Measurement Equivalence/Invariance

Gordon W. Cheung, Ph.D

Professor of Organisational Behaviour, The University of Auckland
Distinguished Chair Professor, National Taiwan University
Adjunct Professor, Centre for Work, Organisation and Wellbeing, Griffith University

Australian & New Zealand Academy of Management (ANZAM) – Life Fellow

Please send your comments and suggestions to:
gordon.cheung@auckland.ac.nz

Measurement Equivalence/ Invariance (ME/I)

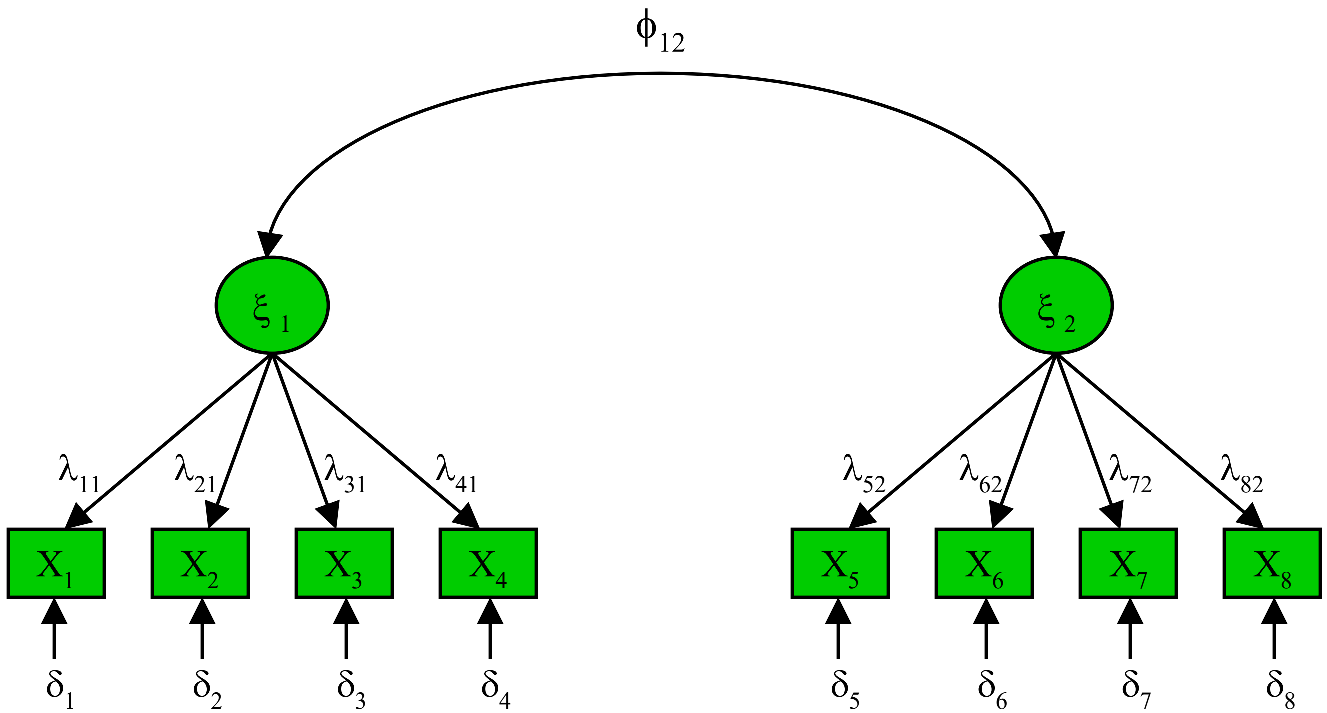
ME/I is a general term that can be applied to the comparison of the various components of measurement models and can sometimes be extended to structural models and mean structures.

Basic Types of ME/I

- Configural Equivalence
- Metric Equivalence
- Scalar Equivalence
- Uniqueness Equivalence
- Latent Mean Equivalence

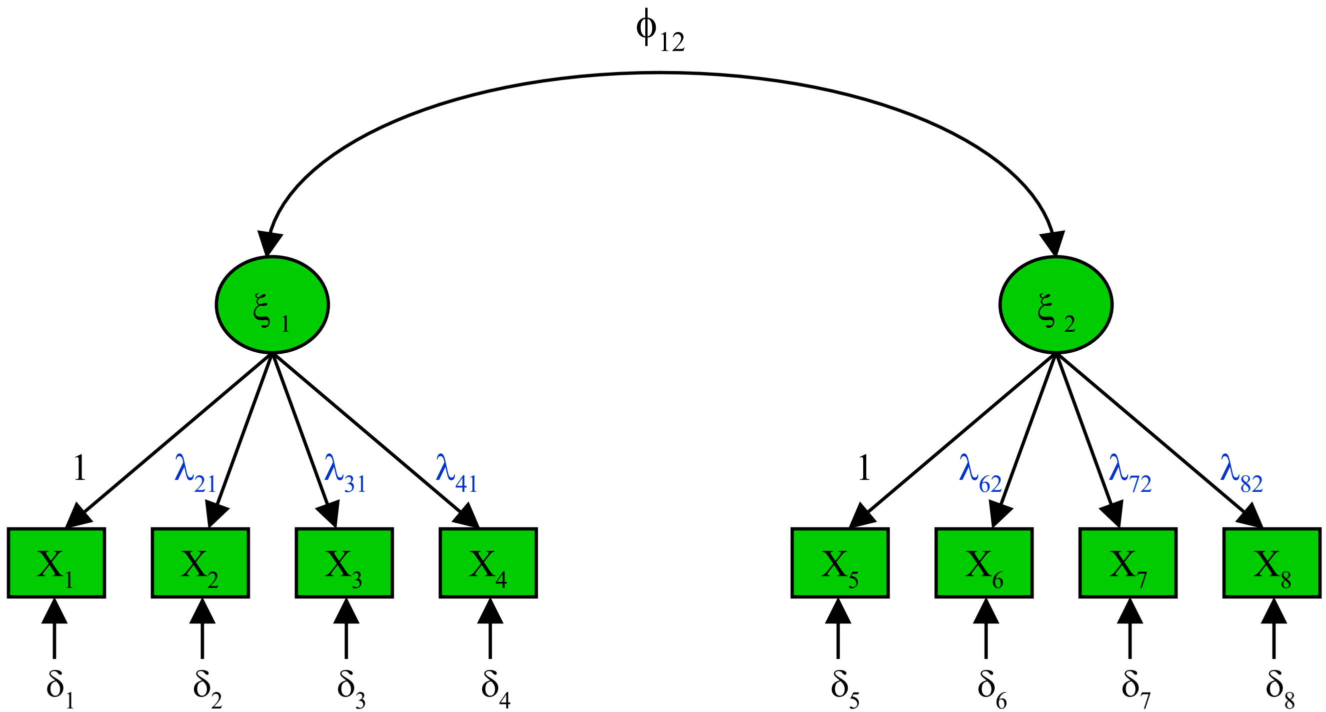
Configural Equivalence

- All groups associate the same subsets of items with the same constructs (the cognitive domains are the same)



Metric Equivalence

- Overall, the strength of the relationships between items and their underlying constructs is the same across groups. (The constructs are manifested in the same way)



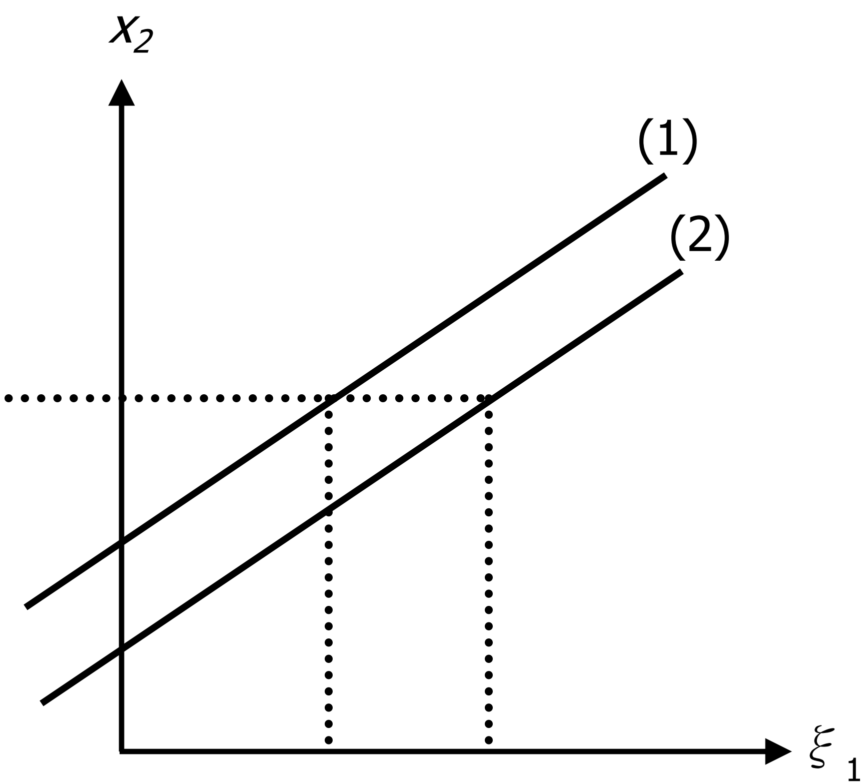
Scalar Equivalence

- Intercepts are the same across groups.
- The cross-group differences indicated by the items are the same across items. Alternatively, all items indicate the same cross-group differences.

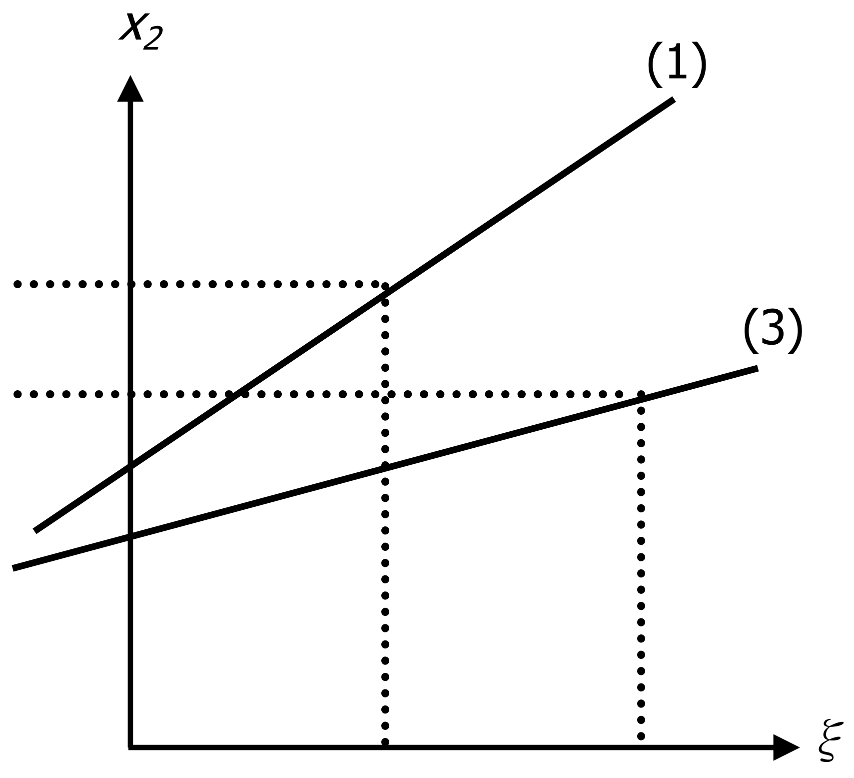
$$x_i = \tau_i + \lambda_{ij}\xi_j + \delta_i$$

