Measurement Invariance

A Course in MplusAutomation

Adam Garber

Refresher (measurement invariance involves....)

- Comparing the fit of a measurement model across two or more known groups (i.e., the **grouping** variable)
- This is done by first comparing the fit of the unconditional model using the full sample with models fit to each sub-sample seperately.
- Next a series of multiple-group models are specified where each groups model parameters are estimated simultaneously.
- This series of multiple-group models begins with the **least restrictive** and proceeds to the **most restrictive** which is implemented by restricting sets of parameters to equality across groups.
 - 1. Configural invariance: Free item loadings, intercepts, and residuals
 - 2. Metric invariance: Loadings fixed to equality
 - 3. Scalar invariance: Loadings and intercepts fixed to equality
 - 4. Strict invariance: Loadings, intercepts, and residuals fixed to equality
- Finally, if the solution is determined to be invariant (at least at the Scalar level) than factor mean differences can be evaluated and interpreted (i.e., **Structural invariance**).

Outline

- 1. Prepare data
- 2. Estimate unconditional CFA model (cfa_m0)
- 3. Estimate group 1 sample CFA model (cfa_m1)
- 4. Estimate group 2 sample CFA model (cfa_m2)
- 5. Estimate configural invariance model (cfa m3)
- 6. Estimate metric invariance model (cfa_m4)
- 7. Estimate scalar invariance model (cfa m5)
- 8. Estimate strict invariance model (cfa_m6)
- 9. Estimate structural invariance A model (cfa m7)
- 10. Estimate structural invariance B model(cfa m8)
- 11. Tabulate summary fit statistics to compare models
- 12. Calculate Satora-Bentler scaled Chi-square difference tests

DATA SOURCE: This exercise utilizes the NCES public-use dataset: Education Longitudinal Study of 2002 (Lauff & Ingels, 2014) See website: nces.ed.gov

Load packages

```
library(MplusAutomation)
library(tidyverse)
library(here)
library(gt)
```

Read in data

```
els_data <- read_csv("https://garberadamc.github.io/project-site/data/els_sub5_data.csv")</pre>
```

Prepare data (subset, reorder, rename, and re-code the grouping variabe)

Are there differences in levels of victimization and school climate among 10th grade students attending schools with low or high proportions of free lunch service participation?

- low_free_lunch: Less than 11% of students have service in their school (proxy for low SES)
- high_free_lunch:Greater than or equal to 11% of students have service in their school (proxy for high SES)

See frequency of grouping variable (freelnch)

```
##
## 0 1
## 338 388
```

Reverse code factor 2 items for clear interpretation

```
cols = c("unsafe", "disrupt", "gangs", "rac_fght")
invar_data[ ,cols] <- 5 - invar_data[ ,cols]</pre>
```

Factor names and interpretation:

- VICTIM: Student reports being a victim of injury to self or property
 - scale range: Never, Once or twice, More than twice
 - higher values indicate greater frequency of victimization events
- NEG_CLIM: Student reports on negative school climate attributes
 - scale range: Strongly Disagree Strongly Agree
 - higher values indicate a more negative climate

Estimate the Unconditional Confirmatory Factor Analysis (CFA) model

- 10 item loadings
- 10 intercepts
- 10 residual variances
- 01 factor co-variances

```
cfa_m0 <- mplusObject(</pre>
  TITLE = "model0 - unconditional CFA model",
  VARIABLE =
    "usevar = stolen-rac_fght;",
  ANALYSIS =
    "estimator = mlr;",
  MODEL =
    "!!! Unit Variance Identification !!!
    VICTIM by stolen* t_hurt p_fight hit damaged bullied;
    VICTIM@1;
    NEG_CLIM by unsafe* disrupt gangs rac_fght;
    NEG_CLIM@1; ",
  PLOT = "type = plot3;",
  OUTPUT = "sampstat standardized residual modindices (3.84);",
  usevariables = colnames(invar_data),
  rdata = invar_data)
cfa m0 fit <- mplusModeler(cfa m0,
              dataout=here("07-invariance", "invar_mplus", "invar_data.dat"),
```

```
modelout=here("07-invariance", "invar_mplus", "m0_cfa_fullsample.inp"),
    check=TRUE, run = TRUE, hashfilename = FALSE)
```

