Measurement Invariance

Latent Trait Measurement and
Structural Equation Models
Lecture #13

April 24, 2013

Measurement Invariance in CFA

- Today's topics:
  - What is measurement invariance?

  - 2 major types of invariance
    - Measurement and Structural

  - Sequence of tests for invariance
    - Metric, Scalar, Residual…. then Structural

MEASUREMENT INVARIANCE

What is 'Measurement Invariance'?

- aka 'factorial invariance' and 'measurement equivalence'

- Concerns the extent to which are the psychometric properties of the observed indicators are transportable (generalizable) across groups or over time/condition
  - In other words, that we are measuring *the same construct in the same way* in different groups or over time/condition
  - In other words, observed scores should depend *only on latent construct scores*, and not on group membership or occasion
  - In other words, that observed differences between groups reflect TRUE differences in the amount or variability of the construct

- Relevant concern in many applied settings
  - e.g., across cultures, language, age, modality

2 Major Types of Factorial Invariance

- **Measurement Invariance** concerns how the items measure the latent construct across groups or over time
  - Measurement model invariance: same factor loadings, same item intercepts, (possibly) same residual (co)variances
  - **Measurement model invariance is a precursor to ANY group comparison (whether it is tested/acknowledged or not)**

- Measurement invariance is often assumed, not tested
  - Even a t-test assumes measurement invariance
  - Modeling change over time assumes measurement invariance
  - People tend to accept this assumption unless you try to use a factor model… then they usually insist on testing invariance

2 Major Types of Factorial Invariance

- **Structural Invariance** concerns how the latent factors are distributed and related in the separate populations
  - Structural model invariance: same factor variances and covariances (or same higher-order structure) and factor means

- Structural invariance may not hold… and that's ok
  - Assuming measurement invariance holds, structural invariance represents '**real**' differences in the construct across groups/time
  - Structural non-invariance does **not** indicate a problem with your instrument – group structural differences may be of interest
    - e.g., real growth of factors over time
    - e.g., differentiation or de-differentiation of latent traits

Today's Example

- We return to our familiar gambling data: our one factor model…but with 10 items
- To demonstrate our invariance example, we will seek to test factorial invariance of the two groups of people we have: students versus experienced gamblers
- We expect these groups will be different, but we will seek to determine this statistically
  - We will use the robust ML correction to the model Chi-Square, which will alter the way we do the likelihood ratio tests

Overall Syntax for 10-Item/1-Factor Model

```
TITLE:
    Gambling Research Instrument Items
    Data from 1192 College Students/144 Gamblers
    41 Likert Items (1-6): GRI1-GRI41
    12 SOGS items (SOGS4-SOGS15), mostly dichotomous
    ================================================================
    Identification: Marker Item Factor Variance, Zero Factor Mean
    ================================================================
    One-Factor GAMBLING tendencies model with 10 GRI items
    MLR is used to adjust the model Chi-Square
DATA:
    FILE = gamblingdata.csv;

VARIABLE:
    NAMES = GRI1-GRI41 SOGS4-SOGS15 Student ID;
    USEVARIABLES = GRI1 GRI3 GRI5 GRI9 GRI10 GRI13
                   GRI14 GRI18 GRI21 GRI23;
    IDVARIABLE = ID;
    MISSING = ALL(99);

ANALYSIS:
    ESTIMATOR = MLR;

MODEL:
    GAMBLING by GRI1 GRI3 GRI5 GRI9 GRI10 GRI13 GRI14 GRI18 GRI21 GRI23;

OUTPUT:
    STANDARDIZED MODINDICES(ALL 0) RESIDUAL;

PLOT:
    TYPE = PLOT1 PLOT2 PLOT3;
```

Overall Analysis Results: Model Fit

```
MODEL FIT INFORMATION

Number of Free Parameters                         30

Loglikelihood

          H0 Value                        -16648.054
          H0 Scaling Correction Factor         2.366
            for MLR
          H1 Value                        -16567.417
          H1 Scaling Correction Factor         1.938
            for MLR
```

```
RMSEA (Root Mean Square Error Of Approximation)

          Estimate                           0.038
          90 Percent C.I.                    0.030  0.047
          Probability RMSEA <= .05           0.989

CFI/TLI

          CFI                                0.969
          TLI                                0.960
```