Metric and Scalar Invariance

$$x = \tau_x + \Lambda_x \xi + \varepsilon$$



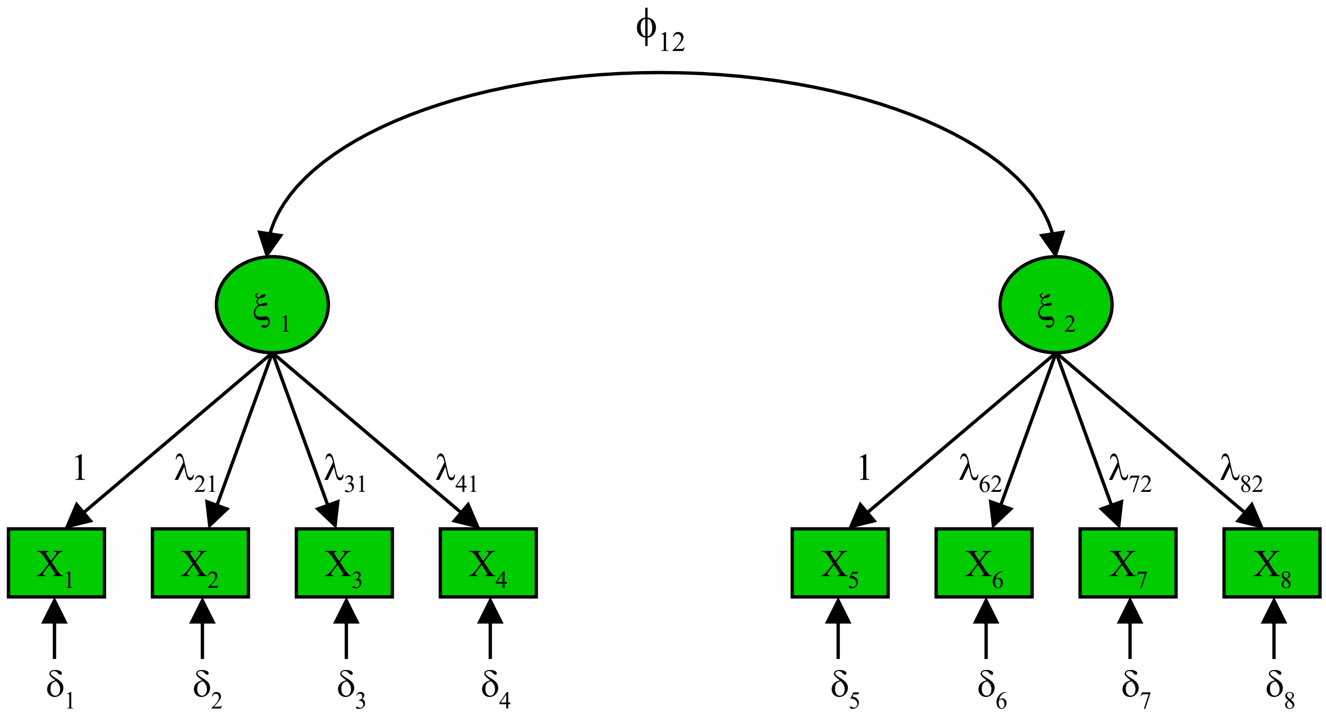
Equal factor loading,
different intercept



Different factor loading,
different intercept

Uniqueness Equivalence

- Items demonstrate the same size of measurement error across groups. Alternatively, Items have the same quality as measures of the underlying construct across groups.



Latent Mean Equivalence

- The mean level of each construct is the same across groups

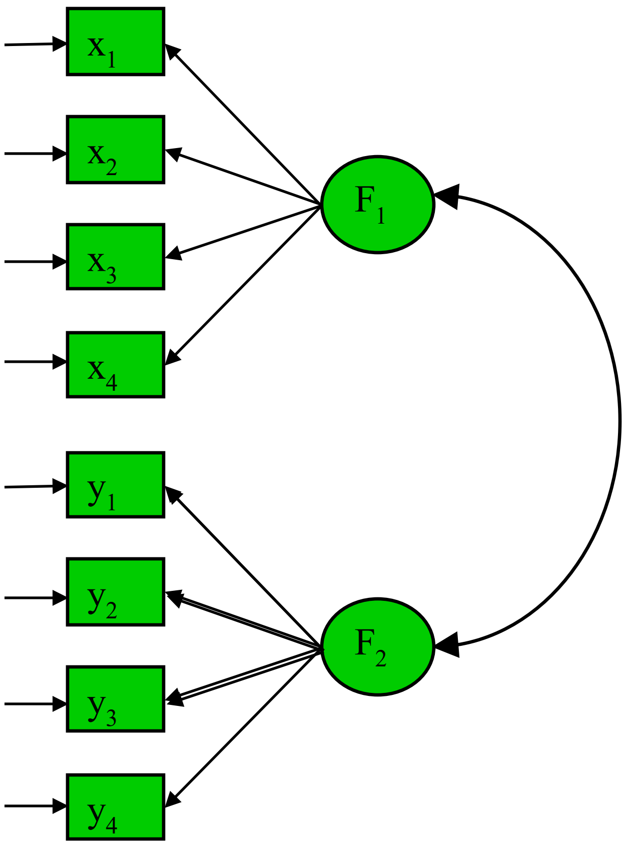
$$E(x_i) = \tau_i + \lambda_{ij}\kappa_j$$

$$\kappa_j^{(1)} = \kappa_j^{(2)}$$

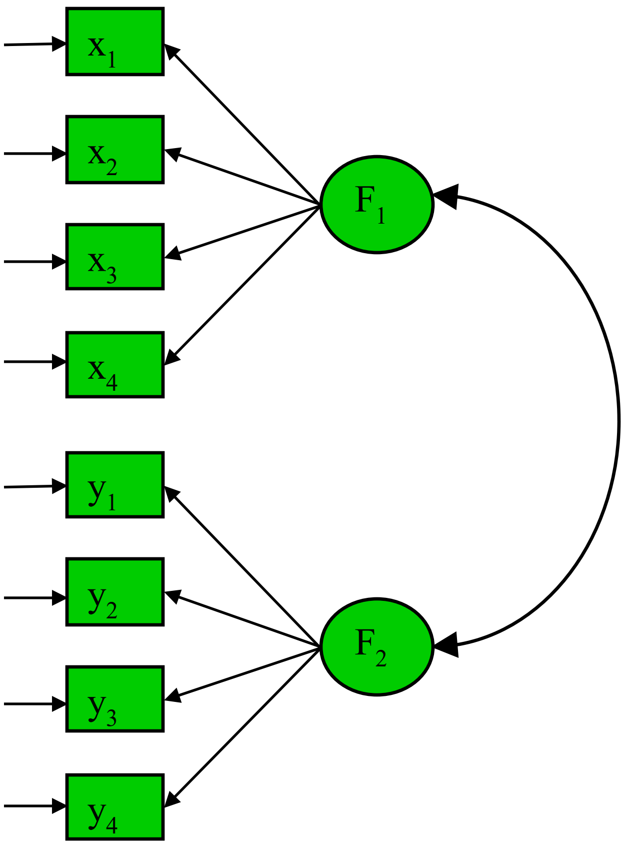
Comparison of Latent Means Across Groups

- *If Item 1 is the reference item, factor loading is fixed at 1 and intercept at 0*
  - $\bar{X}_1^{(1)} - \bar{X}_1^{(2)} = \kappa_1^{(1)} - \kappa_1^{(2)}$
- *Item 2*
  - $\bar{\mathrm{X}}_2^{(1)} - \bar{\mathrm{X}}_2^{(2)} = \left(\tau_2^{(1)} + \lambda_{21}^{(1)}\kappa_1^{(1)}\right) - \left(\tau_2^{(2)} + \lambda_{21}^{(2)}\kappa_1^{(2)}\right)$
  - $\bar{\mathrm{X}}_2^{(1)} - \bar{\mathrm{X}}_2^{(2)} = \left(\tau_2^{(1)} - \tau_2^{(2)}\right) + \left(\lambda_{21}^{(1)}\kappa_1^{(1)} - \lambda_{21}^{(2)}\kappa_1^{(2)}\right)$
- *If metric invariance:* $\lambda_{21}^{(1)} = \lambda_{21}^{(2)} = \lambda_{21}$
  - $\bar{\mathrm{X}}_2^{(1)} - \bar{\mathrm{X}}_2^{(2)} = \left(\tau_2^{(1)} - \tau_2^{(2)}\right) + \lambda_{21}\left(\kappa_1^{(1)} - \kappa_1^{(2)}\right)$
- *If scalar invariance:* $\tau_2^{(1)} = \tau_2^{(2)}$
  - $\bar{X}_2^{(1)} - \bar{X}_2^{(2)} = \lambda_{21}\left(\kappa_1^{(1)} - \kappa_1^{(2)}\right) = \lambda_{21}\left(\bar{X}_1^{(1)} - \bar{X}_1^{(2)}\right)$
- *Each item implies the same mean difference after adjusting the scale (factor loadings)*

Multiple-Group Comparisons



Male



Female

Nested Models

- When one or more free parameters of a model are constrained, the model thus constrained is said to be nested in the one from which it was derived.

- The constrained model usually has a worse fit.

- Does the improvement in fit provided by a more comprehensive model warrant preferring it to a more parsimonious model nested in it?

Comparing Nested Models

- Likelihood Ratio Test (LRT; Bollen, 1989)
  - $\Delta\chi^2 = \chi^2_{(\text{constrained})} - \chi^2_{(\text{unconstrained})}$
  - $\Delta \text{df} = \text{df}_{(\text{constrained})} - \text{df}_{(\text{unconstrained})}$
- Changes in CFI less than -0.01
  - ΔCFI (Cheung & Rensvold, 2002; Meade, Johnson, & Braddy, 2008)

Common Methods for Identifying Non-invariant Items

- Modification Index (Byrne et al., 1989; Yoon & Millsap, 2007)
  - Byrne, B. M., Shavelson, R. J., & Muthén, B. 1989. Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105: 456-466.
  - Yoon, M., & Millsap, R. E. 2007. Detecting violations of factorial invariance using data-based specification searches: A Monte Carlo study. *Structural Equation Modeling: A Multidisciplinary Journal*, 14: 435-463.
- Alignment Method for multiple groups (Asparouhov & Muthén, 2014)
  - Asparouhov, T., & Muthén, B. O. 2014. Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, 21: 495–508.
- Direct Comparison (Cheung & Lau, 2012)
  - Cheung & Lau (2012). A Direct Comparison Approach for Testing Measurement Invariance. *Organizational Research Methods*, 15: 167-198

Modification Index (MI) Approach

- MI shows the expected reduction in the chi-square value for overall model fit when a constrained parameter in the model is freely estimated
- Researchers can refer to the MI for each equivalent constraint on the factor loadings in a full metric invariance model (in which all factor loadings are constrained to be equal across groups) to find out if relaxing a constraint can improve model fit substantially
- Relaxing the invariant constraint on the factor loadings sequentially until the overall model fit is adequate

Shortcomings of the MI Approach

- MI is estimated assuming all other constraints are valid; that is, factor loadings of all other items are invariant across groups. If there is more than one non-invariant item, this assumption is violated, such that the MI method may incorrectly identify the set of non-invariant items
- Some MI values may be very similar, but choosing a different sequence of relaxing the constraints may result in a different set of non-invariant items
- Ineffective in identifying non-invariant items
- Not practical when there are many groups with large sample sizes, such as in a cross-cultural study, because relaxing the invariant constraints with many large modification indices sequentially to identify all non-invariant items could easily result in a wrong model
- An example: "example a mi.out" from Mplus

Alignment Method Approach

- Transforms the values of factor loadings and intercepts in a configural invariance model with an algorithm that seeks to minimize any group differences in both factor loadings and item intercepts without altering model fit -- analogous to rotations in exploratory factor analysis
- Compares each estimated factor loading and intercept between every pair of groups. An initial set of non-invariant groups with statistically non-significant differences in the estimated parameter is generated using a Type I error of 0.01
- The non-invariant set is refined by comparing a group's estimated parameter with the average in the non-invariant set. A group is added to the set of invariant groups if the difference is not statistically significant using a Type I error rate of 0.001
- Presents the results of pairwise comparisons of the latent means without using equality constraints

Shortcomings of the Alignment Method (I)

- Results are only reliable when no more than 25% of the items are non-invariant

- Separately identifies non-invariant groups for the intercepts and the factor loadings -- groups with significantly different factor loadings may have non-significantly different intercepts