Run separate CFA models for each sub-sample

Group freelnch = 0 (low) CFA

```
cfa_m1 <- mplusObject(</pre>
 TITLE = "CFA model1 - group is 0 for freelnch",
    "usevar = stolen-rac_fght;
    !!! USE SAMPLE GROUP 1 (LOW FREE LUNCH) !!!
    USEOBS = freelnch == 0; ",
  ANALYSIS =
    "estimator = mlr;",
  MODEL =
   "!!! Unit Variance Identification !!!
     VICTIM by stolen* t_hurt p_fight hit damaged bullied;
    VICTIM@1;
    NEG_CLIM by unsafe* disrupt gangs rac_fght;
    NEG CLIM@1; ",
 PLOT = "type = plot3;",
  OUTPUT = "sampstat standardized residual modindices (3.84);",
 usevariables = colnames(invar_data),
 rdata = invar_data)
cfa_m1_fit <- mplusModeler(cfa_m1,</pre>
              dataout=here("07-invariance", "invar_mplus", "invar_data.dat"),
              modelout=here("07-invariance", "invar_mplus", "m1_cfa_low_free_lunch.inp"),
              check=TRUE, run = TRUE, hashfilename = FALSE)
```

Group freelnch = 1 (moderate to high) CFA

```
cfa_m2 <- mplusObject(

TITLE = "CFA model2 - group is 1 for freelnch",</pre>
```

```
VARIABLE =
    "usevar = stolen-rac_fght;
     !!! USE SAMPLE GROUP 2 (HIGH FREE LUNCH) !!!
     USEOBS = freelnch == 1; ",
  ANALYSIS =
    "estimator = mlr;",
  MODEL =
    "!!! Unit Variance Identification !!!
     VICTIM by stolen* t_hurt p_fight hit damaged bullied;
     VICTIM@1; ! UVI identification
     NEG_CLIM by unsafe* disrupt gangs rac_fght;
     NEG_CLIM@1; ",
  PLOT = "type = plot3;",
  OUTPUT = "sampstat standardized residual modindices (3.84);",
  usevariables = colnames(invar_data),
 rdata = invar_data)
cfa_m2_fit <- mplusModeler(cfa_m2,</pre>
               dataout=here("07-invariance", "invar_mplus", "invar_data.dat"),
modelout=here("07-invariance", "invar_mplus", "m2_cfa_high_free_lunch.inp"),
               check=TRUE, run = TRUE, hashfilename = FALSE)
```

Multi-Group Invariance Models

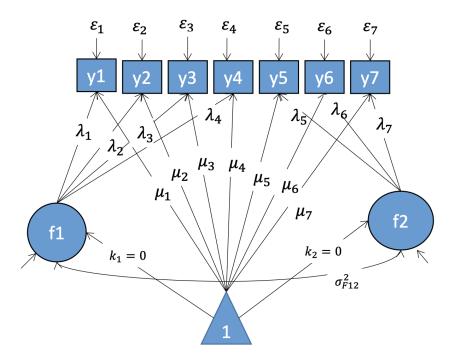


Figure: Picture depicting mean structure from slide by Dr. Karen Nylund-Gibson

Configural invariance

- free item loadings, intercepts, and residuals
- factor means fixed to zero
- factor variances fixed to 1

