

Lab10.1 Factor Analysis - Weighted Least Squares (WLS)

Adam Garber

Factor Analysis ED 216B - Instructor: Karen Nylund-Gibson

March 10, 2020

Contents

1	Getting started: Rprojects, Rmarkdown, Git-Github	2
2	Steps to download repositories from Github and create a version controlled R-project	2
3	Outline	2
3.1	When to use the Weighted Least Squared (WSL) method?	3
3.2	How to specify the Robust WSL estimator (WLSMV) in Mplus	3
3.3	Interpretation of χ^2 with WLS estimator	3
4	Prepare data	4
4.1	Read in data	4
4.2	Reverse code for factor interpretation	4
4.3	Check correlations of science indicators	4
5	Model 0 - Exploratory Factor Analysis (EFA) with WLS Estimator	5
6	BEGIN: SINGLE GROUP INVARIANCE MODELS	6
7	Model 1 - Confirmatory Factor Analysis (CFA) - Full Sample	6
8	Model 2 - CFA for group SEX = 0 (Male)	7
9	Model 3 - CFA for group SEX = 1 (Female)	8
10	BEGIN: MULTI-GROUP INVARIANCE MODELS	9
10.1	How does <code>update()</code> work with MplusAutomation?	9
11	Model 4 - Configural Invariance:	9
12	Model 5 - Metric Invariance:	11

13 Model 6 - Scalar Invariance:	13
14 Model 7 - Strict Invariance:	15
15 Model 8 - Structural Invariance (using scalar model):	17
16 Create and customize tables using the package {gt}	19
16.1 Read into R summary of all models	19
16.2 Extract relevant data and generate table	19
16.3 Save table	21
16.4 How to use DIFFTEST?	21
17 End of Lab 10	21
18 References	21

1 Getting started: Rprojects, Rmarkdown, Git-Github

- [R-studio, Projects, Scripts: Go here](#)
 - [Rmarkdown basics tutorial: Go here here](#)
 - [Connect Git-Hithub with R-studio and download Repositories: Go here](#)
-

2 Steps to download repositories from Github and create a version controlled R-project

0. Create a Github account and connect R-Studio with Git
 1. Go to the Lab10 repository link to **Fork** and **Clone** (copy address) the repository:
 2. Within R-studio create a **New project** and choose the **Version Control** Option (Git)
 3. Paste the repository address copied (cloned) from Github and save locally on your computer
 4. After making changes in your branch of the repository to update the version on Github follow the following sequence of steps: **Stage**, **Commit** (add commit message), **Pull**, and then **Push**
-

3 Outline

- Prepare data
- EFA (model 0)
- CFA (models 1-3)
- Invariance (models 4-7)

- Generate summary table for model comparison

DATA SOURCE: This lab exercise utilizes a subset of the HSLs public-use dataset: High School Longitudinal Study of 2009 (Ingels et al., 2011) [See website: nces.ed.gov](https://nces.ed.gov/ipeds/data/hsls/)

FURTHER READING: A description of the models run in this lab exercise can be found in chapter 16 (page 411) in Kline, 2016. All specifications for invariance with the WLS estimator follow the recommendations found in this book. [Example Mplus Syntax](#)

3.1 When to use the Weighted Least Squared (WSL) method?

1. you have ordinal indicators (e.g., Likert scales)
2. you have categorical indicators (e.g., binary, or nominal)
3. you have a combination of ordinal, categorical, and/or continuous indicators
4. you have skewed continuous indicators

3.2 How to specify the Robust WSL estimator (WLSMV) in Mplus

Include the following lines of syntax:

- `categorical = X1 X2 X3... ;` *identify any categorical indicators*
- `estimator = wlsmv;`
- `parameterization = theta;` *to allow for error variances to be freely estimated across groups*

3.3 Interpretation of χ^2 with WLS estimator

- Cannot do normal chi-square difference tests (use the DIFFTEST option)
- The scaled Satorra-Bentler correction is required

```
# load packages
library(MplusAutomation)
library(haven)
library(rhdf5)
library(tidyverse)
library(here)
library(corrplot)
library(kableExtra)
library(reshape2)
library(semPlot)
library(gt)
```