  - Interpretation of the latent mean differences for groups with non-invariant factor loadings and invariant intercepts will be misleading

Shortcomings of the Alignment Method (II)

- Identifies only the set with the largest number of invariant groups but ignores any invariance in the remaining groups -- does not maximize the number of comparisons available in a sample
  - A B C D *E F G*
- Invariant groups identified by the alignment method may be inconsistent with the pairwise comparison results—two groups with significantly different intercepts may be classified as invariant, or two groups with non-significantly different intercepts may be identified as non-invariant
- An Example: "example a alignment free.out"

Direct Comparison Approach

- BC bootstrap confidence intervals of differences in parameters (factor loadings and intercepts) for testing ME/I, an extension of the procedure for testing mediation effects (Lau & Cheung, 2012)
- Factor-Ratio Test
- List-and-Delete Method

Factor-Ratio Test

	When first item is referent:	When second item is referent:
λ_{11}	$\lambda_{11}/\lambda_{11} = 1$	
λ_{21}	$\lambda_{21}/\lambda_{11} = \lambda_{21}^{*}$	
λ_{31}	$\lambda_{31}/\lambda_{11} = \lambda_{31}^{*}$	$\dfrac{\lambda_{31}}{\lambda_{21}} = \dfrac{\frac{\lambda_{31}}{\lambda_{11}}}{\frac{\lambda_{21}}{\lambda_{11}}} = \dfrac{\lambda_{31}^{*}}{\lambda_{21}^{*}}$
λ_{41}	$\lambda_{41}/\lambda_{11} = \lambda_{41}^{*}$	

List-and-Delete Method (4 items)

1	2	3	4		1	2		
1	2	3			1		3	
1	2		4		1			4
1		3	4			2	3	
	2	3	4			2		4
							3	4

Shortcomings of the Direct Comparison Approach

- Transforming parameters, particularly the item intercepts, is cumbersome and prone to human error when shifting the reference item
- Bootstrapping procedure to generate confidence intervals of parameter differences can only be applied to non-nested data, so it does not apply to multilevel studies
- Can only be used to compare two groups -- cannot be applied to longitudinal studies with multiple waves of data

MEI Package

- Extension of Cheung and Lau's (2012) direct comparison method to multi-group, multi-wave, and multi-level data
- Estimate multiple models, each with a different item as the reference item
- Using Bootstrapping or Monte Carlo simulations to estimate confidence intervals of differences in estimated parameters. Bootstrapping 2,000 samples, or one million sets of factor loadings and intercepts are simulated based on the parameter estimates and the variance-covariance matrix of the freely estimated factor loadings and intercepts in the configural invariance model (or partial metric invariance model), and the confidence intervals of differences in factor loadings and intercepts between groups are calculated

MEI Package

- When more than two groups are tested for ME/I, for each combination of reference item and argument, first compares each pair of groups using a Type I error rate of 0.01 or 0.001, then applies the list-and-delete method to identify the set of invariant groups
- After identifying the sets of groups with invariant factor loadings or intercepts in each pair of reference and argument items, the model with the least number of freely estimated factor loadings and intercepts (the most parsimonious model with the largest number of invariant items) will be selected. When multiple models have the same number of estimated parameters, the model with the best fit should be selected

What the MEI Package can do

- Test configural invariance (including model fit for each group), metric and scalar invariance
- Identify items with non-invariant factor loadings for the partial metric invariance model
- Identify items with non-invariant intercepts for the partial scalar invariance model
- Compare latent means across groups
- Compare defined parameters (e.g., direct, indirect and total effects) across groups
- Identify the partial metric invariance model, the partial scalar invariance model and compare latent means across times or across sources
- Identify items with non-invariant factor loadings across levels

Install the MEI Package

From the RStudio

```
install.packages("devtools")
devtools::install_github("gche935/MEI")
```

- After installation, type `library(MEI)` and you will see a welcome message if you have successfully installed MEI

Measurement Invariance Tests: Example A

- Simulated dataset (400 observations for each region) based on the 6th European Working Conditions Survey (EWCS 2015; Eurofound, 2024)
- Examined the life and working conditions of 43,850 respondents in 35 European countries, which were divided into seven regions: North West, Nordic, Continental, Southern, Central East, North East and South East
- Compare the latent means and indirect effect of work-life conflict on psychological well-being through engagement across the seven regions after controlling for organizational size and tenure
- All items were reverse-coded such that a higher value represents a higher level of the constructs
    - Work-life conflict -- 5 items (Q45a-e; “How often have you - Kept worrying about work when you were not working?”) on a 5-point scale (1 = always; 5 = never)
    - Engagement -- 3 items (Q90a-c; “How often you feel this way - At my work, I feel full of energy?”) on a 5-point scale (1 = always; 5 = never).
    - Psychological wellbeing -- 5 items (Q87a-e; “Been feeling over the last two weeks - I have felt cheerful and in good spirits?”) on a 6-point scale (1 = all of the time; 6 = at no time)

Measurement Invariance Tests: Example A

load the MEI package (only needed once per session)

```
library(MEI)
```

** Welcome to MEI Measurement Equivalence/Invariance Test (version 1.0.10) **

load data file (Example A.csv)

```
Example.A <- read.csv(file = "Example A.csv")
```

Specify the measurement model - Model.A1

```
Model.A <- '
     WorkLifeConflict =~ R45a + R45b + R45c + R45d + R45e
     Engagement =~ R90a + R90b + R90c
     Wellbeing =~ R87a + R87b + R87c + R87d + R87e
'
```

Run Full Measurement Invariance Test

```
Full_MEI(Model.A, Example.A, "Region")  # Region is the grouping variable
```

Example A1: Fit indices

MODEL FIT INDICES

	chisq.scaled	df.scaled	pvalue.scaled	rmsea.robust	cfi.robust	tli.robust	srmr	aic	bic
Model.config	1146.3219	434	0.0000	0.0672	0.9428	0.9281	0.0544	92546.6460	94292.2341
Model.metric	1290.5884	494	0.0000	0.0658	0.9376	0.9311	0.0609	92542.0905	93931.4362
Model.scalar	1715.9588	554	0.0000	0.0747	0.9098	0.9111	0.0670	92863.1581	93896.2613

DIFFERENCES IN FIT INDICES

	cfi.robust	rmsea.robust	srmr
Model.metric - Model.config	-0.0052	-0.0014	0.0064
Model.scalar - Model.metric	-0.0278	0.0089	0.0061

########## Model Fit Indices for each Group in Configural Invariance Model ##########

Group	Sample Size	chisq.scaled	df.scaled	pvalue.scaled	rmsea.robust	cfi.robust	tli.robust	srmr	aic	bic
NorthWest	400	174.6067	62	0.0000	0.0674	0.9360	0.9195	0.0509	13571.6215	13739.2630
Nordic	400	139.8673	62	0.0000	0.0560	0.9469	0.9332	0.0521	12133.5978	12301.2393
Continental	400	118.4582	62	0.0000	0.0477	0.9654	0.9565	0.0489	13469.0496	13636.6911
Southern	400	144.6688	62	0.0000	0.0577	0.9472	0.9336	0.0477	13750.1259	13917.7674
CentralEast	400	162.2901	62	0.0000	0.0636	0.9518	0.9393	0.0531	12918.5567	13086.1982
NorthEast	400	219.4161	62	0.0000	0.0797	0.9166	0.8950	0.0720	12835.7526	13003.3941
SouthEast	400	190.0394	62	0.0000	0.0719	0.9353	0.9186	0.0563	13867.9418	14035.5834

Example A2: Metric Invariance

Run function CompareLoadings in MEI

```
CompareLoadings(Model.A, Example.A, Groups="Region", alpha=0.001)
```

alpha can be changed based on the number of groups. If there
are 5 groups, there will be 10 pair-wise comparisons on each
factor loading/intercept, alpha can be set at 0.05/10 = 0.005

Example A2: Metric Invariance

```
################### Nested Model Comparison #########################

Scaled Chi-Squared Difference Test (method = “satorra.bentler.2001”)

lavaan NOTE:
    The “Chisq” column contains standard test statistics, not the
    robust test that should be reported per model. A robust difference
    test is a function of two standard (not robust) statistics.

               Df   AIC   BIC  Chisq Chisq diff Df diff Pr(>Chisq)
Model.config  434  92547  94292 1262.2
PMI.Model.fit 494  92542  93931 1377.6     140.07      60  2.381e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

####################### Model Fit Indices ###########################
              chisq.scaled df.scaled pvalue.scaled rmsea.robust cfi.robust tli.robust  srmr        aic        bic
Model.config     1146.322†       434          .000         .067       .943†       .928  .054† 92546.646  94292.234
PMI.Model.fit    1290.588        494          .000         .066†      .938        .931† .061  92542.090† 93931.436†

################## Differences in Fit Indices #######################
                             df.scaled rmsea.robust cfi.robust tli.robust  srmr    aic      bic
PMI.Model.fit - Model.config        60       -0.001     -0.005      0.003 0.006 -4.555 -360.798
```

Example A2: PMI.txt

===== Recommended Model =====

```
PMI.Model.R <- '
  WorkLifeConflict =~ c(1,1,1,1,1,1,1)*R45a +
  c(F1L1,F1L1,F1L1,F1L1,F1L1,F1L1,F1L1)*R45b +
  c(F1L2,F1L2,F1L2,F1L2,F1L2,F1L2,F1L2)*R45c +
  c(F1L3,F1L3,F1L3,F1L3,F1L3,F1L3,F1L3)*R45d +
  c(F1L4,F1L4,F1L4,F1L4,F1L4,F1L4,F1L4)*R45e
  Engagement =~ c(1,1,1,1,1,1,1)*R90a +
  c(F2L1,F2L1,F2L1,F2L1,F2L1,F2L1,F2L1)*R90b +
  c(F2L2,F2L2,F2L2,F2L2,F2L2,F2L2,F2L2)*R90c
  Wellbeing =~ c(F3L1,F3L1,F3L1,F3L1,F3L1,F3L1,F3L1)*R87a +
  c(F3L2,F3L2,F3L2,F3L2,F3L2,F3L2,F3L2)*R87b +
  c(1,1,1,1,1,1,1)*R87c +
  c(F3L3,F3L3,F3L3,F3L3,F3L3,F3L3,F3L3)*R87d +
  c(F3L4,F3L4,F3L4,F3L4,F3L4,F3L4,F3L4)*R87e
  '
```

Example A2: Factor Loadings in PMI.Model.R

```
## === Factor Loadings in Final Model === ##
                  Group NorthWest   Group Nordic   Group Continental   Group Southern   Group CentralEast   Group NorthEast   Group SouthEast
 Item R45a                 1.0000         1.0000              1.0000           1.0000              1.0000            1.0000            1.0000
 Item R45b                 1.0951         1.0951              1.0951           1.0951              1.0951            1.0951            1.0951
 Item R45c                 1.4316         1.4316              1.4316           1.4316              1.4316            1.4316            1.4316
 Item R45d                 1.0543         1.0543              1.0543           1.0543              1.0543            1.0543            1.0543
 Item R45e                 0.9299         0.9299              0.9299           0.9299              0.9299            0.9299            0.9299
 Item R90a                 1.0000         1.0000              1.0000           1.0000              1.0000            1.0000            1.0000
 Item R90b                 0.9477         0.9477              0.9477           0.9477              0.9477            0.9477            0.9477
 Item R90c                 0.6376         0.6376              0.6376           0.6376              0.6376            0.6376            0.6376
 Item R87a                 0.8967         0.8967              0.8967           0.8967              0.8967            0.8967            0.8967
 Item R87b                 0.9932         0.9932              0.9932           0.9932              0.9932            0.9932            0.9932
 Item R87c                 1.0000         1.0000              1.0000           1.0000              1.0000            1.0000            1.0000
 Item R87d                 1.0473         1.0473              1.0473           1.0473              1.0473            1.0473            1.0473
 Item R87e                 0.8554         0.8554              0.8554           0.8554              0.8554            0.8554            0.8554
```