- 20 item loadings (10items*2groups)
- 20 intercepts
- 20 residual variances
- 02 factor co-variances (1 for each group)

```
cfa_m3 <- mplusObject(</pre>
 TITLE = "CFA model3 - configural invariance",
  VARIABLE =
    "usevar = stolen-rac_fght;
    grouping = freelnch (0=low_free_lunch 1=high_free_lunch); ",
  ANALYSIS =
   "estimator = mlr;",
   "!!! MODEL SPECIFICATION FOR GROUP 1 (LOW FREE LUNCH) !!!
    VICTIM by stolen* t_hurt p_fight hit damaged bullied;
    VICTIM01; ! UVI identification
    NEG_CLIM by unsafe* disrupt gangs rac_fght;
    NEG_CLIM@1;
     [VICTIM-NEG_CLIM@0]; !!! factor means set to zero !!!
     !!! MODEL SPECIFICATION FOR GROUP 2 (HIGH FREE LUNCH) !!!
    MODEL high free lunch:
    VICTIM by stolen* t_hurt p_fight hit damaged bullied; !!! free loadings !!!
    VICTIM@1;
     [stolen t_hurt p_fight hit damaged bullied]; !!! free intercepts !!!
    NEG_CLIM by unsafe* disrupt gangs rac_fght;
    NEG_CLIM@1;
     [unsafe disrupt gangs rac_fght];
     [VICTIM-NEG_CLIM00]; !!! factor means set to zero !!!",
  PLOT = "type = plot3;",
  OUTPUT = "sampstat standardized residual modindices (3.84);",
 usevariables = colnames(invar_data),
 rdata = invar_data)
cfa_m3_fit <- mplusModeler(cfa_m3,</pre>
              dataout=here("07-invariance", "invar_mplus", "invar_data.dat"),
              modelout=here("07-invariance", "invar_mplus", "m3_configural.inp"),
              check=TRUE, run = TRUE, hashfilename = FALSE)
```

Metric invariance

- item loadings (set to equal)
- free intercepts and residuals
- factor means fixed to zero
- free factor variances in group 2

Number of parameters = 54

- 10 item loadings (set to equal)
- 20 intercepts
- 20 residual variances
- 02 factor variances

cfa_m4 <- mplusObject(</pre>

NEG_CLIM;

TITLE = "CFA model4 - metric invariance",

• 02 factor co-variances

VARIABLE =
 "usevar = stolen-rac_fght;

grouping = freelnch (O=low_free_lunch 1=high_free_lunch); ",

ANALYSIS =
 "estimator = mlr;",

MODEL =
 "!!! MODEL SPECIFICATION FOR GROUP 1 (LOW FREE LUNCH) !!!

VICTIM by stolen* t_hurt p_fight hit damaged bullied;
VICTIM@1; ! UVI identification

NEG_CLIM by unsafe* disrupt gangs rac_fght;
NEG_CLIM@1;

[VICTIM-NEG_CLIM@0]; !!! factor means set to zero !!!

!!! MODEL SPECIFICATION FOR GROUP 2 (HIGH FREE LUNCH) !!!

MODEL high_free_lunch:

VICTIM; !!! free factor variances for group 2 !!!

[stolen t_hurt p_fight hit damaged bullied]; !!! free intercepts !!!

Scalar invariance

- item loadings (set to equal)
- intercepts (set to equal)
- free residuals
- free factor variances and means in group $2\,$

- 10 item loadings (set to equal)
- 10 intercepts (set to equal)
- 20 residual variances
- 02 factor variances
- 02 factor co-variances
- 02 factor means

```
cfa_m5 <- mplusObject(
  TITLE = "model5 - scalar invariance",
  VARIABLE =
    "usevar = stolen-rac_fght;
    grouping = freelnch (O=low_free_lunch 1=high_free_lunch); ",

ANALYSIS =
    "estimator = mlr;",</pre>
```

```
"!!! MODEL SPECIFICATION FOR GROUP 1 (LOW FREE LUNCH) !!!
    VICTIM by stolen* t_hurt p_fight hit damaged bullied;
     VICTIM@1;
     NEG_CLIM by unsafe* disrupt gangs rac_fght;
     NEG_CLIM@1;
     [VICTIM-NEG_CLIM@O];
     !!! MODEL SPECIFICATION FOR GROUP 2 (HIGH FREE LUNCH) !!!
    MODEL high_free_lunch:
    VICTIM;
    NEG_CLIM;
     [VICTIM-NEG_CLIM]; !!! free factor means !!!",
  PLOT = "type = plot3;",
  OUTPUT = "sampstat standardized residual modindices (3.84);",
  usevariables = colnames(invar_data),
 rdata = invar_data)
cfa_m5_fit <- mplusModeler(cfa_m5,</pre>
              dataout=here("07-invariance", "invar_mplus", "invar_data.dat"),
              modelout=here("07-invariance", "invar_mplus", "m5_scalar.inp"),
              check=TRUE, run = TRUE, hashfilename = FALSE)
```

