0.778	0.025	0.790	0.038	0.793	0.035	0.794	0.036
Model 2	0.786	0.039	0.788	0.018	0.780	0.048	0.791	0.031	0.787	0.039
Model 3	0.794	0.036	0.791	0.015	0.792	0.034	0.794	0.031	0.795	0.036
Model 4	0.796	0.035	0.795	0.024	0.783	0.038	0.798	0.028	0.795	0.036

Model	4	SE 4	5	SE 5	6	SE 6	7	SE 7	8	SE 8
Model 1	0.792	0.035	0.790	0.033	0.798	0.033	0.786	0.040	0.839	NA
Model 2	0.787	0.039	0.783	0.042	0.791	0.041	0.775	0.042	0.812	0.042
Model 3	0.797	0.037	0.789	0.036	0.798	0.036	0.775	0.058	0.819	0.034
Model 4	0.796	0.036	0.796	0.031	0.800	0.035	0.781	0.044	0.785	0.008

Table 4.15
Attribute 3 Kappa Statistic

Model	All	All SE	0	SE 0	1	SE 1	2	SE 2	3	SE 3
Model 1	0.583	0.070	0.555	0.050	0.578	0.076	0.583	0.071	0.585	0.072
Model 2	0.573	0.077	0.576	0.035	0.561	0.093	0.582	0.062	0.574	0.076
Model 3	0.586	0.073	0.582	0.030	0.584	0.066	0.585	0.063	0.588	0.072
Model 4	0.589	0.069	0.590	0.048	0.564	0.076	0.595	0.057	0.589	0.071

Model	4	SE 4	5	SE 5	6	SE 6	7	SE 7	8	SE 8
Model 1	0.583	0.071	0.577	0.066	0.592	0.066	0.572	0.079	0.677	NA
Model 2	0.574	0.076	0.567	0.080	0.582	0.080	0.551	0.081	0.623	0.086
Model 3	0.591	0.075	0.576	0.072	0.594	0.073	0.549	0.109	0.638	0.068
Model 4	0.589	0.072	0.590	0.062	0.598	0.071	0.557	0.090	0.571	0.015

Table 4.16
Pattern Classification Rate

Model	All	All SE	0	SE 0	1	SE 1	2	SE 2	3	SE 3
Model 1	0.506	0.038	0.499	0.031	0.509	0.037	0.512	0.037	0.506	0.042
Model 2	0.483	0.071	0.509	0.027	0.491	0.063	0.500	0.052	0.486	0.070
Model 3	0.505	0.049	0.520	0.018	0.520	0.033	0.506	0.056	0.513	0.044
Model 4	0.511	0.041	0.508	0.013	0.504	0.055	0.517	0.037	0.514	0.039

Model	4	SE 4	5	SE 5	6	SE 6	7	SE 7	8	SE 8
Model 1	0.503	0.037	0.507	0.034	0.501	0.043	0.504	0.027	0.556	NA
Model 2	0.477	0.077	0.480	0.071	0.475	0.082	0.470	0.065	0.529	0.040
Model 3	0.504	0.052	0.499	0.051	0.500	0.045	0.479	0.044	0.499	0.086
Model 4	0.508	0.041	0.512	0.042	0.514	0.036	0.492	0.055	0.497	0.024

Table 4.17
Pattern Kappa Rate

Model	All	All SE	0	SE 0	1	SE 1	2	SE 2	3	SE 3
Model 1	0.394	0.041	0.375	0.042	0.372	0.048	0.388	0.048	0.401	0.039
Model 2	0.400	0.038	0.407	0.040	0.408	0.010	0.408	0.062	0.401	0.042
Model 3	0.402	0.037	0.414	0.016	0.372	NA	0.411	0.056	0.402	0.042
Model 4	0.409	0.040	0.416	0.019	0.401	0.036	0.405	0.042	0.406	0.045

Model	4	SE 4	5	SE 5	6	SE 6	7	SE 7	8	SE 8
Model 1	0.376	0.040	0.410	0.040	0.413	0.030	0.387	0.038	NA	NA
Model 2	0.395	0.020	0.403	0.038	0.410	0.032	0.368	0.030	NA	NA
Model 3	0.396	0.019	0.402	0.034	0.414	0.033	0.348	NA	0.323	NA
Model 4	0.398	0.033	0.425	0.039	0.429	0.021	0.356	0.012	NA	NA

Item Parameter Estimation

Table 4.18 contains the average item parameter estimates and standard errors both with and without DIF for each of the four models. Table 4.19 contains the RMSE for these parameter estimates indicating the variability across replications. For Model 1, in which all parameters were freely estimated, the interaction parameter estimates were highly varied with diverse standard errors and large RMSE indicating a high degree of variability in this parameter estimate across replications. Intercept and main effect estimates for Model 1 cases without DIF show little similarity with other models; however, when DIF is present, Model 1 and Model 4 intercept, main effect and interaction term estimates are fairly similar which is what we would expect for these conditions.

In Model 2, all items and the structural model were held equal across groups, this level of constraint resulted in parameter estimates that look unlike any other model for both DIF and no DIF conditions. Similarly, for Model 3 in which item parameters were held constant but the structural model was freely estimated across groups, we see parameter estimates that do not resemble other model parameter estimates.

Structural Parameter Estimation

Table 4.20 contains the structural parameter estimates for each model separated by DIF and no DIF parameters. Table 4.21 contains the RMSE for the structural parameters indicating the degree of variability across replications. As expected, models in which the structural portion of the model was freely estimated across groups show similar parameter estimates. This is particularly true for the DIF cases. Model 2, the only model which constrained the structural parameters to be equal across groups, is meaningfully different from the other models.

Table 4.18
Item Parameter Estimate Bias

Model	Intercept No DIF	SE Intercept No DIF	Main Effect No DIF	SE Main Effect No DIF	Interaction No DIF	SE Interaction No DIF
Model 1	-0.034	0.092	0.060	0.114	0.601	4.179
Model 2	-0.064	0.102	0.021	0.136	0.131	0.488
Model 3	0.009	0.083	0.039	0.244	0.130	2.116
Model 4	-0.012	0.076	0.033	0.103	0.152	2.164

Model	Intercept DIF	SE Intercept DIF	Main Effect DIF	SE Main Effect DIF	Interaction DIF	SE Interaction DIF
Model 1	-0.041	0.178	0.056	0.338	0.231	2.220
Model 2	-0.114	0.683	0.071	0.771	0.282	2.157
Model 3	-0.058	0.266	0.127	1.681	0.203	1.757
Model 4	-0.035	0.216	0.053	0.335	0.293	2.525

Table 4.19
Item Parameter Estimate RMSE

Model	Intercept No DIF	SE Intercept No DIF	Main Effect No DIF	SE Main Effect No DIF	Interaction No DIF	SE Interaction No DIF
Model 1	0.279	0.216	0.527	0.258	2.841	13.109
Model 2	0.229	0.204	0.406	0.298	0.841	0.943
Model 3	0.175	0.079	0.446	0.888	1.130	5.067
Model 4	0.182	0.080	0.371	0.129	1.144	4.990

Model	Intercept DIF	SE Intercept DIF	Main Effect DIF	SE Main Effect DIF	Interaction DIF	SE Interaction DIF
Model 1	0.232	0.172	0.449	0.322	1.154	3.038
Model 2	0.392	0.658	0.596	0.830	1.043	1.913
Model 3	0.340	0.213	0.614	1.637	1.003	1.462
Model 4	0.226	0.252	0.426	0.314	1.253	3.266

Table 4.20
Structural Parameter Estimate Bias

Model	Main Effect No DIF	SE Main Effect No DIF	Interaction No DIF	SE Interaction No DIF
Model 1	0.472	0.168	-0.471	0.179
Model 2	0.306	1.223	-0.174	3.449
Model 3	0.409	0.412	-0.386	0.633
Model 4	0.578	0.352	-0.538	0.255

Model	Main Effect DIF	SE Main Effect DIF	Interaction DIF	SE Interaction DIF
Model 1	0.573	0.285	-0.500	0.364
Model 2	0.261	4.339	-0.555	4.845
Model 3	0.503	0.625	-0.493	0.820
Model 4	0.572	0.259	-0.611	0.553

Table 4.21
Structural Parameter Estimated RMSE

Model	Main Effect No DIF	SE Main Effect No DIF	Interaction No DIF	SE Interaction No DIF
Model 1	0.979	0.820	1.093	1.076
Model 2	0.641	2.016	0.859	4.590
Model 3	1.094	1.592	1.339	1.600
Model 4	0.976	0.475	1.117	0.580

Model	Main Effect DIF	SE Main Effect DIF	Interaction DIF	SE Interaction DIF
Model 1	1.129	0.515	1.099	1.549
Model 2	0.877	4.259	0.812	4.809
Model 3	1.285	1.645	1.369	2.948
Model 4	1.085	0.399	1.024	0.671

Real Data Analysis

For the real data analysis, the original five-item assessment was reduced to three so as to minimize the impact of MPlus and its instability in estimation. Additionally, three items allows for perfect item fit in the free baseline model thereby establishing a solid baseline for model comparisons.

Three models were applied to the real data and evaluated for differences in model fit. Table 4.22 reports the model fit statistics for Model 1, the free baseline model, Model 2, the fixed item and freely estimated structural model, and Model 3, the fully constrained model. As expected, the fully estimated Model 1 shows the best fit for these data. Conversely, Model 3 shows the poorest fit with *Chi-Square Categorical Pearson Value* (*44.894, 0.00*) indicating this model does not fit the data at all.

Table 4.22
Real Data: Model Estimates

Model	Parameters	ChiSq Categorical Pearson Value	ChiSq Categorical Pearson DF	ChiSq Categorical Pearson PValue	LL
Free Baseline	15	0	0	1	-1564.068
Free Structural/ Fixed Items	9	4.093	6	0.6641	-1566.13
Fixed Structural/ Fixed Items	8	44.894	7	0	-1587.081

Model	AIC	BIC	aBIC	Entropy	AICC
Free Baseline	3158.136	3225.767	3178.141	0.822	3158.8688
Free Structural/ Fixed Items	3150.259	3190.838	3162.263	0.797	3150.5313
Fixed Structural/ Fixed Items	3190.161	3226.232	3200.831	0.777	3190.3785

Table 4.23 contains the results of the log-likelihood ratio model comparison tests. These results show Model 2 is not significantly different from Model 1, X^2 (6) = 4.124, p= .66; however Model 3 is statistically significantly poorer fit than Model 1, X^2 (1) = 41.902, p<.001. These results indicate the groups do not differ on the items but do possess different levels of mastery of the measured attribute because the structural model must vary across groups in order for the model to fit the data. Model 2 is the perfect model for the real data because it's not significantly different from Model 1 but it is significantly different from Model 3. For Model 2, the items are fixed yet there is no difference from Model 1 which shows there is no difference between groups on the items. The only difference is in the structural model or the proportion of masters that are in each group.

Table 4.23
Real Data: Model Comparisons

Comparison	X^2	*df*	*Pvalue*
Free Baseline vs. Fixed Structural/Fixed Items	46.026	7	<.0001
Free Baseline vs. Free Structural/Fixed Items	4.124	6	.6599
Free Structural/Fixed Items vs. Fixed Structural/Fixed Items	41.902	1	<.0001

All parameter estimates for the real data are reported in Table 4.24 and further indicate the differences in model fit across conditions.

Table 4.24
Real Data: All Parameter Estimates

	Free Baseline		Structural Free		All Fixed	
Parameter	Estimate	Std Error	Estimate	Std Error	Estimate	Std Error
Mean Difference for Attribute	1.673	0.587			1.419	0.233
Mean for District	0.039	0.077	0.039	0.08	0.039	0.077
Mean for Attribute	-1.366	0.441	-0.436	0.263	-1.359	0.27
Intercept for I33488 District 1	0.29	0.238	0.29	0.21	0.189	0.187
Main Effect for I33488 District 1	2.91	1.05	2.665	0.448	2.688	0.378
Intercept for I33488 District 2	0.32	0.421				
Main Effect for I33488 District 2	2.492	0.538				
Intercept for I17880 District 1	-0.436	0.27	-0.513	0.256	-0.708	0.228
Main Effect for I17880 District 1	2.248	0.485	2.499	0.343	2.709	0.32
Intercept for I17880 District 2	-0.802	0.639				
Main Effect for I17880 District 2	2.932	0.666				
Intercept for I17852 District 1	-1.27	0.416	-1.237	0.357	-1.211	0.291
Main Effect for I17852 District 1	4.793	3.946	3.165	0.463	2.868	0.339
Intercept for I17852 District 2	-1.14	0.686				
Main Effect for I17852 District 2	2.652	0.682				

The Wald tests for parameter differences evaluate parameter estimates from the free baseline model for significant differences or DIF. These results show a statistically significant difference in attribute means ($p<0.01$) and no statistically significant differences for item parameters. These results further support the best fit of Model 2 and the fixed item and free structural model.

Table 4.25
Real Data: Wald Tests for Parameter Difference Estimates

Parameter	Estimate	Std Error	Pvalue
Mean Difference for Attribute	3.04	3.155	0.002
Intercept Difference for I33488	-0.029	-0.061	0.952
Main Effect Difference for I33488	0.419	0.355	0.723
Intercept Difference for I17880	0.366	0.528	0.598
Main Effect Difference for I17880	-0.684	-0.83	0.407
Intercept Difference for I17852	-0.13	-0.162	0.871
Main Effect Difference for I17852	2.141	0.535	0.593

Conclusions

Analysis of the simulated data revealed that MPlus is unstable for complex data and is more likely to experience errors or fail to converge when the number of parameters being estimated is high or interaction terms are present. In this case, the free baseline model, having the largest number of parameters to estimate, experienced the greatest number of errors in estimation.

Evaluation of the model fit statistics from the simulation also showed that allowing the structural model to be freely estimated while constraining the item parameters to be equal is not effective if DIF may be present in the items. This is further validated by the real data which showed when no DIF is present in the items, constraining item parameters to be equal across groups while allowing the structural model to vary can yield the best fitting model.

Finally, the D parameters which assessed the mean differences in parameters with DIF, were effective at identifying significant differences when models were effectively estimated by MPlus.

CHAPTER 5

DISCUSSION

This dissertation presented a method for the identification of Differential Item Functioning (DIF) for the Log-linear Cognitive Diagnosis Model (LCDM). By adapting an existing framework for DIF detection frequently implemented for Confirmatory Factor Analysis (CFA) models, a new omnibus test method was proposed and evaluated for the LCDM. This omnibus test attempted to evaluate all item and structural parameters simultaneously for the presence of DIF and was applied to both the simulated and real data samples.

Model 1 Discussion

The free-baseline model encountered frequent errors in MPlus. This was most likely due to issues with MPlus more than the freebaseline estimation itself. High number of parameters to be estimated could have contributed to the instability. Many samples failed to converge in this case, thereby limiting the number of available samples for model comparison and Type I Error and Power analysis. Regardless of these issues, there was sufficient data for model comparison.