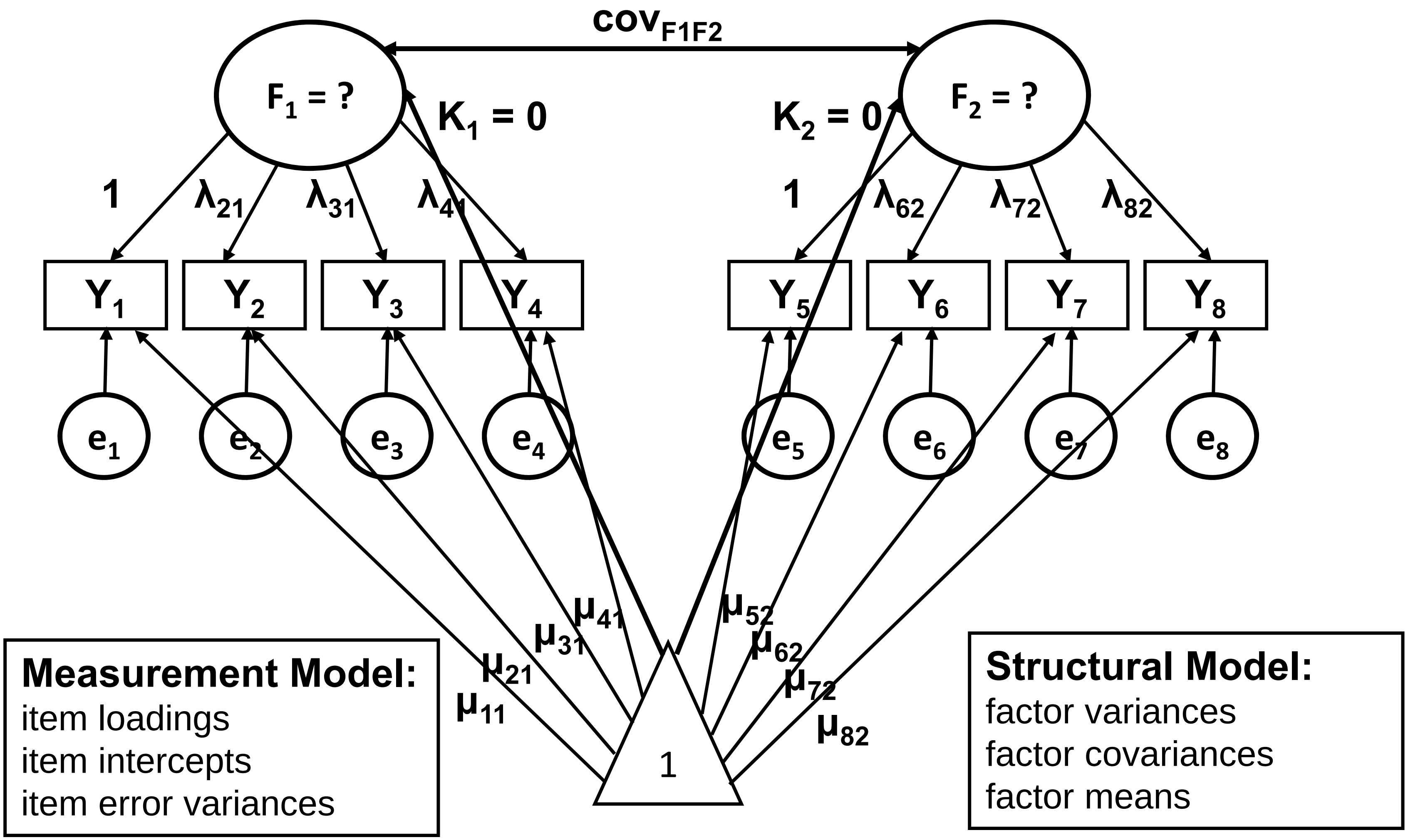
```
Chi-Square Test of Model Fit

          Value                            102.635*
          Degrees of Freedom                    35
          P-Value                           0.0000
          Scaling Correction Factor          1.571
            for MLR

*   The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used
    for chi-square difference testing in the regular way.  MLM, MLR and WLSM
    chi-square difference testing is described on the Mplus website.  MLMV, WLSMV,
    and ULSMV difference testing is done using the DIFFTEST option.
```

CFA Configural Baseline Model: *Marker Item for Factor Variance*

For a two-factor model:



Levels of Invariance across Groups

- **Step 0: Omnibus test of equality of the overall indicator covariance matrix across groups**
  - Do the matrices differ between groups, on the whole?
  - If not, game over. You are done. You have invariance. Congratulations.
  - Many people disagree with the necessity or usefulness of this test to begin testing invariance… why might that be?
    - People also differ in whether invariance should go from top-down or bottom-up directions… I favor bottom-up for the same reason.

- Let’s proceed with an example with our one-factor gambling data

Logic and Statistics of the Omnibus Test

- The omnibus test compares the full **saturated covariance matrix** for both groups
  - This test is likely to be rejected due to the strictness of its constraints – but if not, we can stop!
- The idea is to treat each group independently
  - Independent groups means that we can represent the covariance matrix for each in a block diagonal form

Omnibus Test: Block Diagonal Covariance Matrix

For both groups, we are building the following:

$$\boldsymbol{\Sigma}_B = \begin{bmatrix} \boldsymbol{\Sigma}_{g_1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{g_2} \end{bmatrix}$$

$\boldsymbol{\Sigma}_{g_1}$ is the 10 items x 10 items covariance matrix for group 1 (the students)

$\boldsymbol{\Sigma}_{g_2}$ is the 10 items x 10 items covariance matrix for group 2 (the non-students)

$\mathbf{0}$ in the off diagonal represents independent groups

The omnibus hypothesis test is then:

$$H_0\colon \boldsymbol{\Sigma}_{g_1} = \boldsymbol{\Sigma}_{g_2}$$
$$H_1\colon \boldsymbol{\Sigma}_{g_1} \neq \boldsymbol{\Sigma}_{g_2}$$

Omnibus Test: Null Model Syntax

```
ANALYSIS:
    ESTIMATOR = MLR;

VARIABLE:
    NAMES = GRI1-GRI41 SOGS4-SOGS15 Student ID;
    USEVARIABLES = GRI1 GRI3 GRI5 GRI9 GRI10 GRI13
                   GRI14 GRI18 GRI21 GRI23;
    IDVARIABLE = ID;
    MISSING = ALL(99);

MODEL:
    GRI1 GRI3 GRI5 GRI9 GRI10 GRI13 GRI14 GRI18 GRI21 GRI23 WITH
    GRI1 GRI3 GRI5 GRI9 GRI10 GRI13 GRI14 GRI18 GRI21 GRI23;
```

Omnibus Test: Alternative Model Syntax

```
ANALYSIS:
    ESTIMATOR = MLR;

VARIABLE:
    NAMES = GRI1-GRI41 SOGS4-SOGS15 Student ID;
    USEVARIABLES = GRI1 GRI3 GRI5 GRI9 GRI10 GRI13
                   GRI14 GRI18 GRI21 GRI23;
    IDVARIABLE = ID;
    MISSING = ALL(99);

    GROUPING = student (1 = student 0 = nonstudent);

MODEL:
    [GRI1] (mean1); [GRI3] (mean3); [GRI5] (mean5);
    [GRI9] (mean9); [GRI10] (mean10); [GRI13] (mean13);
    [GRI14] (mean14); [GRI18] (mean18); [GRI21] (mean21);
    [GRI23] (mean23);

    GRI1 GRI3 GRI5 GRI9 GRI10 GRI13 GRI14 GRI18 GRI21 GRI23 WITH
    GRI1 GRI3 GRI5 GRI9 GRI10 GRI13 GRI14 GRI18 GRI21 GRI23;

MODEL nonstudent:
    [GRI1] (mean1); [GRI3] (mean3); [GRI5] (mean5);
    [GRI9] (mean9); [GRI10] (mean10); [GRI13] (mean13);
    [GRI14] (mean14); [GRI18] (mean18); [GRI21] (mean21);
    [GRI23] (mean23);

    GRI1 GRI3 GRI5 GRI9 GRI10 GRI13 GRI14 GRI18 GRI21 GRI23 WITH
    GRI1 GRI3 GRI5 GRI9 GRI10 GRI13 GRI14 GRI18 GRI21 GRI23;
```