Strict invariance

- item loadings (set to equal)
- intercepts (set to equal)
- residuals (set to equal)
- free factor variances and means in group 2

- 10 item loadings (set to equal)
- 10 intercepts (set to equal)
- 10 residual variances
- 02 factor variances
- 02 factor co-variances
- 02 factor means

```
cfa_m6 <- mplusObject(</pre>
 TITLE = "model6 - strict invariance",
 VARIABLE =
   "usevar = stolen-rac_fght;
    grouping = freelnch (0=low_free_lunch 1=high_free_lunch); ",
  ANALYSIS =
   "estimator = mlr;",
 MODEL =
   "!!! MODEL SPECIFICATION FOR GROUP 1 (LOW FREE LUNCH) !!!
    VICTIM by stolen* t_hurt p_fight hit damaged bullied;
    VICTIM@1;
    NEG_CLIM by unsafe* disrupt gangs rac_fght;
    NEG CLIM@1;
     [VICTIM-NEG_CLIM@O];
    stolen-rac fght(1-10); !!! set residuals to be equal across groups !!!
    !!! MODEL SPECIFICATION FOR GROUP 2 (HIGH FREE LUNCH) !!!
    MODEL high_free_lunch:
    VICTIM;
    NEG_CLIM;
     [VICTIM-NEG_CLIM]; !!! free factor means !!!
     stolen-rac_fght(1-10); !!! set residuals to be equal across groups !!!",
  PLOT = "type = plot3;",
  OUTPUT = "sampstat standardized residual modindices (3.84);",
 usevariables = colnames(invar_data),
 rdata = invar_data)
cfa_m6_fit <- mplusModeler(cfa_m6,</pre>
              dataout=here("07-invariance", "invar_mplus", "invar_data.dat"),
              modelout=here("07-invariance", "invar_mplus", "m6_strict.inp"),
              check=TRUE, run = TRUE, hashfilename = FALSE)
```

Structural invariance A (fixed factor variances)

Demonstration of structural invariance using the Scalar model

- item loadings (set to equal)
- intercepts (set to equal)
- free residuals (Scalar)
- factor means free in group 2
- factor variances (set to 1)
- free factor covariances

- 10 item loadings (set to equal)
- 10 intercepts (set to equal)
- 20 residual variances
- 00 factor variances
- 02 factor co-variances
- 02 factor means

```
# fixed factor variances

cfa_m7 <- mplusObject(

TITLE = "model7 - structural invariance A" ,

VARIABLE =
    "usevar = stolen-rac_fght;
    grouping = freelnch (O=low_free_lunch 1=high_free_lunch); ",

ANALYSIS =
    "estimator = mlr;",

MODEL =
    "!!! MODEL SPECIFICATION FOR GROUP 1 (LOW FREE LUNCH) !!!

VICTIM by stolen* t_hurt p_fight hit damaged bullied;
VICTIM01;

NEG_CLIM by unsafe* disrupt gangs rac_fght;
NEG_CLIM01;

[VICTIM-NEG_CLIM00];</pre>
```

Structural invariance B (fixed factor variances and equal covariances)

Demonstration of structural invariance using the Scalar model

- item loadings (set to equal)
- intercepts (set to equal)
- free residuals (Scalar)
- \bullet factor means free in group 2
- factor variances (set to equal)
- factor covariances (set to equal)

