Multiple group measurement invariance analysis in Lavaan

Kate Xu

Department of Psychiatry
University of Cambridge

Email: mx212@medschl.cam.ac.uk

Measurement invariance

- In empirical research, comparisons of means or regression coefficients is often drawn from distinct population groups such as culture, gender, language spoken
- Unless explicitly tested, these analysis automatically assumes the measurement of these outcome variables are equivalent across these groups
- Measurement invariance can be tested and it is important to make sure that the variables used in the analysis are indeed comparable constructs across distinct groups

Applications of measurement invariance

- Psychometric validation of new instrument, e.g. mental health questionnaire in patients vs healthy, men vs. women
- Cross cultural comparison research – people from different cultures might have different understandings towards the same questions included in an instrument
- Longitudinal study that look at change of a latent variable across time, e.g. cognition, mental health

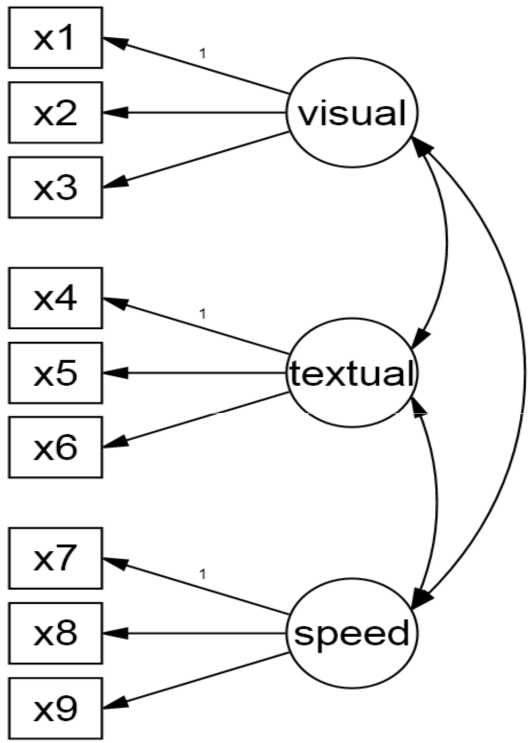
Assessing measurement invariance

- Multiple group confirmatory factor analysis is a popular method for measurement invariance analysis (Meredith, 1993)
  - Evaluation on whether the variables of interest is equivalent across groups, using latent variable modelling method
  - Parameters in the CFA model can be set equal or vary across groups
  - Level of measurement equivalency can be assessed through model fit of a series of **nested** multiple group models

Illustration of MI analysis based on the Holzinger-Swineford study

- Cognitive function tests (n=301)
  - Two school groups: Pasteur=156 Grant-white=145
  - Three factors, 9 indicators

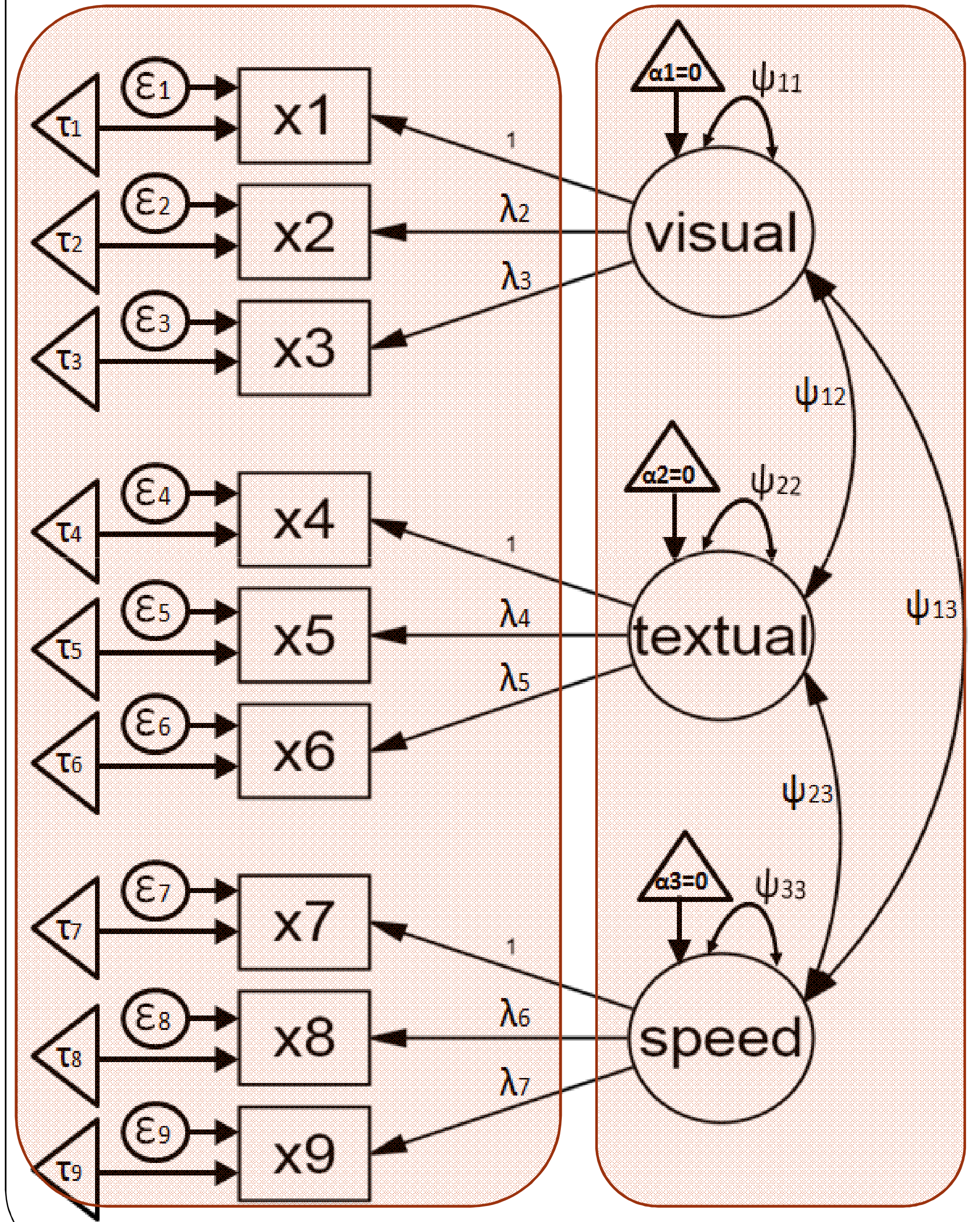
x1	Visual perception
x2	Cubes
x3	Lozenges
x4	Paragraph comprehension
x5	Sentence completion
x6	Word meaning
x7	Addition speed
x8	Speed of counting of dots
x9	Discrimination speed between straight and curved capitals



- Some indicators might show measurement non-invariance due to different backgrounds of the students or the specific teaching style of the type of schools