4 Prepare data

4.1 Read in data

```
data_raw <- read_csv(here("data", "hsls_fa_data_subset.csv"))
```

4.2 Reverse code for factor interpretation

```
hsls_data <- data_raw

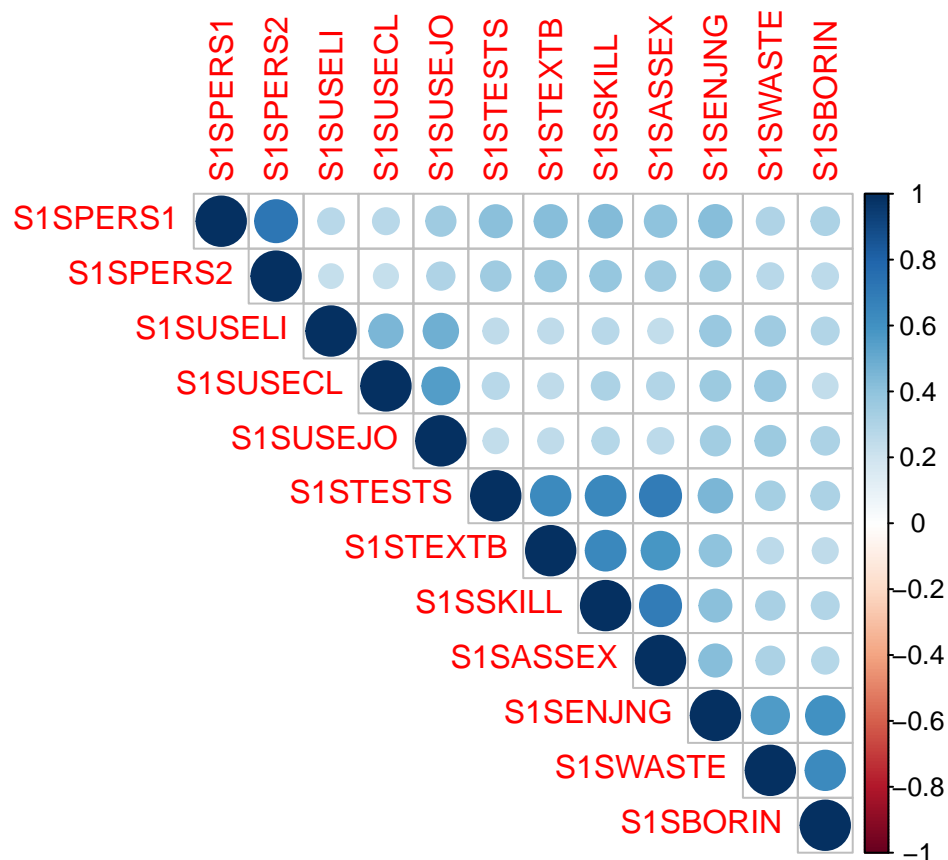
cols = c("S1MPERS1", "S1MPERS2", "S1MUSELI", "S1MUSECL", "S1MUSEJO",
         "S1MTESTS", "S1MTEXTB", "S1MSKILL", "S1MASSEX", "S1MENJNG",
         "S1SPERS1", "S1SPERS2", "S1SUSELI", "S1SUSECL", "S1SUSEJO",
         "S1STESTS", "S1STEXTB", "S1SSKILL", "S1SASSEX", "S1SENJNG")

hsls_data[,cols] <- 5 - hsls_data[,cols]
```

4.3 Check correlations of science indicators

```
sci_cor <- cor(hsls_data[27:38], use = "pairwise.complete.obs")
#sci_cor

corrplot(sci_cor,
         method = "circle",
         type = "upper")
```



NOTE: correlations should be reported from Mplus output (polychoric)
 # the matrix estimated using "cor()" are not polychoric correlations

5 Model 0 - Exploratory Factor Analysis (EFA) with WLS Estimator

```
m.step0 <- mplusObject(
  TITLE = " WSL FACTOR ANALYSIS EFA - HSLs SCIENCE DEMO",
  VARIABLE =
    "usevar =
    S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
    S1SUSEJO S1STESTS S1STEXTB S1SSKILL
    S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

    categorical =
    S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
    S1SUSEJO S1STESTS S1STEXTB S1SSKILL
    S1SASSEX S1SENJNG S1SWASTE S1SBORIN;"
```

```

ANALYSIS =
"type = efa 1 7;
estimator=wlsmv;
parameterization=theta;";

MODEL = "" ,

PLOT = "type = plot3;";
OUTPUT = "sampstat standardized residual modindices (3.84);";

usevariables = colnames(hsIs_data),
rdata = hsIs_data)

m.step0.fit <- mplusModeler(m.step0,
                           dataout=here("wls_efa", "efa_sci_HSLs_wls.dat"),
                           modelout=here("wls_efa", "efa_sci_HSLs_wls.inp"),
                           check=TRUE, run = TRUE, hashfilename = FALSE)

```

6 BEGIN: SINGLE GROUP INVARIANCE MODELS

7 Model 1 - Confirmatory Factor Analysis (CFA) - Full Sample

```

m.step1 <- mplusObject(
  TITLE = " WSL FACTOR ANALYSIS CFA - HSLs SCIENCE DEMO",
  VARIABLE =
  "usevar =
  S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
  S1SUSEJO S1TESTS S1TEXTB S1SSKILL
  S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

  categorical =
  S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
  S1SUSEJO S1TESTS S1TEXTB S1SSKILL
  S1SASSEX S1SENJNG S1SWASTE S1SBORIN;";

  ANALYSIS =
  "estimator=wlsmv;
  parameterization=theta;";

  MODEL =
  "SCI_ID BY S1SPERS1* S1SPERS2;
  SCI_ID@1;

  SCI_UT BY S1SUSELI* S1SUSECL S1SUSEJO;

```

```

SCI_UT@1;

SCI_EFF BY S1TESTS* S1STEXTB S1SSKILL S1SASSEX ;
SCI_EFF@1;

SCI_INT BY S1SENJNG* S1SWASTE S1SBORIN;
SCI_INT@1;" ,

PLOT = "type = plot3;",
OUTPUT = "sampstat standardized residual modindices (3.84);",

usevariables = colnames(hs1s_data),
rdata = hs1s_data)

m.step1.fit <- mplusModeler(m.step1,
                           dataout=here("wls_invar", "m1_cfa_wls.dat"),
                           modelout=here("wls_invar", "m1_cfa_wls.inp"),
                           check=TRUE, run = TRUE, hashfilename = FALSE)

## END: CONFIRMATORY FACTOR ANALYSIS OF SCIENCE INDICATORS

```

8 Model 2 - CFA for group SEX = 0 (Male)

```

m.step2 <- mplusObject(
  TITLE = " WSL FACTOR ANALYSIS CFA MALE GROUP SCIENCE",
  VARIABLE =
    "usevar =
    S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
    S1SUSEJO S1TESTS S1STEXTB S1SSKILL
    S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

    categorical =
    S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
    S1SUSEJO S1TESTS S1STEXTB S1SSKILL
    S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

    !X1SEX (1 = MALE, 2 = FEMALE)
    USEOBS = X1SEX == 1;",

  ANALYSIS =
    "estimator=wlsmv; parameterization=theta;",

  MODEL =
    "SCI_ID BY S1SPERS1* S1SPERS2;
    SCI_ID@1;

    SCI_UT BY S1SUSELI* S1SUSECL S1SUSEJO;