- 10 item loadings (set to equal)
- 10 intercepts (set to equal)
- 20 residual variances
- 00 factor variances
- 01 factor co-variances
- 02 factor means

```
cfa_m8 <- mplusObject(</pre>
  TITLE = "model8 - structural invariance B" ,
  VARIABLE =
   "usevar = stolen-rac_fght;
    grouping = freelnch (0=low_free_lunch 1=high_free_lunch); ",
  ANALYSIS =
    "estimator = mlr;",
 MODEL =
    "!!! MODEL SPECIFICATION FOR GROUP 1 (LOW FREE LUNCH) !!!
     VICTIM by stolen* t_hurt p_fight hit damaged bullied;
    VICTIM@1;
    NEG_CLIM by unsafe* disrupt gangs rac_fght;
    NEG_CLIM@1;
     [VICTIM-NEG_CLIM@O];
    VICTIM with NEG_CLIM (11); !!! set covariances to equality !!!
     !!! MODEL SPECIFICATION FOR GROUP 2 (HIGH FREE LUNCH) !!!
    MODEL high_free_lunch:
     [VICTIM-NEG_CLIM]; !!! free factor means !!!
    VICTIM01; NEG_CLIM01; !!! fix factor variance to 1 !!!
    VICTIM with NEG_CLIM (11); !!! set covariances to equality !!!",
  PLOT = "type = plot3;",
  OUTPUT = "sampstat standardized residual modindices (3.84);",
 usevariables = colnames(invar_data),
 rdata = invar_data)
cfa_m8_fit <- mplusModeler(cfa_m8,</pre>
              dataout=here("07-invariance", "invar_mplus", "invar_data.dat"),
              modelout=here("07-invariance", "invar_mplus", "m8_structuralB.inp"),
              check=TRUE, run = TRUE, hashfilename = FALSE)
```

Latent Factor Means differences:

```
(model: Step_07_STRUCTURAL)
```

Mean differences: Students in sub-sample high_free_lunch have...

```
VICTIM -0.026 0.091 -0.291 0.771 (not significant)
NEG_CLIM 0.632 0.104 6.104 0.000 (higher scores for "NEG_CLIM")
```

Comparing Fit Across Models

Guidlines: for loadings & fit indices

- Simple structure: "0.4 0.3 0.2" rule Howard (2016) (primary loadings > 0.4 / cross-loadings < 0.3 / minimum difference = 0.2)
- RMSEA: < .05 indicates "good" fit Brown (2015)
- **CFI:** > .95 indicates "good" fit Brown (2015)
- SRMR: < .08 indicates "good" fit Hu and Bentler (1999)
- Invariance: Changes in CFI less than or equal to -0.01 are acceptable

Read into R summary of all models

```
all_models <- readModels(here("07-invariance", "invar_mplus"), quiet = TRUE)</pre>
```

Extract fit statistics and sort rows by Filename

Create model fit summary table comparing all models

```
invar_summary %>%
  mutate(Filename = str_remove(Filename, ".out")) %>%
  gt() %>%
  cols_label(
    Filename = "Model",
    Parameters = "Par",
    ChiSqM_Value = "ChiSq",
    CFI = "CFI", TLI = "TLI", SRMR = "SRMR",
    RMSEA_Estimate = "RMSEA",
    RMSEA_90CI_LB = "Lower CI",
    RMSEA_90CI_UB = "Upper CI")
```

Model	Par	ChiSq	CFI	TLI	SRMR	RMSEA	Lower CI	Upper CI
m0_cfa_fullsample	31	121.460	0.898	0.865	0.043	0.060	0.048	0.071
$m1_cfa_freelnch_0$	31	67.305	0.907	0.877	0.053	0.057	0.036	0.076
$m2_cfa_freelnch_1$	31	82.253	0.893	0.858	0.049	0.063	0.046	0.081
$m3$ _configural	62	149.315	0.899	0.867	0.051	0.060	0.047	0.073
$m4_metric$	54	163.312	0.892	0.872	0.061	0.059	0.047	0.072
$m5$ _scalar	46	179.176	0.882	0.874	0.062	0.059	0.047	0.071
$m6_strict$	36	180.933	0.892	0.897	0.067	0.053	0.041	0.065
$m7$ _structuralA	44	177.512	0.887	0.881	0.063	0.057	0.045	0.069
$m8_structuralB$	43	177.144	0.888	0.885	0.063	0.056	0.044	0.068

Compare nested models using χ^2 difference test

Calculate Satora-Bentler scaled χ^2 difference test (use with MLR estimator)