Parameter annotations

UNIVERSITY OF CAMBRIDGE



- Measurement parameters
  - 6 factor loadings

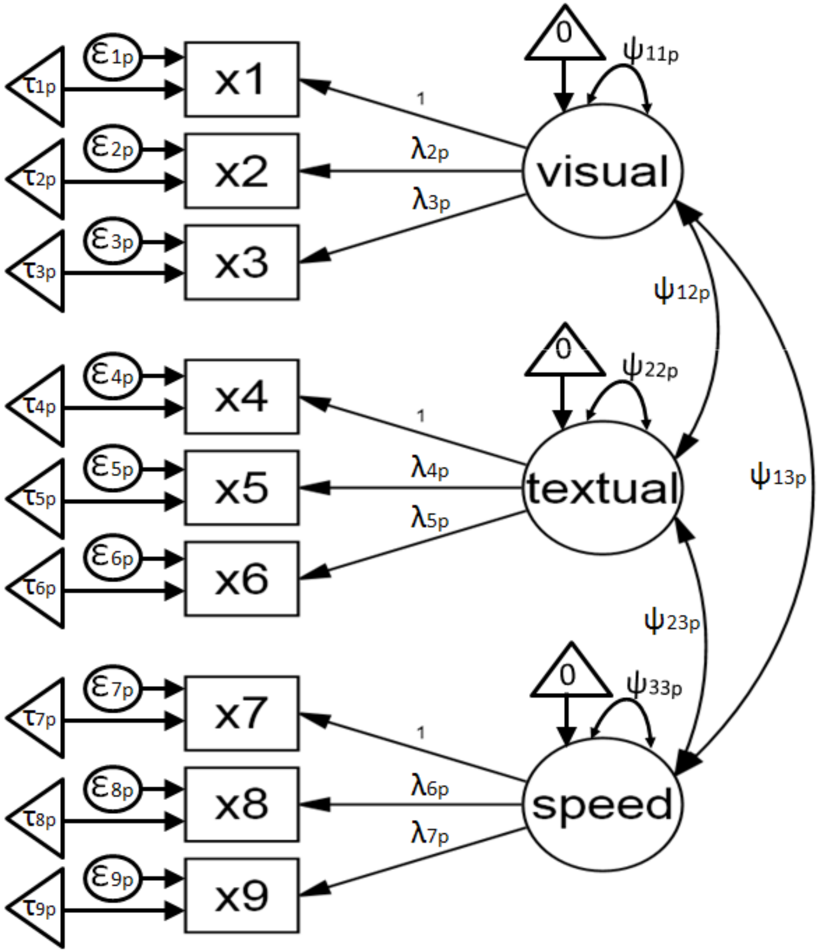
    $\lambda2$, $\lambda3$, $\lambda4$, $\lambda5$, $\lambda6$, $\lambda7$
  - 9 factor intercepts

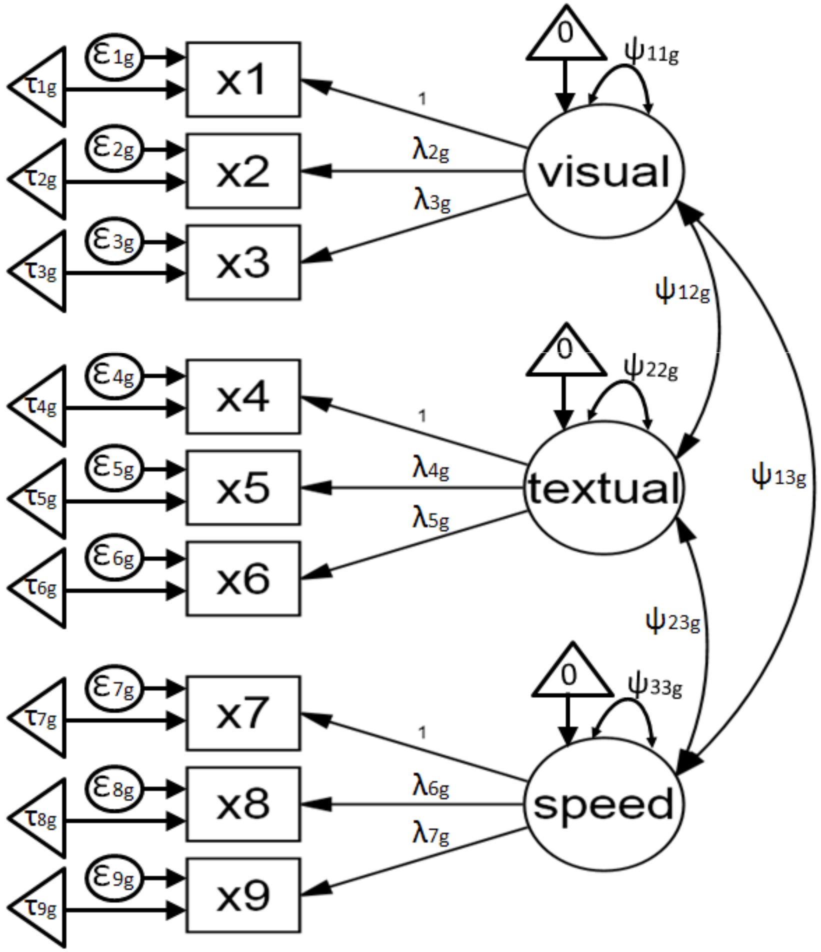
    $\tau1$, $\tau2$, $\tau3$, $\tau4$, $\tau5$, $\tau6$, $\tau7$, $\tau8$, $\tau9$
  - 9 Item residuals

    $\varepsilon1$, $\varepsilon2$, $\varepsilon3$, $\varepsilon4$, $\varepsilon5$, $\varepsilon6$, $\varepsilon7$, $\varepsilon8$, $\varepsilon9$
- Structural parameters
  - latent means
  - $\alpha1$, $\alpha1$, $\alpha3$ (set to 0)
  - 3 factor variances
  - $\psi11$ $\psi22$, $\psi33$
  - 3 factor covariances
  - $\psi12$ $\psi13$, $\psi23$

Multiple group CFA

- Pasteur (n=156)
- Grand-white (n=145)





Summary of steps in measurement invariance tests

	Constrained parameters	Free parameters	comparison model
configural	FMean (=0)	fl+inter+res+var	
Weak/loading invariance	fl+Fmean (=0)	inter+res+var	configural
Strong/scalar invariance	fl+inter	res+var+Fmean*	Weak/loading invariance
strict invariance	fl+inter+res	Fmean*+var	Strong/scalar invariance

Note. fl= factor loadings, inter = item intercepts, res = item residual variances, Fmean = mean of latent variable, var = variance of latent variable

*Fmean is fixed to 0 in group 1 and estimated in the other group(s)

Evaluating measurement invariance using fit indices

- Substantial decrease in goodness of fit indicates non-invariance
- It is a good practise to look at several model fit indices rather than relying on a single one
  - $\Delta\chi^2$
  - ΔRMSEA
  - ΔCFI
  - ΔTLI
  - ΔBIC
  - ΔAIC
  - …