In the real data case, the number of items in the dataset were reduced to three. This allowed for a more stable estimation of the free baseline model, and far fewer issues occurred.

Model 2 Discussion

The fully constrained model was much more stable with regard to Mplus estimation, and the relationship between the fully constrained model and free baseline models was as expected. The freebaseline model, when it converged, yielded a statistically significantly better fit than the fully constrained model. Forcing group equivalence in model 2 produced item parameter estimates and structural parameter estimates that were meaningfully different from the estimates produced by other models.

The fully constrained model was applied to the real data and also proved to be a statistically significantly poorer fit than the fully estimated model. This model was also compared to the free structural/fixed item model in condition 3.

Model 3 Discussion

Model 3 explored the case when the structural model is allowed to vary between groups but item parameters are fixed across groups. This represents the theoretical case that a different proportion of masters are present for one group and is a fairly common practice in real data analyses. For the simulated data, differences in items became manifest in the freely estimated structural model and produced very different structural parameter estimates across replications. When DIF is present in the items, allowing the structural model to vary while constraining items to be equal produces inaccurate structural estimates.

For the real data sample, this model represented an ideal fit. The free baseline comparison verified that item parameters were not statistically significantly different across groups. This gave justification to the use of model 3 as a means of assessing the presence of different proportions of masters in each group or structural model DIF. When model 3 fit was compared to the fully constrained model, it was verified that allowing the structural model to vary did improve overall fit. This indicates that different proportions of masters are present between groups in the real data sample.

Model 4 Discussion

Model 4 was intended to be the ideal case in which the items with DIF are known and are uniquely estimated for each group. This condition was also the testing ground for implementing ‘difference parameters” or values that are created in the model estimation to assess the difference in parameter estimates between groups. Difference parameters were created for the known DIF items. Unfortunately, this model was another source of estimation frustration with MPlus. Though not quite as unstable as model 1, it is likely that the volume of estimated parameters contributed to the challenges experienced with MPlus. When this model did converge, the difference parameters were fairly stable. This indicates the use of difference parameters could be a means of efficiently calculating item parameter differences if another estimation tool is used.

This model did not work with the real dataset as there were no “known” items or parameters with DIF.

Conclusions & Future Research

The simulated data analysis was not as successful as hoped in that MPlus estimation errors frequently occurred. Because this was particularly problematic for the free baseline model, it would be worth future research investigating whether complexity of the data (items measuring multiple attributes and incorporating interaction terms) or volume of parameters generated these estimation errors.

This analysis stopped short of the step-by-step identification of item biases. Future research should incorporate these multiple comparisons tests to locate item bias when the omnibus test (fully constrained vs free baseline) is found to yield significantly different fit statistics.

The real data analysis was more successful than the simulated data analysis and proves that the omnibus tests for DIF work when MPlus remains stable. The omnibus test of item parameter differences was successfully applied, and based on these results, the item parameters were found to hold no differences between groups. The real data analysis was also successful in testing the structural model for DIF. In this case, DIF was found to be present at the structural level.

The comparison of model 3 (freely estimated structural model/constrained items) produced inflated Type I error rates for the simulated data but was successful for the real data. The inflated Type I error in the simulation study indicates that maybe a more stepwise model comparison process needs to occur. The simulation shows that model 3 is limited when DIF is present in items. Item level DIF must be controlled before testing for DIF at the structural level. Differences in groups that are based on ability/trait may be spurious if DIF in items is not controlled for first.

Finally, the success of the real data analysis lends support to the idea that the simulated data were too complex for MPlus. In the real data, items measured only one dimension and there were no interaction terms. Given the success of the real data analysis, we can infer that complex data are not easily estimated in MPlus.

REFERENCES

Bozard, Jennifer L. (2010*). Invariance testing in diagnostic classification models*. Unpublished master's thesis, University of Georgia.

Bradley, James V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology. 31*, 144-152.

Bradshaw, L., Izsák, A., Templin, J., & Jacobsen, E. (in press). Diagnosing teachers' understandings of rational number: Building a multidimensional test within the diagnostic classification framework. Unpublished article, University of Georgia.

Bradshaw, L. (2011). *Combining scaling and classification: A psychometric model for scaling ability and diagnosing misconceptions*. Unpublished doctoral dissertation, University of Georgia.

Byrne, B. M., Shavelson, R. J., & Muthen, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456-466.

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine, 19*, 3127-3131.

Cochran, William G. (1952). The X^2 test of goodness of fit. *The Annals of Mathematical Statistics, 3*(23), 315-34.

*5*Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Wadsworth Group/ Thomson Learning.

DeCoster, J. (1998). *Overview of factor analysis*. Retrieved January 27, 2013 from http://www.stat-help.com/factor.pdf

Finkelman, M. (n.d.). Multidimensional IRT defined. Retrieved March 10, 2013 from http://www.measuredprogress.org/documents/10157/19213/MultidimensionalIRT.pdf

Glockner-Rist, A. & Hoijtink, H. (2003). The best of both worlds: factor analysis of dichotomous data using item response theory and structural equation modeling. Structural Equation Modeling: A Multidisciplinary Journal, 544-565.

Gregorich, S. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care, 44*(11), S78-S94.

Hambleton, R., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of item response theory.* Newbury Park, CA: Sage Publications, Inc.

Henson, R., Templin, J., & Willse, J. (2009). Defining a family of cognitive diagnosis models. *Psychometrika, 74*(2), 191-210.

Hill, H. (2007). Mathematical knowledge of middle school teachers: Implications for the No Child Left Behind Act policy initiative. *Educational Evaluation and Policy Analysis*, 29, 95–114.

Hoyle, R. (2012). *Handbook of Structural Equation Modeling.* New York, NY: Guilford Press.

Junker, B., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement, 25*, 258-272.

Kansas State Department of Education, 2017 Kansas Mathematics Standards. Retrieved July 21, 2018 from the Kansas State Department of Education http://community.ksde.org/Default.aspx?tabid=5276.

Kim, E.S & Yoon, M. (2011): Testing Measurement Invariance: A Comparison of Multiple-Group Categorical CFA and IRT, *Structural Equation Modeling: A Multidisciplinary Journal*, *18*(2), 212-228.

Kim, S. & Cohen, A. S. (1995). A comparison of Lord's chi-square, Raju's area measures, and the likelihood ratio test on detection of differential item functioning. *Applied Measurement in Education*, *8*(4), 291-312.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: The Guilford Press.

Li, F. (2008). A modified higher-order DINA model for detecting differential item functioning and differential attribute functioning. Unpublished doctoral dissertation, University of Georgia.

Lord, F. M. (1980). *Application of item response theory to practical testing problems.* Hillsdale, NJ: Lawrence Erlbaum.

MacCallum, R .C. (2012). *Studying Measurement Invariance Using Confirmatory Factor Analysis.* Retrieved December 25, 2012 from http://www.unc.edu/~rcm/psy236/measinv.pdf

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

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

Millsap, R. E. (2011). *Statistical Approaches to Measurement Invariance*. New York: Taylor and Francis Group.

Millsap, R.E. & Meredith, W. (2007). Factorial Invariance: Historical Perspectives and New Problems. In Cudeck, R. & MacCallum, R.C. (Eds.) *Factor Analysis at 100: Historical Developments and Future Directions.* Lawrence Erlbaum Associates: Mahwah, New Jersey.

Muthén, L. K., & Muthén, B. O. (1998-2013). Mplus user's guide. (6th ed.). Los Angeles, CA: Muthén & Muthén.

Raju, N. S. (1990). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement*, 14, 197-207.

Reis, S. P. Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114, 552-566.*

Rupp, A.A., & Templin, J. (2008). Unique characteristics of cognitive diagnostic models: A comprehensive review of the current state-of-the-art. *Measurement: Interdisciplinary Research and Perspectives, 6,* 219-262.

Rupp, A. A., Templin, J., & Henson, R. (2010). *Diagnostic measurement: Theory, methods, and applications*. New York: Guilford.

Serlin, Ronald C. (2000). Testing for robustness in Monte Carlo studies. *Psychological Methods, 5, 230-240.*

Stark, S., Chernyshenko, O.S., & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. *Journal of Applied Psychology, 6,* 1292-1306.

Steinmetz, H., Schmidt, P., Tina-Booh, A., Wieczorek, S., & Schwartz, S. (2009). Testing measurement invariance using multigroup CFA: Differences between educational groups in human values measurement. *Qual Quant, 43*, 599-616.

Thissen, D., Steinberg, L., & Wainer, H. (1988). Detection of differential item functioning using the parameters of item response models. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 67-113). Hillsdale, NJ: Erlbaum.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4-70.