Example A3: Scalar Invariance and Comparison of Latent Means

Run function CompareMeans

```
CompareMeans(PMI.Model.R, Example.A, Groups="Region", alpha=0.001)


# PMI.Model.R is the partial metric invariance model from the
# previous step
```

Example A3: PSI.txt

```
## =====  Recommended Model  ===== ##
PSI.Model.R <- '
 WorkLifeConflict =~ c(F1L1,F1L1,F1L1,F1L1,F1L1,F1L1,F1L1)*R45a +
   c(F1L2,F1L2,F1L2,F1L2,F1L2,F1L2,F1L2)*R45b +
   c(1,1,1,1,1,1,1)*R45c +
   c(F1L3,F1L3,F1L3,F1L3,F1L3,F1L3,F1L3)*R45d +
   c(F1L4,F1L4,F1L4,F1L4,F1L4,F1L4,F1L4)*R45e
   R45a ~ c(F1I25,F1I26,F1I25,F1I25,F1I25,F1I25,F1I25)*1
   R45b ~ c(F1I33,F1I34,F1I34,F1I33,F1I34,F1I33,F1I33)*1
   R45c ~ c(0,0,0,0,0,0,0)*1
   R45d ~ c(F1I10,F1I10,F1I10,F1I10,F1I10,F1I10,F1I10)*1
   R45e ~ c(F1I11,F1I11,F1I11,F1I11,F1I11,F1I11,F1I11)*1
 Engagement =~ c(F2L1,F2L1,F2L1,F2L1,F2L1,F2L1,F2L1)*R90a +
   c(F2L2,F2L2,F2L2,F2L2,F2L2,F2L2,F2L2)*R90b +
   c(1,1,1,1,1,1,1)*R90c
   R90a ~ c(F2I24,F2I24,F2I24,F2I25,F2I24,F2I24,F2I24)*1
   R90b ~ c(F2I30,F2I30,F2I26,F2I30,F2I26,F2I26,F2I26)*1
   R90c ~ c(0,0,0,0,0,0,0)*1
```

Example A3: PSI.txt

```
Wellbeing =~ c(F3L1,F3L1,F3L1,F3L1,F3L1,F3L1,F3L1)*R87a +
  c(F3L2,F3L2,F3L2,F3L2,F3L2,F3L2,F3L2)*R87b +
  c(1,1,1,1,1,1,1)*R87c +
  c(F3L3,F3L3,F3L3,F3L3,F3L3,F3L3,F3L3)*R87d +
  c(F3L4,F3L4,F3L4,F3L4,F3L4,F3L4,F3L4)*R87e
  R87a ~ c(F3I37,F3I37,F3I37,F3I36,F3I36,F3I36,F3I36)*1
  R87b ~ c(F3I38,F3I38,F3I38,F3I41,F3I38,F3I41,F3I38)*1
  R87c ~ c(0,0,0,0,0,0,0)*1
  R87d ~ c(F3I42,F3I42,F3I42,F3I42,F3I43,F3I42,F3I42)*1
  R87e ~ c(F3I44,F3I44,F3I44,F3I44,F3I44,F3I45,F3I44)*1
 '
```

Example A3: Scalar Invariance

```
################### Nested Model Comparison ########################

Scaled Chi-Squared Difference Test (method = "satorra.bentler.2001")

lavaan NOTE:
    The "Chisq" column contains standard test statistics, not the
    robust test that should be reported per model. A robust difference
    test is a function of two standard (not robust) statistics.

               Df   AIC   BIC  Chisq Chisq diff Df diff Pr(>Chisq)
PMI.Model.fit 494 92542 93931 1377.6
PSI.Model.fit 546 92594 93675 1533.5     156.76      52  1.939e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

###################### Model Fit Indices ###########################
              chisq.scaled df.scaled pvalue.scaled rmsea.robust cfi.robust tli.robust  srmr        aic         bic
PMI.Model.fit    1290.588†       494          .000         .066†      .938†      .931† .061† 92542.090† 93931.436
PSI.Model.fit    1446.049        546          .000         .066       .930       .930  .063  92594.002  93674.604†

################## Differences in Fit Indices #######################
                              df.scaled rmsea.robust cfi.robust tli.robust  srmr    aic      bic
PSI.Model.fit - PMI.Model.fit        52            0     -0.008     -0.001 0.002 51.911 -256.832
```

Each group has 10 estimated intercepts (13 items – 3 fixed intercepts)
In 7 groups, a maximum of 10 x (7-1) = 60 equally constrained intercepts
60 – 52 = 8 additional freely estimated intercepts

Example A3: Scalar Invariance

```
## === Factor Loadings in Final Model === ##
           Group NorthWest  Group Nordic  Group Continental  Group Southern  Group CentralEast  Group NorthEast  Group SouthEast
 Item R45a          0.7058        0.7058             0.7058          0.7058             0.7058           0.7058           0.7058
 Item R45b          0.7714        0.7714             0.7714          0.7714             0.7714           0.7714           0.7714
 Item R45c          1.0000        1.0000             1.0000          1.0000             1.0000           1.0000           1.0000
 Item R45d          0.7404        0.7404             0.7404          0.7404             0.7404           0.7404           0.7404
 Item R45e          0.6515        0.6515             0.6515          0.6515             0.6515           0.6515           0.6515
 Item R90a          1.5495        1.5495             1.5495          1.5495             1.5495           1.5495           1.5495
 Item R90b          1.4907        1.4907             1.4907          1.4907             1.4907           1.4907           1.4907
 Item R90c          1.0000        1.0000             1.0000          1.0000             1.0000           1.0000           1.0000
 Item R87a          0.8931        0.8931             0.8931          0.8931             0.8931           0.8931           0.8931
 Item R87b          0.9894        0.9894             0.9894          0.9894             0.9894           0.9894           0.9894
 Item R87c          1.0000        1.0000             1.0000          1.0000             1.0000           1.0000           1.0000
 Item R87d          1.0440        1.0440             1.0440          1.0440             1.0440           1.0440           1.0440
 Item R87e          0.8551        0.8551             0.8551          0.8551             0.8551           0.8551           0.8551

## === Intercepts in Final Model === ##
           Group NorthWest  Group Nordic  Group Continental  Group Southern  Group CentralEast  Group NorthEast  Group SouthEast
 Item R45a          0.7982        0.9560             0.7982          0.7982             0.7982           0.7982           0.7982
 Item R45b          1.1413        0.9148             0.9148          1.1413             0.9148           1.1413           1.1413
 Item R45c          0.0000        0.0000             0.0000          0.0000             0.0000           0.0000           0.0000
 Item R45d          0.1073        0.1073             0.1073          0.1073             0.1073           0.1073           0.1073
 Item R45e          0.1591        0.1591             0.1591          0.1591             0.1591           0.1591           0.1591
 Item R90a         -2.6760       -2.6760            -2.6760         -2.3626            -2.6760          -2.6760          -2.6760
 Item R90b         -2.1047       -2.1047            -2.4128         -2.1047            -2.4128          -2.4128          -2.4128
 Item R90c          0.0000        0.0000             0.0000          0.0000             0.0000           0.0000           0.0000
 Item R87a          0.7585        0.7585             0.7585          0.5429             0.5429           0.5429           0.5429
 Item R87b          0.0294        0.0294             0.0294         -0.1535             0.0294          -0.1535           0.0294
 Item R87c          0.0000        0.0000             0.0000          0.0000             0.0000           0.0000           0.0000
 Item R87d         -0.4637       -0.4637            -0.4637         -0.4637            -0.2445          -0.4637          -0.4637
 Item R87e          0.7598        0.7598             0.7598          0.7598             0.7598           0.4987           0.7598

## === Latent Means in Final Model === ##
                 Group NorthWest  Group Nordic  Group Continental  Group Southern  Group CentralEast  Group NorthEast  Group SouthEast
WorkLifeConflict          2.2751        2.2690             2.1713          2.2856             2.1106           2.1734           2.3700
Engagement                4.0925        4.1434             4.2139          4.0158             4.0652           4.1760           4.0831
Wellbeing                 4.2298        4.5005             4.3743          4.5358             4.3196           4.4777           4.3478
```

Example A3: Latent Means Comparisons – Work Life Conflict

===== Latent Variable: WorkLifeConflict =====

== Intercepts ==

	Group NorthWest	Group Nordic	Group Continental	Group Southern	Group CentralEast	Group NorthEast	Group SouthEast
Item R45a	0.7982	0.9559	0.7982	0.7982	0.7982	0.7982	0.7982
Item R45b	1.1413	0.9148	0.9148	1.1413	0.9148	1.1413	1.1413
Item R45c	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Item R45d	0.1073	0.1073	0.1073	0.1073	0.1073	0.1073	0.1073
Item R45e	0.1591	0.1591	0.1591	0.1591	0.1591	0.1591	0.1591

== Latent Means ==

Group NorthWest	Group Nordic	Group Continental	Group Southern	Group CentralEast	Group NorthEast	Group SouthEast
2.2751	2.2690	2.1714	2.2856	2.1106	2.1734	2.3700

== Pairwise Difference in Latent Means ==

	Group NorthWest	Group Nordic	Group Continental	Group Southern	Group CentralEast	Group NorthEast	Group SouthEast
Group NorthWest	0.0000	0.0062	0.1038	-0.0105	0.1645	0.1018	-0.0948
Group Nordic	0.0062	0.0000	0.0976	-0.0166	0.1584	0.0956	-0.1010
Group Continental	0.1038	0.0976	0.0000	-0.1143	0.0607	-0.0020	-0.1986
Group Southern	-0.0105	-0.0166	-0.1143	0.0000	0.1750	0.1122	-0.0844
Group CentralEast	0.1645	0.1584	0.0607	0.1750	0.0000	-0.0628	-0.2594
Group NorthEast	0.1018	0.0956	-0.0020	0.1122	-0.0628	0.0000	-0.1966
Group SouthEast	-0.0948	-0.1010	-0.1986	-0.0844	-0.2594	-0.1966	0.0000

== p-values for pairwise comparisons ==

	Group NorthWest	Group Nordic	Group Continental	Group Southern	Group CentralEast	Group NorthEast	Group SouthEast
Group NorthWest	0.0000	0.9205	0.1083	0.8724	0.0099	0.1100	0.1643
Group Nordic	0.9205	0.0000	0.1020	0.7853	0.0073	0.1134	0.1169
Group Continental	0.1083	0.1020	0.0000	0.0749	0.3354	0.9739	0.0031
Group Southern	0.8724	0.7853	0.0749	0.0000	0.0058	0.0761	0.2088
Group CentralEast	0.0099	0.0073	0.3354	0.0058	0.0000	0.3153	0.0001
Group NorthEast	0.1100	0.1134	0.9739	0.0761	0.3153	0.0000	0.0030
Group SouthEast	0.1643	0.1169	0.0031	0.2088	0.0001	0.0030	0.0000

NorthWest & CentralEast have 4 invariant items, latent means can be compared with partial scalar invariance.
Nordic & CentralEast have 4 invariant items, latent means can be compared with partial scalar invariance.
Continental & SouthEast have 4 invariant items, latent means can be compared with partial scalar invariance.
Southern & CentralEast have 4 invariant items, latent means can be compared with partial scalar invariance.
CentralEast & SouthEast have 4 invariant items, latent means can be compared with full scalar invariance.
NorthEast & SouthEast have 5 invariant items, latent means can be compared with full scalar invariance.