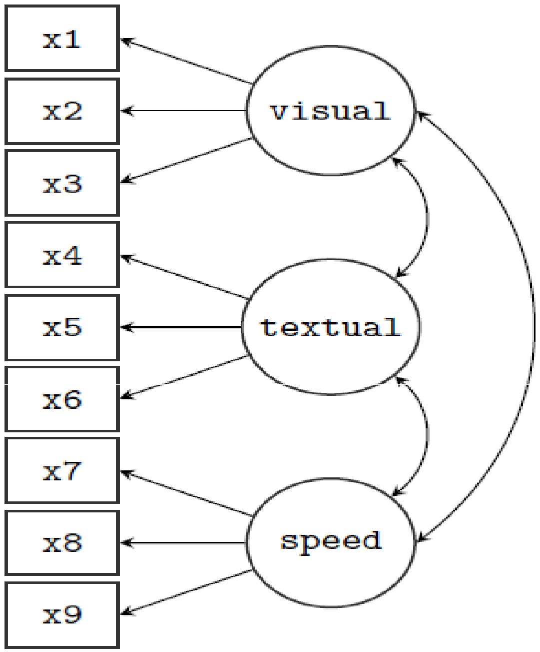
Identifying non-invariance

- Modification index (MI)
  - MI indicates the expected decrease in chi-square if a restricted parameter is to be freed in a less restrictive model
  - Usually look for the largest MI value in the MI output, and free one parameter at a time through an iterative process
  - The usual cut-off value is 3.84, but this needs to be adjusted based on sample size (chi-square is sensitive to sample size) and number of tests conducted (type I error)

Lavaan: Measurement invariance analysis

- Data: *HolzingerSwineford1939*
- School type:
  - 1=Pasteur (156)
  - 2=Grand-white (145)
- Define the CFA model

```
library(lavaan)
HS.model <-
        'visual =~ x1 + x2 + x3
        textual =~ x4 + x5 + x6
        speed =~ x7 + x8 + x9'
```



- semTools fits a series of increasingly restrictive models in one command:

```
library(semTools)
measurementInvariance(HS.model,data=HolzingerSwineford1939,
        group="school")
```

```
measurementInvariance(HS.model,data=HolzingerSwineford1939,
        group="school")
```

```
Measurement invariance tests:

Model 1: configural invariance:
   chisq        df    pvalue       cfi     rmsea       bic
 115.851    48.000     0.000     0.923     0.097  7706.822

Model 2: weak invariance (equal loadings):
   chisq        df    pvalue       cfi     rmsea       bic
 124.044    54.000     0.000     0.921     0.093  7680.771

[Model 1 versus model 2]
  delta.chisq      delta.df delta.p.value      delta.cfi
        8.192         6.000         0.224          0.002

Model 3: strong invariance (equal loadings + intercepts)
   chisq        df    pvalue       cfi     rmsea       bic
 164.103    60.000     0.000     0.882     0.107  7686.588

[Model 1 versus model 3]
  delta.chisq      delta.df delta.p.value      delta.cfi
       48.251        12.000         0.000          0.041

[Model 2 versus model 3]
  delta.chisq      delta.df delta.p.value      delta.cfi
       40.059         6.000         0.000          0.038

Model 4: equal loadings + intercepts + means:
   chisq        df    pvalue       cfi     rmsea       bic
 204.605    63.000     0.000     0.840     0.122  7709.969

[Model 1 versus model 4]
  delta.chisq      delta.df delta.p.value      delta.cfi
       88.754        15.000         0.000          0.083
```

<-configural model (Model 1)

<-metric MI model (Model 2)

<- Metric MI achieved: non-significant chi-square change

<-scalar MI model (Model 3)

<- Scalar MI failed

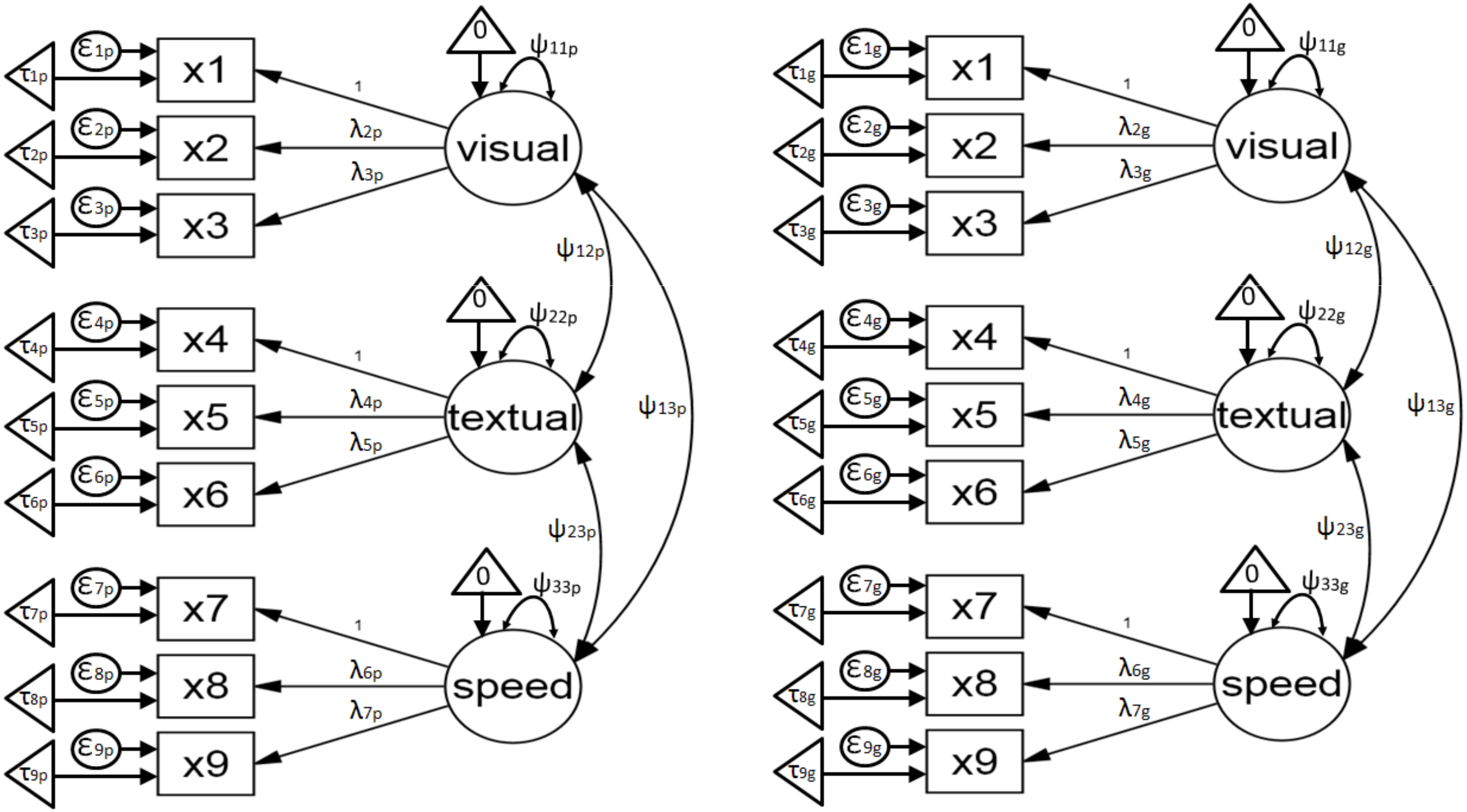
<- Constrain latent means equal across groups, but this is no longer meaningful because of non-MI in Model 3.