APPENDIX

Model 1 Estimation Code

```
TITLE:  ! Section that appears in header of output file

DATA:  ! Location of free format data file
   FILE = data.csv;

VARIABLE:
   NAMES = id group class mitem1-mitem12;        ! List of variables in data file
   USEVARIABLE = mitem1-mitem12;     ! Variables to be analyzed
   CATEGORICAL = mitem1-mitem12;     ! Binary outcomes
   CLASSES =sex(2) c(8);        !classes and group
   KNOWNCLASS = sex(group = 0 group = 1);
   IDVARIABLE = id;
   AUXILIARY = class;

ANALYSIS:
   TYPE = MIXTURE;                  ! Estimates latent classes
   STARTS = 0;                    ! Turn off multiple random start feature
   PROCESSORS = 16;                 ! Number of processors available
   MCITERATIONS = 2;
   MUITERATIONS = 2;
   MITERATIONS = 1000;
   MCONVERGENCE = .0001;

MODEL:

%OVERALL%

c#1 on sex (m12);
c#2 on sex (m22);
c#3 on sex (m32);
c#4 on sex (m42);
c#5 on sex (m52);
c#6 on sex (m62);
c#7 on sex (m72);

[c#1] (m11);
```

```
[c#2] (m21);
[c#3] (m31);
[c#4] (m41);
[c#5] (m51);
[c#6] (m61);
[c#7] (m71);

%sex#1.c#1% ! Model for Class 1
   [mitem1$1] (T1_1_1);      ! Item 1 Thresh 1
   [mitem2$1] (T2_1_1);      ! Item 2 Thresh 1
   [mitem3$1] (T3_1_1);      ! Item 3 Thresh 1
   [mitem4$1] (T4_1_1);      ! Item 4 Thresh 1
   [mitem5$1] (T5_1_1);      ! Item 5 Thresh 1
   [mitem6$1] (T6_1_1);      ! Item 6 Thresh 1
   [mitem7$1] (T7_1_1);      ! Item 7 Thresh 1
   [mitem8$1] (T8_1_1);      ! Item 8 Thresh 1
   [mitem9$1] (T9_1_1);      ! Item 9 Thresh 1
   [mitem10$1] (T10_1_1);    ! Item 10 Thresh 1
   [mitem11$1] (T11_1_1);    ! Item 11 Thresh 1
   [mitem12$1] (T12_1_1);    ! Item 12 Thresh 1

%sex#1.c#2%  ! Model for Class 2
   [mitem1$1] (T1_1_1);      ! Item 1 Thresh 1
   [mitem2$1] (T2_1_1);      ! Item 2 Thresh 1
   [mitem3$1] (T3_2_1);      ! Item 3 Thresh 2
   [mitem4$1] (T4_1_1);      ! Item 4 Thresh 1
   [mitem5$1] (T5_2_1);      ! Item 5 Thresh 2
   [mitem6$1] (T6_2_1);      ! Item 6 Thresh 2
   [mitem7$1] (T7_1_1);      ! Item 7 Thresh 1
   [mitem8$1] (T8_1_1);      ! Item 8 Thresh 1
   [mitem9$1] (T9_2_1);      ! Item 9 Thresh 2
   [mitem10$1] (T10_1_1);    ! Item 10 Thresh 1
   [mitem11$1] (T11_2_1);    ! Item 11 Thresh 2
   [mitem12$1] (T12_2_1);    ! Item 12 Thresh 2

%sex#1.c#3%  ! Model for Class 3
   [mitem1$1] (T1_1_1);      ! Item 1 Thresh 1
   [mitem2$1] (T2_2_1);      ! Item 2 Thresh 2
   [mitem3$1] (T3_1_1);      ! Item 3 Thresh 1
   [mitem4$1] (T4_2_1);      ! Item 4 Thresh 2
   [mitem5$1] (T5_3_1);      ! Item 5 Thresh 3
   [mitem6$1] (T6_1_1);      ! Item 6 Thresh 1
   [mitem7$1] (T7_1_1);      ! Item 7 Thresh 1
   [mitem8$1] (T8_2_1);      ! Item 8 Thresh 2
   [mitem9$1] (T9_1_1);      ! Item 9 Thresh 1
   [mitem10$1] (T10_2_1);    ! Item 10 Thresh 2
```

```
   [mitem11$1] (T11_3_1);     ! Item 11 Thresh 3
   [mitem12$1] (T12_1_1);     ! Item 12 Thresh 1

%sex#1.c#4%  ! Model for Class 4
   [mitem1$1] (T1_1_1);       ! Item 1 Thresh 1
   [mitem2$1] (T2_2_1);       ! Item 2 Thresh 2
   [mitem3$1] (T3_2_1);       ! Item 3 Thresh 2
   [mitem4$1] (T4_2_1);       ! Item 4 Thresh 2
   [mitem5$1] (T5_4_1);       ! Item 5 Thresh 4
   [mitem6$1] (T6_2_1);       ! Item 6 Thresh 2
   [mitem7$1] (T7_1_1);       ! Item 7 Thresh 1
   [mitem8$1] (T8_2_1);       ! Item 8 Thresh 2
   [mitem9$1] (T9_2_1);       ! Item 9 Thresh 2
   [mitem10$1] (T10_2_1);     ! Item 10 Thresh 2
   [mitem11$1] (T11_4_1);     ! Item 11 Thresh 4
   [mitem12$1] (T12_2_1);     ! Item 12 Thresh 2

%sex#1.c#5%  ! Model for Class 5
   [mitem1$1] (T1_2_1);       ! Item 1 Thresh 2
   [mitem2$1] (T2_1_1);       ! Item 2 Thresh 1
   [mitem3$1] (T3_1_1);       ! Item 3 Thresh 1
   [mitem4$1] (T4_3_1);       ! Item 4 Thresh 3
   [mitem5$1] (T5_1_1);       ! Item 5 Thresh 1
   [mitem6$1] (T6_3_1);       ! Item 6 Thresh 3
   [mitem7$1] (T7_2_1);       ! Item 7 Thresh 2
   [mitem8$1] (T8_1_1);       ! Item 8 Thresh 1
   [mitem9$1] (T9_1_1);       ! Item 9 Thresh 1
   [mitem10$1] (T10_3_1);     ! Item 10 Thresh 3
   [mitem11$1] (T11_1_1);     ! Item 11 Thresh 1
   [mitem12$1] (T12_3_1);     ! Item 12 Thresh 3

%sex#1.c#6%  ! Model for Class 6
   [mitem1$1] (T1_2_1);       ! Item 1 Thresh 2
   [mitem2$1] (T2_1_1);       ! Item 2 Thresh 1
   [mitem3$1] (T3_2_1);       ! Item 3 Thresh 2
   [mitem4$1] (T4_3_1);       ! Item 4 Thresh 3
   [mitem5$1] (T5_2_1);       ! Item 5 Thresh 2
   [mitem6$1] (T6_4_1);       ! Item 6 Thresh 4
   [mitem7$1] (T7_2_1);       ! Item 7 Thresh 2
   [mitem8$1] (T8_1_1);       ! Item 8 Thresh 1
   [mitem9$1] (T9_2_1);       ! Item 9 Thresh 2
   [mitem10$1] (T10_3_1);     ! Item 10 Thresh 3
   [mitem11$1] (T11_2_1);     ! Item 11 Thresh 2
   [mitem12$1] (T12_4_1);     ! Item 12 Thresh 4

%sex#1.c#7%  ! Model for Class 7
```

```
   [mitem1$1] (T1_2_1);       ! Item 1 Thresh 2
   [mitem2$1] (T2_2_1);       ! Item 2 Thresh 2
   [mitem3$1] (T3_1_1);       ! Item 3 Thresh 1
   [mitem4$1] (T4_4_1);       ! Item 4 Thresh 4
   [mitem5$1] (T5_3_1);       ! Item 5 Thresh 3
   [mitem6$1] (T6_3_1);       ! Item 6 Thresh 3
   [mitem7$1] (T7_2_1);       ! Item 7 Thresh 2
   [mitem8$1] (T8_2_1);       ! Item 8 Thresh 2
   [mitem9$1] (T9_1_1);       ! Item 9 Thresh 1
   [mitem10$1] (T10_4_1);     ! Item 10 Thresh 4
   [mitem11$1] (T11_3_1);     ! Item 11 Thresh 3
   [mitem12$1] (T12_3_1);     ! Item 12 Thresh 3

%sex#1.c#8%  ! Model for Class 8
   [mitem1$1] (T1_2_1);       ! Item 1 Thresh 2
   [mitem2$1] (T2_2_1);       ! Item 2 Thresh 2
   [mitem3$1] (T3_2_1);       ! Item 3 Thresh 2
   [mitem4$1] (T4_4_1);       ! Item 4 Thresh 4
   [mitem5$1] (T5_4_1);       ! Item 5 Thresh 4
   [mitem6$1] (T6_4_1);       ! Item 6 Thresh 4
   [mitem7$1] (T7_2_1);       ! Item 7 Thresh 2
   [mitem8$1] (T8_2_1);       ! Item 8 Thresh 2
   [mitem9$1] (T9_2_1);       ! Item 9 Thresh 2
   [mitem10$1] (T10_4_1);     ! Item 10 Thresh 4
   [mitem11$1] (T11_4_1);     ! Item 11 Thresh 4
   [mitem12$1] (T12_4_1);     ! Item 12 Thresh 4

%sex#2.c#1% ! Model for Class 1
   [mitem1$1] (T1_1_2);       ! Item 1 Thresh 1
   [mitem2$1] (T2_1_2);       ! Item 2 Thresh 1
   [mitem3$1] (T3_1_2);       ! Item 3 Thresh 1
   [mitem4$1] (T4_1_2);       ! Item 4 Thresh 1
   [mitem5$1] (T5_1_2);       ! Item 5 Thresh 1
   [mitem6$1] (T6_1_2);       ! Item 6 Thresh 1
   [mitem7$1] (T7_1_2);       ! Item 7 Thresh 1
   [mitem8$1] (T8_1_2);       ! Item 8 Thresh 1
   [mitem9$1] (T9_1_2);       ! Item 9 Thresh 1
   [mitem10$1] (T10_1_2);     ! Item 10 Thresh 1
   [mitem11$1] (T11_1_2);     ! Item 11 Thresh 1
   [mitem12$1] (T12_1_2);     ! Item 12 Thresh 1

%sex#2.c#2%  ! Model for Class 2
   [mitem1$1] (T1_1_2);       ! Item 1 Thresh 1
   [mitem2$1] (T2_1_2);       ! Item 2 Thresh 1
   [mitem3$1] (T3_2_2);       ! Item 3 Thresh 2
   [mitem4$1] (T4_1_2);       ! Item 4 Thresh 1
```

```
   [mitem5$1] (T5_2_2);       ! Item 5 Thresh 2
   [mitem6$1] (T6_2_2);       ! Item 6 Thresh 2
   [mitem7$1] (T7_1_2);       ! Item 7 Thresh 1
   [mitem8$1] (T8_1_2);       ! Item 8 Thresh 1
   [mitem9$1] (T9_2_2);       ! Item 9 Thresh 2
   [mitem10$1] (T10_1_2);     ! Item 10 Thresh 1
   [mitem11$1] (T11_2_2);     ! Item 11 Thresh 2
   [mitem12$1] (T12_2_2);     ! Item 12 Thresh 2

%sex#2.c#3%  ! Model for Class 3
   [mitem1$1] (T1_1_2);       ! Item 1 Thresh 1
   [mitem2$1] (T2_2_2);       ! Item 2 Thresh 2
   [mitem3$1] (T3_1_2);       ! Item 3 Thresh 1
   [mitem4$1] (T4_2_2);       ! Item 4 Thresh 2
   [mitem5$1] (T5_3_2);       ! Item 5 Thresh 3
   [mitem6$1] (T6_1_2);       ! Item 6 Thresh 1
   [mitem7$1] (T7_1_2);       ! Item 7 Thresh 1
   [mitem8$1] (T8_2_2);       ! Item 8 Thresh 2
   [mitem9$1] (T9_1_2);       ! Item 9 Thresh 1
   [mitem10$1] (T10_2_2);     ! Item 10 Thresh 2
   [mitem11$1] (T11_3_2);     ! Item 11 Thresh 3
   [mitem12$1] (T12_1_2);     ! Item 12 Thresh 1

%sex#2.c#4%  ! Model for Class 4
   [mitem1$1] (T1_1_2);       ! Item 1 Thresh 1
   [mitem2$1] (T2_2_2);       ! Item 2 Thresh 2
   [mitem3$1] (T3_2_2);       ! Item 3 Thresh 2
   [mitem4$1] (T4_2_2);       ! Item 4 Thresh 2
   [mitem5$1] (T5_4_2);       ! Item 5 Thresh 4
   [mitem6$1] (T6_2_2);       ! Item 6 Thresh 2
   [mitem7$1] (T7_1_2);       ! Item 7 Thresh 1
   [mitem8$1] (T8_2_2);       ! Item 8 Thresh 2
   [mitem9$1] (T9_2_2);       ! Item 9 Thresh 2
   [mitem10$1] (T10_2_2);     ! Item 10 Thresh 2
   [mitem11$1] (T11_4_2);     ! Item 11 Thresh 4
   [mitem12$1] (T12_2_2);     ! Item 12 Thresh 2

%sex#2.c#5%  ! Model for Class 5
   [mitem1$1] (T1_2_2);       ! Item 1 Thresh 2
   [mitem2$1] (T2_1_2);       ! Item 2 Thresh 1
   [mitem3$1] (T3_1_2);       ! Item 3 Thresh 1
   [mitem4$1] (T4_3_2);       ! Item 4 Thresh 3
   [mitem5$1] (T5_1_2);       ! Item 5 Thresh 1
   [mitem6$1] (T6_3_2);       ! Item 6 Thresh 3
   [mitem7$1] (T7_2_2);       ! Item 7 Thresh 2
   [mitem8$1] (T8_1_2);       ! Item 8 Thresh 1
```

```
    [mitem9$1] (T9_1_2);      ! Item 9 Thresh 1
    [mitem10$1] (T10_3_2);    ! Item 10 Thresh 3
    [mitem11$1] (T11_1_2);    ! Item 11 Thresh 1
    [mitem12$1] (T12_3_2);    ! Item 12 Thresh 3


%sex#2.c#6%  ! Model for Class 6
    [mitem1$1] (T1_2_2);      ! Item 1 Thresh 2
    [mitem2$1] (T2_1_2);      ! Item 2 Thresh 1
    [mitem3$1] (T3_2_2);      ! Item 3 Thresh 2
    [mitem4$1] (T4_3_2);      ! Item 4 Thresh 3
    [mitem5$1] (T5_2_2);      ! Item 5 Thresh 2
    [mitem6$1] (T6_4_2);      ! Item 6 Thresh 4
    [mitem7$1] (T7_2_2);      ! Item 7 Thresh 2
    [mitem8$1] (T8_1_2);      ! Item 8 Thresh 1
    [mitem9$1] (T9_2_2);      ! Item 9 Thresh 2
    [mitem10$1] (T10_3_2);    ! Item 10 Thresh 3
    [mitem11$1] (T11_2_2);    ! Item 11 Thresh 2
    [mitem12$1] (T12_4_2);    ! Item 12 Thresh 4

%sex#2.c#7%  ! Model for Class 7
    [mitem1$1] (T1_2_2);      ! Item 1 Thresh 2
    [mitem2$1] (T2_2_2);      ! Item 2 Thresh 2
    [mitem3$1] (T3_1_2);      ! Item 3 Thresh 1
    [mitem4$1] (T4_4_2);      ! Item 4 Thresh 4
    [mitem5$1] (T5_3_2);      ! Item 5 Thresh 3
    [mitem6$1] (T6_3_2);      ! Item 6 Thresh 3
    [mitem7$1] (T7_2_2);      ! Item 7 Thresh 2
    [mitem8$1] (T8_2_2);      ! Item 8 Thresh 2
    [mitem9$1] (T9_1_2);      ! Item 9 Thresh 1
    [mitem10$1] (T10_4_2);    ! Item 10 Thresh 4
    [mitem11$1] (T11_3_2);    ! Item 11 Thresh 3
    [mitem12$1] (T12_3_2);    ! Item 12 Thresh 3

%sex#2.c#8%  ! Model for Class 8
    [mitem1$1] (T1_2_2);      ! Item 1 Thresh 2
    [mitem2$1] (T2_2_2);      ! Item 2 Thresh 2
    [mitem3$1] (T3_2_2);      ! Item 3 Thresh 2
    [mitem4$1] (T4_4_2);      ! Item 4 Thresh 4
    [mitem5$1] (T5_4_2);      ! Item 5 Thresh 4
    [mitem6$1] (T6_4_2);      ! Item 6 Thresh 4
    [mitem7$1] (T7_2_2);      ! Item 7 Thresh 2
    [mitem8$1] (T8_2_2);      ! Item 8 Thresh 2
    [mitem9$1] (T9_2_2);      ! Item 9 Thresh 2
    [mitem10$1] (T10_4_2);    ! Item 10 Thresh 4
    [mitem11$1] (T11_4_2);    ! Item 11 Thresh 4
```

```
   [mitem12$1] (T12_4_2);     ! Item 12 Thresh 4

MODEL CONSTRAINT:  ! Used to define LCDM parameters
! Mplus uses P(X=0) rather than P(X=1) so multiply by -1

! STRUCTURAL MODEL 1
NEW(G_11_1*-1 G_12_1*-1 G_13_1*-1 G_212_1*1 G_213_1*1 G_223_1*1);