```

```

SCI_UT@1;

SCI_EFF BY S1TESTS* S1STEXTB S1SSKILL S1SASSEX ;
SCI_EFF@1;

SCI_INT BY S1SENJNG* S1SWASTE S1SBORIN;
SCI_INT@1;" ,

PLOT = "type = plot3;",
OUTPUT = "sampstat standardized residual modindices (3.84);",

usevariables = colnames(hs1s_data),
rdata = hs1s_data)

m.step2.fit <- mplusModeler(m.step2,
                           dataout=here("wls_invar", "m2_cfa_males_wls.dat"),
                           modelout=here("wls_invar", "m2_cfa_males_wls.inp"),
                           check=TRUE, run = TRUE, hashfilename = FALSE)

```

9 Model 3 - CFA for group SEX = 1 (Female)

```

m.step3 <- mplusObject(
  TITLE = " WSL FACTOR ANALYSIS CFA - FEMALE GROUP SCIENCE",
  VARIABLE =
    "usevar =
    S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
    S1SUSEJO S1TESTS S1STEXTB S1SSKILL
    S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

    categorical =
    S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
    S1SUSEJO S1TESTS S1STEXTB S1SSKILL
    S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

    !X1SEX (1 = MALE, 2 = FEMALE)
    USEOBS = X1SEX==2;",

  ANALYSIS = "estimator=wlsmv; parameterization=theta;",

  MODEL =
    "SCI_ID BY S1SPERS1* S1SPERS2;
    SCI_ID@1;

    SCI_UT BY S1SUSELI* S1SUSECL S1SUSEJO;
    SCI_UT@1;

    SCI_EFF BY S1TESTS* S1STEXTB S1SSKILL S1SASSEX ;

```



```

SCI_EFF@1;

SCI_INT BY S1SENJNG* S1SWASTE S1SBORIN;
SCI_INT@1;" ,

PLOT = "type = plot3;",
OUTPUT = "sampstat standardized residual modindices (3.84);",

usevariables = colnames(hs1s_data),
rdata = hs1s_data)

m.step3.fit <- mplusModeler(m.step3,
                           dataout=here("wls_invar", "m3_cfa_females_wls.dat"),
                           modelout=here("wls_invar", "m3_cfa_females_wls.inp"),
                           check=TRUE, run = TRUE, hashfilename = FALSE)

```

10 BEGIN: MULTI-GROUP INVARIANCE MODELS

10.1 How does `update()` work with `MplusAutomation`?

- Instead of copying the whole model above and altering one line of syntax we can use the `update()` model function.
- A key advantage of this is that it highlights changes in code which otherwise might be easy to overlook.
- **tilde** (~) will replace everything in that section of the input.
- **tilde-dot-plus** (~.+) will update the section by adding the specified code into that section.

This coding strategy although space-efficient was not utilized for following models. Rather, code is made explicit for pedagogy.

11 Model 4 - Configural Invariance:

- free item loadings
- free intercepts
- free residuals (group 2)

Number of parameters for configural model = 108

- 16 item loadings (8free loadings by 2groups = 16)
- 56 intercepts (36*2groups - 16 fixed = 56)
- 12 residual variances

- 12 factor co-variances
- 04 factor means
- 08 factor variances

```

m.step4 <- mplusObject(
  TITLE = " WSL FACTOR ANALYSIS - CONFIGURAL",
  VARIABLE =
"usevar =
S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
S1SUSEJO S1TESTS S1TEXTB S1SSKILL
S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

categorical =
S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
S1SUSEJO S1TESTS S1TEXTB S1SSKILL
S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

grouping = X1SEX (1=MALE 2=FEMALE);",

  ANALYSIS = "estimator=wlsmv; parameterization=theta;",

  MODEL =
"SCI_ID BY S1SPERS1@1 S1SPERS2;
SCI_UT BY S1SUSELI@1 S1SUSECL S1SUSEJO;
SCI_EFF BY S1TESTS@1 S1TEXTB S1SSKILL S1SASSEX ;
SCI_INT BY S1SENJNG@1 S1SWASTE S1SBORIN;

! threshold 1 for each item is constrained across groups
! one additional threshold constrained across groups for the identification variable (ULI)
[S1SPERS1$1] (f1_11); [S1SPERS1$2] (f1_12); [S1SPERS1$3] ;
[S1SPERS2$1] (f1_21); [S1SPERS2$2] ; [S1SPERS2$3] ;
[S1SUSELI$1] (f2_11); [S1SUSELI$2] (f2_12); [S1SUSELI$3] ;
[S1SUSECL$1] (f2_21); [S1SUSECL$2] ; [S1SUSECL$3] ;
[S1SUSEJO$1] (f2_31); [S1SUSEJO$2] ; [S1SUSEJO$3] ;
[S1TESTS$1] (f3_11); [S1TESTS$2] (f3_12); [S1TESTS$3] ;
[S1TEXTB$1] (f3_21); [S1TEXTB$2] ; [S1TEXTB$3] ;
[S1SSKILL$1] (f3_31); [S1SSKILL$2] ; [S1SSKILL$3] ;
[S1SASSEX$1] (f3_41); [S1SASSEX$2] ; [S1SASSEX$3] ;
[S1SENJNG$1] (f4_11); [S1SENJNG$2] (f4_12); [S1SENJNG$3] ;
[S1SWASTE$1] (f4_21); [S1SWASTE$2] ; [S1SWASTE$3] ;
[S1SBORIN$1] (f4_31); [S1SBORIN$2] ; [S1SBORIN$3] ;

[SCI_ID-SCI_INT@0]; ! factor means are fixed to 0 in group 1
SCI_ID-SCI_INT; ! factor variances are free in group 1
S1SPERS1-S1SBORIN@1; ! residual variances are fixed to 1 in group 1

MODEL FEMALE:

SCI_ID BY S1SPERS1@1 S1SPERS2;
SCI_UT BY S1SUSELI@1 S1SUSECL S1SUSEJO;
SCI_EFF BY S1TESTS@1 S1TEXTB S1SSKILL S1SASSEX ;
SCI_INT BY S1SENJNG@1 S1SWASTE S1SBORIN;