Measurement invariance:
Step 1: Configural invariance

- Same factor structure in each group
- First, fit model separately in each group
- Second, fit model in multiple group but let all parameters vary freely in each group
- No latent mean difference is estimated

Configural invariance

- Constrained = none



Lavaan: Model 1 configural model

UNIVERSITY OF CAMBRIDGE

```
model1<- cfa(HS.model, data=HolzingerSwineford1939, group="school")
summary(model1,fit.measures=TRUE)
```

All parameters are different across groups

chisq	df	pvalue	cfi	rmsea	bic
115.851	48.000	0.000	0.923	0.097	7706.822

```
Group 1 [Pasteur]:

                   Estimate  Std.err  Z-value  P(>|z|)
Latent variables:
  visual =~
    x1                1.000
    x2                0.394    0.122    3.220    0.001
    x3                0.570    0.140    4.076    0.000
  textual =~
    x4                1.000
    x5                1.183    0.102   11.613    0.000
    x6                0.875    0.077   11.421    0.000
  speed =~
    x7                1.000
    x8                1.125    0.277    4.057    0.000
    x9                0.922    0.225    4.104    0.000
Intercepts:
    x1                4.941    0.095   52.249    0.000
    x2                5.984    0.098   60.949    0.000
    x3                2.487    0.093   26.778    0.000
    x4                2.823    0.092   30.689    0.000
    x5                3.995    0.105   38.183    0.000
    x6                1.922    0.079   24.321    0.000
    x7                4.432    0.087   51.181    0.000
    x8                5.563    0.078   71.214    0.000
    x9                5.418    0.079   68.440    0.000
    visual            0.000
    textual           0.000
    speed             0.000
```

```
Group 2 [Grant-White]:

                   Estimate  Std.err  Z-value  P(>|z|)
Latent variables:
  visual =~
    x1                1.000
    x2                0.736    0.155    4.760    0.000
    x3                0.925    0.166    5.583    0.000
  textual =~
    x4                1.000
    x5                0.990    0.087   11.418    0.000
    x6                0.963    0.085   11.377    0.000
  speed =~
    x7                1.000
    x8                1.226    0.187    6.569    0.000
    x9                1.058    0.165    6.429    0.000
Intercepts:
    x1                4.930    0.095   51.696    0.000
    x2                6.200    0.092   67.416    0.000
    x3                1.996    0.086   23.195    0.000
    x4                3.317    0.093   35.625    0.000
    x5                4.712    0.096   48.986    0.000
    x6                2.469    0.094   26.277    0.000
    x7                3.921    0.086   45.819    0.000
    x8                5.488    0.087   63.174    0.000
    x9                5.327    0.085   62.571    0.000
    visual            0.000
    textual           0.000
    speed             0.000
```

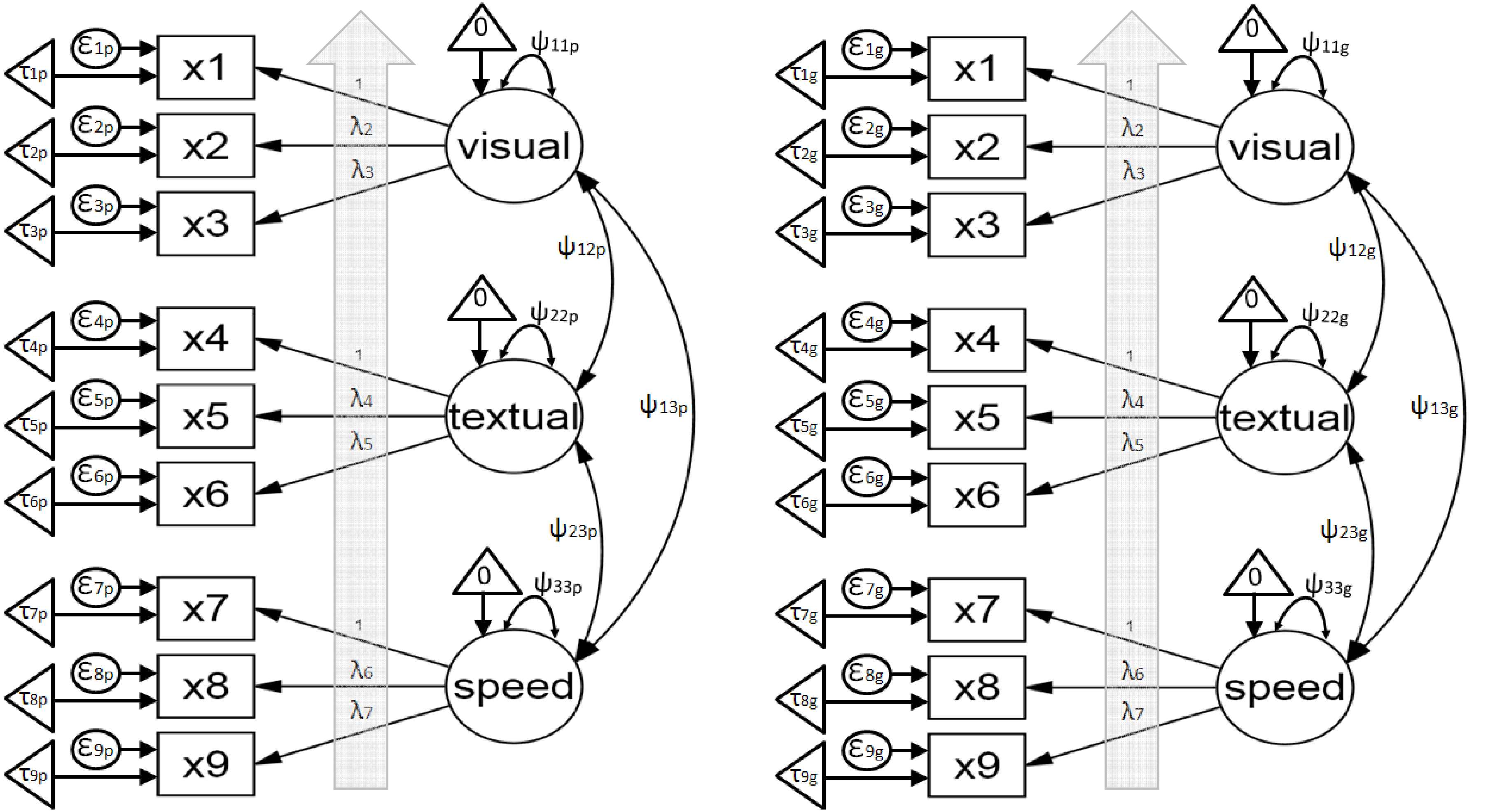
Measurement invariance: Step 2: Weak/metric invariance

UNIVERSITY OF CAMBRIDGE

- Constrain factor loadings equal across groups
- This shows that the construct has the same meaning across groups
- In case of partial invariance of factor loadings, constrain the invariant loadings and set free the non-invariant loadings (Byrne, Shavelson, et al.;1989)
- Based on separation of error variance of the items, one can assess invariance of latent factor variances, covariances, SEM regression paths
- No latent mean difference is estimated