m11=  - (G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m21= G_13_1 - (G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m31= G_12_1 - (G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m41= G_12_1+G_13_1+G_223_1 -
(G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m51= G_11_1 - (G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m61= G_11_1+G_13_1+G_213_1 -
(G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m71= G_11_1+G_12_1+G_212_1 -
(G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);

! STRUCTURAL MODEL 2
NEW(G_11_2*-1 G_12_2*-1 G_13_2*-1 G_212_2*1 G_213_2*1 G_223_2*1);
m12=  - (G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m22= G_13_2  - (G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m32= G_12_2 - (G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m42= G_12_2+G_13_2+G_223_2  -
(G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m52= G_11_2  - (G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m62= G_11_2+G_13_2+G_213_2 -
(G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m72= G_11_2+G_12_2+G_212_2 -
(G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);


! Item 1: Define LCDM parameters present for item 1
NEW(L1_0_1 L1_11_1 L1_0_2 L1_11_2);
T1_1_1=-(L1_0_1);                               ! Item 1 Thresh 1
T1_2_1=-(L1_0_1+L1_11_1);                             ! Item 1 Thresh 2
! Main effect order constraints
L1_11_1>0;

T1_1_2=-(L1_0_2);                               ! Item 1 Thresh 1
T1_2_2=-(L1_0_2+L1_11_2);                             ! Item 1 Thresh 2
! Main effect order constraints
L1_11_2>0;

!NEW(D1_0 D1_11);
!D1_0 = L1_0_2-L1_0_1;
```

```
!D1_11 = L1_11_2-L1_11_1;

! Item 2: Define LCDM parameters present for item 2
NEW(L2_0_1 L2_12_1 L2_0_2 L2_12_2);
T2_1_1=-(L2_0_1);                               ! Item 2 Thresh 1
T2_2_1=-(L2_0_1+L2_12_1);                       ! Item 2 Thresh 2
! Main effect order constraints
L2_12_1>0;

T2_1_2=-(L2_0_2);                               ! Item 2 Thresh 1
T2_2_2=-(L2_0_2+L2_12_2);                       ! Item 2 Thresh 2
! Main effect order constraints
L2_12_2>0;

!NEW(D2_0 D2_12);
!D2_0 = L2_0_2-L2_0_1;
!D2_12 = L2_12_2-L2_12_1;

! Item 3: Define LCDM parameters present for item 3
NEW(L3_0_1 L3_13_1 L3_0_2 L3_13_2);
T3_1_1=-(L3_0_1);                               ! Item 27 Thresh 1
T3_2_1=-(L3_0_1+L3_13_1);                       ! Item 27 Thresh 2
! Main effect order constraints
L3_13_1>0;

T3_1_2=-(L3_0_2);                               ! Item 27 Thresh 1
T3_2_2=-(L3_0_2+L3_13_2);                       ! Item 27 Thresh 2
! Main effect order constraints
L3_13_2>0;

!NEW(D3_0 D3_13);
!D3_0 = L3_0_2-L3_0_1;
!D3_13 = L3_13_2-L3_13_1;


! Item 4: Define LCDM parameters present for item 4
NEW(L4_0_1 L4_11_1 L4_12_1 L4_212_1 L4_0_2 L4_11_2 L4_12_2 L4_212_2);
T4_1_1=-(L4_0_1);                               ! Item 4 Thresh 1
T4_2_1=-(L4_0_1+L4_12_1);                       ! Item 4 Thresh 2
T4_3_1=-(L4_0_1+L4_11_1);                       ! Item 4 Thresh 3
T4_4_1=-(L4_0_1+L4_11_1+L4_12_1+L4_212_1);      ! Item 4 Thresh 4
! Main effect order constraints
L4_11_1>0; L4_12_1>0;
! Two-way interaction order constraints
L4_212_1>-L4_11_1;
L4_212_1>-L4_12_1;
```

```
T4_1_2=-(L4_0_2);                                ! Item 4 Thresh 1
T4_2_2=-(L4_0_2+L4_12_2);                        ! Item 4 Thresh 2
T4_3_2=-(L4_0_2+L4_11_2);                        ! Item 4 Thresh 3
T4_4_2=-(L4_0_2+L4_11_2+L4_12_2+L4_212_2);       ! Item 4 Thresh 4
! Main effect order constraints
L4_11_2>0; L4_12_2>0;
! Two-way interaction order constraints
L4_212_2>-L4_11_2;
L4_212_2>-L4_12_2;

!NEW(D4_0 D4_11 D4_12 D4_212);
!D4_0=L4_0_2-L4_0_1;
!D4_11=L4_11_2-L4_11_1;
!D4_12=L4_12_2-L4_12_1;
!D4_212=L4_212_2-L4_212_1;


! Item 5: Define LCDM parameters present for item 5
NEW(L5_0_1 L5_12_1 L5_13_1 L5_223_1 L5_0_2 L5_12_2 L5_13_2 L5_223_2);
T5_1_1=-(L5_0_1);                                ! Item 29 Thresh 1
T5_2_1=-(L5_0_1+L5_13_1);                        ! Item 29 Thresh 2
T5_3_1=-(L5_0_1+L5_12_1);                        ! Item 29 Thresh 3
T5_4_1=-(L5_0_1+L5_12_1+L5_13_1+L5_223_1);       ! Item 29 Thresh 4
! Main effect order constraints
L5_12_1>0; L5_13_1>0;
! Two-way interaction order constraints
L5_223_1>-L5_12_1;
L5_223_1>-L5_13_1;

T5_1_2=-(L5_0_2);                                ! Item 29 Thresh 1
T5_2_2=-(L5_0_2+L5_13_2);                        ! Item 29 Thresh 2
T5_3_2=-(L5_0_2+L5_12_2);                        ! Item 29 Thresh 3
T5_4_2=-(L5_0_2+L5_12_2+L5_13_2+L5_223_2);       ! Item 29 Thresh 4
! Main effect order constraints
L5_12_2>0; L5_13_2>0;
! Two-way interaction order constraints
L5_223_2>-L5_12_2;
L5_223_2>-L5_13_2;

!NEW(D5_0 D5_12 D5_13 D5_223);
!D5_0=L5_0_2-L5_0_1;
!D5_12=L5_12_2-L5_12_1;
!D5_13=L5_13_2-L5_13_1;
!D5_223=L5_223_2-L5_223_1;
```

```
! Item 6: Define LCDM parameters present for item 6
NEW(L6_0_1 L6_11_1 L6_13_1 L6_213_1 L6_0_2 L6_11_2 L6_13_2 L6_213_2);
T6_1_1=-(L6_0_1);                                  ! Item 6 Thresh 1
T6_2_1=-(L6_0_1+L6_13_1);                          ! Item 6 Thresh 2
T6_3_1=-(L6_0_1+L6_11_1);                          ! Item 6 Thresh 3
T6_4_1=-(L6_0_1+L6_11_1+L6_13_1+L6_213_1);         ! Item 6 Thresh 4
! Main effect order constraints
L6_11_1>0; L6_13_1>0;
! Two-way interaction order constraints
L6_213_1>-L6_11_1;
L6_213_1>-L6_13_1;

T6_1_2=-(L6_0_2);                                  ! Item 6 Thresh 1
T6_2_2=-(L6_0_2+L6_13_2);                          ! Item 6 Thresh 2
T6_3_2=-(L6_0_2+L6_11_2);                          ! Item 6 Thresh 3
T6_4_2=-(L6_0_2+L6_11_2+L6_13_2+L6_213_2);         ! Item 6 Thresh 4
! Main effect order constraints
L6_11_2>0; L6_13_2>0;
! Two-way interaction order constraints
L6_213_2>-L6_11_2;
L6_213_2>-L6_13_2;

!NEW(D6_0 D6_11 D6_13 D6_213);
!D6_0=L6_0_2-L6_0_1;
!D6_11=L6_11_2-L6_11_1;
!D6_13=L6_13_2-L6_13_1;
!D6_213=L6_213_2-L6_213_1;

! Item 7: Define LCDM parameters present for item 7
NEW(L7_0_1 L7_11_1 L7_0_2 L7_11_2);
T7_1_1=-(L7_0_1);                                  ! Item 1 Thresh 1
T7_2_1=-(L7_0_1+L7_11_1);                          ! Item 1 Thresh 2
! Main effect order constraints
L7_11_1>0;

T7_1_2=-(L7_0_2);                                  ! Item 1 Thresh 1
T7_2_2=-(L7_0_2+L7_11_2);                          ! Item 1 Thresh 2
! Main effect order constraints
L7_11_2>0;

!NEW(D7_0 D7_11);
!D7_0 = L7_0_2-L7_0_1;
!D7_11 = L7_11_2-L7_11_1;

! Item 8: Define LCDM parameters present for item 8
NEW(L8_0_1 L8_12_1 L8_0_2 L8_12_2);
```

```
T8_1_1=-(L8_0_1);                                   ! Item 2 Thresh 1
T8_2_1=-(L8_0_1+L8_12_1);                           ! Item 2 Thresh 2
! Main effect order constraints
L8_12_1>0;

T8_1_2=-(L8_0_2);                                   ! Item 2 Thresh 1
T8_2_2=-(L8_0_2+L8_12_2);                           ! Item 2 Thresh 2
! Main effect order constraints
L8_12_2>0;

!NEW(D8_0 D8_12);
!D8_0 = L8_0_2-L8_0_1;
!D8_12 = L8_12_2-L8_12_1;

! Item 9: Define LCDM parameters present for item 9
NEW(L9_0_1 L9_13_1 L9_0_2 L9_13_2);
T9_1_1=-(L9_0_1);                                   ! Item 27 Thresh 1
T9_2_1=-(L9_0_1+L9_13_1);                           ! Item 27 Thresh 2
! Main effect order constraints
L9_13_1>0;

T9_1_2=-(L9_0_2);                                   ! Item 27 Thresh 1
T9_2_2=-(L9_0_2+L9_13_2);                           ! Item 27 Thresh 2
! Main effect order constraints
L9_13_2>0;

!NEW(D9_0 D9_13);
!D9_0 = L9_0_2-L9_0_1;
!D9_13 = L9_13_2-L9_13_1;

! Item 10: Define LCDM parameters present for item 10
NEW(L10_0_1 L10_11_1 L10_12_1 L10_2121 L10_0_2 L10_11_2 L10_12_2 L10_2122);
T10_1_1=-(L10_0_1);                                 ! Item 4 Thresh 1
T10_2_1=-(L10_0_1+L10_12_1);                        ! Item 4 Thresh 2
T10_3_1=-(L10_0_1+L10_11_1);                        ! Item 4 Thresh 3
T10_4_1=-(L10_0_1+L10_11_1+L10_12_1+L10_2121);      ! Item 4
Thresh 4
! Main effect order constraints
L10_11_1>0; L10_12_1>0;
! Two-way interaction order constraints
L10_2121>-L10_11_1;
L10_2121>-L10_12_1;

T10_1_2=-(L10_0_2);                                 ! Item 4 Thresh 1
T10_2_2=-(L10_0_2+L10_12_2);                        ! Item 4 Thresh 2
T10_3_2=-(L10_0_2+L10_11_2);                        ! Item 4 Thresh 3
```

```
T10_4_2=-(L10_0_2+L10_11_2+L10_12_2+L10_2122);                    ! Item 4
Thresh 4
! Main effect order constraints
L10_11_2>0; L10_12_2>0;
! Two-way interaction order constraints
L10_2122>-L10_11_2;
L10_2122>-L10_12_2;

!NEW(D10_0 D10_11 D10_12 D10_212);
!D10_0=L10_0_2-L10_0_1;
!D10_11=L10_11_2-L10_11_1;
!D10_12=L10_12_2-L10_12_1;
!D10_212=L10_212_2-L10_212_1;

! Item 11: Define LCDM parameters present for item 11
NEW(L11_0_1 L11_12_1 L11_13_1 L11_2231 L11_0_2 L11_12_2 L11_13_2 L11_2232);
T11_1_1=-(L11_0_1);                             ! Item 29 Thresh 1
T11_2_1=-(L11_0_1+L11_13_1);                         ! Item 29 Thresh 2
T11_3_1=-(L11_0_1+L11_12_1);                         ! Item 29 Thresh 3
T11_4_1=-(L11_0_1+L11_12_1+L11_13_1+L11_2231);                 ! Item 29 Thresh
4
! Main effect order constraints
L11_12_1>0; L11_13_1>0;
! Two-way interaction order constraints
L11_2231>-L11_12_1;
L11_2231>-L11_13_1;

T11_1_2=-(L11_0_2);                             ! Item 29 Thresh 1
T11_2_2=-(L11_0_2+L11_13_2);                         ! Item 29 Thresh 2
T11_3_2=-(L11_0_2+L11_12_2);                         ! Item 29 Thresh 3
T11_4_2=-(L11_0_2+L11_12_2+L11_13_2+L11_2232);                 ! Item 29 Thresh
4
! Main effect order constraints
L11_12_2>0; L11_13_2>0;
! Two-way interaction order constraints
L11_2232>-L11_12_2;
L11_2232>-L11_13_2;

!NEW(D11_0 D11_12 D11_13 D11_223);
!D11_0=L11_0_2-L11_0_1;
!D11_12=L11_12_2-L11_12_1;
!D11_13=L11_13_2-L11_13_1;
!D11_223=L11_223_2-L11_223_1;

! Item 12: Define LCDM parameters present for item 12
NEW(L12_0_1 L12_11_1 L12_13_1 L12_2131 L12_0_2 L12_11_2 L12_13_2 L12_2132);
```

```
T12_1_1=-(L12_0_1);                                    ! Item 6 Thresh 1
T12_2_1=-(L12_0_1+L12_13_1);                           ! Item 6 Thresh 2
T12_3_1=-(L12_0_1+L12_11_1);                           ! Item 6 Thresh 3
T12_4_1=-(L12_0_1+L12_11_1+L12_13_1+L12_2131);                   ! Item 6
Thresh 4
! Main effect order constraints
L12_11_1>0; L12_13_1>0;
! Two-way interaction order constraints
L12_2131>-L12_11_1;
L12_2131>-L12_13_1;

T12_1_2=-(L12_0_2);                                    ! Item 6 Thresh 1
T12_2_2=-(L12_0_2+L12_13_2);                           ! Item 6 Thresh 2
T12_3_2=-(L12_0_2+L12_11_2);                           ! Item 6 Thresh 3
T12_4_2=-(L12_0_2+L12_11_2+L12_13_2+L12_2132);                   ! Item 6
Thresh 4
! Main effect order constraints
L12_11_2>0; L12_13_2>0;
! Two-way interaction order constraints
L12_2132>-L12_11_2;
L12_2132>-L12_13_2;

!NEW(D12_0 D12_11 D12_13 D12_213);
!D12_0=L12_0_2-L12_0_1;
!D12_11=L12_11_2-L12_11_1;
!D12_13=L12_13_2-L12_13_1;
!D12_213=L12_213_2-L12_213_1;

OUTPUT:
   TECH10;  ! Request additional model fit statistics

SAVEDATA: ! Format, name of posterior probabilities of class membership file
   FORMAT = F10.5;
   FILE = model01_exam.dat;
   SAVE = CPROBABILITIES;
```