```

```

[S1SPERS1$1] (f1_11); [S1SPERS1$2] (f1_12); [S1SPERS1$3] ;
[S1SPERS2$1] (f1_21); [S1SPERS2$2] ; [S1SPERS2$3] ;
[S1SUSELI$1] (f2_11); [S1SUSELI$2] (f2_12); [S1SUSELI$3] ;
[S1SUSECL$1] (f2_21); [S1SUSECL$2] ; [S1SUSECL$3] ;
[S1SUSEJO$1] (f2_31); [S1SUSEJO$2] ; [S1SUSEJO$3] ;
[S1STESTS$1] (f3_11); [S1STESTS$2] (f3_12); [S1STESTS$3] ;
[S1STEXTB$1] (f3_21); [S1STEXTB$2] ; [S1STEXTB$3] ;
[S1SSKILL$1] (f3_31); [S1SSKILL$2] ; [S1SSKILL$3] ;
[S1SASSEX$1] (f3_41); [S1SASSEX$2] ; [S1SASSEX$3] ;
[S1SENJNG$1] (f4_11); [S1SENJNG$2] (f4_12); [S1SENJNG$3] ;
[S1SWASTE$1] (f4_21); [S1SWASTE$2] ; [S1SWASTE$3] ;
[S1SBORIN$1] (f4_31); [S1SBORIN$2] ; [S1SBORIN$3] ;

[SCI_ID-SCI_INT]; ! factor means are free in group 2
SCI_ID-SCI_INT; ! factor variances are free in group 2
S1SPERS1-S1SBORIN; ! residual variances are free in group 2

savedata: DIFFTEST=diff_1_2.dat;",

usevariables = colnames(hs1s_data),
rdata = hs1s_data)

m.step4.fit <- mplusModeler(m.step4,
                           dataout=here("wls_invar", "m4_configural_wls.dat"),
                           modelout=here("wls_invar", "m4_configural_wls.inp"),
                           check=TRUE, run = TRUE, hashfilename = FALSE)

```

12 Model 5 - Metric Invariance:

- item loadings (set to equal)
- free intercepts
- free residuals (group 2)

Number of parameters for metric model = 100

- 08 item loadings (8 free loadings)
 - 56 intercepts (36*2groups - 16 fixed = 56)
 - 12 residual variances
 - 12 factor co-variances
 - 04 factor means
 - 08 factor variances
-

```

m.step5 <- mplusObject(
  TITLE = "WSL FACTOR ANALYSIS - METRIC",
  VARIABLE =
    "usevar =
    S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
    S1SUSEJO S1TESTS S1TEXTB S1SSKILL
    S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

    categorical =
    S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
    S1SUSEJO S1TESTS S1TEXTB S1SSKILL
    S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

    grouping = X1SEX (1=MALE 2=FEMALE);",
  ANALYSIS = "estimator=wlsmv; parameterization=theta;",
  MODEL =
    "SCI_ID BY S1SPERS1@1 S1SPERS2;
    SCI_UT BY S1SUSELI@1 S1SUSECL S1SUSEJO;
    SCI_EFF BY S1TESTS@1 S1TEXTB S1SSKILL S1SASSEX ;
    SCI_INT BY S1SENJNG@1 S1SWASTE S1SBORIN;

    ! threshold 1 for each item is constrained across groups
    ! one additional threshold constrained across groups for the identification variable (ULI)
    [S1SPERS1$1](f1_11); [S1SPERS1$2](f1_12); [S1SPERS1$3] ;
    [S1SPERS2$1](f1_21); [S1SPERS2$2] ; [S1SPERS2$3] ;
    [S1SUSELI$1](f2_11); [S1SUSELI$2](f2_12); [S1SUSELI$3] ;
    [S1SUSECL$1](f2_21); [S1SUSECL$2] ; [S1SUSECL$3] ;
    [S1SUSEJO$1](f2_31); [S1SUSEJO$2] ; [S1SUSEJO$3] ;
    [S1TESTS$1](f3_11); [S1TESTS$2](f3_12); [S1TESTS$3] ;
    [S1TEXTB$1](f3_21); [S1TEXTB$2] ; [S1TEXTB$3] ;
    [S1SSKILL$1](f3_31); [S1SSKILL$2] ; [S1SSKILL$3] ;
    [S1SASSEX$1](f3_41); [S1SASSEX$2] ; [S1SASSEX$3] ;
    [S1SENJNG$1](f4_11); [S1SENJNG$2](f4_12); [S1SENJNG$3] ;
    [S1SWASTE$1](f4_21); [S1SWASTE$2] ; [S1SWASTE$3] ;
    [S1SBORIN$1](f4_31); [S1SBORIN$2] ; [S1SBORIN$3] ;

    [SCI_ID-SCI_INT@0]; ! factor means are fixed to 0 in group 1
    SCI_ID-SCI_INT; ! factor variances are free in group 1
    S1SPERS1-S1SBORIN@1; ! residual variances are fixed to 1 in group 1

    MODEL FEMALE:

    ! ***NEW TO METRIC INV*** (comment out to hold loadings equal)
    !SCI_ID BY S1SPERS1@1 S1SPERS2;
    !SCI_UT BY S1SUSELI@1 S1SUSECL S1SUSEJO;
    !SCI_EFF BY S1TESTS@1 S1TEXTB S1SSKILL S1SASSEX ;
    !SCI_INT BY S1SENJNG@1 S1SWASTE S1SBORIN;

    [S1SPERS1$1](f1_11); [S1SPERS1$2](f1_12); [S1SPERS1$3] ;
    [S1SPERS2$1](f1_21); [S1SPERS2$2] ; [S1SPERS2$3] ;
    [S1SUSELI$1](f2_11); [S1SUSELI$2](f2_12); [S1SUSELI$3] ;