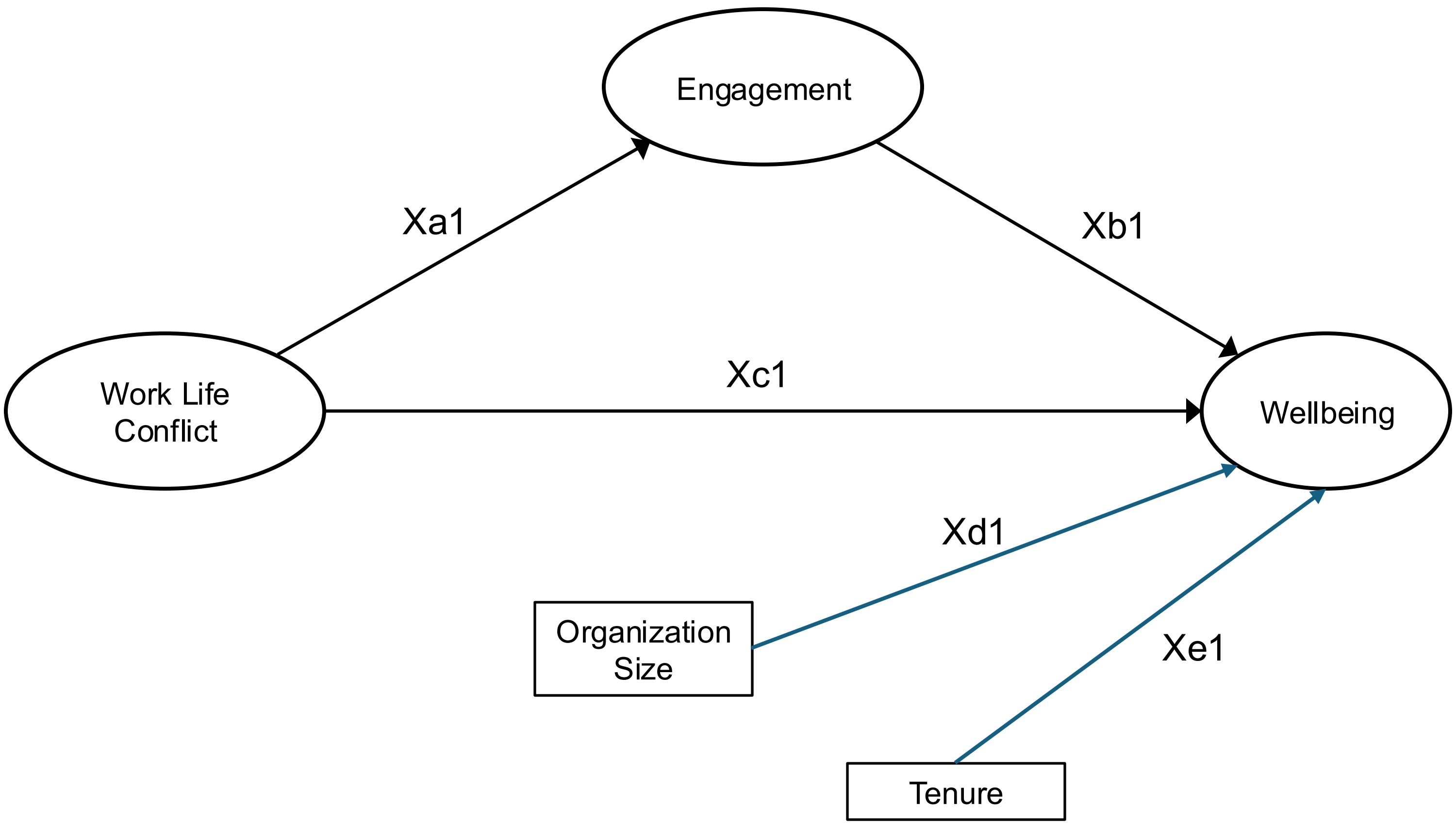
Example A3: Latent Means Comparisons – Engagement

```
## ===== Latent Variable:  Engagement  ===== ##
== Intercepts ==
            Group NorthWest    Group Nordic    Group Continental    Group Southern    Group CentralEast    Group NorthEast    Group SouthEast
 Item R90a    -2.6760           -2.6760          -2.6760              -2.3626            -2.6760              -2.6760            -2.6760
 Item R90b    -2.1047           -2.1047          -2.4128              -2.1047            -2.4128              -2.4128            -2.4128
 Item R90c     0.0000            0.0000           0.0000               0.0000             0.0000               0.0000             0.0000

== Latent Means ==
     Group NorthWest        Group Nordic     Group Continental      Group Southern     Group CentralEast      Group NorthEast      Group SouthEast
              4.0925              4.1434                4.2139              4.0158                4.0652               4.1760               4.0831

== Pairwise Difference in Latent Means ==
                     Group NorthWest    Group Nordic    Group Continental    Group Southern    Group CentralEast    Group NorthEast    Group SouthEast
 Group NorthWest      0.0000            -0.0509          -0.1214              0.0767            0.0273              -0.0835             0.0094
 Group Nordic        -0.0509             0.0000          -0.0705              0.1276            0.0783              -0.0325             0.0604
 Group Continental   -0.1214            -0.0705           0.0000              0.1981            0.1488               0.0379             0.1308
 Group Southern       0.0767             0.1276           0.1981              0.0000           -0.0493              -0.1601            -0.0672
 Group CentralEast    0.0273             0.0783           0.1488             -0.0493            0.0000              -0.1108            -0.0179
 Group NorthEast     -0.0835            -0.0325           0.0379             -0.1601           -0.1108               0.0000             0.0929
 Group SouthEast      0.0094             0.0604           0.1308             -0.0672           -0.0179               0.0929             0.0000

== p-values for pairwise comparisons ==
                     Group NorthWest    Group Nordic    Group Continental    Group Southern    Group CentralEast    Group NorthEast    Group SouthEast
 Group NorthWest     0.0000             0.1075          0.0005               0.0627            0.4720               0.0191             0.8077
 Group Nordic        0.1075             0.0000          0.0244               0.0009            0.0221               0.3145             0.0859
 Group Continental   0.0005             0.0244          0.0000               0.0000            0.0000               0.2569             0.0006
 Group Southern      0.0627             0.0009          0.0000               0.0000            0.2621               0.0002             0.1347
 Group CentralEast   0.4720             0.0221          0.0000               0.2621            0.0000               0.0030             0.6517
 Group NorthEast     0.0191             0.3145          0.2569               0.0002            0.0030               0.0000             0.0174
 Group SouthEast     0.8077             0.0859          0.0006               0.1347            0.6517               0.0174             0.0000
```

NorthWest & Continental have 2 invariant items, latent means can be compared with partial scalar invariance.
Nordic & Southern have 2 invariant items, latent means can be compared with partial scalar invariance.
Continental & Southern have 1 invariant items, latent means cannot be compared.
Continental & CentralEast have 3 invariant items, latent means can be compared with full scalar invariance.
Continental & SouthEast have 3 invariant items, latent means can be compared with full scalar invariance.
Southern & NorthEast have 1 invariant items, latent means cannot be compared.
CantralEast & NorthEast have 3 invariant items, latent means can be compared with full scalar invariance.

Testing Moderated Mediation
Moderator is Categorical

Mediation Model



Example A4: Compare Parameters

- Based on partial metric invariance model (PMI.Model.R)

== Example A4 - Compare Parameters ==

```
Model.DP <- '
  Wellbeing ~ Xb1*Engagement+Xc1*WorkLifeConflict+Xd1*OrgSize+Xe1*Tenure
  Engagement ~ Xa1*WorkLifeConflict
  # Defined Parameters #
  IndirectP := Xa1*Xb1
  DirectP := Xc1
  Total := Xa1*Xb1 + Xc1
'

CompareParameters(PMI.Model.R, Model.DP, Example.A, Groups = "Region")
```

Example A4: Fit Indices

```
lavaan 0.6.20 ended normally after 140 iterations

Model Test User Model:
                                              Standard      Scaled
  Test Statistic                              1540.016    1470.461
  Degrees of freedom                               662         662
  P-value (Chi-square)                           0.000       0.000
  Scaling correction factor                                  1.047
    Yuan-Bentler correction (Mplus variant)
  Test statistic for each group:
    NorthWest                                  240.471     229.611
    Nordic                                     182.024     173.803
    Continental                                155.811     148.774
    Southern                                   182.598     174.351
    CentralEast                                222.774     212.712
    NorthEast                                  296.476     283.086
    SouthEast                                  259.862     248.125

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    0.936       0.936
  Tucker-Lewis Index (TLI)                       0.930       0.930


Root Mean Square Error of Approximation:

  RMSEA                                          0.058       0.055
  90 Percent confidence interval - lower         0.054       0.052
  90 Percent confidence interval - upper         0.061       0.059
  P-value H_0: RMSEA <= 0.050                    0.001       0.010
  P-value H_0: RMSEA >= 0.080                    0.000       0.000

Standardized Root Mean Square Residual:

  SRMR                                           0.058       0.058
```

Example A4: Compare Parameters

Defined Parameters:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
GP1IndirectP	-0.241	0.078	-3.088	0.002	-0.139	-0.139
GP1DirectP	-0.541	0.085	-6.351	0.000	-0.311	-0.311
GP1Total	-0.783	0.107	-7.310	0.000	-0.450	-0.450
GP2IndirectP	-0.207	0.090	-2.292	0.022	-0.117	-0.117
GP2DirectP	-0.702	0.084	-8.360	0.000	-0.397	-0.397
GP2Total	-0.909	0.112	-8.104	0.000	-0.514	-0.514
GP3IndirectP	-0.321	0.090	-3.575	0.000	-0.172	-0.172
GP3DirectP	-0.618	0.090	-6.879	0.000	-0.332	-0.332
GP3Total	-0.939	0.107	-8.751	0.000	-0.504	-0.504
GP4IndirectP	-0.319	0.090	-3.540	0.000	-0.182	-0.182
GP4DirectP	-0.353	0.090	-3.909	0.000	-0.202	-0.202
GP4Total	-0.672	0.127	-5.284	0.000	-0.384	-0.384
GP5IndirectP	-0.373	0.090	-4.162	0.000	-0.205	-0.205
GP5DirectP	-0.420	0.092	-4.572	0.000	-0.230	-0.230
GP5Total	-0.793	0.115	-6.917	0.000	-0.435	-0.435
GP6IndirectP	-0.298	0.097	-3.073	0.002	-0.174	-0.174
GP6DirectP	-0.459	0.079	-5.819	0.000	-0.268	-0.268
GP6Total	-0.756	0.113	-6.682	0.000	-0.442	-0.442
GP7IndirectP	-0.382	0.098	-3.907	0.000	-0.227	-0.227
GP7DirectP	-0.252	0.082	-3.095	0.002	-0.150	-0.150
GP7Total	-0.634	0.111	-5.715	0.000	-0.378	-0.378

Example A4: Compare Parameters

Number of Successful Simulated Samples = 1000000

===== Defined Parameter: IndirectP =====

	Group NorthWest	Group Nordic	Group Continental	Group Southern	Group CentralEast	Group NorthEast	Group SouthEast
Estimated Defined Parameter	-0.2414	-0.2069	-0.3209	-0.3191	-0.3730	-0.2978	-0.3820
Lower Limit 99% CI	-0.4621	-0.4540	-0.5679	-0.5756	-0.6286	-0.5626	-0.6558
Lower Limit 95% CI	-0.4045	-0.3907	-0.5043	-0.5084	-0.5621	-0.4947	-0.5854
Lower Limit 90% CI	-0.3760	-0.3591	-0.4726	-0.4753	-0.5291	-0.4615	-0.5502
Upper Limit 90% CI	-0.1182	-0.0611	-0.1768	-0.1778	-0.2338	-0.1418	-0.2282
Upper Limit 95% CI	-0.0961	-0.0337	-0.1507	-0.1532	-0.2099	-0.1132	-0.2014
Upper Limit 99% CI	-0.0534	0.0207	-0.0992	-0.1064	-0.1652	-0.0572	-0.1491

== p-values for pairwise comparisons ==

	Group NorthWest	Group Nordic	Group Continental	Group Southern	Group CentralEast	Group NorthEast	Group SouthEast
Group NorthWest	0.0000	0.7672	0.5090	0.5176	0.2641	0.6547	0.2582
Group Nordic	0.7672	0.0000	0.3717	0.3757	0.1844	0.4919	0.1816
Group Continental	0.5090	0.3717	0.0000	0.9886	0.6703	0.8619	0.6346
Group Southern	0.5176	0.3757	0.9886	0.0000	0.6597	0.8730	0.6253
Group CentralEast	0.2641	0.1844	0.6703	0.6597	0.0000	0.5597	0.9421
Group NorthEast	0.6547	0.4919	0.8619	0.8730	0.5597	0.0000	0.5309
Group SouthEast	0.2582	0.1816	0.6346	0.6253	0.9421	0.5309	0.0000