Model 2 Estimation Code

```
TITLE:  ! Section that appears in header of output file
   DCM for ExampleData with 3 attributes and 2-order structural model,
   30 items, and maximum 2-order item model,
   Saturated structural model (Mplus default).

DATA:  ! Location of free format data file
   FILE = data.csv;

VARIABLE:
   NAMES = id group class mitem1-mitem12;        ! List of variables in data file
   USEVARIABLE = mitem1-mitem12;      ! Variables to be analyzed
   CATEGORICAL = mitem1-mitem12;      ! Binary outcomes
   CLASSES = c(8);        !classes and group
   IDVARIABLE = id;
   AUXILIARY = class;

ANALYSIS:
   TYPE = MIXTURE;                  ! Estimates latent classes
   STARTS = 0;                  ! Turn off multiple random start feature
   PROCESSORS =16;                  ! Number of processors available
   MCITERATIONS = 2;
   MUITERATIONS = 2;
   MITERATIONS = 1000;
   MCONVERGENCE = .0001;

MODEL:

%OVERALL%

[c#1] (m1);  ! Latent variable mean for class 1
[c#2] (m2);  ! Latent variable mean for class 2
[c#3] (m3);  ! Latent variable mean for class 3
[c#4] (m4);  ! Latent variable mean for class 4
[c#5] (m5);  ! Latent variable mean for class 5
[c#6] (m6);  ! Latent variable mean for class 6
[c#7] (m7);  ! Latent variable mean for class 7


%c#1% ! Model for Class 1
   [mitem1$1] (T1_1);       ! Item 1 Thresh 1
   [mitem2$1] (T2_1);       ! Item 2 Thresh 1
```

```
    [mitem3$1] (T3_1);      ! Item 3 Thresh 1
    [mitem4$1] (T4_1);      ! Item 4 Thresh 1
    [mitem5$1] (T5_1);      ! Item 5 Thresh 1
    [mitem6$1] (T6_1);      ! Item 6 Thresh 1
    [mitem7$1] (T7_1);      ! Item 7 Thresh 1
    [mitem8$1] (T8_1);      ! Item 8 Thresh 1
    [mitem9$1] (T9_1);      ! Item 9 Thresh 1
    [mitem10$1] (T10_1);    ! Item 10 Thresh 1
    [mitem11$1] (T11_1);    ! Item 11 Thresh 1
    [mitem12$1] (T12_1);    ! Item 12 Thresh 1

%c#2%  ! Model for Class 2
    [mitem1$1] (T1_1);      ! Item 1 Thresh 1
    [mitem2$1] (T2_1);      ! Item 2 Thresh 1
    [mitem3$1] (T3_2);      ! Item 3 Thresh 2
    [mitem4$1] (T4_1);      ! Item 4 Thresh 1
    [mitem5$1] (T5_2);      ! Item 5 Thresh 2
    [mitem6$1] (T6_2);      ! Item 6 Thresh 2
    [mitem7$1] (T7_1);      ! Item 7 Thresh 1
    [mitem8$1] (T8_1);      ! Item 8 Thresh 1
    [mitem9$1] (T9_2);      ! Item 9 Thresh 2
    [mitem10$1] (T10_1);    ! Item 10 Thresh 1
    [mitem11$1] (T11_2);    ! Item 11 Thresh 2
    [mitem12$1] (T12_2);    ! Item 12 Thresh 2


%c#3%  ! Model for Class 3
    [mitem1$1] (T1_1);      ! Item 1 Thresh 1
    [mitem2$1] (T2_2);      ! Item 2 Thresh 1
    [mitem3$1] (T3_1);      ! Item 3 Thresh 2
    [mitem4$1] (T4_2);      ! Item 4 Thresh 1
    [mitem5$1] (T5_3);      ! Item 5 Thresh 2
    [mitem6$1] (T6_1);      ! Item 6 Thresh 2
    [mitem7$1] (T7_1);      ! Item 7 Thresh 1
    [mitem8$1] (T8_2);      ! Item 8 Thresh 1
    [mitem9$1] (T9_1);      ! Item 9 Thresh 2
    [mitem10$1] (T10_2);    ! Item 10 Thresh 1
    [mitem11$1] (T11_3);    ! Item 11 Thresh 2
    [mitem12$1] (T12_1);    ! Item 12 Thresh 2

%c#4%  ! Model for Class 4
    [mitem1$1] (T1_1);      ! Item 1 Thresh 1
    [mitem2$1] (T2_2);      ! Item 2 Thresh 1
    [mitem3$1] (T3_2);      ! Item 3 Thresh 2
    [mitem4$1] (T4_2);      ! Item 4 Thresh 1
    [mitem5$1] (T5_4);      ! Item 5 Thresh 2
```

```
    [mitem6$1] (T6_2);        ! Item 6 Thresh 2
    [mitem7$1] (T7_1);        ! Item 7 Thresh 1
    [mitem8$1] (T8_2);        ! Item 8 Thresh 1
    [mitem9$1] (T9_2);        ! Item 9 Thresh 2
    [mitem10$1] (T10_2);      ! Item 10 Thresh 1
    [mitem11$1] (T11_4);      ! Item 11 Thresh 2
    [mitem12$1] (T12_2);      ! Item 12 Thresh 2

%c#5%  ! Model for Class 5
    [mitem1$1] (T1_2);        ! Item 1 Thresh 2
    [mitem2$1] (T2_1);        ! Item 2 Thresh 1
    [mitem3$1] (T3_1);        ! Item 3 Thresh 1
    [mitem4$1] (T4_3);        ! Item 4 Thresh 3
    [mitem5$1] (T5_1);        ! Item 5 Thresh 1
    [mitem6$1] (T6_3);        ! Item 6 Thresh 3
    [mitem7$1] (T7_2);        ! Item 7 Thresh 2
    [mitem8$1] (T8_1);        ! Item 8 Thresh 1
    [mitem9$1] (T9_1);        ! Item 9 Thresh 1
    [mitem10$1] (T10_3);      ! Item 10 Thresh 3
    [mitem11$1] (T11_1);      ! Item 11 Thresh 1
    [mitem12$1] (T12_3);      ! Item 12 Thresh 3

%c#6%  ! Model for Class 6
    [mitem1$1] (T1_2);        ! Item 1 Thresh 2
    [mitem2$1] (T2_1);        ! Item 2 Thresh 1
    [mitem3$1] (T3_2);        ! Item 3 Thresh 2
    [mitem4$1] (T4_3);        ! Item 4 Thresh 3
    [mitem5$1] (T5_2);        ! Item 5 Thresh 2
    [mitem6$1] (T6_4);        ! Item 6 Thresh 4
    [mitem7$1] (T7_2);        ! Item 7 Thresh 2
    [mitem8$1] (T8_1);        ! Item 8 Thresh 1
    [mitem9$1] (T9_2);        ! Item 9 Thresh 2
    [mitem10$1] (T10_3);      ! Item 10 Thresh 3
    [mitem11$1] (T11_2);      ! Item 11 Thresh 2
    [mitem12$1] (T12_4);      ! Item 12 Thresh 4

%c#7%  ! Model for Class 7
    [mitem1$1] (T1_2);        ! Item 1 Thresh 2
    [mitem2$1] (T2_2);        ! Item 2 Thresh 2
    [mitem3$1] (T3_1);        ! Item 3 Thresh 1
    [mitem4$1] (T4_4);        ! Item 4 Thresh 4
    [mitem5$1] (T5_3);        ! Item 5 Thresh 3
    [mitem6$1] (T6_3);        ! Item 6 Thresh 3
    [mitem7$1] (T7_2);        ! Item 7 Thresh 2
    [mitem8$1] (T8_2);        ! Item 8 Thresh 2
    [mitem9$1] (T9_1);        ! Item 9 Thresh 1
```

```
   [mitem10$1] (T10_4);     ! Item 10 Thresh 4
   [mitem11$1] (T11_3);     ! Item 11 Thresh 3
   [mitem12$1] (T12_3);     ! Item 12 Thresh 3

%c#8%  ! Model for Class 8
   [mitem1$1] (T1_2);      ! Item 1 Thresh 2
   [mitem2$1] (T2_2);      ! Item 2 Thresh 2
   [mitem3$1] (T3_2);      ! Item 3 Thresh 2
   [mitem4$1] (T4_4);      ! Item 4 Thresh 4
   [mitem5$1] (T5_4);      ! Item 5 Thresh 4
   [mitem6$1] (T6_4);      ! Item 6 Thresh 4
   [mitem7$1] (T7_2);      ! Item 7 Thresh 2
   [mitem8$1] (T8_2);      ! Item 8 Thresh 2
   [mitem9$1] (T9_2);      ! Item 9 Thresh 2
   [mitem10$1] (T10_4);     ! Item 10 Thresh 4
   [mitem11$1] (T11_4);     ! Item 11 Thresh 4
   [mitem12$1] (T12_4);     ! Item 12 Thresh 4


MODEL CONSTRAINT:  ! Used to define LCDM parameters
! Mplus uses P(X=0) rather than P(X=1) so multiply by -1

! STRUCTURAL MODEL
NEW(G_11 G_12 G_13 G_212 G_213 G_223);

m1=  - (G_11+G_12+G_13+G_212+G_213+G_223);
m2= G_13 - (G_11+G_12+G_13+G_212+G_213+G_223);
m3= G_12 - (G_11+G_12+G_13+G_212+G_213+G_223);
m4= G_12+G_13+G_223 - (G_11+G_12+G_13+G_212+G_213+G_223);
m5= G_11 - (G_11+G_12+G_13+G_212+G_213+G_223);
m6= G_11+G_13+G_213 - (G_11+G_12+G_13+G_212+G_213+G_223);
m7= G_11+G_12+G_212 - (G_11+G_12+G_13+G_212+G_213+G_223);

! Item 1: Define LCDM parameters present for item 1
NEW(L1_0 L1_11);
T1_1=-(L1_0);                                    ! Item 1 Thresh 1
T1_2=-(L1_0+L1_11);                                  ! Item 1 Thresh 2
! Main effect order constraints
L1_11>0;

! Item 2: Define LCDM parameters present for item 2
NEW(L2_0 L2_12);
T2_1=-(L2_0);                                    ! Item 2 Thresh 1
T2_2=-(L2_0+L2_12);                                  ! Item 2 Thresh 2
! Main effect order constraints
L2_12>0;
```

```
! Item 3: Define LCDM parameters present for item 3
NEW(L3_0 L3_13);
T3_1=-(L3_0);                          ! Item 3 Thresh 1
T3_2=-(L3_0+L3_13);                    ! Item 3 Thresh 2
! Main effect order constraints
L3_13>0;


! Item 4: Define LCDM parameters present for item 4
NEW(L4_0 L4_11 L4_12 L4_212);
T4_1=-(L4_0);                          ! Item 4 Thresh 1
T4_2=-(L4_0+L4_12);                    ! Item 4 Thresh 2
T4_3=-(L4_0+L4_11);                    ! Item 4 Thresh 3
T4_4=-(L4_0+L4_11+L4_12+L4_212);       ! Item 4 Thresh 4
! Main effect order constraints
L4_11>0; L4_12>0;
! Two-way interaction order constraints
L4_212>-L4_11;
L4_212>-L4_12;


! Item 5: Define LCDM parameters present for item 5
NEW(L5_0 L5_12 L5_13 L5_223);
T5_1=-(L5_0);                          ! Item 5 Thresh 1
T5_2=-(L5_0+L5_13);                    ! Item 5 Thresh 2
T5_3=-(L5_0+L5_12);                    ! Item 5 Thresh 3
T5_4=-(L5_0+L5_12+L5_13+L5_223);       ! Item 5 Thresh 4
! Main effect order constraints
L5_12>0; L5_13>0;
! Two-way interaction order constraints
L5_223>-L5_12;
L5_223>-L5_13;


! Item 6: Define LCDM parameters present for item 6
NEW(L6_0 L6_11 L6_13 L6_213);
T6_1=-(L6_0);                          ! Item 6 Thresh 1
T6_2=-(L6_0+L6_13);                    ! Item 6 Thresh 2
T6_3=-(L6_0+L6_11);                    ! Item 6 Thresh 3
T6_4=-(L6_0+L6_11+L6_13+L6_213);       ! Item 6 Thresh 4
! Main effect order constraints
L6_11>0; L6_13>0;
! Two-way interaction order constraints
L6_213>-L6_11;
```

```
L6_213>-L6_13;

! Item 7: Define LCDM parameters present for item 7
NEW(L7_0 L7_11);
T7_1=-(L7_0);                      ! Item 7 Thresh 1
T7_2=-(L7_0+L7_11);                    ! Item 7 Thresh 2
! Main effect order constraints
L7_11>0;

! Item 8: Define LCDM parameters present for item 8
NEW(L8_0 L8_12);
T8_1=-(L8_0);                      ! Item 8 Thresh 1
T8_2=-(L8_0+L8_12);                    ! Item 8 Thresh 2
! Main effect order constraints
L8_12>0;

! Item 9: Define LCDM parameters present for item 9
NEW(L9_0 L9_13);
T9_1=-(L9_0);                      ! Item 9 Thresh 1
T9_2=-(L9_0+L9_13);                    ! Item 9 Thresh 2
! Main effect order constraints
L9_13>0;

! Item 10: Define LCDM parameters present for item 10
NEW(L10_0 L10_11 L10_12 L10_212);
T10_1=-(L10_0);                    ! Item 10 Thresh 1
T10_2=-(L10_0+L10_12);                 ! Item 10 Thresh 2
T10_3=-(L10_0+L10_11);                 ! Item 10 Thresh 3
T10_4=-(L10_0+L10_11+L10_12+L10_212);          ! Item 10 Thresh 4
! Main effect order constraints
L10_11>0; L10_12>0;
! Two-way interaction order constraints
L10_212>-L10_11;
L10_212>-L10_12;

! Item 11: Define LCDM parameters present for item 11
NEW(L11_0 L11_12 L11_13 L11_223);
T11_1=-(L11_0);                    ! Item 11 Thresh 1
T11_2=-(L11_0+L11_13);                 ! Item 11 Thresh 2
T11_3=-(L11_0+L11_12);                 ! Item 11 Thresh 3
T11_4=-(L11_0+L11_12+L11_13+L11_223);          ! Item 11 Thresh 4
```

```
! Main effect order constraints
L11_12>0; L11_13>0;
! Two-way interaction order constraints
L11_223>-L11_12;
L11_223>-L11_13;


! Item 12: Define LCDM parameters present for item 12
NEW(L12_0 L12_11 L12_13 L12_213);
T12_1=-(L12_0);                          ! Item 12 Thresh 1
T12_2=-(L12_0+L12_13);                      ! Item 12 Thresh 2
T12_3=-(L12_0+L12_11);                      ! Item 12 Thresh 3
T12_4=-(L12_0+L12_11+L12_13+L12_213);                ! Item 12 Thresh 4
! Main effect order constraints
L12_11>0; L12_13>0;
! Two-way interaction order constraints
L12_213>-L12_11;
L12_213>-L12_13;



OUTPUT:
   TECH10;  ! Request additional model fit statistics

SAVEDATA: ! Format, name of posterior probabilities of class membership file
   FORMAT = F10.5;
   FILE = model02_exam.dat;
   SAVE = CPROBABILITIES;
```