```

```

[S1SUSECL$1] (f2_21); [S1SUSECL$2] ; [S1SUSECL$3] ;
[S1SUSEJO$1] (f2_31); [S1SUSEJO$2] ; [S1SUSEJO$3] ;
[S1TESTS$1] (f3_11); [S1TESTS$2] (f3_12); [S1TESTS$3] ;
[S1STEXTB$1] (f3_21); [S1STEXTB$2] ; [S1STEXTB$3] ;
[S1SSKILL$1] (f3_31); [S1SSKILL$2] ; [S1SSKILL$3] ;
[S1SASSEX$1] (f3_41); [S1SASSEX$2] ; [S1SASSEX$3] ;
[S1SENJNG$1] (f4_11); [S1SENJNG$2] (f4_12); [S1SENJNG$3] ;
[S1SWASTE$1] (f4_21); [S1SWASTE$2] ; [S1SWASTE$3] ;
[S1SBORIN$1] (f4_31); [S1SBORIN$2] ; [S1SBORIN$3] ;

[SCI_ID-SCI_INT]; ! factor means are free in group 2
SCI_ID-SCI_INT; ! factor variances are free in group 2
S1SPERS1-S1SBORIN; ! residual variances are free in group 2

savedata: DIFFTEST=diff_2_3.dat;",

usevariables = colnames(hs1s_data),
rdata = hs1s_data)

m.step5.fit <- mplusModeler(m.step5,
                           dataout=here("wls_invar", "m5_metric_wls.dat"),
                           modelout=here("wls_invar", "m5_metric_wls.inp"),
                           check=TRUE, run = TRUE, hashfilename = FALSE)

```

13 Model 6 - Scalar Invariance:

- item loadings (set to equal)
- intercepts (set to equal)
- free residuals (group 2)

Number of parameters for metric model = 80

- 08 item loadings (8 free loadings)
 - 36 intercepts
 - 12 residual variances
 - 12 factor co-variances
 - 04 factor means
 - 08 factor variances
-

```

m.step6 <- mplusObject(
  TITLE = "WSL FACTOR ANALYSIS - SCALAR",
  VARIABLE =
  "usevar =

```

```

S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
S1SUSEJO S1STESTS S1STEXTB S1SSKILL
S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

categorical =
S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
S1SUSEJO S1STESTS S1STEXTB S1SSKILL
S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

grouping = X1SEX (1=MALE 2=FEMALE);",

ANALYSIS = "estimator=wlsmv; parameterization=theta;",

MODEL =
"SCI_ID BY S1SPERS1@1 S1SPERS2;
SCI_UT BY S1SUSELI@1 S1SUSECL S1SUSEJO;
SCI_EFF BY S1STESTS@1 S1STEXTB S1SSKILL S1SASSEX ;
SCI_INT BY S1SENJNG@1 S1SWASTE S1SBORIN;

! ***NEW TO SCALAR INV*** (all thresholds set to equality)
[S1SPERS1$1](f1_11); [S1SPERS1$2](f1_12); [S1SPERS1$3](f1_13);
[S1SPERS2$1](f1_21); [S1SPERS2$2](f1_22); [S1SPERS2$3](f1_23);
[S1SUSELI$1](f2_11); [S1SUSELI$2](f2_12); [S1SUSELI$3](f2_13);
[S1SUSECL$1](f2_21); [S1SUSECL$2](f2_22); [S1SUSECL$3](f2_23);
[S1SUSEJO$1](f2_31); [S1SUSEJO$2](f2_32); [S1SUSEJO$3](f2_33);
[S1STESTS$1](f3_11); [S1STESTS$2](f3_12); [S1STESTS$3](f3_13);
[S1STEXTB$1](f3_21); [S1STEXTB$2](f3_22); [S1STEXTB$3](f3_23);
[S1SSKILL$1](f3_31); [S1SSKILL$2](f3_32); [S1SSKILL$3](f3_33);
[S1SASSEX$1](f3_41); [S1SASSEX$2](f3_42); [S1SASSEX$3](f3_43);
[S1SENJNG$1](f4_11); [S1SENJNG$2](f4_12); [S1SENJNG$3](f4_13);
[S1SWASTE$1](f4_21); [S1SWASTE$2](f4_22); [S1SWASTE$3](f4_23);
[S1SBORIN$1](f4_31); [S1SBORIN$2](f4_32); [S1SBORIN$3](f4_33);

[SCI_ID-SCI_INT@0]; ! factor means are fixed to 0 in group 1
SCI_ID-SCI_INT; ! factor variances are free in group 1
S1SPERS1-S1SBORIN@1; ! residual variances are fixed to 1 in group 1

MODEL FEMALE:

!SCI_ID BY S1SPERS1@1 S1SPERS2;
!SCI_UT BY S1SUSELI@1 S1SUSECL S1SUSEJO;
!SCI_EFF BY S1STESTS@1 S1STEXTB S1SSKILL S1SASSEX ;
!SCI_INT BY S1SENJNG@1 S1SWASTE S1SBORIN;

! ***NEW TO SCALAR INV*** (all thresholds set to equality)
[S1SPERS1$1](f1_11); [S1SPERS1$2](f1_12); [S1SPERS1$3](f1_13);
[S1SPERS2$1](f1_21); [S1SPERS2$2](f1_22); [S1SPERS2$3](f1_23);
[S1SUSELI$1](f2_11); [S1SUSELI$2](f2_12); [S1SUSELI$3](f2_13);
[S1SUSECL$1](f2_21); [S1SUSECL$2](f2_22); [S1SUSECL$3](f2_23);
[S1SUSEJO$1](f2_31); [S1SUSEJO$2](f2_32); [S1SUSEJO$3](f2_33);
[S1STESTS$1](f3_11); [S1STESTS$2](f3_12); [S1STESTS$3](f3_13);
[S1STEXTB$1](f3_21); [S1STEXTB$2](f3_22); [S1STEXTB$3](f3_23);
[S1SSKILL$1](f3_31); [S1SSKILL$2](f3_32); [S1SSKILL$3](f3_33);