Weak/metric Invariance

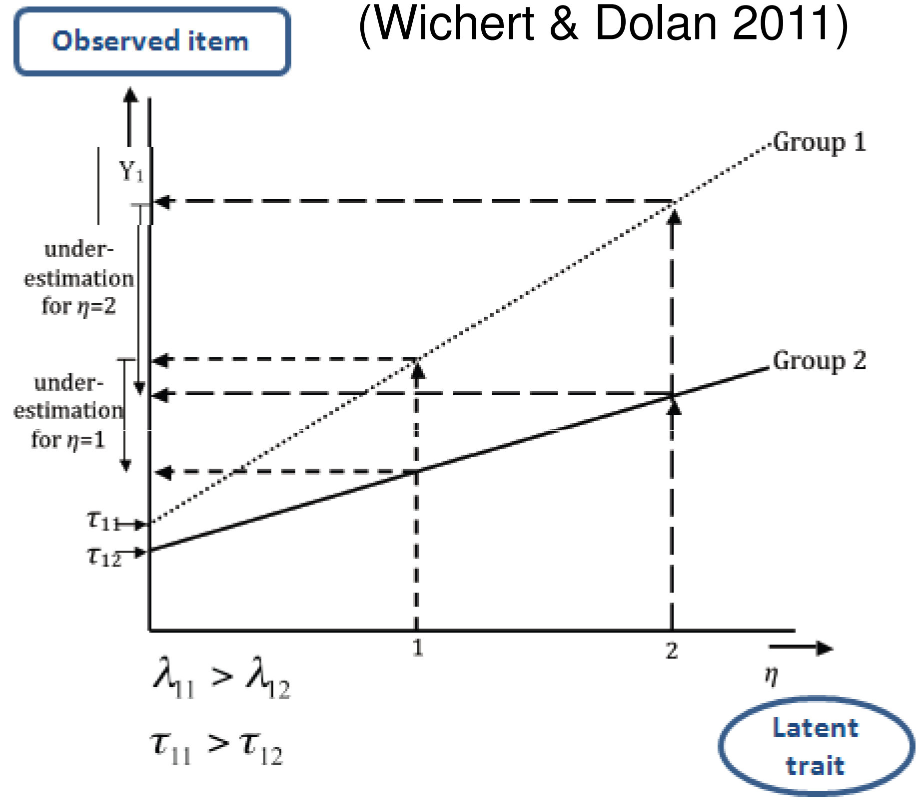
- Constrained = **factor loadings**



Weak/metric non-invariance

- Meaning of the items are different across groups
- Extreme response style might be present for some items
  - E.g. More likely to say “yes” in a group valuing decisiveness
  - Or more likely to choose a middle point in a group valuing humility
- One shouldn’t compare variances and covariances of the scale based on observed scores that contain non-invariant items

(Wichert & Dolan 2011)



- Non-invariant loading
- Non-invariant intercept

Lavaan: Model 2 metric MI

UNIVERSITY OF CAMBRIDGE

```
model2 <- cfa(HS.model, data=HolzingerSwineford1939, group="school",
          group.equal=c("loadings") )
summary(model2,fit.measures=TRUE)
```

Model 1: configural invariance:

chisq	df	pvalue	cfi	rmsea	bic
115.851	48.000	0.000	0.923	0.097	7706.822

Model 2: weak invariance (equal loadings):

chisq	df	pvalue	cfi	rmsea	bic
124.044	54.000	0.000	0.921	0.093	7680.771

```
anova(model1, model2)
```

```
Chi Square Difference Test

       Df    AIC    BIC  Chisq Chisq diff Df diff Pr(>Chisq)
model1 48 7484.4 7706.8 115.85
model2 54 7480.6 7680.8 124.04     8.1922       6     0.2244
```

- Model fit index changes are minimal, hence, metric invariance is established.

Lavaan: Model 2 metric MI

UNIVERSITY OF CAMBRIDGE

```
model2 <- cfa(HS.model, data=HolzingerSwineford1939, group="school",
            group.equal=c("loadings") )
```

Loadings are the same across groups, but intercepts are freely estimated

```
Group 1 [Pasteur]:

                   Estimate  Std.err  Z-value  P(>|z|)
Latent variables:
  visual =~
    x1                1.000
    x2                0.599    0.100    5.979    0.000
    x3                0.784    0.108    7.267    0.000
  textual =~
    x4                1.000
    x5                1.083    0.067   16.049    0.000
    x6                0.912    0.058   15.785    0.000
  speed =~
    x7                1.000
    x8                1.201    0.155    7.738    0.000
    x9                1.038    0.136    7.629    0.000

Intercepts:
    x1                4.941    0.093   52.991    0.000
    x2                5.984    0.100   60.096    0.000
    x3                2.487    0.094   26.465    0.000
    x4                2.823    0.093   30.371    0.000
    x5                3.995    0.101   39.714    0.000
    x6                1.922    0.081   23.711    0.000
    x7                4.432    0.086   51.540    0.000
    x8                5.563    0.078   71.087    0.000
    x9                5.418    0.079   68.153    0.000
    visual            0.000
    textual           0.000
    speed             0.000
```

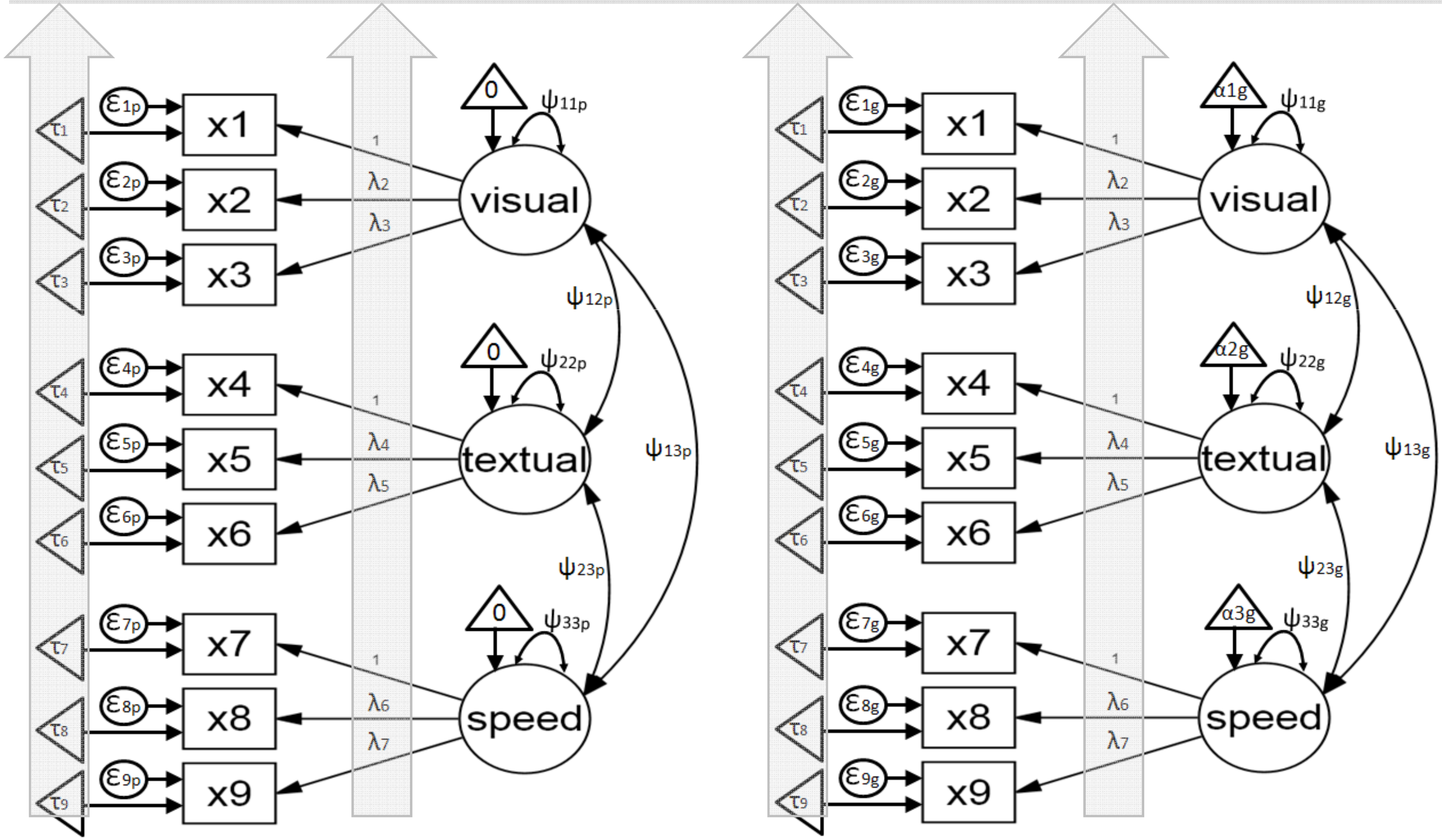
```
Group 2 [Grant-White]:

                   Estimate  Std.err  Z-value  P(>|z|)
Latent variables:
  visual =~
    x1                1.000
    x2                0.599    0.100    5.979    0.000
    x3                0.784    0.108    7.267    0.000
  textual =~
    x4                1.000
    x5                1.083    0.067   16.049    0.000
    x6                0.912    0.058   15.785    0.000
  speed =~
    x7                1.000
    x8                1.201    0.155    7.738    0.000
    x9                1.038    0.136    7.629    0.000

Intercepts:
    x1                4.930    0.097   50.763    0.000
    x2                6.200    0.091   68.379    0.000
    x3                1.996    0.085   23.455    0.000
    x4                3.317    0.092   35.950    0.000
    x5                4.712    0.100   47.173    0.000
    x6                2.469    0.091   27.248    0.000
    x7                3.921    0.086   45.555    0.000
    x8                5.488    0.087   63.257    0.000
    x9                5.327    0.085   62.786    0.000
    visual            0.000
    textual           0.000
    speed             0.000
```