Omnibus Test: Log-likelihood Results

Null Model:

```
MODEL FIT INFORMATION

Number of Free Parameters                          65

Loglikelihood

          H0 Value                         -16567.417
          H0 Scaling Correction Factor          1.938
            for MLR
          H1 Value                         -16567.417
          H1 Scaling Correction Factor          1.938
            for MLR
```

Alternative Model:

```
MODEL FIT INFORMATION

Number of Free Parameters                         120

Loglikelihood

          H0 Value                         -15985.313
          H0 Scaling Correction Factor          1.729
            for MLR
          H1 Value                         -15909.738
          H1 Scaling Correction Factor          1.661
            for MLR
```

Null Model Parameters: 10 means + $\frac{10(10+1)}{2} = 55$ variances/covariances = 65

Alternative Model Parameters: 10 means + $2 * \frac{10(10+1)}{2} = 110$ variance/covariances = 120

Omnibus Test Statistic Under MLR

- To use MLR, we must first get our scale factor:

$$c_{LR} = \left| \frac{(q_{restricted})(c_{restricted}) - (q_{full})(c_{full})}{(q_{restricted} - q_{full})} \right|$$

$$= \left| \frac{65 * 1.938 - 120 * 1.729}{65 - 120} \right| = 1.482$$

- Then, we compute our LR test statistic

$$LR_{RS} = \frac{-2\left(\log L_{restricted} - \log L_{full}\right)}{c_{LR}} = \frac{-2(-16{,}567.417 - -15{,}985.313)}{1.482} = 785.565$$

- The p-value (using a Chi-Square with 55 DF) is < .001
  - Therefore, these data fail the omnibus test
  - We must now further investigate the invariance of the groups

Levels of Invariance across Groups

- **Step 1: Test "configural" invariance**
  - Do the groups have the **same factor structure**, broadly construed?
  - Same number of factors, same pattern of free/0 loadings
    → same conceptual definition of constructs being measured
    - We will assume this is the case for our two groups – otherwise we would have to examine the fit of one-, two-, and three- factor solutions
  - Test factor structure within each group separately, pray they are 'close enough' (if not, game over, pretty much)
  - Then estimate a combined model in which all model parameters are allowed to differ across groups
    - This will be the baseline model for further comparisons
    - Model χ^2 and df will be additive across groups
      - Keep in mind that different sample sizes across groups will result in differential weighting of the obtained χ^2 across groups

Configural Invariance Model: Same Factor Structure; All Parameters Separate

Group 1 (subscript = item, group):

- **DF = 65 – 30 = 25**
- $Y_{11} = \mu_{11} + \quad 1F_1 + e_{11}$
- $Y_{21} = \mu_{21} + \lambda_{21}F_1 \; + e_{21}$
- $Y_{31} = \mu_{31} + \lambda_{31}F_1 \; + e_{31}$
- $Y_{41} = \mu_{41} + \lambda_{41}F_1 \; + e_{41}$
- $Y_{51} = \mu_{51} + \lambda_{51}F_1 + e_{51}$
- $Y_{61} = \mu_{61} + \lambda_{61}F_1 \; + e_{61}$
- $Y_{71} = \mu_{71} + \lambda_{71}F_1 \; + e_{71}$
- $Y_{81} = \mu_{81} + \lambda_{81}F_1 \; + e_{81}$
- $Y_{91} = \mu_{91} + \lambda_{91}F_1 \; + e_{91}$
- $Y_{101} = \mu_{101} + \lambda_{101}F_1 \; + e_{101}$

- Factor 1 has own variance, but mean fixed to 0

Group 2 (subscript = item, group):

- **DF = 65 – 30 = 25**
- $Y_{12} = \mu_{12} + \quad 1F_1 + e_{12}$
- $Y_{22} = \mu_{22} + \lambda_{22}F_1 \; + e_{22}$
- $Y_{32} = \mu_{32} + \lambda_{32}F_1 \; + e_{32}$
- $Y_{42} = \mu_{42} + \lambda_{42}F_1 \; + e_{42}$
- $Y_{52} = \mu_{52} + \lambda_{51}F_1 + e_{52}$
- $Y_{62} = \mu_{62} + \lambda_{62}F_1 \; + e_{62}$
- $Y_{72} = \mu_{72} + \lambda_{72}F_1 \; + e_{72}$
- $Y_{82} = \mu_{82} + \lambda_{82}F_1 \; + e_{82}$
- $Y_{92} = \mu_{92} + \lambda_{92}F_1 \; + e_{92}$
- $Y_{102} = \mu_{102} + \lambda_{102}F_1 \; + e_{102}$

- Factor 1 has own variance, but mean fixed to 0

Configural Invariance Model in Matrices

For both groups, we are building the following:

$$\boldsymbol{\Sigma}_B = \begin{bmatrix} \boldsymbol{\Sigma}_{g_1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{g_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Lambda}_{g_1}\boldsymbol{\Phi}_{g_1}\boldsymbol{\Lambda}_{g_1}^T + \boldsymbol{\Psi}_{g_1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}_{g_2}\boldsymbol{\Phi}_{g_2}\boldsymbol{\Lambda}_{g_2}^T + \boldsymbol{\Psi}_{g_2} \end{bmatrix}$$

And, now the mean vector gets involved:

$$\boldsymbol{\mu}_B = \begin{bmatrix} \boldsymbol{\mu}_{g_1} \\ \boldsymbol{\mu}_{g_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_{I_{g_1}} \\ \boldsymbol{\mu}_{I_{g_2}} \end{bmatrix}$$

The configural model will essentially become our alternative model in future model comparisons