```

```

[S1SASSEX$1](f3_41); [S1SASSEX$2](f3_42); [S1SASSEX$3](f3_43);
[S1SENJNG$1](f4_11); [S1SENJNG$2](f4_12); [S1SENJNG$3](f4_13);
[S1SWASTE$1](f4_21); [S1SWASTE$2](f4_22); [S1SWASTE$3](f4_23);
[S1SBORIN$1](f4_31); [S1SBORIN$2](f4_32); [S1SBORIN$3](f4_33);

[SCI_ID-SCI_INT];      ! factor means are free in group 2
SCI_ID-SCI_INT;        ! factor variances are free in group 2
S1SPERS1-S1SBORIN;     ! residual variances are free in group 2

savedata: DIFFTEST=diff_3_4.dat;",

usevariables = colnames(hsIs_data),
rdata = hsIs_data)

m.step6.fit <- mplusModeler(m.step6,
                           dataout=here("wls_invar", "m6_scalar_wls.dat"),
                           modelout=here("wls_invar", "m6_scalar_wls.inp"),
                           check=TRUE, run = TRUE, hashfilename = FALSE)

```

14 Model 7 - Strict Invariance:

- item loadings (set to equal)
- intercepts (set to equal)
- residuals (fixed to 1)

Number of parameters for metric model = 68

- 08 item loadings (8 free loadings)
 - 36 intercepts
 - 00 residual variances
 - 12 factor co-variances
 - 04 factor means
 - 08 factor variances
-

```

m.step7 <- mplusObject(
  TITLE = "WSL FACTOR ANALYSIS - STRICT",
  VARIABLE =
    "usevar =
    S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
    S1SUSEJO S1TESTS S1TEXTB S1SSKILL
    S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

    categorical =

```

```

S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
S1SUSEJO S1STESTS S1STEXTB S1SSKILL
S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

grouping = X1SEX (1=MALE 2=FEMALE);",

ANALYSIS = "estimator=wlsmv; parameterization=theta;",

MODEL =
"SCI_ID BY S1SPERS1@1 S1SPERS2;
SCI_UT BY S1SUSELI@1 S1SUSECL S1SUSEJO;
SCI_EFF BY S1STESTS@1 S1STEXTB S1SSKILL S1SASSEX ;
SCI_INT BY S1SENJNG@1 S1SWASTE S1SBORIN;

[S1SPERS1$1] (f1_11); [S1SPERS1$2] (f1_12); [S1SPERS1$3] (f1_13);
[S1SPERS2$1] (f1_21); [S1SPERS2$2] (f1_22); [S1SPERS2$3] (f1_23);
[S1SUSELI$1] (f2_11); [S1SUSELI$2] (f2_12); [S1SUSELI$3] (f2_13);
[S1SUSECL$1] (f2_21); [S1SUSECL$2] (f2_22); [S1SUSECL$3] (f2_23);
[S1SUSEJO$1] (f2_31); [S1SUSEJO$2] (f2_32); [S1SUSEJO$3] (f2_33);
[S1STESTS$1] (f3_11); [S1STESTS$2] (f3_12); [S1STESTS$3] (f3_13);
[S1STEXTB$1] (f3_21); [S1STEXTB$2] (f3_22); [S1STEXTB$3] (f3_23);
[S1SSKILL$1] (f3_31); [S1SSKILL$2] (f3_32); [S1SSKILL$3] (f3_33);
[S1SASSEX$1] (f3_41); [S1SASSEX$2] (f3_42); [S1SASSEX$3] (f3_43);
[S1SENJNG$1] (f4_11); [S1SENJNG$2] (f4_12); [S1SENJNG$3] (f4_13);
[S1SWASTE$1] (f4_21); [S1SWASTE$2] (f4_22); [S1SWASTE$3] (f4_23);
[S1SBORIN$1] (f4_31); [S1SBORIN$2] (f4_32); [S1SBORIN$3] (f4_33);

[SCI_ID-SCI_INT@0]; ! factor means are fixed to 0 in group 1
SCI_ID-SCI_INT; ! factor variances are free in group 1
S1SPERS1-S1SBORIN@1; ! residual variances are fixed to 1 in group 1

MODEL FEMALE:

!SCI_ID BY S1SPERS1@1 S1SPERS2;
!SCI_UT BY S1SUSELI@1 S1SUSECL S1SUSEJO;
!SCI_EFF BY S1STESTS@1 S1STEXTB S1SSKILL S1SASSEX ;
!SCI_INT BY S1SENJNG@1 S1SWASTE S1SBORIN;

[S1SPERS1$1] (f1_11); [S1SPERS1$2] (f1_12); [S1SPERS1$3] (f1_13);
[S1SPERS2$1] (f1_21); [S1SPERS2$2] (f1_22); [S1SPERS2$3] (f1_23);
[S1SUSELI$1] (f2_11); [S1SUSELI$2] (f2_12); [S1SUSELI$3] (f2_13);
[S1SUSECL$1] (f2_21); [S1SUSECL$2] (f2_22); [S1SUSECL$3] (f2_23);
[S1SUSEJO$1] (f2_31); [S1SUSEJO$2] (f2_32); [S1SUSEJO$3] (f2_33);
[S1STESTS$1] (f3_11); [S1STESTS$2] (f3_12); [S1STESTS$3] (f3_13);
[S1STEXTB$1] (f3_21); [S1STEXTB$2] (f3_22); [S1STEXTB$3] (f3_23);
[S1SSKILL$1] (f3_31); [S1SSKILL$2] (f3_32); [S1SSKILL$3] (f3_33);
[S1SASSEX$1] (f3_41); [S1SASSEX$2] (f3_42); [S1SASSEX$3] (f3_43);
[S1SENJNG$1] (f4_11); [S1SENJNG$2] (f4_12); [S1SENJNG$3] (f4_13);
[S1SWASTE$1] (f4_21); [S1SWASTE$2] (f4_22); [S1SWASTE$3] (f4_23);
[S1SBORIN$1] (f4_31); [S1SBORIN$2] (f4_32); [S1SBORIN$3] (f4_33);

[SCI_ID-SCI_INT]; ! factor means are free in group 2
SCI_ID-SCI_INT; ! factor variances are free in group 2