Measurement invariance:
Step 3: Strong/scalar invariance

- Constrain item intercepts equal across groups
- Constrain factor loadings
- This is important for assessing mean difference of the latent variable across groups
- In case of partial invariance of item intercepts, constrain the invariant intercepts and set free the non-invariant intercepts (Byrne, Shavelson, et al.;1989)
- Latent mean difference is estimated

Strong/scalar invariance

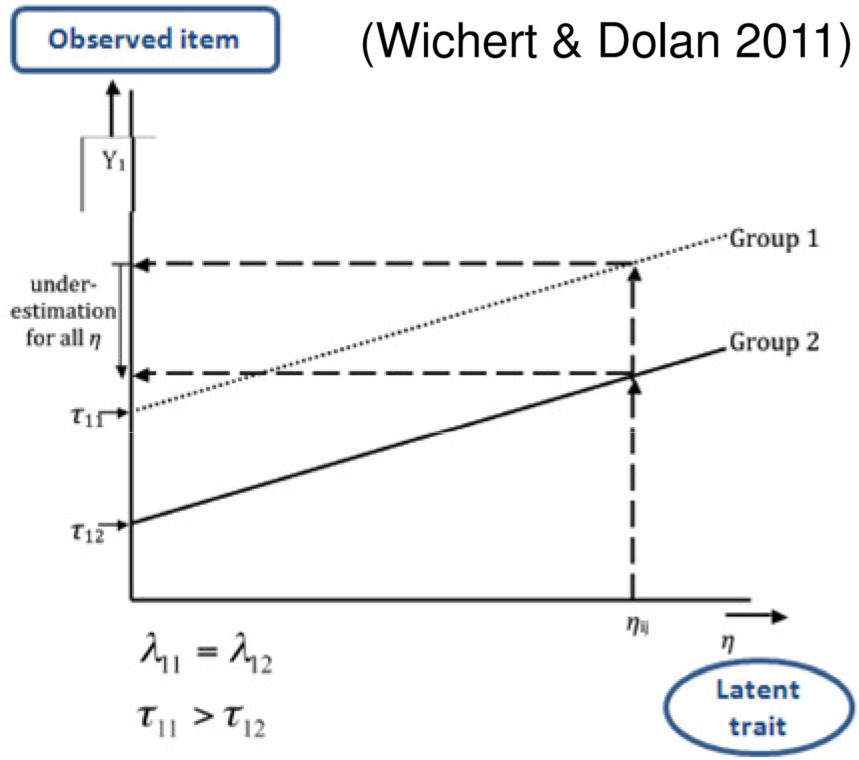
- Constrained = Factor loadings+ **item intercepts**



Strong/scalar non-invariance

- A group tend to systematically give higher or lower item response
- This might be caused by a norm specific to that group
  - For instance in name learning tests that involve unfamiliar names for a group
- This is an additive effect. It affects the means of the observed item, hence affects the mean of the scale and the latent variable

(Wichert & Dolan 2011)



- Invariant loading
- Non-invariant intercept

Lavaan: Model 3 scalar invariance

UNIVERSITY OF CAMBRIDGE

```
model3 <- cfa(HS.model, data=HolzingerSwineford1939,
             group="school", group.equal=c("loadings", "intercepts"))
summary(model3,fit.measures=TRUE)
```

Model 2: weak invariance (equal loadings):

chisq	df	pvalue	cfi	rmsea	bic
124.044	54.000	0.000	0.921	0.093	7680.771

Model 3: strong invariance (equal loadings + intercepts):

chisq	df	pvalue	cfi	rmsea	bic
164.103	60.000	0.000	0.882	0.107	7686.588

```
anova(model1, model2)
```

```
Chi Square Difference Test

       Df    AIC    BIC  Chisq Chisq diff Df diff Pr(>Chisq)
model2 54 7480.6 7680.8 124.04
model3 60 7508.6 7686.6 164.10     40.059       6  4.435e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Significant χ^2 change indicates intercepts non-invariance
- Modification index can be used to identify which item intercepts are non-invariant

Lavaan: Model 3 scalar invariance

UNIVERSITY OF CAMBRIDGE

```
model3 <- cfa(HS.model, data=HolzingerSwineford1939, group="school",
              group.equal=c("loadings", "intercepts"))
```

Both intercepts and loadings are constrained across groups, but latent means are estimated

```
Group 1 [Pasteur]:

                   Estimate  Std.err  Z-value  P(>|z|)
Latent variables:
  visual =~
    x1                1.000
    x2                0.576    0.101    5.713    0.000
    x3                0.798    0.112    7.146    0.000
  textual =~
    x4                1.000
    x5                1.120    0.066   16.965    0.000
    x6                0.932    0.056   16.608    0.000
  speed =~
    x7                1.000
    x8                1.130    0.145    7.786    0.000
    x9                1.009    0.132    7.667    0.000