Configural Model Syntax

```
MODEL:
    !intercept labels
    [GRI1] (INT1_1); [GRI3] (INT3_1); [GRI5] (INT5_1);
    [GRI9] (INT9_1); [GRI10] (INT10_1); [GRI13] (INT13_1);
    [GRI14] (INT14_1); [GRI18] (INT18_1); [GRI21] (INT21_1);
    [GRI23] (INT23_1);

    !loading labels
    GAMBLING BY GRI1@1 (L1_1)
                GRI3   (L3_1)
                GRI5   (L5_1)
                GRI9   (L9_1)
                GRI10  (L10_1)
                GRI13  (L13_1)
                GRI14  (L14_1)
                GRI18  (L18_1)
                GRI21  (L21_1)
                GRI23  (L23_1);

    !unique variance labels
    GRI1 (U1_1); GRI3 (U3_1); GRI5 (U5_1);
    GRI9 (U9_1); GRI10 (U10_1); GRI13 (U13_1);
    GRI14 (U14_1); GRI18 (U18_1); GRI21 (U21_1);
    GRI23 (U23_1);

    [GAMBLING@0] (Fmean_1);
    GAMBLING (Fvar_1);
```

```
MODEL nonstudent:
    !intercept labels
    [GRI1] (INT1_2); [GRI3] (INT3_2); [GRI5] (INT5_2);
    [GRI9] (INT9_2); [GRI10] (INT10_2); [GRI13] (INT13_2);
    [GRI14] (INT14_2); [GRI18] (INT18_2); [GRI21] (INT21_2);
    [GRI23] (INT23_2);

    !loading labels
    GAMBLING BY GRI1@1 (L1_2)
                GRI3   (L3_2)
                GRI5   (L5_2)
                GRI9   (L9_2)
                GRI10  (L10_2)
                GRI13  (L13_2)
                GRI14  (L14_2)
                GRI18  (L18_2)
                GRI21  (L21_2)
                GRI23  (L23_2);

    !unique variance labels
    GRI1 (U1_2); GRI3 (U3_2); GRI5 (U5_2);
    GRI9 (U9_2); GRI10 (U10_2); GRI13 (U13_2);
    GRI14 (U14_2); GRI18 (U18_2); GRI21 (U21_2);
    GRI23 (U23_2);

    [GAMBLING@0] (Fmean_2);
    GAMBLING (Fvar_2);
```

Levels of Invariance across Groups

- **Step 2: Test "metric" invariance**
  - Also called "**weak** factorial invariance"
  - Do the groups have the **same factor loadings**?
    - Each congeneric item is still allowed to have a different loading (i.e., this is *not* a tau-equivalent model)
    - Loadings for same item are constrained to equality across *groups*
  - **Marker items** (that are fixed=1 for identification) are assumed invariant – because they are already fixed, they cannot be tested
    - For this reason, I suggest moving to an alternative specification: **Estimate all factor loadings, but fix the factor variance(s) to 1 in the reference group only** (still free them in the alternative group)
    - This allows us to evaluate ALL loadings and still identify the model (see Yoon & Millsap, 2007)

Metric Invariance Model:
Same Factor Loadings (saves +9 df here)

Group 1 (subscript = item, group):

- $Y_{11} = \mu_{11} + \boldsymbol{\lambda_1 F_1} + e_{11}$
- $Y_{21} = \mu_{21} + \boldsymbol{\lambda_2 F_1} + e_{21}$
- $Y_{31} = \mu_{31} + \boldsymbol{\lambda_3 F_1} + e_{31}$
- $Y_{41} = \mu_{41} + \boldsymbol{\lambda_4 F_1} + e_{41}$
- $Y_{51} = \mu_{51} + \boldsymbol{\lambda_5 F_1} + e_{51}$
- $Y_{61} = \mu_{61} + \boldsymbol{\lambda_6 F_1} + e_{61}$
- $Y_{71} = \mu_{71} + \boldsymbol{\lambda_7 F_1} + e_{71}$
- $Y_{81} = \mu_{81} + \boldsymbol{\lambda_8 F_1} + e_{81}$
- $Y_{91} = \mu_{91} + \boldsymbol{\lambda_9 F_1} + e_{91}$
- $Y_{101} = \mu_{101} + \boldsymbol{\lambda_{10} F_1} + e_{101}$

- **Factor 1 has variance fixed to 1**, but mean fixed to 0

Group 2 (subscript = item, group):

- $Y_{12} = \mu_{12} + \boldsymbol{\lambda_1 F_1} + e_{12}$
- $Y_{22} = \mu_{22} + \boldsymbol{\lambda_2 F_1} + e_{22}$
- $Y_{32} = \mu_{32} + \boldsymbol{\lambda_3 F_1} + e_{32}$
- $Y_{42} = \mu_{42} + \boldsymbol{\lambda_4 F_1} + e_{42}$
- $Y_{52} = \mu_{52} + \boldsymbol{\lambda_5 F_1} + e_{52}$
- $Y_{62} = \mu_{62} + \boldsymbol{\lambda_6 F_1} + e_{62}$
- $Y_{72} = \mu_{72} + \boldsymbol{\lambda_7 F_1} + e_{72}$
- $Y_{82} = \mu_{82} + \boldsymbol{\lambda_8 F_1} + e_{82}$
- $Y_{92} = \mu_{92} + \boldsymbol{\lambda_9 F_1} + e_{92}$
- $Y_{102} = \mu_{102} + \boldsymbol{\lambda_{10} F_1} + e_{102}$

- **Factor 1 has freely estimated variance,** but mean fixed to 0 (for now)

Metric Invariance Model in Matrices

For both groups, we are building the following:

$$\boldsymbol{\Sigma}_B = \begin{bmatrix} \boldsymbol{\Sigma}_{g_1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{g_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Lambda}\boldsymbol{\Phi}_{g_1}\boldsymbol{\Lambda}^T + \boldsymbol{\Psi}_{g_1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}\boldsymbol{\Phi}_{g_2}\boldsymbol{\Lambda}^T + \boldsymbol{\Psi}_{g_2} \end{bmatrix}$$

And, now the mean vector gets involved:

$$\boldsymbol{\mu}_B = \begin{bmatrix} \boldsymbol{\mu}_{g_1} \\ \boldsymbol{\mu}_{g_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_{I_{g_1}} \\ \boldsymbol{\mu}_{I_{g_2}} \end{bmatrix}$$

The metric model hypothesis test is:

H_0: $\boldsymbol{\Lambda}_{g_1} = \boldsymbol{\Lambda}_{g_2} = \boldsymbol{\Lambda}$ (metric invariance model)

H_1: $\boldsymbol{\Lambda}_{g_1} \neq \boldsymbol{\Lambda}_{g_2}$ (configural invariance model)