```



```

! ***NEW TO STRICT INV*** (all residual variances fixed to 1 in group 2)
S1SPERS1-S1SBORIN@1; ! residual variances are fixed in group 2

savedata: DIFFTEST=diff_4_5.dat;",

usevariables = colnames(hsIs_data),
rdata = hsIs_data)

m.step7.fit <- mplusModeler(m.step7,
                           dataout=here("wls_invar", "m7_strict_wls.dat"),
                           modelout=here("wls_invar", "m7_strict_wls.inp"),
                           check=TRUE, run = TRUE, hashfilename = FALSE)

```

15 Model 8 - Structural Invariance (using scalar model):

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

Number of parameters for metric model = 70

- 08 item loadings (8 free loadings)
- 36 intercepts
- 12 residual variances
- 06 factor co-variances (set to equal)
- 04 factor means
- 04 factor variances (set to equal)

```

m.step8 <- mplusObject(
  TITLE = "WSL FACTOR ANALYSIS - STRICT",
  VARIABLE =
"usevar =
S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
S1SUSEJO S1STESTS S1STEXTB S1SSKILL
S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

  categorical =
S1SPERS1 S1SPERS2 S1SUSELI S1SUSECL
S1SUSEJO S1STESTS S1STEXTB S1SSKILL
S1SASSEX S1SENJNG S1SWASTE S1SBORIN;

```

```

grouping = X1SEX (1=MALE 2=FEMALE);",

ANALYSIS = "estimator=wlsmv; parameterization=theta;",

MODEL =
"SCI_ID BY S1SPERS1@1 S1SPERS2;
SCI_UT BY S1SUSELI@1 S1SUSECL S1SUSEJO;
SCI_EFF BY S1STESTS@1 S1STEXTB S1SSKILL S1SASSEX ;
SCI_INT BY S1SENJNG@1 S1SWASTE S1SBORIN;

[S1SPERS1$1](f1_11); [S1SPERS1$2](f1_12); [S1SPERS1$3](f1_13);
[S1SPERS2$1](f1_21); [S1SPERS2$2](f1_22); [S1SPERS2$3](f1_23);
[S1SUSELI$1](f2_11); [S1SUSELI$2](f2_12); [S1SUSELI$3](f2_13);
[S1SUSECL$1](f2_21); [S1SUSECL$2](f2_22); [S1SUSECL$3](f2_23);
[S1SUSEJO$1](f2_31); [S1SUSEJO$2](f2_32); [S1SUSEJO$3](f2_33);
[S1STESTS$1](f3_11); [S1STESTS$2](f3_12); [S1STESTS$3](f3_13);
[S1STEXTB$1](f3_21); [S1STEXTB$2](f3_22); [S1STEXTB$3](f3_23);
[S1SSKILL$1](f3_31); [S1SSKILL$2](f3_32); [S1SSKILL$3](f3_33);
[S1SASSEX$1](f3_41); [S1SASSEX$2](f3_42); [S1SASSEX$3](f3_43);
[S1SENJNG$1](f4_11); [S1SENJNG$2](f4_12); [S1SENJNG$3](f4_13);
[S1SWASTE$1](f4_21); [S1SWASTE$2](f4_22); [S1SWASTE$3](f4_23);
[S1SBORIN$1](f4_31); [S1SBORIN$2](f4_32); [S1SBORIN$3](f4_33);

[SCI_ID-SCI_INT@0]; ! factor means are fixed to 0 in group 1
SCI_ID-SCI_INT (V1-V4); ! factor variances (set to equality)
S1SPERS1-S1SBORIN@1; ! residual variances are fixed to 1 in group 1

! factor covariances set to equality
SCI_ID with SCI_UT-SCI_INT(C1-C3);
SCI_EFF with SCI_ID(C4);
SCI_EFF with SCI_UT(C5);
SCI_UT with SCI_ID (C6);

MODEL FEMALE:

[S1SPERS1$1](f1_11); [S1SPERS1$2](f1_12); [S1SPERS1$3](f1_13);
[S1SPERS2$1](f1_21); [S1SPERS2$2](f1_22); [S1SPERS2$3](f1_23);
[S1SUSELI$1](f2_11); [S1SUSELI$2](f2_12); [S1SUSELI$3](f2_13);
[S1SUSECL$1](f2_21); [S1SUSECL$2](f2_22); [S1SUSECL$3](f2_23);
[S1SUSEJO$1](f2_31); [S1SUSEJO$2](f2_32); [S1SUSEJO$3](f2_33);
[S1STESTS$1](f3_11); [S1STESTS$2](f3_12); [S1STESTS$3](f3_13);
[S1STEXTB$1](f3_21); [S1STEXTB$2](f3_22); [S1STEXTB$3](f3_23);
[S1SSKILL$1](f3_31); [S1SSKILL$2](f3_32); [S1SSKILL$3](f3_33);
[S1SASSEX$1](f3_41); [S1SASSEX$2](f3_42); [S1SASSEX$3](f3_43);
[S1SENJNG$1](f4_11); [S1SENJNG$2](f4_12); [S1SENJNG$3](f4_13);
[S1SWASTE$1](f4_21); [S1SWASTE$2](f4_22); [S1SWASTE$3](f4_23);
[S1SBORIN$1](f4_31); [S1SBORIN$2](f4_32); [S1SBORIN$3](f4_33);

[SCI_ID-SCI_INT]; ! factor means are free in group 2
SCI_ID-SCI_INT (V1-V4); ! factor variances (set to equality)
S1SPERS1-S1SBORIN; ! residual variances are free group 2

! factor covariances set to equality