===== Defined Parameter: DirectP =====

	Group NorthWest	Group Nordic	Group Continental	Group Southern	Group CentralEast	Group NorthEast	Group SouthEast
Estimated Defined Parameter	-0.5412	-0.7022	-0.6176	-0.3532	-0.4203	-0.4586	-0.2523
Lower Limit 99% CI	-0.7598	-0.9185	-0.8485	-0.5856	-0.6570	-0.6615	-0.4621
Lower Limit 95% CI	-0.7080	-0.8668	-0.7936	-0.5303	-0.6005	-0.6131	-0.4119
Lower Limit 90% CI	-0.6813	-0.8403	-0.7654	-0.5019	-0.5715	-0.5883	-0.3864
Upper Limit 90% CI	-0.4011	-0.5641	-0.4698	-0.2043	-0.2693	-0.3292	-0.1183
Upper Limit 95% CI	-0.3739	-0.5375	-0.4414	-0.1760	-0.2404	-0.3043	-0.0927
Upper Limit 99% CI	-0.3214	-0.4860	-0.3865	-0.1199	-0.1844	-0.2556	-0.0421

== p-values for pairwise comparisons ==

	Group NorthWest	Group Nordic	Group Continental	Group Southern	Group CentralEast	Group NorthEast	Group SouthEast
Group NorthWest	0.0000	0.1736	0.5355	0.1299	0.3328	0.4724	0.0139
Group Nordic	0.1736	0.0000	0.4826	0.0044	0.0214	0.0297	0.0001
Group Continental	0.5355	0.4826	0.0000	0.0375	0.1214	0.1752	0.0023
Group Southern	0.1299	0.0044	0.0375	0.0000	0.6024	0.3759	0.4062
Group CentralEast	0.3328	0.0214	0.1214	0.6024	0.0000	0.7484	0.1691
Group NorthEast	0.4724	0.0297	0.1752	0.3759	0.7484	0.0000	0.0659
Group SouthEast	0.0139	0.0001	0.0023	0.4062	0.1691	0.0659	0.0000

===== Defined Parameter: Total =====

	Group NorthWest	Group Nordic	Group Continental	Group Southern	Group CentralEast	Group NorthEast	Group SouthEast
Estimated Defined Parameter	-0.7826	-0.9091	-0.9385	-0.6723	-0.7933	-0.7564	-0.6343
Lower Limit 99% CI	-1.0580	-1.1957	-1.2100	-1.0044	-1.0916	-1.0445	-0.9252
Lower Limit 95% CI	-0.9926	-1.1270	-1.1462	-0.9229	-1.0197	-0.9759	-0.8549
Lower Limit 90% CI	-0.9591	-1.0925	-1.1128	-0.8825	-0.9831	-0.9411	-0.8192
Upper Limit 90% CI	-0.6055	-0.7219	-0.7591	-0.4626	-0.6057	-0.5681	-0.4532
Upper Limit 95% CI	-0.5706	-0.6845	-0.7236	-0.4220	-0.5692	-0.5308	-0.4186
Upper Limit 99% CI	-0.5014	-0.6117	-0.6527	-0.3417	-0.4972	-0.4566	-0.3507

== p-values for pairwise comparisons ==

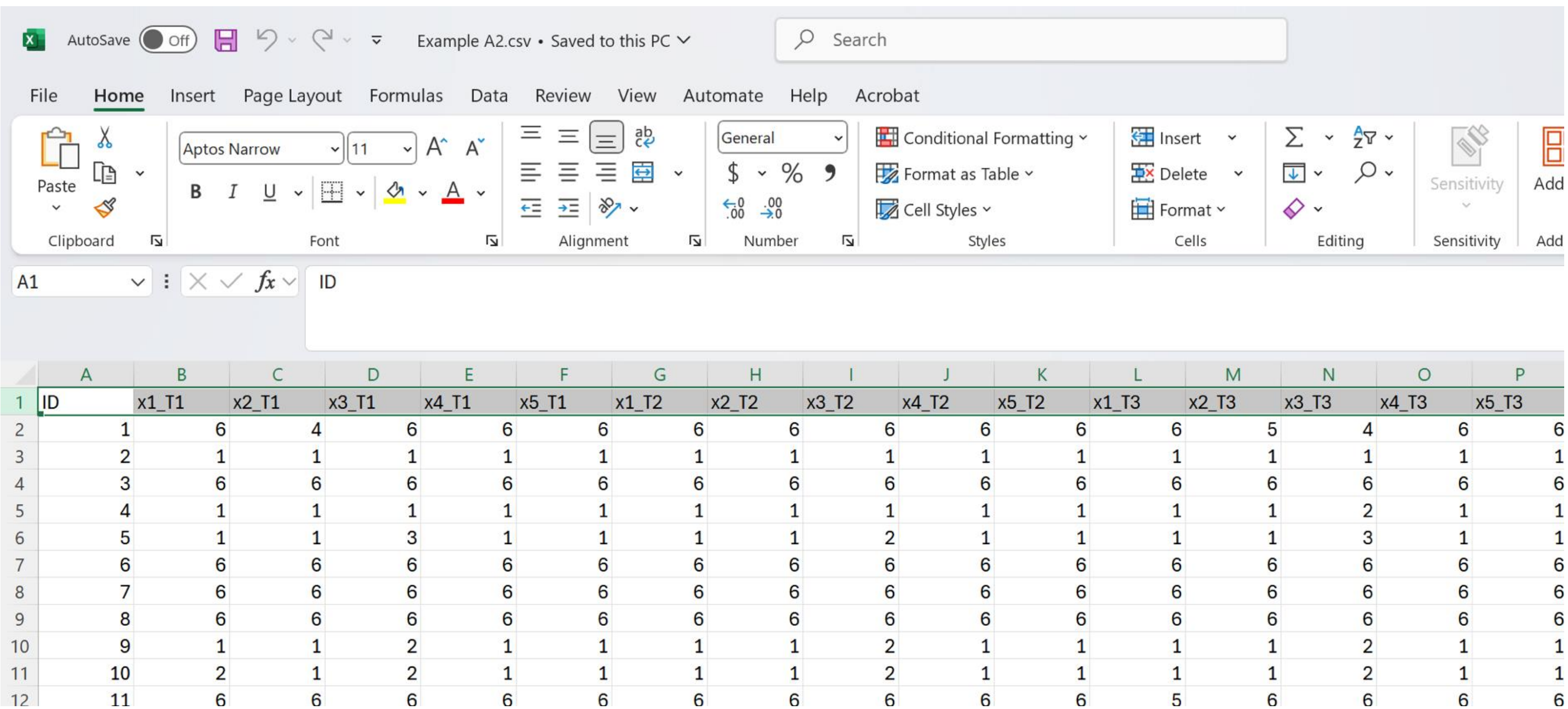
	Group NorthWest	Group Nordic	Group Continental	Group Southern	Group CentralEast	Group NorthEast	Group SouthEast
Group NorthWest	0.0000	0.4206	0.3088	0.5053	0.9430	0.8642	0.3325
Group Nordic	0.4206	0.0000	0.8483	0.1608	0.4624	0.3286	0.0770
Group Continental	0.3088	0.8483	0.0000	0.1079	0.3451	0.2329	0.0452
Group Southern	0.5053	0.1608	0.1079	0.0000	0.4691	0.6162	0.8219
Group CentralEast	0.9430	0.4624	0.3451	0.4691	0.0000	0.8088	0.3001
Group NorthEast	0.8642	0.3286	0.2329	0.6162	0.8088	0.0000	0.4304
Group SouthEast	0.3325	0.0770	0.0452	0.8219	0.3001	0.4304	0.0000

Example B – Testing MEI in Longitudinal Studies

- Simulated dataset with 242 observations across three time points based on Study 2 in Wu, Chen and Tsai (2009: 300)
  - Wu, C., Chen, L. H., & Tsai, Y. 2009. Longitudinal invariance analysis of the satisfaction with life scale. *Personality and Individual Differences*, 46: 396-401.
- Examined the longitudinal measurement invariance of Diener et al.'s (1985) satisfaction with life scale (SWLS), measured with 5 items using a 6-point Likert scale across three time points.
- Compare the latent means across time

Example B – Testing MEI in Longitudinal Studies

- MEI only requires defining the measurement model once. All items should have the suffix _T1, _T2 and _T3 (e.g., x1_T1, x1_T2, x1_T3) in the data file to indicate when the item was measured.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	ID	x1_T1	x2_T1	x3_T1	x4_T1	x5_T1	x1_T2	x2_T2	x3_T2	x4_T2	x5_T2	x1_T3	x2_T3	x3_T3	x4_T3	x5_T3
2	1	6	4	6	6	6	6	6	6	6	6	6	5	4	6	6
3	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	3	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
5	4	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1
6	5	1	1	3	1	1	1	1	2	1	1	1	1	3	1	1
7	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
8	7	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
9	8	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
10	9	1	1	2	1	1	1	1	2	1	1	1	1	2	1	1
11	10	2	1	2	1	1	1	1	2	1	1	1	1	2	1	1
12	11	6	6	6	6	6	6	6	6	6	6	5	6	6	6	6

Example B1

load the MEI package (only needed once per session)

```
library(MEI)
```

** Welcome to MEI Measurement Equivalence/Invariance Test (version 1.0.10) **

load data file (Example B.csv)

```
Example.B <- read.csv(file = "Example B.csv")
```

Specify the measurement model - Model.B

```
Model.B <- 'SWLC =~ x1 + x2 + x3 + x4 + x5'
```

Metric Invariance Test (specify the number of waves with no.waves)

```
LGCompareLoadings(Model.B, Example.B, no.waves=3, alpha=0.01)
```

Example B1: PMI.txt

```
## =====  Recommended Model  ===== ##
PMI.Model.R <- '
    SWLC_T1 =~ 1*x1_T1 + F1L1*x2_T1 + F1L2*x3_T1 + F1L3*x4_T1 + F1L4*x5_T1
    SWLC_T2 =~ 1*x1_T2 + F1L1*x2_T2 + F1L2*x3_T2 + F1L3*x4_T2 + F1L4*x5_T2
    SWLC_T3 =~ 1*x1_T3 + F1L1*x2_T3 + F1L2*x3_T3 + F1L3*x4_T3 + F1L5*x5_T3

    x1_T1  ~~  x1_T2
    x1_T1  ~~  x1_T3
    x1_T2  ~~  x1_T3
    x2_T1  ~~  x2_T2
    x2_T1  ~~  x2_T3
    x2_T2  ~~  x2_T3
    x3_T1  ~~  x3_T2
    x3_T1  ~~  x3_T3
    x3_T2  ~~  x3_T3
    x4_T1  ~~  x4_T2
    x4_T1  ~~  x4_T3
    x4_T2  ~~  x4_T3
    x5_T1  ~~  x5_T2
    x5_T1  ~~  x5_T3
    x5_T2  ~~  x5_T3

   '
```

Example B1: Metric Invariance

```
## Compare fit indices across models ##
                              chisq.scaled  df.scaled  pvalue.scaled  rmsea.scaled  cfi.scaled  tlicaled   srmr        aic        bic
Model.config                      121.8175         72         0.0002        0.0535      0.9645    0.9483 0.0505 10626.6743 10846.4774
PMI.Model.fit                     132.3242         79         0.0002        0.0528      0.9620    0.9495 0.0543 10618.8582 10814.2387
PMI.Model.fit - Model.config       10.5067          7         0.1616       -0.0007     -0.0025    0.0013 0.0038    -7.8161   -32.2387
```

```
## === Factor Loadings in Final Model === ##
             Time 1 Time 2 Time 3
 Item x1     1.0000 1.0000 1.0000
 Item x2     1.1010 1.1010 1.1010
 Item x3     1.1491 1.1491 1.1491
 Item x4     1.1956 1.1956 1.1956
 Item x5     0.8015 0.8015 1.1589
```