Model 3 Estimation Code

```
TITLE:  ! Section that appears in header of output file


DATA:  ! Location of free format data file
   FILE = data.csv;

VARIABLE:
   NAMES = id group class mitem1-mitem12;        ! List of variables in data file
   USEVARIABLE = mitem1-mitem12;      ! Variables to be analyzed
   CATEGORICAL = mitem1-mitem12;      ! Binary outcomes
   CLASSES =sex(2)  c(8) ;        !classes and group
   KNOWNCLASS = sex(group = 0 group = 1);
   IDVARIABLE = id;
   AUXILIARY = class;

ANALYSIS:
   TYPE = MIXTURE;                  ! Estimates latent classes
   STARTS = 0;                 ! Turn off multiple random start feature
   PROCESSORS = 16;                 ! Number of processors available

MODEL:

%OVERALL%

c#1 on sex (m12);
c#2 on sex (m22);
c#3 on sex (m32);
c#4 on sex (m42);
c#5 on sex (m52);
c#6 on sex (m62);
c#7 on sex (m72);

[c#1] (m11);
[c#2] (m21);
[c#3] (m31);
[c#4] (m41);
[c#5] (m51);
[c#6] (m61);
[c#7] (m71);

%sex#1.c#1% ! Model for Class 1
   [mitem1$1] (T1_1);       ! Item 1 Thresh 1
```

```
  [mitem2$1] (T2_1);      ! Item 2 Thresh 1
  [mitem3$1] (T3_1);      ! Item 3 Thresh 1
  [mitem4$1] (T4_1);      ! Item 4 Thresh 1
  [mitem5$1] (T5_1);      ! Item 5 Thresh 1
  [mitem6$1] (T6_1);      ! Item 6 Thresh 1
  [mitem7$1] (T7_1);      ! Item 7 Thresh 1
  [mitem8$1] (T8_1);      ! Item 8 Thresh 1
  [mitem9$1] (T9_1);      ! Item 9 Thresh 1
  [mitem10$1] (T10_1);    ! Item 10 Thresh 1
  [mitem11$1] (T11_1);    ! Item 11 Thresh 1
  [mitem12$1] (T12_1);    ! Item 12 Thresh 1

%sex#1.c#2%  ! Model for Class 2
  [mitem1$1] (T1_1);      ! Item 1 Thresh 1
  [mitem2$1] (T2_1);      ! Item 2 Thresh 1
  [mitem3$1] (T3_2);      ! Item 3 Thresh 2
  [mitem4$1] (T4_1);      ! Item 4 Thresh 1
  [mitem5$1] (T5_2);      ! Item 5 Thresh 2
  [mitem6$1] (T6_2);      ! Item 6 Thresh 2
  [mitem7$1] (T7_1);      ! Item 7 Thresh 1
  [mitem8$1] (T8_1);      ! Item 8 Thresh 1
  [mitem9$1] (T9_2);      ! Item 9 Thresh 2
  [mitem10$1] (T10_1);    ! Item 10 Thresh 1
  [mitem11$1] (T11_2);    ! Item 11 Thresh 2
  [mitem12$1] (T12_2);    ! Item 12 Thresh 2


%sex#1.c#3%  ! Model for Class 3
  [mitem1$1] (T1_1);      ! Item 1 Thresh 1
  [mitem2$1] (T2_2);      ! Item 2 Thresh 1
  [mitem3$1] (T3_1);      ! Item 3 Thresh 2
  [mitem4$1] (T4_2);      ! Item 4 Thresh 1
  [mitem5$1] (T5_3);      ! Item 5 Thresh 2
  [mitem6$1] (T6_1);      ! Item 6 Thresh 2
  [mitem7$1] (T7_1);      ! Item 7 Thresh 1
  [mitem8$1] (T8_2);      ! Item 8 Thresh 1
  [mitem9$1] (T9_1);      ! Item 9 Thresh 2
  [mitem10$1] (T10_2);    ! Item 10 Thresh 1
  [mitem11$1] (T11_3);    ! Item 11 Thresh 2
  [mitem12$1] (T12_1);    ! Item 12 Thresh 2

%sex#1.c#4%  ! Model for Class 4
  [mitem1$1] (T1_1);      ! Item 1 Thresh 1
  [mitem2$1] (T2_2);      ! Item 2 Thresh 1
  [mitem3$1] (T3_2);      ! Item 3 Thresh 2
  [mitem4$1] (T4_2);      ! Item 4 Thresh 1
```

```
   [mitem5$1] (T5_4);      ! Item 5 Thresh 2
   [mitem6$1] (T6_2);      ! Item 6 Thresh 2
   [mitem7$1] (T7_1);      ! Item 7 Thresh 1
   [mitem8$1] (T8_2);      ! Item 8 Thresh 1
   [mitem9$1] (T9_2);      ! Item 9 Thresh 2
   [mitem10$1] (T10_2);    ! Item 10 Thresh 1
   [mitem11$1] (T11_4);    ! Item 11 Thresh 2
   [mitem12$1] (T12_2);    ! Item 12 Thresh 2

%sex#1.c#5%  ! Model for Class 5
   [mitem1$1] (T1_2);      ! Item 1 Thresh 2
   [mitem2$1] (T2_1);      ! Item 2 Thresh 1
   [mitem3$1] (T3_1);      ! Item 3 Thresh 1
   [mitem4$1] (T4_3);      ! Item 4 Thresh 3
   [mitem5$1] (T5_1);      ! Item 5 Thresh 1
   [mitem6$1] (T6_3);      ! Item 6 Thresh 3
   [mitem7$1] (T7_2);      ! Item 7 Thresh 2
   [mitem8$1] (T8_1);      ! Item 8 Thresh 1
   [mitem9$1] (T9_1);      ! Item 9 Thresh 1
   [mitem10$1] (T10_3);    ! Item 10 Thresh 3
   [mitem11$1] (T11_1);    ! Item 11 Thresh 1
   [mitem12$1] (T12_3);    ! Item 12 Thresh 3

%sex#1.c#6%  ! Model for Class 6
   [mitem1$1] (T1_2);      ! Item 1 Thresh 2
   [mitem2$1] (T2_1);      ! Item 2 Thresh 1
   [mitem3$1] (T3_2);      ! Item 3 Thresh 2
   [mitem4$1] (T4_3);      ! Item 4 Thresh 3
   [mitem5$1] (T5_2);      ! Item 5 Thresh 2
   [mitem6$1] (T6_4);      ! Item 6 Thresh 4
   [mitem7$1] (T7_2);      ! Item 7 Thresh 2
   [mitem8$1] (T8_1);      ! Item 8 Thresh 1
   [mitem9$1] (T9_2);      ! Item 9 Thresh 2
   [mitem10$1] (T10_3);    ! Item 10 Thresh 3
   [mitem11$1] (T11_2);    ! Item 11 Thresh 2
   [mitem12$1] (T12_4);    ! Item 12 Thresh 4

%sex#1.c#7%  ! Model for Class 7
   [mitem1$1] (T1_2);      ! Item 1 Thresh 2
   [mitem2$1] (T2_2);      ! Item 2 Thresh 2
   [mitem3$1] (T3_1);      ! Item 3 Thresh 1
   [mitem4$1] (T4_4);      ! Item 4 Thresh 4
   [mitem5$1] (T5_3);      ! Item 5 Thresh 3
   [mitem6$1] (T6_3);      ! Item 6 Thresh 3
   [mitem7$1] (T7_2);      ! Item 7 Thresh 2
   [mitem8$1] (T8_2);      ! Item 8 Thresh 2
```

```
   [mitem9$1] (T9_1);       ! Item 9 Thresh 1
   [mitem10$1] (T10_4);     ! Item 10 Thresh 4
   [mitem11$1] (T11_3);     ! Item 11 Thresh 3
   [mitem12$1] (T12_3);     ! Item 12 Thresh 3

%sex#1.c#8%  ! Model for Class 8
   [mitem1$1] (T1_2);       ! Item 1 Thresh 2
   [mitem2$1] (T2_2);       ! Item 2 Thresh 2
   [mitem3$1] (T3_2);       ! Item 3 Thresh 2
   [mitem4$1] (T4_4);       ! Item 4 Thresh 4
   [mitem5$1] (T5_4);       ! Item 5 Thresh 4
   [mitem6$1] (T6_4);       ! Item 6 Thresh 4
   [mitem7$1] (T7_2);       ! Item 7 Thresh 2
   [mitem8$1] (T8_2);       ! Item 8 Thresh 2
   [mitem9$1] (T9_2);       ! Item 9 Thresh 2
   [mitem10$1] (T10_4);     ! Item 10 Thresh 4
   [mitem11$1] (T11_4);     ! Item 11 Thresh 4
   [mitem12$1] (T12_4);     ! Item 12 Thresh 4

%sex#2.c#1% ! Model for Class 1
   [mitem1$1] (T1_1);       ! Item 1 Thresh 1
   [mitem2$1] (T2_1);       ! Item 2 Thresh 1
   [mitem3$1] (T3_1);       ! Item 3 Thresh 1
   [mitem4$1] (T4_1);       ! Item 4 Thresh 1
   [mitem5$1] (T5_1);       ! Item 5 Thresh 1
   [mitem6$1] (T6_1);       ! Item 6 Thresh 1
   [mitem7$1] (T7_1);       ! Item 7 Thresh 1
   [mitem8$1] (T8_1);       ! Item 8 Thresh 1
   [mitem9$1] (T9_1);       ! Item 9 Thresh 1
   [mitem10$1] (T10_1);     ! Item 10 Thresh 1
   [mitem11$1] (T11_1);     ! Item 11 Thresh 1
   [mitem12$1] (T12_1);     ! Item 12 Thresh 1

%sex#2.c#2%  ! Model for Class 2
   [mitem1$1] (T1_1);       ! Item 1 Thresh 1
   [mitem2$1] (T2_1);       ! Item 2 Thresh 1
   [mitem3$1] (T3_2);       ! Item 3 Thresh 2
   [mitem4$1] (T4_1);       ! Item 4 Thresh 1
   [mitem5$1] (T5_2);       ! Item 5 Thresh 2
   [mitem6$1] (T6_2);       ! Item 6 Thresh 2
   [mitem7$1] (T7_1);       ! Item 7 Thresh 1
   [mitem8$1] (T8_1);       ! Item 8 Thresh 1
   [mitem9$1] (T9_2);       ! Item 9 Thresh 2
   [mitem10$1] (T10_1);     ! Item 10 Thresh 1
   [mitem11$1] (T11_2);     ! Item 11 Thresh 2
   [mitem12$1] (T12_2);     ! Item 12 Thresh 2
```

```
%sex#2.c#3%  ! Model for Class 3
   [mitem1$1] (T1_1);       ! Item 1 Thresh 1
   [mitem2$1] (T2_2);       ! Item 2 Thresh 1
   [mitem3$1] (T3_1);       ! Item 3 Thresh 2
   [mitem4$1] (T4_2);       ! Item 4 Thresh 1
   [mitem5$1] (T5_3);       ! Item 5 Thresh 2
   [mitem6$1] (T6_1);       ! Item 6 Thresh 2
   [mitem7$1] (T7_1);       ! Item 7 Thresh 1
   [mitem8$1] (T8_2);       ! Item 8 Thresh 1
   [mitem9$1] (T9_1);       ! Item 9 Thresh 2
   [mitem10$1] (T10_2);     ! Item 10 Thresh 1
   [mitem11$1] (T11_3);     ! Item 11 Thresh 2
   [mitem12$1] (T12_1);     ! Item 12 Thresh 2

%sex#2.c#4%  ! Model for Class 4
   [mitem1$1] (T1_1);       ! Item 1 Thresh 1
   [mitem2$1] (T2_2);       ! Item 2 Thresh 1
   [mitem3$1] (T3_2);       ! Item 3 Thresh 2
   [mitem4$1] (T4_2);       ! Item 4 Thresh 1
   [mitem5$1] (T5_4);       ! Item 5 Thresh 2
   [mitem6$1] (T6_2);       ! Item 6 Thresh 2
   [mitem7$1] (T7_1);       ! Item 7 Thresh 1
   [mitem8$1] (T8_2);       ! Item 8 Thresh 1
   [mitem9$1] (T9_2);       ! Item 9 Thresh 2
   [mitem10$1] (T10_2);     ! Item 10 Thresh 1
   [mitem11$1] (T11_4);     ! Item 11 Thresh 2
   [mitem12$1] (T12_2);     ! Item 12 Thresh 2

%sex#2.c#5%  ! Model for Class 5
   [mitem1$1] (T1_2);       ! Item 1 Thresh 2
   [mitem2$1] (T2_1);       ! Item 2 Thresh 1
   [mitem3$1] (T3_1);       ! Item 3 Thresh 1
   [mitem4$1] (T4_3);       ! Item 4 Thresh 3
   [mitem5$1] (T5_1);       ! Item 5 Thresh 1
   [mitem6$1] (T6_3);       ! Item 6 Thresh 3
   [mitem7$1] (T7_2);       ! Item 7 Thresh 2
   [mitem8$1] (T8_1);       ! Item 8 Thresh 1
   [mitem9$1] (T9_1);       ! Item 9 Thresh 1
   [mitem10$1] (T10_3);     ! Item 10 Thresh 3
   [mitem11$1] (T11_1);     ! Item 11 Thresh 1
   [mitem12$1] (T12_3);     ! Item 12 Thresh 3

%sex#2.c#6%  ! Model for Class 6
   [mitem1$1] (T1_2);       ! Item 1 Thresh 2
   [mitem2$1] (T2_1);       ! Item 2 Thresh 1
```

```
   [mitem3$1] (T3_2);       ! Item 3 Thresh 2
   [mitem4$1] (T4_3);       ! Item 4 Thresh 3
   [mitem5$1] (T5_2);       ! Item 5 Thresh 2
   [mitem6$1] (T6_4);       ! Item 6 Thresh 4
   [mitem7$1] (T7_2);       ! Item 7 Thresh 2
   [mitem8$1] (T8_1);       ! Item 8 Thresh 1
   [mitem9$1] (T9_2);       ! Item 9 Thresh 2
   [mitem10$1] (T10_3);     ! Item 10 Thresh 3
   [mitem11$1] (T11_2);     ! Item 11 Thresh 2
   [mitem12$1] (T12_4);     ! Item 12 Thresh 4

%sex#2.c#7%  ! Model for Class 7
   [mitem1$1] (T1_2);       ! Item 1 Thresh 2
   [mitem2$1] (T2_2);       ! Item 2 Thresh 2
   [mitem3$1] (T3_1);       ! Item 3 Thresh 1
   [mitem4$1] (T4_4);       ! Item 4 Thresh 4
   [mitem5$1] (T5_3);       ! Item 5 Thresh 3
   [mitem6$1] (T6_3);       ! Item 6 Thresh 3
   [mitem7$1] (T7_2);       ! Item 7 Thresh 2
   [mitem8$1] (T8_2);       ! Item 8 Thresh 2
   [mitem9$1] (T9_1);       ! Item 9 Thresh 1
   [mitem10$1] (T10_4);     ! Item 10 Thresh 4
   [mitem11$1] (T11_3);     ! Item 11 Thresh 3
   [mitem12$1] (T12_3);     ! Item 12 Thresh 3

%sex#2.c#8%  ! Model for Class 8
   [mitem1$1] (T1_2);       ! Item 1 Thresh 2
   [mitem2$1] (T2_2);       ! Item 2 Thresh 2
   [mitem3$1] (T3_2);       ! Item 3 Thresh 2
   [mitem4$1] (T4_4);       ! Item 4 Thresh 4
   [mitem5$1] (T5_4);       ! Item 5 Thresh 4
   [mitem6$1] (T6_4);       ! Item 6 Thresh 4
   [mitem7$1] (T7_2);       ! Item 7 Thresh 2
   [mitem8$1] (T8_2);       ! Item 8 Thresh 2
   [mitem9$1] (T9_2);       ! Item 9 Thresh 2
   [mitem10$1] (T10_4);     ! Item 10 Thresh 4
   [mitem11$1] (T11_4);     ! Item 11 Thresh 4
   [mitem12$1] (T12_4);     ! Item 12 Thresh 4


MODEL CONSTRAINT:  ! Used to define LCDM parameters
! Mplus uses P(X=0) rather than P(X=1) so multiply by -1

! STRUCTURAL MODEL 1
NEW(G_11_1*-1 G_12_1*-1 G_13_1*-1 G_212_1*1 G_213_1*1 G_223_1*1);
m11=  - (G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
```

```
m21= G_13_1 - (G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m31= G_12_1 - (G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m41= G_12_1+G_13_1+G_223_1 -
(G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m51= G_11_1 - (G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m61= G_11_1+G_13_1+G_213_1 -
(G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m71= G_11_1+G_12_1+G_212_1 -
(G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);

! STRUCTURAL MODEL 2
NEW(G_11_2*-1 G_12_2*-1 G_13_2*-1 G_212_2*1 G_213_2*1 G_223_2*1);
m12=  - (G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m22= G_13_2  - (G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m32= G_12_2 - (G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m42= G_12_2+G_13_2+G_223_2  -
(G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m52= G_11_2  - (G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m62= G_11_2+G_13_2+G_213_2 -
(G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m72= G_11_2+G_12_2+G_212_2 -
(G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);

! Item 1: Define LCDM parameters present for item 1
NEW(L1_0 L1_11);
T1_1=-(L1_0);                     ! Item 1 Thresh 1
T1_2=-(L1_0+L1_11);                  ! Item 1 Thresh 2
! Main effect order constraints
L1_11>0;

! Item 2: Define LCDM parameters present for item 2
NEW(L2_0 L2_12);
T2_1=-(L2_0);                     ! Item 2 Thresh 1
T2_2=-(L2_0+L2_12);                  ! Item 2 Thresh 2
! Main effect order constraints
L2_12>0;


! Item 3: Define LCDM parameters present for item 3
NEW(L3_0 L3_13);
T3_1=-(L3_0);                     ! Item 3 Thresh 1
T3_2=-(L3_0+L3_13);                  ! Item 3 Thresh 2
! Main effect order constraints
L3_13>0;
```

```
! Item 4: Define LCDM parameters present for item 4
NEW(L4_0 L4_11 L4_12 L4_212);
T4_1=-(L4_0);                                 ! Item 4 Thresh 1
T4_2=-(L4_0+L4_12);                                ! Item 4 Thresh 2
T4_3=-(L4_0+L4_11);                                ! Item 4 Thresh 3
T4_4=-(L4_0+L4_11+L4_12+L4_212);                        ! Item 4 Thresh 4
! Main effect order constraints
L4_11>0; L4_12>0;
! Two-way interaction order constraints
L4_212>-L4_11;
L4_212>-L4_12;


! Item 5: Define LCDM parameters present for item 5
NEW(L5_0 L5_12 L5_13 L5_223);
T5_1=-(L5_0);                                 ! Item 5 Thresh 1
T5_2=-(L5_0+L5_13);                                ! Item 5 Thresh 2
T5_3=-(L5_0+L5_12);                                ! Item 5 Thresh 3
T5_4=-(L5_0+L5_12+L5_13+L5_223);                        ! Item 5 Thresh 4
! Main effect order constraints
L5_12>0; L5_13>0;
! Two-way interaction order constraints
L5_223>-L5_12;
L5_223>-L5_13;


! Item 6: Define LCDM parameters present for item 6
NEW(L6_0 L6_11 L6_13 L6_213);
T6_1=-(L6_0);                                 ! Item 6 Thresh 1
T6_2=-(L6_0+L6_13);                                ! Item 6 Thresh 2
T6_3=-(L6_0+L6_11);                                ! Item 6 Thresh 3
T6_4=-(L6_0+L6_11+L6_13+L6_213);                        ! Item 6 Thresh 4
! Main effect order constraints
L6_11>0; L6_13>0;
! Two-way interaction order constraints
L6_213>-L6_11;
L6_213>-L6_13;


! Item 7: Define LCDM parameters present for item 7
NEW(L7_0 L7_11);
T7_1=-(L7_0);                                 ! Item 7 Thresh 1
T7_2=-(L7_0+L7_11);                                ! Item 7 Thresh 2
! Main effect order constraints
L7_11>0;
```

```
! Item 8: Define LCDM parameters present for item 8
NEW(L8_0 L8_12);
T8_1=-(L8_0);                          ! Item 8 Thresh 1
T8_2=-(L8_0+L8_12);                    ! Item 8 Thresh 2
! Main effect order constraints
L8_12>0;


! Item 9: Define LCDM parameters present for item 9
NEW(L9_0 L9_13);
T9_1=-(L9_0);                          ! Item 9 Thresh 1
T9_2=-(L9_0+L9_13);                    ! Item 9 Thresh 2
! Main effect order constraints
L9_13>0;


! Item 10: Define LCDM parameters present for item 10
NEW(L10_0 L10_11 L10_12 L10_212);
T10_1=-(L10_0);                        ! Item 10 Thresh 1
T10_2=-(L10_0+L10_12);                 ! Item 10 Thresh 2
T10_3=-(L10_0+L10_11);                 ! Item 10 Thresh 3
T10_4=-(L10_0+L10_11+L10_12+L10_212);  ! Item 10 Thresh 4
! Main effect order constraints
L10_11>0; L10_12>0;
! Two-way interaction order constraints
L10_212>-L10_11;
L10_212>-L10_12;


! Item 11: Define LCDM parameters present for item 11
NEW(L11_0 L11_12 L11_13 L11_223);
T11_1=-(L11_0);                        ! Item 11 Thresh 1
T11_2=-(L11_0+L11_13);                 ! Item 11 Thresh 2
T11_3=-(L11_0+L11_12);                 ! Item 11 Thresh 3
T11_4=-(L11_0+L11_12+L11_13+L11_223);  ! Item 11 Thresh 4
! Main effect order constraints
L11_12>0; L11_13>0;
! Two-way interaction order constraints
L11_223>-L11_12;
L11_223>-L11_13;


! Item 12: Define LCDM parameters present for item 12
NEW(L12_0 L12_11 L12_13 L12_213);
T12_1=-(L12_0);                        ! Item 12 Thresh 1
```

```
T12_2=-(L12_0+L12_13);                        ! Item 12 Thresh 2
T12_3=-(L12_0+L12_11);                        ! Item 12 Thresh 3
T12_4=-(L12_0+L12_11+L12_13+L12_213);                 ! Item 12 Thresh 4
! Main effect order constraints
L12_11>0; L12_13>0;
! Two-way interaction order constraints
L12_213>-L12_11;
L12_213>-L12_13;

OUTPUT:
   TECH10;  ! Request additional model fit statistics

SAVEDATA: ! Format, name of posterior probabilities of class membership file
   FORMAT = F10.5;
   FILE = model03_exam.dat;
   SAVE = CPROBABILITIES;
```