Intercepts:
    x1                5.001    0.090   55.760    0.000
    x2                6.151    0.077   79.905    0.000
    x3                2.271    0.083   27.387    0.000
    x4                2.778    0.087   31.954    0.000
    x5                4.035    0.096   41.858    0.000
    x6                1.926    0.079   24.426    0.000
    x7                4.242    0.073   57.975    0.000
    x8                5.630    0.072   78.531    0.000
    x9                5.465    0.069   79.016    0.000
    visual            0.000
    textual           0.000
    speed             0.000
```

```
Group 2 [Grant-White]:

                   Estimate  Std.err  Z-value  P(>|z|)
Latent variables:
  visual =~
    x1                1.000
    x2                0.576    0.101    5.713    0.000
    x3                0.798    0.112    7.146    0.000
  textual =~
    x4                1.000
    x5                1.120    0.066   16.965    0.000
    x6                0.932    0.056   16.608    0.000
  speed =~
    x7                1.000
    x8                1.130    0.145    7.786    0.000
    x9                1.009    0.132    7.667    0.000

Intercepts:
    x1                5.001    0.090   55.760    0.000
    x2                6.151    0.077   79.905    0.000
    x3                2.271    0.083   27.387    0.000
    x4                2.778    0.087   31.954    0.000
    x5                4.035    0.096   41.858    0.000
    x6                1.926    0.079   24.426    0.000
    x7                4.242    0.073   57.975    0.000
    x8                5.630    0.072   78.531    0.000
    x9                5.465    0.069   79.016    0.000
    visual           -0.148    0.122   -1.211    0.226
    textual           0.576    0.117    4.918    0.000
    speed            -0.177    0.090   -1.968    0.049
```

Lavaan: Modification index

UNIVERSITY OF CAMBRIDGE

```
model3 <- cfa(HS.model, data=HolzingerSwineford1939,
            group="school",    group.equal=c("loadings","intercepts"))
modindices(model3)
```

	lhs	op	group	mi	epc	sepc.lv	sepc.all	sepc.nox
81	x3	~1	1	17.717	0.248	0.248	0.206	0.206
85	x7	~1	1	13.681	0.205	0.205	0.186	0.186
171	x3	~1	2	17.717	-0.248	-0.248	-0.238	-0.238
175	x7	~1	2	13.681	-0.205	-0.205	-0.193	-0.193

- Modification index showed that item 3 and item 7 have intercept estimates that are non-invariant across groups.

- In the next model, we allow partial invariance of item intercept, freeing the intercepts of item 3 and item 7.

Lavaan: Model 3a scalar invariance with partial invariance

```
model3a <- cfa(HS.model, data=HolzingerSwineford1939, group="school",
            group.equal=c("loadings", "intercepts"), group.partial=c("x3~1", "x7~1"))
summary(model3a,fit.measures=TRUE)
```

Model 2: weak invariance (equal loadings):

chisq	df	pvalue	cfi	rmsea	bic
124.044	54.000	0.000	0.921	0.093	7680.771

Model 3a: strong invariance (equal loadings + intercepts), allowing intercepts of item 3 and item 7 to vary:

chisq	df	pvalue	cfi	rmsea	bic
129.422	58.000	0.000	0.919	0.090	7663.322

```
anova(model3a, model2)
```

```
Chi Square Difference Test

        Df    AIC    BIC  Chisq Chisq diff Df diff Pr(>Chisq)
model2  54 7480.6 7680.8 124.04
model3a 58 7478.0 7663.3 129.42     5.3789       4     0.2506
```

➢ The scalar invariance model now has partial invariance, thus latent means can be compared

Lavaan: Model 3a scalar invariance with partial invariance (x3, x7)

Group 1 [Pasteur]:

Intercepts:	Estimate	Std.err	Z-value	P(>|z|)
x1	4.914	0.092	53.538	0.000
x2	6.087	0.079	76.999	0.000
x3	2.487	0.094	26.474	0.000
x4	2.778	0.087	31.953	0.000
x5	4.035	0.096	41.861	0.000
x6	1.926	0.079	24.425	0.000
x7	4.432	0.086	51.533	0.000
x8	5.569	0.074	75.328	0.000
x9	5.409	0.070	77.182	0.000
visual	0.000			
textual	0.000			
speed	0.000			

Group 2 [Grant-White]:

Intercepts:	Estimate	Std.err	Z-value	P(>|z|)
x1	4.914	0.092	53.538	0.000
x2	6.087	0.079	76.999	0.000
x3	1.955	0.108	18.170	0.000
x4	2.778	0.087	31.953	0.000
x5	4.035	0.096	41.861	0.000
x6	1.926	0.079	24.425	0.000
x7	3.992	0.094	42.478	0.000
x8	5.569	0.074	75.328	0.000
x9	5.409	0.070	77.182	0.000
visual	0.051	0.129	0.393	0.695
textual	0.576	0.117	4.918	0.000
speed	-0.071	0.089	-0.800	0.424

- Grant-White school students does better on textual factor as compared to Pasteur school students
- After allowing for partial invariance, there is no difference in speed between Grant-While school and Pasteur school

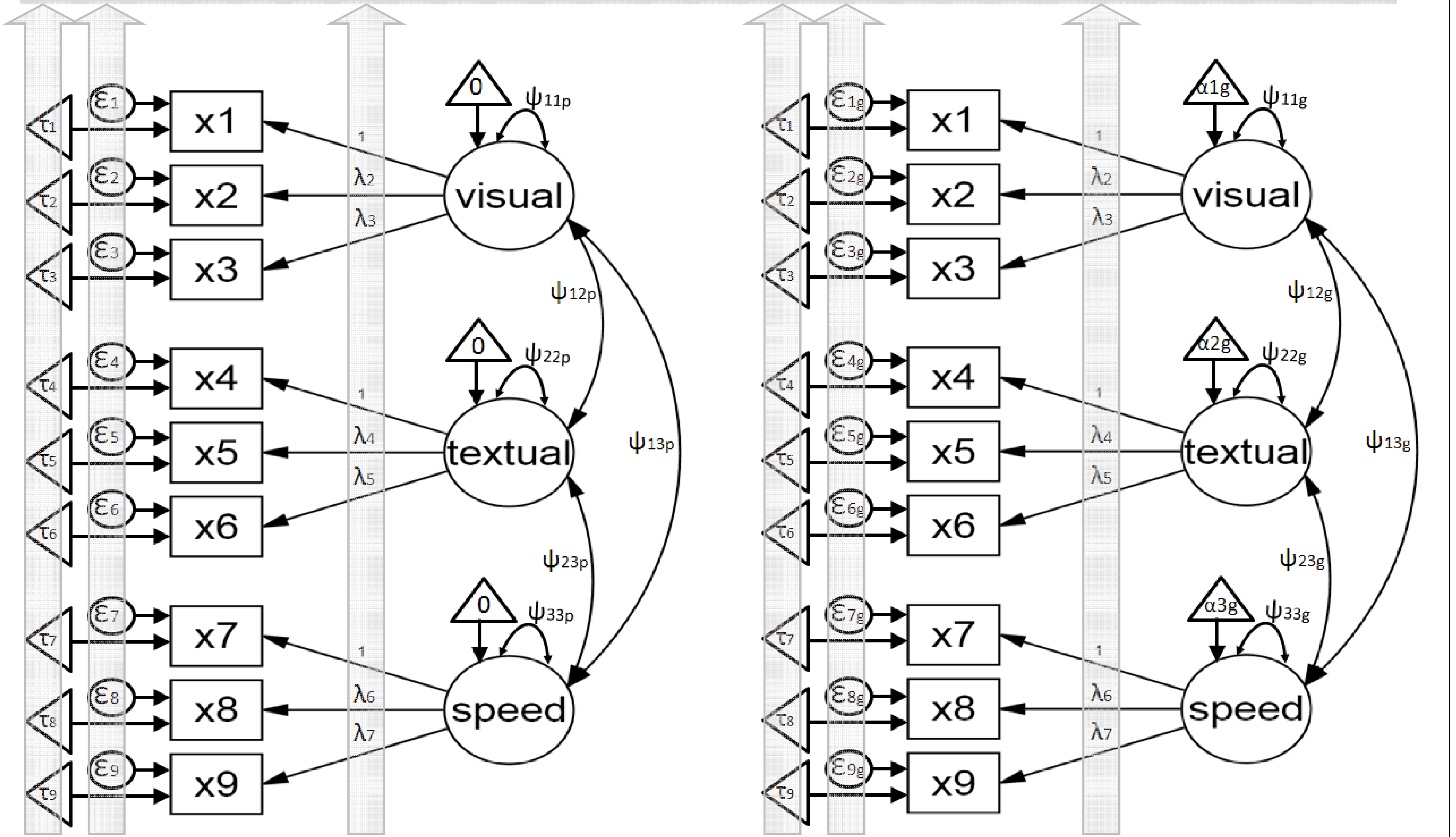
Measurement invariance: Step 4: Strict invariance

UNIVERSITY OF CAMBRIDGE

- Constrain item residual variances to be equal across groups
- Constrain item factor loadings and intercepts equal across groups. In case of partial invariance constrain the invariant parameters and set free the non-invariant parameters
- Strict invariance is important for group comparisons based on the sum of observed item scores, because observed variance is a combination of true score variance and residual variance
- Latent mean difference is estimated

Strict invariance

- Constrained = factor loadings + item intercepts + **residual variances**



Lavaan: Model 4 strict invariance

UNIVERSITY OF CAMBRIDGE