Each wave has 4 estimated factor loadings (5 items – 1 fixed factor loading)
In 3 waves, a maximum of 4 x (3-1) = 8 equally constrained factor loadings
8 – 7 = 1 additional freely estimated factor loading

Example B2: Scalar Invariance and Comparison of Latent Means

Run function LGCompareMeans

```
LGCompareMeans(PMI.Model.R, Example.B, no.waves=3, alpha=0.01)
```

Example B2: Scalar Invariance

```
                Df   AIC   BIC  Chisq Chisq diff Df diff Pr(>Chisq)
PMI.Model.fit 79 10619 10814 136.53
PSI.Model.fit 86 10623 10794 154.35     17.718       7    0.01331 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

###################### Model Fit Indices ###########################
              chisq.scaled df.scaled pvalue.scaled rmsea.robust cfi.robust tli.robust  srmr        aic        bic
PMI.Model.fit     132.324†        79          .000        .054†      .963†      .951† .054† 10618.858† 10814.239
PSI.Model.fit     149.904         86          .000        .057       .956       .946  .056  10622.679  10793.637†

################# Differences in Fit Indices ######################
                              df.scaled rmsea.robust cfi.robust tli.robust  srmr   aic     bic
PSI.Model.fit - PMI.Model.fit         7        0.003     -0.007     -0.005 0.002 3.821 -20.601


## === Intercepts in Final Model === ##
         Time 1  Time 2  Time 3
 Item x1  0.1352  0.1352  0.1352
 Item x2  0.0000  0.0000  0.0000
 Item x3  0.1640  0.1640  0.1640
 Item x4 -1.0803 -1.0803 -1.0803
 Item x5  0.7230  0.7230 -0.4733

## === Latent Means in Final Model === ##
     Time 1 Time 2 Time 3
 SWLC 3.8779 3.8132  3.906
```

Each wave has 4 estimated intercepts (5 items – 1 fixed intercept)
In 3 waves, a maximum of 4 x (3-1) = 8 equally constrained intercepts
8 – 7 = 1 additional freely estimated intercepts

Example B2: Latent Mean Comparisons

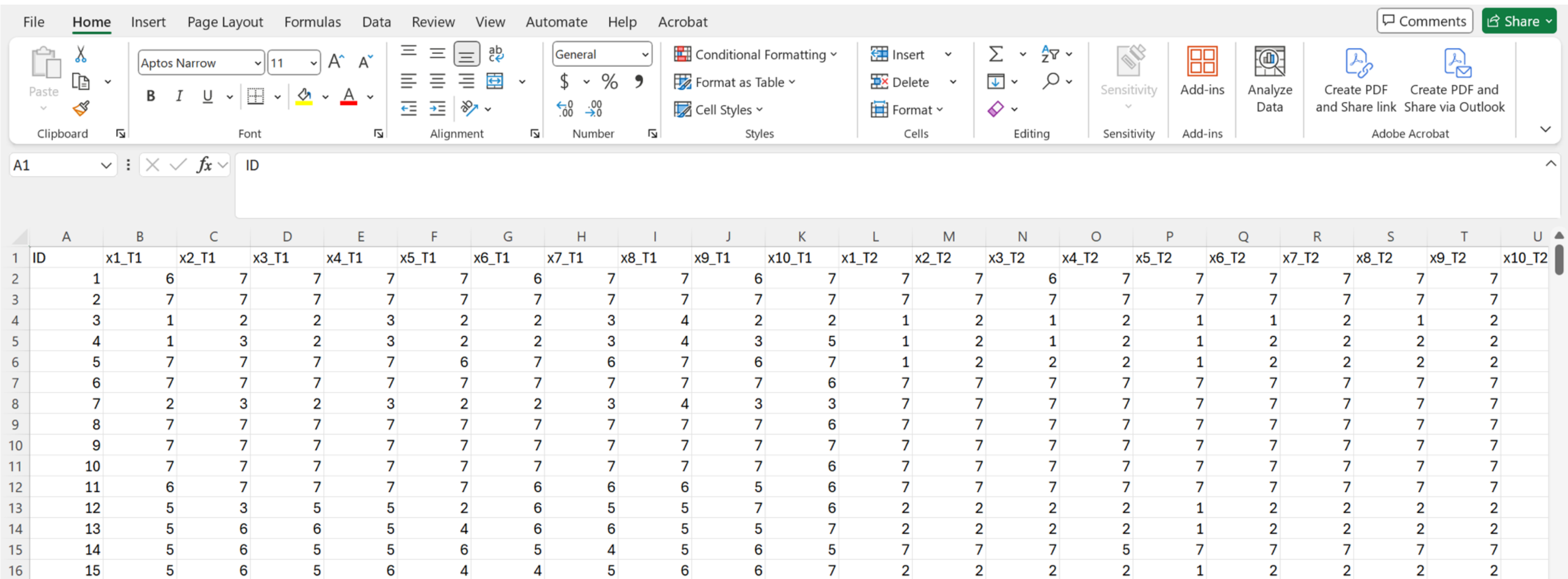
```
## ======= COMPARISON OF LATENT MEANS ======= ##

Number of Successful Simulated Samples = 1000000


## ===== Latent Variable:  SWLC  ===== ##
== Intercepts ==
         Time 1  Time 2  Time 3
 Item x1  0.1352  0.1352  0.1352
 Item x2  0.0000  0.0000  0.0000
 Item x3  0.1640  0.1640  0.1640
 Item x4 -1.0803 -1.0803 -1.0803
 Item x5  0.7230  0.7230 -0.4733

== Latent Means ==
Time 1 Time 2 Time 3
3.8779 3.8132 3.9060

== Pairwise Difference in Latent Means ==
       Time 1  Time 2  Time 3
Time 1  0.0000  0.0646 -0.0282
Time 2  0.0646  0.0000 -0.0928
Time 3 -0.0282 -0.0928  0.0000

== p-values for pairwise comparisons ==
       Time 1 Time 2 Time 3
Time 1 0.0000 0.2483 0.6806
Time 2 0.2483 0.0000 0.1614
Time 3 0.6806 0.1614 0.0000
```

Example C – Testing MEI in Congruence Studies

- Simulated dataset with 332 dyads of mid-level executives and their supervisors based on Ashford and Tsui (1991)
  - Ashford, S. J., & Tsui A. S. 1991. Self-regulation for managerial effectiveness: The role of active feedback seeking. *Academy of Management Journal*, 34: 251-280.
- Mid-level executives and their supervisors rated the executives' performance using Mintzberg's (1973) 10 managerial roles using a 7-point scale, with a higher score representing more effective performance. The 10 roles were classified into internal and external (Cheung, 1999).
- Examined the measurement invariance and compared latent means across sources

Example C – Testing MEI in Congruence Studies

- MEI only requires defining the measurement model once. All items should have the suffix _T1 and _T2 (e.g., x1_T1 and x1_T2) in the data file to indicate the source of the item



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	ID	x1_T1	x2_T1	x3_T1	x4_T1	x5_T1	x6_T1	x7_T1	x8_T1	x9_T1	x10_T1	x1_T2	x2_T2	x3_T2	x4_T2	x5_T2	x6_T2	x7_T2	x8_T2	x9_T2	x10_T2
2	1	6	7	7	7	7	6	7	7	6	7	7	7	6	7	7	7	7	7	7	
3	2	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	
4	3	1	2	2	3	2	2	3	4	2	2	1	2	1	2	1	1	2	1	2	
5	4	1	3	2	3	2	2	3	4	3	5	1	2	1	2	1	2	2	2	2	
6	5	7	7	7	7	6	7	6	7	6	7	1	2	2	2	1	2	2	2	2	
7	6	7	7	7	7	7	7	7	7	7	6	7	7	7	7	7	7	7	7	7	
8	7	2	3	2	3	2	2	3	4	3	3	7	7	7	7	7	7	7	7	7	
9	8	7	7	7	7	7	7	7	7	7	6	7	7	7	7	7	7	7	7	7	
10	9	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	
11	10	7	7	7	7	7	7	7	7	7	6	7	7	7	7	7	7	7	7	7	
12	11	6	7	7	7	7	6	6	6	5	6	7	7	7	7	7	7	7	7	7	
13	12	5	3	5	5	2	6	5	5	7	6	2	2	2	2	1	2	2	2	2	
14	13	5	6	6	5	4	6	6	5	5	7	2	2	2	2	1	2	2	2	2	
15	14	5	6	5	5	6	5	4	5	6	5	7	7	7	5	7	7	7	7	7	
16	15	5	6	5	6	4	4	5	6	6	7	2	2	2	2	1	2	2	2	2	

Example C

load the MEI package (only needed once per session)

```
library(MEI)
```

** Welcome to MEI Measurement Equivalence/Invariance Test (version 1.0.10) **

load data file (Example C.csv)

```
Example.C <- read.csv(file = "Example C.csv")
```

Specify the measurement model - Model.C

```
Model.C <- 'External =~ x1 + x2 + x3 + x4 + x5 + x6
            Internal =~ x7 + x8 + x9 + x10'
```

Metric Invariance Test

```
LGCompareLoadings(Model.C, Example.C, no.waves=2, alpha=0.01)
```

Example C: PMI.txt

```
## =====  Recommended Model  ===== ##
PMI.Model.R <- '
    External_T1 =~ 1*x1_T1 + F1L1*x2_T1 + F1L2*x3_T1 + F1L3*x4_T1 + F1L4*x5_T1 + F1L5*x6_T1
    External_T2 =~ 1*x1_T2 + F1L1*x2_T2 + F1L2*x3_T2 + F1L3*x4_T2 + F1L4*x5_T2 + F1L5*x6_T2

    x1_T1  ~~  x1_T2
    x2_T1  ~~  x2_T2
    x3_T1  ~~  x3_T2
    x4_T1  ~~  x4_T2
    x5_T1  ~~  x5_T2
    x6_T1  ~~  x6_T2

    Internal_T1 =~ 1*x7_T1 + F2L1*x8_T1 + F2L2*x9_T1 + F2L3*x10_T1
    Internal_T2 =~ 1*x7_T2 + F2L1*x8_T2 + F2L2*x9_T2 + F2L3*x10_T2

    x7_T1  ~~  x7_T2
    x8_T1  ~~  x8_T2
    x9_T1  ~~  x9_T2
    x10_T1  ~~  x10_T2

   '
```

Example C: Metric Invariance

```
## Compare fit indices across models ##
                             chisq.scaled   df.scaled   pvalue.scaled   rmsea.scaled   cfi.scaled   tlicaled   srmr         aic        bic
Model.config                     245.9743         154          0.0000         0.0424       0.9561     0.9459 0.0457 18872.8393 19162.030
PMI.Model.fit                    253.8679         162          0.0000         0.0413       0.9562     0.9486 0.0474 18862.4764 19121.226
PMI.Model.fit - Model.config       7.8935           8          0.4439        -0.0011       0.0001     0.0027 0.0017   -10.3629   -40.804
```

```
## === Factor Loadings in Final Model === ##
             Time 1 Time 2
 Item x1     1.0000 1.0000
 Item x2     1.0120 1.0120
 Item x3     0.9691 0.9691
 Item x4     0.9305 0.9305
 Item x5     1.1584 1.1584
 Item x6     0.7736 0.7736
 Item x7     1.0000 1.0000
 Item x8     1.1034 1.1034
 Item x9     0.9599 0.9599
 Item x10    1.0356 1.0356
```

Each source has 8 estimated factor loadings (10 items – 2 fixed factor loadings)
In 2 sources, a maximum of 8 x (2-1) = 8 equally constrained factor loadings
Full Metric Invariance

Example C: Scalar Invariance and Comparison of Latent Means

Run function LGCompareMeans

```
LGCompareMeans(PMI.Model.R, Example.C, no.waves=2, alpha=0.01)
```