Model 4 Estimation Code

```
TITLE:  ! Section that appears in header of output file


DATA:  ! Location of free format data file
   FILE = data.csv;

VARIABLE:
   NAMES = id group class mitem1-mitem12;        ! List of variables in data file
   USEVARIABLE = mitem1-mitem12;      ! Variables to be analyzed
   CATEGORICAL = mitem1-mitem12;      ! Binary outcomes
   CLASSES = sex(2) c(8);        !classes and group
   KNOWNCLASS = sex(group = 0 group = 1);
   IDVARIABLE = id;
   AUXILIARY = class;

ANALYSIS:
   TYPE = MIXTURE;                  ! Estimates latent classes
   STARTS = 0;                  ! Turn off multiple random start feature
   PROCESSORS =16;                  ! Number of processors available
   MCITERATIONS = 2;
   MUITERATIONS = 2;
   MITERATIONS = 1000;
   MCONVERGENCE = .0001;

MODEL:

%OVERALL%

c#1 on sex (m12);
c#2 on sex (m22);
c#3 on sex (m32);
c#4 on sex (m42);
c#5 on sex (m52);
c#6 on sex (m62);
c#7 on sex (m72);

[c#1] (m11);
[c#2] (m21);
[c#3] (m31);
[c#4] (m41);
[c#5] (m51);
[c#6] (m61);
[c#7] (m71);
```

```
%sex#1.c#1% ! Model for Class 1
   [mitem1$1] (T1_1);       ! Item 1 Thresh 1
   [mitem2$1] (T2_1);       ! Item 2 Thresh 1
   [mitem3$1] (T3_1);       ! Item 3 Thresh 1
   [mitem4$1] (T4_1);       ! Item 4 Thresh 1
   [mitem5$1] (T5_1_1);       ! Item 5 Thresh 1
   [mitem6$1] (T6_1);       ! Item 6 Thresh 1
   [mitem7$1] (T7_1_1);       ! Item 7 Thresh 1
   [mitem8$1] (T8_1);       ! Item 8 Thresh 1
   [mitem9$1] (T9_1);       ! Item 9 Thresh 1
   [mitem10$1] (T10_1);     ! Item 10 Thresh 1
   [mitem11$1] (T11_1);     ! Item 11 Thresh 1
   [mitem12$1] (T12_1);     ! Item 12 Thresh 1

%sex#1.c#2%  ! Model for Class 2
   [mitem1$1] (T1_1);       ! Item 1 Thresh 1
   [mitem2$1] (T2_1);       ! Item 2 Thresh 1
   [mitem3$1] (T3_2);       ! Item 3 Thresh 2
   [mitem4$1] (T4_1);       ! Item 4 Thresh 1
   [mitem5$1] (T5_2_1);       ! Item 5 Thresh 2
   [mitem6$1] (T6_2);       ! Item 6 Thresh 2
   [mitem7$1] (T7_1_1);       ! Item 7 Thresh 1
   [mitem8$1] (T8_1);       ! Item 8 Thresh 1
   [mitem9$1] (T9_2);       ! Item 9 Thresh 2
   [mitem10$1] (T10_1);     ! Item 10 Thresh 1
   [mitem11$1] (T11_2);     ! Item 11 Thresh 2
   [mitem12$1] (T12_2);     ! Item 12 Thresh 2


%sex#1.c#3%  ! Model for Class 3
   [mitem1$1] (T1_1);       ! Item 1 Thresh 1
   [mitem2$1] (T2_2);       ! Item 2 Thresh 1
   [mitem3$1] (T3_1);       ! Item 3 Thresh 2
   [mitem4$1] (T4_2);       ! Item 4 Thresh 1
   [mitem5$1] (T5_3_1);       ! Item 5 Thresh 2
   [mitem6$1] (T6_1);       ! Item 6 Thresh 2
   [mitem7$1] (T7_1_1);       ! Item 7 Thresh 1
   [mitem8$1] (T8_2);       ! Item 8 Thresh 1
   [mitem9$1] (T9_1);       ! Item 9 Thresh 2
   [mitem10$1] (T10_2);     ! Item 10 Thresh 1
   [mitem11$1] (T11_3);     ! Item 11 Thresh 2
   [mitem12$1] (T12_1);     ! Item 12 Thresh 2

%sex#1.c#4%  ! Model for Class 4
   [mitem1$1] (T1_1);       ! Item 1 Thresh 1
```

```
   [mitem2$1] (T2_2);       ! Item 2 Thresh 1
   [mitem3$1] (T3_2);       ! Item 3 Thresh 2
   [mitem4$1] (T4_2);       ! Item 4 Thresh 1
   [mitem5$1] (T5_4_1);       ! Item 5 Thresh 2
   [mitem6$1] (T6_2);       ! Item 6 Thresh 2
   [mitem7$1] (T7_1_1);       ! Item 7 Thresh 1
   [mitem8$1] (T8_2);       ! Item 8 Thresh 1
   [mitem9$1] (T9_2);       ! Item 9 Thresh 2
   [mitem10$1] (T10_2);     ! Item 10 Thresh 1
   [mitem11$1] (T11_4);     ! Item 11 Thresh 2
   [mitem12$1] (T12_2);     ! Item 12 Thresh 2

%sex#1.c#5%  ! Model for Class 5
   [mitem1$1] (T1_2);       ! Item 1 Thresh 2
   [mitem2$1] (T2_1);       ! Item 2 Thresh 1
   [mitem3$1] (T3_1);       ! Item 3 Thresh 1
   [mitem4$1] (T4_3);       ! Item 4 Thresh 3
   [mitem5$1] (T5_1_1);       ! Item 5 Thresh 1
   [mitem6$1] (T6_3);       ! Item 6 Thresh 3
   [mitem7$1] (T7_2_1);       ! Item 7 Thresh 2
   [mitem8$1] (T8_1);       ! Item 8 Thresh 1
   [mitem9$1] (T9_1);       ! Item 9 Thresh 1
   [mitem10$1] (T10_3);     ! Item 10 Thresh 3
   [mitem11$1] (T11_1);     ! Item 11 Thresh 1
   [mitem12$1] (T12_3);     ! Item 12 Thresh 3

%sex#1.c#6%  ! Model for Class 6
   [mitem1$1] (T1_2);       ! Item 1 Thresh 2
   [mitem2$1] (T2_1);       ! Item 2 Thresh 1
   [mitem3$1] (T3_2);       ! Item 3 Thresh 2
   [mitem4$1] (T4_3);       ! Item 4 Thresh 3
   [mitem5$1] (T5_2_1);       ! Item 5 Thresh 2
   [mitem6$1] (T6_4);       ! Item 6 Thresh 4
   [mitem7$1] (T7_2_1);       ! Item 7 Thresh 2
   [mitem8$1] (T8_1);       ! Item 8 Thresh 1
   [mitem9$1] (T9_2);       ! Item 9 Thresh 2
   [mitem10$1] (T10_3);     ! Item 10 Thresh 3
   [mitem11$1] (T11_2);     ! Item 11 Thresh 2
   [mitem12$1] (T12_4);     ! Item 12 Thresh 4

%sex#1.c#7%  ! Model for Class 7
   [mitem1$1] (T1_2);       ! Item 1 Thresh 2
   [mitem2$1] (T2_2);       ! Item 2 Thresh 2
   [mitem3$1] (T3_1);       ! Item 3 Thresh 1
   [mitem4$1] (T4_4);       ! Item 4 Thresh 4
   [mitem5$1] (T5_3_1);       ! Item 5 Thresh 3
```

```
    [mitem6$1] (T6_3);       ! Item 6 Thresh 3
    [mitem7$1] (T7_2_1);       ! Item 7 Thresh 2
    [mitem8$1] (T8_2);       ! Item 8 Thresh 2
    [mitem9$1] (T9_1);       ! Item 9 Thresh 1
    [mitem10$1] (T10_4);     ! Item 10 Thresh 4
    [mitem11$1] (T11_3);     ! Item 11 Thresh 3
    [mitem12$1] (T12_3);     ! Item 12 Thresh 3

%sex#1.c#8%  ! Model for Class 8
    [mitem1$1] (T1_2);       ! Item 1 Thresh 2
    [mitem2$1] (T2_2);       ! Item 2 Thresh 2
    [mitem3$1] (T3_2);       ! Item 3 Thresh 2
    [mitem4$1] (T4_4);       ! Item 4 Thresh 4
    [mitem5$1] (T5_4_1);       ! Item 5 Thresh 4
    [mitem6$1] (T6_4);       ! Item 6 Thresh 4
    [mitem7$1] (T7_2_1);       ! Item 7 Thresh 2
    [mitem8$1] (T8_2);       ! Item 8 Thresh 2
    [mitem9$1] (T9_2);       ! Item 9 Thresh 2
    [mitem10$1] (T10_4);     ! Item 10 Thresh 4
    [mitem11$1] (T11_4);     ! Item 11 Thresh 4
    [mitem12$1] (T12_4);     ! Item 12 Thresh 4

%sex#2.c#1% ! Model for Class 1
    [mitem1$1] (T1_1);       ! Item 1 Thresh 1
    [mitem2$1] (T2_1);       ! Item 2 Thresh 1
    [mitem3$1] (T3_1);       ! Item 3 Thresh 1
    [mitem4$1] (T4_1);       ! Item 4 Thresh 1
    [mitem5$1] (T5_1_2);       ! Item 5 Thresh 1
    [mitem6$1] (T6_1);       ! Item 6 Thresh 1
    [mitem7$1] (T7_1_2);       ! Item 7 Thresh 1
    [mitem8$1] (T8_1);       ! Item 8 Thresh 1
    [mitem9$1] (T9_1);       ! Item 9 Thresh 1
    [mitem10$1] (T10_1);     ! Item 10 Thresh 1
    [mitem11$1] (T11_1);     ! Item 11 Thresh 1
    [mitem12$1] (T12_1);     ! Item 12 Thresh 1

%sex#2.c#2%
    [mitem1$1] (T1_1);       ! Item 1 Thresh 1
    [mitem2$1] (T2_1);       ! Item 2 Thresh 1
    [mitem3$1] (T3_2);       ! Item 3 Thresh 2
    [mitem4$1] (T4_1);       ! Item 4 Thresh 1
    [mitem5$1] (T5_2_2);       ! Item 5 Thresh 2
    [mitem6$1] (T6_2);       ! Item 6 Thresh 2
    [mitem7$1] (T7_1_2);       ! Item 7 Thresh 1
    [mitem8$1] (T8_1);       ! Item 8 Thresh 1
    [mitem9$1] (T9_2);       ! Item 9 Thresh 2
```

```
    [mitem10$1] (T10_1);     ! Item 10 Thresh 1
    [mitem11$1] (T11_2);     ! Item 11 Thresh 2
    [mitem12$1] (T12_2);     ! Item 12 Thresh 2


%sex#2.c#3% ! Model for Class 1
    [mitem1$1] (T1_1);       ! Item 1 Thresh 1
    [mitem2$1] (T2_2);       ! Item 2 Thresh 1
    [mitem3$1] (T3_1);       ! Item 3 Thresh 2
    [mitem4$1] (T4_2);       ! Item 4 Thresh 1
    [mitem5$1] (T5_3_2);       ! Item 5 Thresh 2
    [mitem6$1] (T6_1);       ! Item 6 Thresh 2
    [mitem7$1] (T7_1_2);       ! Item 7 Thresh 1
    [mitem8$1] (T8_2);       ! Item 8 Thresh 1
    [mitem9$1] (T9_1);       ! Item 9 Thresh 2
    [mitem10$1] (T10_2);     ! Item 10 Thresh 1
    [mitem11$1] (T11_3);     ! Item 11 Thresh 2
    [mitem12$1] (T12_1);     ! Item 12 Thresh 2

%sex#2.c#4%
    [mitem1$1] (T1_1);       ! Item 1 Thresh 1
    [mitem2$1] (T2_2);       ! Item 2 Thresh 1
    [mitem3$1] (T3_2);       ! Item 3 Thresh 2
    [mitem4$1] (T4_2);       ! Item 4 Thresh 1
    [mitem5$1] (T5_4_2);       ! Item 5 Thresh 2
    [mitem6$1] (T6_2);       ! Item 6 Thresh 2
    [mitem7$1] (T7_1_2);       ! Item 7 Thresh 1
    [mitem8$1] (T8_2);       ! Item 8 Thresh 1
    [mitem9$1] (T9_2);       ! Item 9 Thresh 2
    [mitem10$1] (T10_2);     ! Item 10 Thresh 1
    [mitem11$1] (T11_4);     ! Item 11 Thresh 2
    [mitem12$1] (T12_2);     ! Item 12 Thresh 2

%sex#2.c#5%
    [mitem1$1] (T1_2);       ! Item 1 Thresh 2
    [mitem2$1] (T2_1);       ! Item 2 Thresh 1
    [mitem3$1] (T3_1);       ! Item 3 Thresh 1
    [mitem4$1] (T4_3);       ! Item 4 Thresh 3
    [mitem5$1] (T5_1_2);       ! Item 5 Thresh 1
    [mitem6$1] (T6_3);       ! Item 6 Thresh 3
    [mitem7$1] (T7_2_2);       ! Item 7 Thresh 2
    [mitem8$1] (T8_1);       ! Item 8 Thresh 1
    [mitem9$1] (T9_1);       ! Item 9 Thresh 1