```
model4<- cfa(HS.model, data=HolzingerSwineford1939, group="school",
        group.equal=c("loadings", "intercepts", "residuals"),
        group.partial=c("x3~1", "x7~1"))
summary(model4,fit.measures=TRUE)
```

Model 3a: strong invariance (equal loadings + intercepts), allowing intercepts of item 3 and item 7 to vary:

chisq	df	pvalue	cfi	rmsea	bic
129.422	58.000	0.000	0.919	0.090	7663.322

Model 4: strict invariance (equal loadings + intercepts + item residual variances)

chisq	df	pvalue	cfi	rmsea	bic
147.260	67	0.000	0.909	0.089	7629.796

- The chi-square difference is borderline significant (p=0.037), but the BIC and RMSEA showed improvement. Based on the number of tests in the model, it is probably safe to ignore the chi-square significance
- This imply that items are equally reliable across groups. If all items were invariant, it would be valid to use sum scores for data involving mean and regression coefficient comparisons across groups

Structural invariances

- Factor variances
- Factor covariances (if more than one latent factors)
- Regression path coefficients (in multiple group SEM analysis)

Lavaan: Model 5 factor variances and covariances

```
model5 <- cfa(HS.model, data=HolzingerSwineford1939, group="school",
          group.equal=c("loadings", "intercepts", "residuals", "lv.variances",
          "lv.covariances"), group.partial=c("x3~1", "x7~1"))
summary(model5,fit.measures=TRUE)
```

Model 4: strict invariance (equal loadings + intercepts + item residual variances)

chisq	df	pvalue	cfi	rmsea	bic
147.260	67	0.000	0.909	0.089	7629.796

Model 5: factor variance and covariance invariance (equal loadings + intercepts + item residual variances + factor var&cov)

chisq	df	pvalue	cfi	rmsea	bic
153.258	73	0.000	0.909	0.085	7601.551

- The chi-square difference is not significant (p= 0.42), and the RMSEA showed improvement. The variance and covariance of latent factors are invariant across groups
- As a matter of fact, if one does analysis with latent variables, then strict invariance if not really a prerequisite, since measurement errors are taken into account of as part of the model

Summarising the MI analysis

UNIVERSITY OF CAMBRIDGE

Model	χ^2	DF	CFI	RMSEA	BIC	Base	$\Delta\chi 2$	ΔDF	ΔCFI	ΔRMSEA	ΔBIC	
m1	115.851	48	0.923	0.097	7707							inv=none, free=fl+inter+uniq+var+cov
m2	124.044	54	0.921	0.093	7681	m1	8.193	6	-0.002	-0.004	-26	inv=fl, free=inter+uniq+var+cov
m3	164.103	60	0.882	0.107	7687	m2	40.059	6	-0.039	0.014	6	inv=fl+inter, free=Fmean+uniq+var+cov
m3a	129.422	58	0.919	0.090	7663	m2	5.378	4	-0.002	-0.003	-17	inv=fl+inter, free=inter(x3+x7)+uniq+var+cov
m4	147.260	67	0.909	0.089	7630	m3a	17.838	9	-0.010	-0.001	-34	inv=fl+inter+uniq, free=inter(x3+x7)+Fmean+var+cov
m5	153.258	73	0.909	0.085	7602	m4	5.998	6	0.000	-0.004	-28	inv=fl+inter+uniq+var+cov , free=inter(x3+x7)+Fmean

- MI analysis includes a series of nested models with an increasingly restrictive parameter specifications across groups

- The same principle applies for longitudinal data
  - Testing measurement invariance of items over time
  - This is a basis for analysis that compares latent means over time, for instance, in a growth curve model

Measurement invariance – other issues

- Setting of referent indicator
  - Identify the “most non-invariant” item to use as referent indicator
  - Or set factor variance to 1 to avoid selecting a referent item
- Multiple testing issue
- Analysing Likert scale data
  - Number of categories and data skewness (Rhemtulla, Brosseau-Liard, & Savalei; 2012)
  - Robust maximum likelihood
  - Ordinal factor analysis treating data as dichotomous or polytomous (Millsap & Tein, 2004; Muthen & Asparouhov, 2002)

Some references

1. Sass, D. A. (2011). "Testing Measurement Invariance and Comparing Latent Factor Means Within a Confirmatory Factor Analysis Framework." Journal of Psychoeducational Assessment 29(4): 347-363.
2. Wicherts, J. M. and C. V. Dolan (2010). "Measurement invariance in confirmatory factor analysis: An illustration using IQ test performance of minorities." Educational Measurement: Issues and Practice 29(3): 39-47.
3. Gregorich, S. E. (2006). "Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework." Medical Care 44(11 Suppl 3): S78.
4. Byrne, B. M., R. J. Shavelson, et al. (1989). "Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance." Psychological bulletin 105(3): 456-466.
5. Millsap, R. E. and J. Yun-Tein (2004). "Assessing factorial invariance in ordered-categorical measures." Multivariate Behavioral Research 39(3): 479-515.
6. Meredith, W. (1993). "Measurement invariance, factor analysis and factorial invariance." Psychometrika 58(4): 525-543.
7. Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. Psychological Methods, 17(3), 354-373. doi: 10.1037/a0029315

Acknowledgement:

Dr. Adam Wagner provided thoughtful comments on earlier drafts