Metric Invariance Syntax

```
MODEL:
    !intercept labels
    [GRI1] (INT1_1); [GRI3] (INT3_1); [GRI5] (INT5_1);
    [GRI9] (INT9_1); [GRI10] (INT10_1); [GRI13] (INT13_1);
    [GRI14] (INT14_1); [GRI18] (INT18_1); [GRI21] (INT21_1);
    [GRI23] (INT23_1);

    !loading labels
    GAMBLING BY GRI1@1 (L1_1)
                GRI3   (L3_1)
                GRI5   (L5_1)
                GRI9   (L9_1)
                GRI10  (L10_1)
                GRI13  (L13_1)
                GRI14  (L14_1)
                GRI18  (L18_1)
                GRI21  (L21_1)
                GRI23  (L23_1);

    !unique variance labels
    GRI1 (U1_1); GRI3 (U3_1); GRI5 (U5_1);
    GRI9 (U9_1); GRI10 (U10_1); GRI13 (U13_1);
    GRI14 (U14_1); GRI18 (U18_1); GRI21 (U21_1);
    GRI23 (U23_1);

    [GAMBLING@0] (Fmean_1);
    GAMBLING (Fvar_1);
```

```
MODEL nonstudent:
    !intercept labels
    [GRI1] (INT1_2); [GRI3] (INT3_2); [GRI5] (INT5_2);
    [GRI9] (INT9_2); [GRI10] (INT10_2); [GRI13] (INT13_2);
    [GRI14] (INT14_2); [GRI18] (INT18_2); [GRI21] (INT21_2);
    [GRI23] (INT23_2);

    !loading labels
    GAMBLING BY GRI1@1 (L1_2)
                GRI3   (L3_2)
                GRI5   (L5_2)
                GRI9   (L9_2)
                GRI10  (L10_2)
                GRI13  (L13_2)
                GRI14  (L14_2)
                GRI18  (L18_2)
                GRI21  (L21_2)
                GRI23  (L23_2);

    !unique variance labels
    GRI1 (U1_2); GRI3 (U3_2); GRI5 (U5_2);
    GRI9 (U9_2); GRI10 (U10_2); GRI13 (U13_2);
    GRI14 (U14_2); GRI18 (U18_2); GRI21 (U21_2);
    GRI23 (U23_2);

    [GAMBLING@0] (Fmean_2);
    GAMBLING (Fvar_2);
```

Metric Invariance Test: Log-likelihood Results

Null (Metric Invariance) Model:

```
MODEL FIT INFORMATION

Number of Free Parameters                         51

Loglikelihood

          H0 Value                        -16244.476
          H0 Scaling Correction Factor         2.042
            for MLR
          H1 Value                        -15909.738
          H1 Scaling Correction Factor         1.661
            for MLR
```

Alternative (Configural) Model:

```
MODEL FIT INFORMATION

Number of Free Parameters                         60

Loglikelihood

          H0 Value                        -16029.062
          H0 Scaling Correction Factor         1.968
            for MLR
          H1 Value                        -15909.738
          H1 Scaling Correction Factor         1.661
            for MLR
```

Null (Metric Invariance) Model Parameters:

2*10 intercepts + 10 loadings + 2*10 unique variances + 1 Factor Variance = 51

Alternative (Configural) Model Parameters:

2*10 intercepts + 2*9 loadings + 2*10 unique variances + 2 Factor Variance = 60

Metric Invariance Test Statistic Under MLR

- To use MLR, we must first get our scale factor:

$$c_{LR} = \left|\frac{(q_{restricted})(c_{restricted}) - (q_{full})(c_{full})}{(q_{restricted} - q_{full})}\right|$$

$$= \left|\frac{51 * 2.042 - 60 * 1.968}{51 - 60}\right| = 1.549$$

- Then, we compute our LR test statistic

$$LR_{RS} = \frac{-2(\log L_{restricted} - \log L_{full})}{c_{LR}} = \frac{-2(-16{,}244.476 - -16{,}029.062)}{1.549} = 430.828$$

- The p-value (using a Chi-Square with 9 DF) is < .001
  - Therefore, these data fail the metric invariance test
  - <u>Across all items</u>, the groups do not have invariant factor loadings

Metric Invariance

- Compare fit of metric invariance to configural invariance model:
  - Does the model fit not get worse (***-2LL diff not significant***)?
    - We are taking parameters away, so it can only get worse…
    - Don't forget to fix variance=1 in reference group only (free in other group)! Otherwise you are imposing a structural constraint too by accident!
  - Either way, inspect the modification indices to see if there are any items whose loadings want to differ between groups
    - Retest the model as needed after releasing one loading at a time

- Do you have at least partial* metric invariance?
  - Congrats! Your construct is measured in the same way across groups
  - If not, it doesn't make sense to evaluate how relationships involving the factor differ across groups (because the factor itself differs)

* No real consensus on how much is "partial", but at least 1 per factor!

Inspection of Modification Indices for Partial metric Invariance

- The modification indices can help us determine if we have partial metric invariance
  - MIs are about the same for each parameter, in both groups

Non-Students:

Group NONSTUDENT

ON/BY Statements

	M.I.	E.P.C.	Std E.P.C.	StdYX E.P.C.
GRI1 ON GAMBLING / GAMBLING BY GRI1	0.293	0.031	0.058	0.043
GRI3 ON GAMBLING / GAMBLING BY GRI3	13.971	-0.208	-0.395	-0.294
GRI5 ON GAMBLING / GAMBLING BY GRI5	0.287	0.036	0.068	0.044
GRI9 ON GAMBLING / GAMBLING BY GRI9	10.759	0.134	0.254	0.233
GRI10 ON GAMBLING / GAMBLING BY GRI10	0.527	0.039	0.074	0.055
GRI13 ON GAMBLING / GAMBLING BY GRI13	3.490	0.085	0.161	0.121
GRI14 ON GAMBLING / GAMBLING BY GRI14	24.029	-0.390	-0.740	-0.404
GRI18 ON GAMBLING / GAMBLING BY GRI18	0.384	-0.025	-0.048	-0.037
GRI21 ON GAMBLING / GAMBLING BY GRI21	0.529	0.035	0.066	0.053
GRI23 ON GAMBLING / GAMBLING BY GRI23	6.552	-0.175	-0.332	-0.207

Students:

Group STUDENT

ON/BY Statements

GRI1 ON GAMBLING / GAMBLING BY GRI1	0.293	-0.007	-0.007	-0.008
GRI3 ON GAMBLING / GAMBLING BY GRI3	13.972	0.037	0.037	0.046
GRI5 ON GAMBLING / GAMBLING BY GRI5	0.286	-0.004	-0.004	-0.005
GRI9 ON GAMBLING / GAMBLING BY GRI9	10.759	-0.019	-0.019	-0.032
GRI10 ON GAMBLING / GAMBLING BY GRI10	0.526	-0.006	-0.006	-0.008
GRI13 ON GAMBLING / GAMBLING BY GRI13	3.491	-0.020	-0.020	-0.024
GRI14 ON GAMBLING / GAMBLING BY GRI14	24.021	0.117	0.117	0.083
GRI18 ON GAMBLING / GAMBLING BY GRI18	0.384	0.017	0.017	0.014
GRI21 ON GAMBLING / GAMBLING BY GRI21	0.529	-0.006	-0.006	-0.008
GRI23 ON GAMBLING / GAMBLING BY GRI23	6.549	0.020	0.020	0.024

Modification Index Specifications

- The modification indices showed that the following items are likely to be non-invariant (different between groups):
  - 1, 3, 9, 14, 23
- The modification indices also showed that the following items are likely to be invariant (not different):
  - 5, 10, 13, 18, 21
- Therefore, we may have **partial metric invariance**, which is important if we wish to compare factor means
- Because Modification Indices are one-DF tests, we must add each one-at-a-time
  - If some items are not invariant (different) – no further tests can be conducted on these items

Testing Partial Invariance

	Test of -2ΔLL Difference							
Model	Model H0 LL	H0 LL Scale Factor	# Free Parms	Diff in LL * -2	Diff Scaling Correction	Abs Value Scaled Diff	DF Diff	Exact P-Value
2a. Metric Invariance (All Loadings)	-16,069.590	2.090	51					
1.Configural Model	-16,029.062	1.968	60					
				81.056	1.277	63.490	9	0.0000
2b. Metric Invariance (No 14)	-16,052.342	2.068	52					
1.Configural Model	-16,029.062	1.968	60					
				46.560	1.318	35.326	8	0.0000
2c. Metric Invariance (No 14 or 3)	-16,041.131	2.061	53					
1.Configural Model	-16,029.062	1.968	60					
				24.138	1.264	19.099	7	0.0079
2d. Metric Invariance (No 14 or 3 or 23)	-16,033.941	2.039	54					
1.Configural Model	-16,029.062	1.968	60					
				9.758	1.329	7.342	6	0.2903

Levels of Invariance across Groups

- **Step 3: Test “scalar” (“strong”) invariance**
  - Do the groups have the **same item intercepts**?
    - Each congeneric item is allowed to have a different intercept
    - Intercepts for same item are constrained to equality across groups
    - Scalar invariance model says factor mean differences cause the item mean differences (*but the item intercepts should still be the same*)
  - If you use marker intercepts (that are fixed=0 for identification), they are assumed invariant – because they are already fixed
    - So we will estimate all intercepts, **but constrain the factor mean(s) to 0 in the reference group** so we can evaluate all intercepts
  - Some folks might say that scalar invariance is not really necessary unless you plan on comparing mean differences…
    - Scalar invariance doesn’t always get tested as a result
    - Probably better to error on the side of caution and examine it anyway

Scalar Invariance Model:
Same Indicator Intercepts (saves +4 df here)

Group 1 (subscript = item, group):

- $Y_{11} = \mu_{11} + \lambda_{11}F_1 + e_{11}$
- $Y_{21} = \mu_{21} + \lambda_{21}F_1 + e_{21}$
- $Y_{31} = \boldsymbol{\mu_3} + \lambda_3F_1 + e_{31}$
- $Y_{41} = \mu_{41} + \lambda_{41}F_1 + e_{41}$
- $Y_{51} = \boldsymbol{\mu_5} + \lambda_5F_1 + e_{51}$
- $Y_{61} = \boldsymbol{\mu_6} + \lambda_6F_1 + e_{61}$
- $Y_{71} = \mu_{71} + \lambda_{71}F_1 + e_{71}$
- $Y_{81} = \boldsymbol{\mu_8} + \lambda_8F_1 + e_{81}$
- $Y_{91} = \boldsymbol{\mu_9} + \lambda_9F_1 + e_{91}$
- $Y_{101} = \mu_{101} + \lambda_{101}F_1 + e_{101}$