```

```

SCI_ID with SCI_UT-SCI_INT(C1-C3);
SCI_EFF with SCI_ID(C4);
SCI_EFF with SCI_UT(C5);
SCI_UT with SCI_ID (C6); ",

usevariables = colnames(hsIs_data),
rdata = hsIs_data)

m.step8.fit <- mplusModeler(m.step8,
                           dataout=here("wls_invar", "m8_structural_wls.dat"),
                           modelout=here("wls_invar", "m8_structural_wls.inp"),
                           check=TRUE, run = TRUE, hashfilename = FALSE)

```

Mean differences: Females are...

Means				
SCI_ID	-0.866	0.452	-1.914	0.056
SCI_UT	0.010	0.054	0.187	0.852
SCI_EFF	-0.577	0.094	-6.149	0.000
SCI_INT	-0.062	0.094	-0.656	0.512

16 Create and customize tables using the package {gt}

16.1 Read into R summary of all models

```

all_models <- readModels(here("wls_invar"))

## Reading model: /Users/agarber/Desktop/FA_W20/Lab10_FA/wls_invar/m1_cfa_wls.out
## Reading model: /Users/agarber/Desktop/FA_W20/Lab10_FA/wls_invar/m2_cfa_males_wls.out
## Reading model: /Users/agarber/Desktop/FA_W20/Lab10_FA/wls_invar/m3_cfa_females_wls.out
## Reading model: /Users/agarber/Desktop/FA_W20/Lab10_FA/wls_invar/m4_configural_wls.out
## Reading model: /Users/agarber/Desktop/FA_W20/Lab10_FA/wls_invar/m5_metric_wls.out
## Reading model: /Users/agarber/Desktop/FA_W20/Lab10_FA/wls_invar/m6_scalar_wls.out
## Reading model: /Users/agarber/Desktop/FA_W20/Lab10_FA/wls_invar/m7_strict_wls.out
## Reading model: /Users/agarber/Desktop/FA_W20/Lab10_FA/wls_invar/m8_structural_wls.out

```

16.2 Extract relevant data and generate table

```

invar_summary <- LatexSummaryTable(all_models,
                                   keepCols=c("Filename", "Parameters", "ChiSqM_Value", "CFI", "TLI",
                                                "SRMR", "RMSEA_Estimate", "RMSEA_90CI_LB", "RMSEA_90CI_UB"),
                                   sortBy = "Filename")

model_table <- invar_summary %>%
  gt() %>%
  tab_header(
    title = "Fit Indices", # Add a title
    subtitle = ""          # And a subtitle
  ) %>%
  tab_options(
    table.width = pct(80)
  ) %>%
  tab_footnote(
    footnote = "Data from HSLS 2009",
    location = cells_title()
  ) %>%
  data_color(
    # Update cell colors...
    columns = vars(CFI), # ...for column
    alpha = .5,
    autocolor_text = FALSE,
    colors = scales::col_numeric(
      palette = c(
        "yellow", "blue"), # choose color
      domain = c(.98,.99) # Scale endpoints (outside will be gray)
    ) %>%
  cols_label(
    Filename = "Model",
    Parameters = "Par",
    ChiSqM_Value = "ChiSq",
    RMSEA_Estimate = "RMSEA",
    RMSEA_90CI_LB = "Lower CI",
    RMSEA_90CI_UB = "Upper CI")

model_table

```

Fit Indices¹

Model	Par	ChiSq	CFI	TLI	SRMR	RMSEA	Lower CI	Upper CI
m1_cfa_wls.out	54	657.141	0.983	0.977	0.028	0.068	0.063	0.073
m2_cfa_males_wls.out	54	345.034	0.986	0.980	0.027	0.067	0.060	0.074
m3_cfa_females_wls.out	54	354.320	0.981	0.974	0.032	0.068	0.062	0.075
m4_configural_wls.out	108	699.033	0.984	0.978	0.030	0.068	0.063	0.072
m5_metric_wls.out	100	685.895	0.984	0.980	0.030	0.064	0.059	0.068
m6_scalar_wls.out	80	749.934	0.983	0.982	0.031	0.061	0.057	0.065
m7_strict_wls.out	68	764.189	0.983	0.984	0.032	0.058	0.054	0.062
m8_structural_wls.out	72	565.901	0.988	0.988	0.034	0.049	0.045	0.053

¹Data from HSLS 2009

16.3 Save table

```
model_table %>% gtsave("model_fit_table_wls.png", path = here("figures"))
```

16.4 How to use DIFFTEST?

<https://stats.idre.ucla.edu/mplus/faq/how-can-i-compute-a-chi-squared-test-for-nested-models-with-the-mlmv-or-wlsmv-estimators-difftest/>

17 End of Lab 10

18 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.

Horst, A. (2020). Course & Workshop Materials. GitHub Repositories, <https://allisonhorst.github.io/>

Ingels, S. J., Pratt, D. J., Herget, D. R., Burns, L. J., Dever, J. A., Ottem, R., ... & Leinwand, S. (2011). High School Longitudinal Study of 2009 (HSL: 09): Base-Year Data File Documentation. NCES 2011-328. 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 tidyverse. *Journal of Open Source Software*, 4(43), 1686, <https://doi.org/10.21105/joss.01686>

UC SANTA BARBARA