    [mitem10$1] (T10_3);     ! Item 10 Thresh 3
    [mitem11$1] (T11_1);     ! Item 11 Thresh 1
    [mitem12$1] (T12_3);     ! Item 12 Thresh 3
```

```
%sex#2.c#6%
   [mitem1$1] (T1_2);       ! Item 1 Thresh 2
   [mitem2$1] (T2_1);       ! Item 2 Thresh 1
   [mitem3$1] (T3_2);       ! Item 3 Thresh 2
   [mitem4$1] (T4_3);       ! Item 4 Thresh 3
   [mitem5$1] (T5_2_2);       ! Item 5 Thresh 2
   [mitem6$1] (T6_4);       ! Item 6 Thresh 4
   [mitem7$1] (T7_2_2);       ! Item 7 Thresh 2
   [mitem8$1] (T8_1);       ! Item 8 Thresh 1
   [mitem9$1] (T9_2);       ! Item 9 Thresh 2
   [mitem10$1] (T10_3);     ! Item 10 Thresh 3
   [mitem11$1] (T11_2);     ! Item 11 Thresh 2
   [mitem12$1] (T12_4);     ! Item 12 Thresh 4

%sex#2.c#7%
   [mitem1$1] (T1_2);       ! Item 1 Thresh 2
   [mitem2$1] (T2_2);       ! Item 2 Thresh 2
   [mitem3$1] (T3_1);       ! Item 3 Thresh 1
   [mitem4$1] (T4_4);       ! Item 4 Thresh 4
   [mitem5$1] (T5_3_2);       ! Item 5 Thresh 3
   [mitem6$1] (T6_3);       ! Item 6 Thresh 3
   [mitem7$1] (T7_2_2);       ! Item 7 Thresh 2
   [mitem8$1] (T8_2);       ! Item 8 Thresh 2
   [mitem9$1] (T9_1);       ! Item 9 Thresh 1
   [mitem10$1] (T10_4);     ! Item 10 Thresh 4
   [mitem11$1] (T11_3);     ! Item 11 Thresh 3
   [mitem12$1] (T12_3);     ! Item 12 Thresh 3

%sex#2.c#8%
   [mitem1$1] (T1_2);       ! Item 1 Thresh 2
   [mitem2$1] (T2_2);       ! Item 2 Thresh 2
   [mitem3$1] (T3_2);       ! Item 3 Thresh 2
   [mitem4$1] (T4_4);       ! Item 4 Thresh 4
   [mitem5$1] (T5_4_2);       ! Item 5 Thresh 4
   [mitem6$1] (T6_4);       ! Item 6 Thresh 4
   [mitem7$1] (T7_2_2);       ! Item 7 Thresh 2
   [mitem8$1] (T8_2);       ! Item 8 Thresh 2
   [mitem9$1] (T9_2);       ! Item 9 Thresh 2
   [mitem10$1] (T10_4);     ! Item 10 Thresh 4
   [mitem11$1] (T11_4);     ! Item 11 Thresh 4
   [mitem12$1] (T12_4);     ! Item 12 Thresh 4


MODEL CONSTRAINT:  ! Used to define LCDM parameters
! Mplus uses P(X=0) rather than P(X=1) so multiply by -1
```

```
! STRUCTURAL MODEL 1
NEW(G_11_1*-1 G_12_1*-1 G_13_1*-1 G_212_1*1 G_213_1*1 G_223_1*1);
m11=  - (G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m21= G_13_1 - (G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m31= G_12_1 - (G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m41= G_12_1+G_13_1+G_223_1 -
(G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m51= G_11_1 - (G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m61= G_11_1+G_13_1+G_213_1 -
(G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);
m71= G_11_1+G_12_1+G_212_1 -
(G_11_1+G_12_1+G_13_1+G_212_1+G_213_1+G_223_1);

! STRUCTURAL MODEL 2
NEW(G_11_2*-1 G_12_2*-1 G_13_2*-1 G_212_2*1 G_213_2*1 G_223_2*1);
m12=  - (G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m22= G_13_2  - (G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m32= G_12_2 - (G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m42= G_12_2+G_13_2+G_223_2  -
(G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m52= G_11_2  - (G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m62= G_11_2+G_13_2+G_213_2 -
(G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);
m72= G_11_2+G_12_2+G_212_2 -
(G_11_2+G_12_2+G_13_2+G_212_2+G_213_2+G_223_2);

NEW(DG_11*0 DG_12*0 DG_13*0 DG_212*0 DG_213*0 DG_223*0);
DG_11 = G_11_2-G_11_1;
DG_12 = G_12_2-G_12_1;
DG_13 = G_13_2-G_13_1;
DG_212 = G_212_2-G_212_1;
DG_213 = G_213_2-G_213_1;
DG_223 = G_223_2-G_223_1;

! Item 1: Define LCDM parameters present for item 1
NEW(L1_0 L1_11);
T1_1=-(L1_0);                          ! Item 1 Thresh 1
T1_2=-(L1_0+L1_11);                       ! Item 1 Thresh 2
! Main effect order constraints
L1_11>0;

! Item 2: Define LCDM parameters present for item 2
NEW(L2_0 L2_12);
T2_1=-(L2_0);                          ! Item 2 Thresh 1
T2_2=-(L2_0+L2_12);                       ! Item 2 Thresh 2
```

```
! Main effect order constraints
L2_12>0;


! Item 3: Define LCDM parameters present for item 3
NEW(L3_0 L3_13);
T3_1=-(L3_0);                          ! Item 3 Thresh 1
T3_2=-(L3_0+L3_13);                    ! Item 3 Thresh 2
! Main effect order constraints
L3_13>0;


! Item 4: Define LCDM parameters present for item 4
NEW(L4_0 L4_11 L4_12 L4_212);
T4_1=-(L4_0);                          ! Item 4 Thresh 1
T4_2=-(L4_0+L4_12);                    ! Item 4 Thresh 2
T4_3=-(L4_0+L4_11);                    ! Item 4 Thresh 3
T4_4=-(L4_0+L4_11+L4_12+L4_212);       ! Item 4 Thresh 4
! Main effect order constraints
L4_11>0; L4_12>0;
! Two-way interaction order constraints
L4_212>-L4_11;
L4_212>-L4_12;


! Item 5: Define LCDM parameters present for item 5
NEW(L5_0_1 L5_12_1 L5_13_1 L5_223_1 L5_0_2 L5_12_2 L5_13_2 L5_223_2);
T5_1_1=-(L5_0_1);                                  ! Item 29 Thresh 1
T5_2_1=-(L5_0_1+L5_13_1);                          ! Item 29 Thresh 2
T5_3_1=-(L5_0_1+L5_12_1);                          ! Item 29 Thresh 3
T5_4_1=-(L5_0_1+L5_12_1+L5_13_1+L5_223_1);         ! Item 29 Thresh 4
! Main effect order constraints
L5_12_1>0; L5_13_1>0;
! Two-way interaction order constraints
L5_223_1>-L5_12_1;
L5_223_1>-L5_13_1;

T5_1_2=-(L5_0_2);                                  ! Item 29 Thresh 1
T5_2_2=-(L5_0_2+L5_13_2);                          ! Item 29 Thresh 2
T5_3_2=-(L5_0_2+L5_12_2);                          ! Item 29 Thresh 3
T5_4_2=-(L5_0_2+L5_12_2+L5_13_2+L5_223_2);         ! Item 29 Thresh 4
! Main effect order constraints
L5_12_2>0; L5_13_2>0;
! Two-way interaction order constraints
L5_223_2>-L5_12_2;
L5_223_2>-L5_13_2;
```

```
NEW(D5_0 D5_12 D5_13 D5_223);
D5_0=L5_0_2-L5_0_1;
D5_12=L5_12_2-L5_12_1;
D5_13=L5_13_2-L5_13_1;
D5_223=L5_223_2-L5_223_1;

! Item 6: Define LCDM parameters present for item 6
NEW(L6_0 L6_11 L6_13 L6_213);
T6_1=-(L6_0);                       ! Item 6 Thresh 1
T6_2=-(L6_0+L6_13);                 ! Item 6 Thresh 2
T6_3=-(L6_0+L6_11);                 ! Item 6 Thresh 3
T6_4=-(L6_0+L6_11+L6_13+L6_213);    ! Item 6 Thresh 4
! Main effect order constraints
L6_11>0; L6_13>0;
! Two-way interaction order constraints
L6_213>-L6_11;
L6_213>-L6_13;


! Item 7: Define LCDM parameters present for item 7
NEW(L7_0_1 L7_11_1 L7_0_2 L7_11_2);
T7_1_1=-(L7_0_1);                   ! Item 1 Thresh 1
T7_2_1=-(L7_0_1+L7_11_1);           ! Item 1 Thresh 2
! Main effect order constraints
L7_11_1>0;

T7_1_2=-(L7_0_2);                   ! Item 1 Thresh 1
T7_2_2=-(L7_0_2+L7_11_2);           ! Item 1 Thresh 2
! Main effect order constraints
L7_11_2>0;

NEW(D7_0 D7_11);
D7_0 = L7_0_2-L7_0_1;
D7_11 = L7_11_2-L7_11_1;

! Item 8: Define LCDM parameters present for item 8
NEW(L8_0 L8_12);
T8_1=-(L8_0);                       ! Item 8 Thresh 1
T8_2=-(L8_0+L8_12);                 ! Item 8 Thresh 2
! Main effect order constraints
L8_12>0;


! Item 9: Define LCDM parameters present for item 9
NEW(L9_0 L9_13);
```

```
T9_1=-(L9_0);                                   ! Item 9 Thresh 1
T9_2=-(L9_0+L9_13);                                 ! Item 9 Thresh 2
! Main effect order constraints
L9_13>0;


! Item 10: Define LCDM parameters present for item 10
NEW(L10_0 L10_11 L10_12 L10_212);
T10_1=-(L10_0);                                 ! Item 10 Thresh 1
T10_2=-(L10_0+L10_12);                              ! Item 10 Thresh 2
T10_3=-(L10_0+L10_11);                              ! Item 10 Thresh 3
T10_4=-(L10_0+L10_11+L10_12+L10_212);                       ! Item 10 Thresh 4
! Main effect order constraints
L10_11>0; L10_12>0;
! Two-way interaction order constraints
L10_212>-L10_11;
L10_212>-L10_12;


! Item 11: Define LCDM parameters present for item 11
NEW(L11_0 L11_12 L11_13 L11_223);
T11_1=-(L11_0);                                 ! Item 11 Thresh 1
T11_2=-(L11_0+L11_13);                              ! Item 11 Thresh 2
T11_3=-(L11_0+L11_12);                              ! Item 11 Thresh 3
T11_4=-(L11_0+L11_12+L11_13+L11_223);                       ! Item 11 Thresh 4
! Main effect order constraints
L11_12>0; L11_13>0;
! Two-way interaction order constraints
L11_223>-L11_12;
L11_223>-L11_13;


! Item 12: Define LCDM parameters present for item 12
NEW(L12_0 L12_11 L12_13 L12_213);
T12_1=-(L12_0);                                 ! Item 12 Thresh 1
T12_2=-(L12_0+L12_13);                              ! Item 12 Thresh 2
T12_3=-(L12_0+L12_11);                              ! Item 12 Thresh 3
T12_4=-(L12_0+L12_11+L12_13+L12_213);                       ! Item 12 Thresh 4
! Main effect order constraints
L12_11>0; L12_13>0;
! Two-way interaction order constraints
L12_213>-L12_11;
L12_213>-L12_13;

OUTPUT:
  TECH10;  ! Request additional model fit statistics
```

```
SAVEDATA: ! Format, name of posterior probabilities of class membership file
   FORMAT = F10.5;
   FILE = model04_exam.dat;
   SAVE = CPROBABILITIES;
```