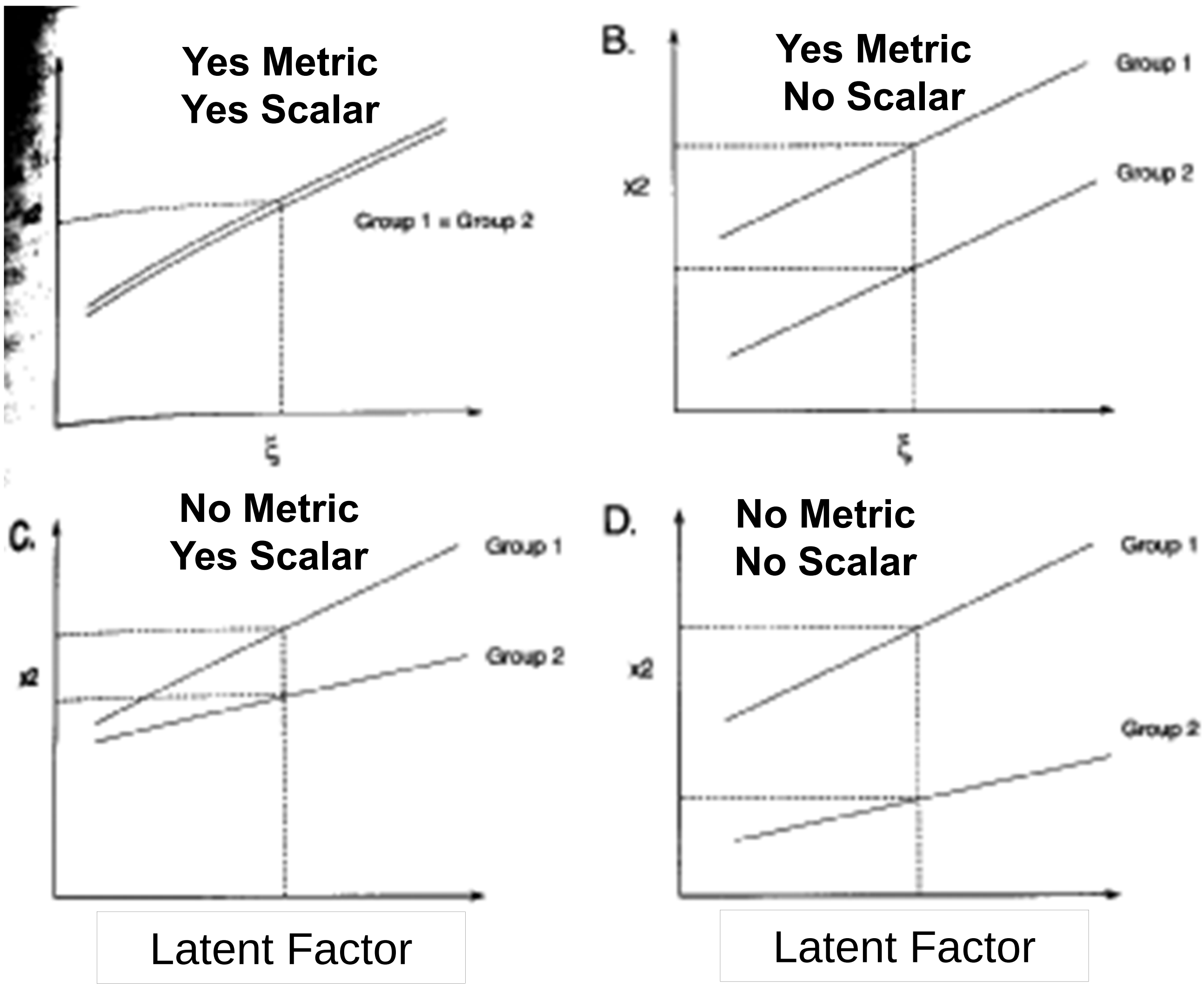
- Factor 1 has variance fixed to 1, but factor mean fixed to 0 (for group 1)

Group 2 (subscript = item, group):

- $Y_{12} = \mu_{12} + \lambda_{11}F_1 + e_{12}$
- $Y_{22} = \mu_{22} + \lambda_{21}F_1 + e_{22}$
- $Y_{32} = \boldsymbol{\mu_3} + \lambda_3F_1 + e_{32}$
- $Y_{42} = \mu_{42} + \lambda_{41}F_1 + e_{42}$
- $Y_{52} = \boldsymbol{\mu_5} + \lambda_5F_1 + e_{52}$
- $Y_{62} = \boldsymbol{\mu_6} + \lambda_6F_1 + e_{62}$
- $Y_{72} = \mu_{72} + \lambda_{71}F_1 + e_{72}$
- $Y_{82} = \boldsymbol{\mu_8} + \lambda_8F_1 + e_{82}$
- $Y_{92} = \boldsymbol{\mu_9} + \lambda_9F_1 + e_{92}$
- $Y_{102} = \mu_{102} + \lambda_{102}F_1 + e_{102}$

- Factor 1 has estimated variance, **but factor mean now free (for group 2) and represents factor mean diffs**

Implications of Non-Invariance



Scalar Invariance

- Only test those intercepts for which metric invariance holds
  - Different slopes can create different intercepts as an artifact

- Compare fit of scalar invariance to metric invariance model:
  - Does the model fit not get worse (***-2LL diff not significant)*****?**
  - Either way, inspect the modification indices to see if there are any items whose intercepts want to differ between groups
    - Retest the model as needed after releasing one intercept at a time

- Do you have at least partial* scalar invariance?
  - Your construct accounts for the item mean differences across groups
  - If not, it doesn't make sense to evaluate mean differences in the factor across groups (because other things create item mean differences)

* No real consensus on what is "partial", but at least 1 per factor!

Testing Scalar Invariance

	Test of -2ΔLL Difference							
Model	Model H0 LL	H0 LL Scale Factor	# Free Parms	Diff in LL * -2	Diff Scaling Correction	Abs Value Scaled Diff	DF Diff	Exact P-Value
3a. Scalar Invariance (All Intercepts)	-16,081.299	2.153	48					
2d. Metric Invariance (No 14 or 3 or 23)	-16,033.941	2.039	54					
				94.716	1.127	84.043	6	0.0000
3b. Scalar Invariance (No 1)	-16,058.009	2.140	49					
2d. Metric Invariance (No 14 or 3 or 23)	-16,033.941	2.039	54					
				48.136	1.049	45.879	5	0.0000
3c. Scalar Invariance (No 1 or 5)	-16,041.250	2.128	50					
2d. Metric Invariance (No 14 or 3 or 23)	-16,033.941	2.039	54					
				14.618	0.927	15.778	4	0.0033
3d. Scalar Invariance (No 1 or 5 or 18)	-16,034.189	2.128	51					
2d. Metric Invariance (No 14 or 3 or 23)	-16,033.941	2.039	54					
				0.496	0.526	0.943	3	0.8150

Levels of Invariance across Groups

- **Step 4: Test "residual variance" invariance**
  - Also called "**strict** factorial invariance"

  - Do the groups have the **same item residual variances**?
    - Each congeneric item is still allowed to have a different residual variance
    - Residual variances for the same item are constrained to equality across groups
    - Testing residual variances is the last step in assessing measurement invariance
      - People disagree as to whether or not this is necessary
      - Note: Equal residual variances are commonly mis-interpreted to mean "equal reliabilities" – this is ONLY the case if the factor variances are the same across groups, too
        - » We test that one next…

Residual Variance Invariance Model: Same Error Variances (saves +4 df here)

Group 1 (subscript = item, group):

- $Y_{11} = \mu_{11} + \lambda_{11}F_1 + e_{11}$
- $Y_{21} = \mu_{21} + \lambda_{21}F_1 + e_{21}$
- $Y_{31} = \mu_{31} + \lambda_{31}F_1 + e_{31}$
- $Y_{41} = \mu_4 + \lambda_4F_1 + \mathbf{e_4}$
- $Y_{51} = \mu_5 + \lambda_5F_1 + \mathbf{e_5}$
- $Y_{61} = \mu_6 + \lambda_6F_1 + \mathbf{e_6}$
- $Y_{71} = \mu_{71} + \lambda_{71}F_1 + e_{71}$
- $Y_{81} = \mu_{81} + \lambda_{81}F_1 + e_{81}$
- $Y_{91} = \mu_9 + \lambda_9F_1 + \mathbf{e_9}$
- $Y_{101} = \mu_{101} + \lambda_{101}F_1 + e_{101}$