See website: stats.idre.ucla.edu

- SB0 = null model Chi-square value
- SB1 = alternate model Chi-square value
- c0 = null model scaling correction factor
- c1 = alternate model scaling correction factor
- d0 = null model degrees of freedom
- \bullet d1 = alternate model degrees of freedom
- \bullet df = Chi-square test degrees of freedom

Evaluate metric invariance

Compare the configural model to metric model

```
SB0 <- all_models[["m4_metric.out"]][["summaries"]][["ChiSqM_Value"]]
SB1 <- all_models[["m3_configural.out"]][["summaries"]][["ChiSqM_ScalingCorrection"]]
c0 <- all_models[["m4_metric.out"]][["summaries"]][["ChiSqM_ScalingCorrection"]]
c1 <- all_models[["m3_configural.out"]][["summaries"]][["ChiSqM_DF"]]
d0 <- all_models[["m4_metric.out"]][["summaries"]][["ChiSqM_DF"]]
d1 <- all_models[["m3_configural.out"]][["summaries"]][["ChiSqM_DF"]]
df <- abs(d0-d1)

# Satora-Bentler scaled Difference test equations
cd <- (((d0*c0)-(d1*c1))/(d0-d1))
t <- (((SB0*c0)-(SB1*c1))/(cd))

# Chi-square and degrees of freedom
t</pre>
```

```
## [1] 14.75942

df

## [1] 8

# Significance test
pchisq(t, df, lower.tail=FALSE)

## [1] 0.06399577
```

Evaluate scalar invariance

Compare the metric model to scalar model

```
SBO <- all_models[["m5_scalar.out"]][["summaries"]][["ChiSqM_Value"]]
SB1 <- all_models[["m4_metric.out"]][["summaries"]][["ChiSqM_Value"]]</pre>
c0 <- all_models[["m5_scalar.out"]][["summaries"]][["ChiSqM_ScalingCorrection"]]</pre>
c1 <- all_models[["m4_metric.out"]][["summaries"]][["ChiSqM_ScalingCorrection"]]</pre>
d0 <- all_models[["m5_scalar.out"]][["summaries"]][["ChiSqM_DF"]]</pre>
d1 <- all_models[["m4_metric.out"]][["summaries"]][["ChiSqM_DF"]]</pre>
df \leftarrow abs(d0-d1)
# Satora-Bentler scaled Difference test equations
cd \leftarrow (((d0*c0)-(d1*c1))/(d0-d1))
t <- (((SB0*c0)-(SB1*c1))/(cd))
# Chi-square and degrees of freedom
## [1] 15.44333
df
## [1] 8
# Significance test
pchisq(t, df, lower.tail=FALSE)
## [1] 0.05107724
```

Invariance short-cut

```
m_short <- mplusObject(</pre>
  TITLE = "INVARIANCE SHORT CUT",
  VARIABLE =
    "usevar = stolen-rac_fght;
     grouping = freelnch (0=low_free_lunch 1=high_free_lunch); ",
  ANALYSIS =
  "Estimator = MLR;
  MODEL = CONFIG METRIC SCALAR; ",
  MODEL =
    "VICTIM by stolen* t hurt p fight hit damaged bullied;
    VICTIM@1;
    NEG_CLIM by unsafe* disrupt gangs rac_fght;
    NEG_CLIM@1;" ,
  PLOT = ""
  OUTPUT = "sampstat residual;",
 usevariables = colnames(invar_data),
 rdata = invar_data)
m_short_fit <- mplusModeler(m_short ,</pre>
          dataout=here("07-invariance", "invar short", "invar short cut.dat"),
          modelout=here("07-invariance", "invar_short", "invar_short_cut.inp"),
          check=TRUE, run = TRUE, hashfilename = FALSE)
```

End

References

Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. Structural equation modeling: a multidisciplinary journal, 25(4), 621-638.

Lauff, E., & Ingels, S. J. (2014). Education Longitudinal Study of 2002 (ELS: 2002): A First Look at 2002 High School Sophomores 10 Years Later. First Look. NCES 2014-363. National Center for Education Statistics.

Muthén, L.K. and Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Wickham et al., (2019). Welcome to the tidy verse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

Further resources & examples here:

https://garberadamc.github.io/project-site/

https://www.adam-garber.com/