Example C: Scalar Invariance

```
                Df   AIC   BIC  Chisq Chisq diff Df diff Pr(>Chisq)
PMI.Model.fit 162 18863 19121 279.41
PSI.Model.fit 166 18859 19103 283.95     4.5777       4     0.3334

####################### Model Fit Indices ###########################
              chisq.scaled df.scaled pvalue.scaled rmsea.robust cfi.robust tli.robust  srmr        aic        bic
PMI.Model.fit     253.868†       162          .000         .044       .955†       .948  .047† 18862.476  19121.226
PSI.Model.fit     258.611        166          .000         .044†      .955        .948† .048  18859.011† 19102.539†

################## Differences in Fit Indices #######################
                                df.scaled rmsea.robust cfi.robust tli.robust srmr    aic     bic
PSI.Model.fit - PMI.Model.fit          4            0          0      0.001    0 -3.466 -18.686


## === Intercepts in Final Model === ##
           Time 1  Time 2
 Item x1  -0.6359 -0.2456
 Item x2   0.0000  0.0000
 Item x3  -0.0723  0.2644
 Item x4   0.3546  0.3546
 Item x5  -1.1884 -1.1884
 Item x6   1.1395  1.4619
 Item x7   0.3929  0.7813
 Item x8   0.0000  0.0000
 Item x9   0.9498  0.9498
 Item x10  0.7459  0.7459

## === Latent Means in Final Model === ##
          Time 1 Time 2
 External 5.2617 4.9858
 Internal 5.8786 5.0769
```

Each source has 8 estimated intercepts (10 items – 2 fixed intercepts)
In 2 sources, a maximum of 8 x (2-1) = 8 equally constrained intercepts
8 – 4 = 4 additional freely estimated intercepts

Example C: Latent Mean Comparisons

```
## ===== Latent Variable:  External  ===== ##
== Intercepts ==
          Time 1  Time 2
 Item x1 -0.6359 -0.2456
 Item x2  0.0000  0.0000
 Item x3 -0.0723  0.2644
 Item x4  0.3546  0.3546
 Item x5 -1.1884 -1.1884
 Item x6  1.1395  1.4619

== Latent Means ==
Time 1 Time 2
5.2617 4.9858

== Pairwise Difference in Latent Means ==
       Time 1 Time 2
Time 1 0.0000 0.2759
Time 2 0.2759 0.0000

== p-values for pairwise comparisons ==
       Time 1 Time 2
Time 1 0.0000 0.0001
Time 2 0.0001 0.0000
```

Since 3 out of 6 items have non-invariant intercepts, latent mean difference should not be interpreted

Example C: Latent Mean Comparisons

```
## ===== Latent Variable:  Internal  ===== ##
== Intercepts ==
          Time 1 Time 2
 Item x7  0.3929 0.7813
 Item x8  0.0000 0.0000
 Item x9  0.9498 0.9498
 Item x10 0.7459 0.7459

== Latent Means ==
Time 1 Time 2
5.8786 5.0769

== Pairwise Difference in Latent Means ==
       Time 1 Time 2
Time 1 0.0000 0.8017
Time 2 0.8017 0.0000

== p-values for pairwise comparisons ==
       Time 1 Time 2
Time 1 0.0000 0.0000
Time 2 0.0000 0.0000
```

Since 3 out of 4 items have invariant intercepts, latent mean difference can be compared with partial scalar invariance. The latent mean difference in Internal roles is 0.8017 ($p < .001$).

Example D – Testing Metric Invariance in Multi-level Studies

- Simulated dataset based on Gruszynska, Basinska and Schaufeli (2021)
  - Gruszczynska, E., Basinska, B. A., & Schaufeli, W. B. 2021. Within- and between-person factor structure of the Oldenburg Burnout Inventory: Analysis of a diary study using multilevel confirmatory factor analysis. *PLoS ONE* 16(5):e0251257. doi: 10.1371/journal.pone.0251257
- Eight negatively-worded items of the Oldenburg Burnout Inventory (OLBI) to measure burnout of 235 employees for 10 consecutive working days
- Since all Level 1 (within-group) variables are group-mean centered, intercepts and latent means of all variables at Level 1 are set to zero. Hence, only configural and metric invariance across levels are tested
- The MEI package requires defining the measurement model only once without specifying the level, and it will compare the model across levels

Example D

load data file (Example D.csv)

```
Example.D <- read.csv(file = "Example D.csv")
```

Specify the measurement model - Model.D

```
Model.D <- 'OLBI =~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8'
```

Metric Invariance Test

```
MLCompareLoadings(Model.D, Example.D, Cluster="ID", alpha=0.05)
```

Example D: PMI.txt

```
## =====  Recommended Model  ===== ##
PMI.Model.R <- '
level: 1
  OLBIw =~ F1L1*x1 +
    1*x2 +
    F1L3*x3 +
    F1L5*x4 +
    F1L6*x5 +
    F1L7*x6 +
    F1L9*x7 +
    F1L10*x8
level: 2
  OLBIb =~ F1L2*x1 +
    1*x2 +
    F1L4*x3 +
    F1L5*x4 +
    F1L6*x5 +
    F1L8*x6 +
    F1L9*x7 +
    F1L11*x8
  '
```

Example D: Metric Invariance

```
 ## Compare fit indices across models ##
                                  chisq   df   pvalue   rmsea     cfi     tli    srmr_within    srmr_between         aic         bic
Model.config                   496.5607   40    0.000  0.0697 0.9462 0.9247          0.036          0.0623 45992.4835 46222.9703
PMI.Model.fit                  497.0873   43    0.000  0.0670 0.9465 0.9303          0.036          0.0620 45988.9885 46202.1888
PMI.Model.fit - Model.config     0.5266    3    0.913 -0.0027 0.0003 0.0056          0.000         -0.0003    -3.4949    -20.7814
```

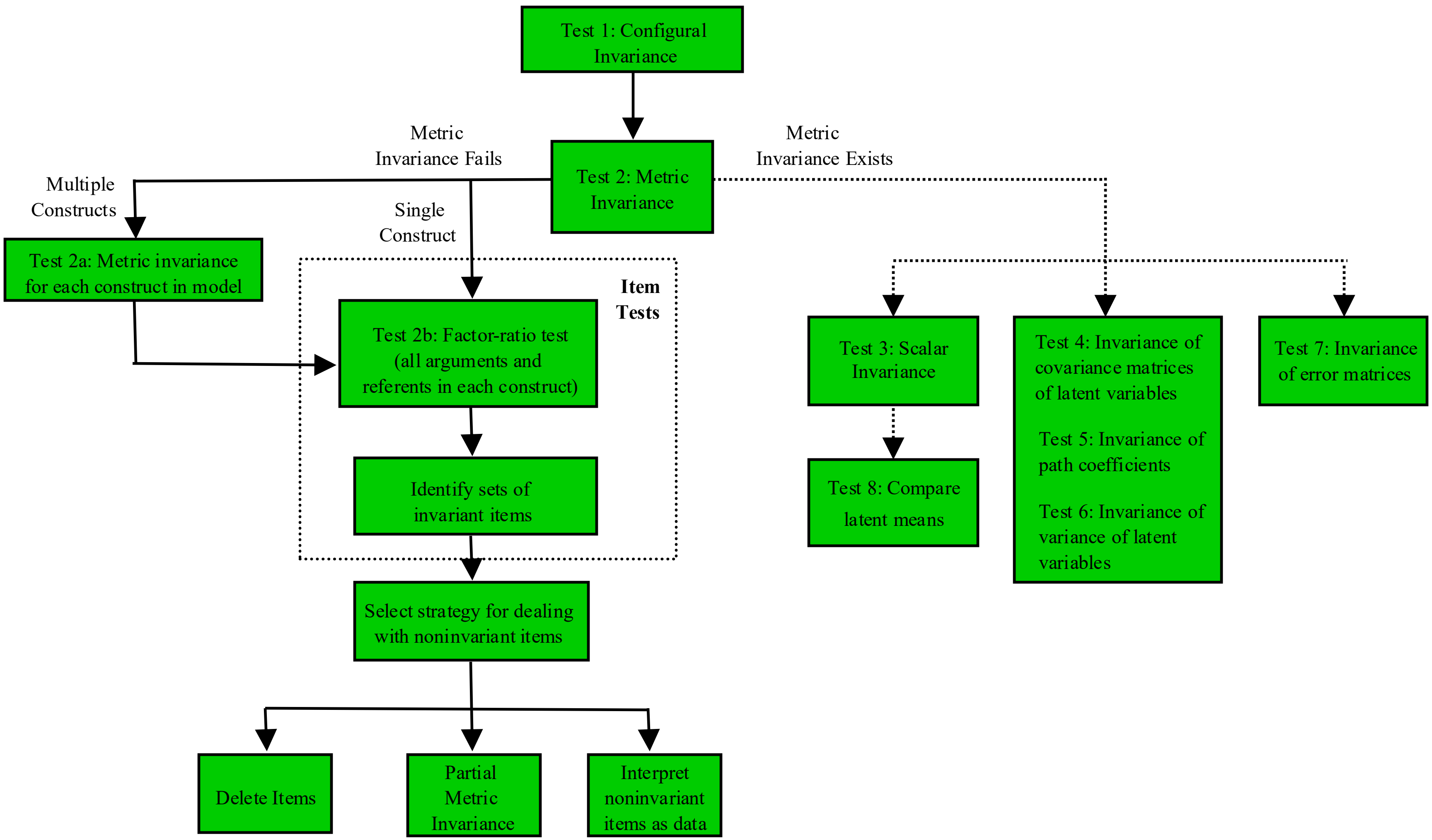
```
## === Factor Loadings in Final Model === ##
              level: 1 level: 2
 Item x1        0.6073   0.9248
 Item x2        1.0000   1.0000
 Item x3        0.4703   0.7187
 Item x4        1.0952   1.0952
 Item x5        0.9269   0.9269
 Item x6        0.5460   0.8174
 Item x7        1.1187   1.1187
 Item x8        0.5929   0.9214
```

Each level has 7 estimated factor loadings (8 items – 1 fixed factor loading)
In 2 levels, a maximum of 7 x (2-1) = 7 equally constrained factor loadings
7 – 3 = 4 additional freely estimated factor loadings

Discussion

- Sequence of ME/I tests
  - Level 1: based on the configural invariance model, test for metric invariance
  - Level 2: based on the full/partial metric invariance model, test for invariance of uniqueness variances, factor variances, factor covariances, item intercepts, or path coefficients
  - Level 3: based on the full/partial scalar invariance model, compare latent means

Measurement Equivalence Tests



Discussion

- Item-level vs. scale-level tests
  - Scale-level tests lack sufficient power to identify non-invariance, especially if testing a whole multifactor model

- Width of BC confidence intervals
  - Wider confidence intervals imply lower precision and power
  - Factors affecting width
    - Sample size, number of items per factor, item communality

Discussion

- The BC confidence interval procedures:
  - Give an estimate of the difference between 2 parameters across groups and a confidence interval for the difference
  - Correct the bias in the sampling distribution
  - Allow all item-level tests for all constructs in one model estimation
- What Invariance tests are required for combining samples:
  - Configural invariance, metric invariance and scalar invariance

Additional References

- Cheung, G. W. & Lau, R. S. (2012). A direct comparison approach for testing measurement invariance. *Organizational Research Methods, 15,* 167-198.
- Cheung, G. W., & Roger B. Rensvold. (2002). "Evaluating Goodness-of-Fit Indices for Testing Measurement Invariance", *Structural Equation Modeling Journal, 9(2),* 233-255.
- Cheung, G. W. & Roger B. Rensvold. (2000) "Assessing Extreme and Acquiescence Response Sets in Cross-Cultural Research using Structural Equations Modeling", *Journal of Cross-Cultural Psychology, 31(2)*, 187-212.
- Cheung, G. W. (1999) "Multifaceted Conceptions of Self-Other Ratings Disagreement", *Personnel Psychology, 52(1),* 1-36.
- Cheung, G. W., & Roger B. Rensvold (1999). "Testing Factorial Invariance Across Groups: A Reconceptualization and Proposed New Method", *Journal of Management, 25(1)*, 1-27.

Additional References

- Vandenberg, R. J. (2002). Toward a further understanding of and improvement in measurement invariance: Methods and procedures. *Organizational Research Methods*, 5, 139-158.
- 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.