- Factor 1 has variance fixed to 1, but factor mean fixed to 0 (for group 1)

Group 2 (subscript = item, group):

- $Y_{12} = \mu_{12} + \lambda_{12}F_1 + e_{12}$
- $Y_{22} = \mu_{22} + \lambda_{22}F_1 + e_{22}$
- $Y_{32} = \mu_{32} + \lambda_{32}F_1 + e_{32}$
- $Y_{42} = \mu_4 + \lambda_4F_1 + \mathbf{e_4}$
- $Y_{52} = \mu_5 + \lambda_5F_1 + \mathbf{e_5}$
- $Y_{62} = \mu_6 + \lambda_6F_1 + \mathbf{e_6}$
- $Y_{72} = \mu_{72} + \lambda_{72}F_1 + e_{72}$
- $Y_{82} = \mu_{82} + \lambda_{82}F_1 + e_{82}$
- $Y_{92} = \mu_9 + \lambda_9F_1 + \mathbf{e_9}$
- $Y_{102} = \mu_{102} + \lambda_{102}F_1 + e_{102}$

- Factor 1 has estimated variance, and factor mean is estimated (for group 2) and represents factor mean diff

Residual Invariance

- Only test those residual variances for which metric and scalar invariance already hold

- Compare fit of residual invariance to scalar invariance model:
  - Does the model fit not get worse (***-2LL diff not significant*)?**
  - Either way, inspect the modification indices to see if there are any items whose residual variances want to differ between groups
    - Retest the model as needed after releasing one residual variance at a time

- Do you have at least partial* residual invariance?
  - Your groups have the same amount of “not the factor” in each item
  - If not??? Ongoing debate about the necessity of this

* No real consensus on what is “partial”, but at least 1 per factor!

Testing Residual Invariance

	Test of -2ΔLL Difference							
Model	Model H0 LL	H0 LL Scale Factor	# Free Parms	Diff in LL * -2	Diff Scaling Correction	Abs Value Scaled Diff	DF Diff	Exact P-Value
4a. Residual Invariance	-16,102.165	2.024	47					
3d. Scalar Invariance (No 1 or 5 or 18)	-16,034.189	2.128	51					
				135.952	3.350	40.583	4	0.0000
4a. Residual Invariance (no 9)	-16,087.266	2.003	48					
3d. Scalar Invariance (No 1 or 5 or 18)	-16,034.189	2.128	51					
				106.154	4.128	25.716	3	0.0000
4a. Residual Invariance (no 9 or 10)	-16,052.902	2.049	49					
3d. Scalar Invariance (No 1 or 5 or 18)	-16,034.189	2.128	51					
				37.426	4.064	9.210	2	0.0100
4a. Residual Invariance (no 9 or 10 or 21)	-16,035.809	2.077	50					
3d. Scalar Invariance (No 1 or 5 or 18)	-16,034.189	2.128	51					
				3.240	4.678	0.693	1	0.4053

Next, Structural Invariance

- Are the **factor variances** the same across groups? (+1 df/factor)
  - Fix the factor variance in the alternative group to 1 (as in the ref group)
  - Did model fit get worse? If so, the groups differ in their factor variances
- Is the **factor covariance** the same across groups? (+1 df per pair)
  - Fix the factor covariances equal across groups, did model fit get worse?
  - Factor correlation will only be the same across groups if the factor variances are the same, too
- Are the **factor means** the same across groups? (+1 df/factor)
  - Fix the factor mean in the alternative group to 0 (as in the ref group)
  - Did model fit get worse? If so, the groups differ in their factor means
- **It is not problematic if structural invariance doesn't hold.**
  - Given measurement invariance, this is a **substantive issue** about differences in the latent trait amounts and relations (and that's ok).

Testing Structural Invariance

	Test of -2ΔLL Difference							
Model	Model H0 LL	H0 LL Scale Factor	# Free Parms	Diff in LL * -2	Diff Scaling Correction	Abs Value Scaled Diff	DF Diff	Exact P-Value
5a. Structural - Factor Variance	-16,105.850	2.093	49					
4a. Residual Invariance (no 9 or 10 or 21)	-16,035.809	2.077	50					
				140.082	1.293	108.339	1	0.0000
5a. Structural - Factor Mean	-16,056.329	2.121	49					
4a. Residual Invariance (no 9 or 10 or 21)	-16,035.809	2.077	50					
				41.040	-0.079	519.494	1	0.0000

Final Results: Factor Mean/Variance

Group	Factor Mean	Factor Mean SE	Factor Variance	Factor Variance SE
Non-Student	1.329	0.215	4.649	0.719
Student	0.000	0.000	1.000	0.000

Setting up the Invariance Model

- Testing invariance across independent groups?
  You need a **multiple-group** model:
  - Estimate factor model for each group at once, but only the variables per group are related within each subgroup model
  - An alternative approach, MIMIC models, in which the grouping variable is entered as a predictor, do not allow testing of equality of factor loadings or factor variances (so MIMIC may be less useful)

- Testing invariance across repeated measures time/condition?
  - Put all the observed indicators into the **SAME MODEL**
    - Correlate errors from same indicators across time (an accepted freebee)
    - Model gets big and complicated quickly
  - Multiple group approach is *not* appropriate because observations from same person are not independent

CONCLUDING REMARKS

Wrapping Up Measurement Invariance in CFA

- The process of testing factorial invariance has two distinct parts:
  - **Measurement invariance**: Is your construct being measured in the same way? Let's hope so!
    - Better hope for at least "partial" invariance… otherwise, game over.
  - **Structural invariance**: Do your groups differ in their distribution and/or means of the construct? Let's find out!
    - Structural differences are real and interpretable differences *given measurement invariance of the constructs*

- Measurement invariance is always assumed in any statistical analysis…
  - But can be tested explicitly in a latent trait modeling framework

Up Next…

- No homework this week!

- Next week: Putting it all together – Measurement models, structural models, and path models