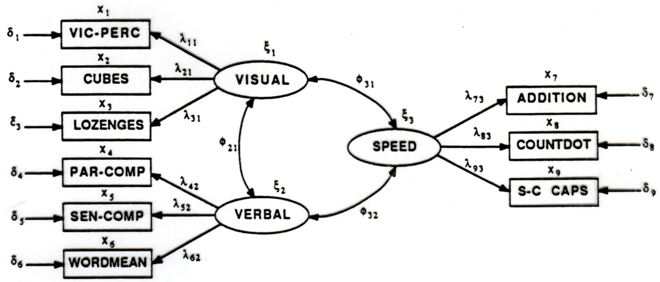
CFA & SEM

Lecture 2: Measurement models and CFA

Michael Friendly

SCS Short Course



Measurement error

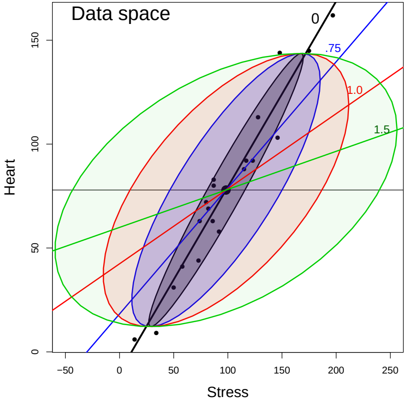
- Path analysis models assume that all exogenous predictors (x) are measured without error
  - The only error terms are the residuals ζ (errors-in-equations) for the endogenous (y) variables
- This is often (at least approximately) true for variables like age, height, income, occupational status, etc.
- It is less likely to be true for constructs of interest in the social sciences: intelligence, depression, mathematical aptitude, need for achievement, etc.
  - Measurement error has severe consequences— reduced precision, but much worse: bias
  - CFA & SEM handle this by introducing a measurement model, using latent variables

Measurement error: Example

Data on the relationship between Heart (y) damage and Stress (x)

$$\text{Heart} = \beta_0 + \beta_1 \text{Stress}$$

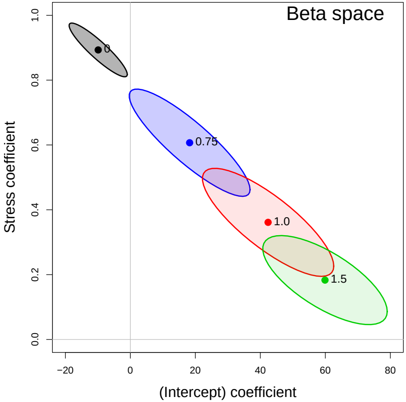
What happens if we add random error, $\mathcal{N}(0, \delta \times SD_{Stress})$ to each x-value ($\delta = \{0.75, 1.0, 1.5\}$)?



- The grey ellipse and the regression line “0” show the original data
- Increasing measurement error makes the data ellipses wider
- Increasing measurement error biases β_1 towards zero!
- NB: Adding random error to Heart (y) would decrease precision but not introduce bias.

Measurement error: Example

These effects can also be seen in parameter (β) space



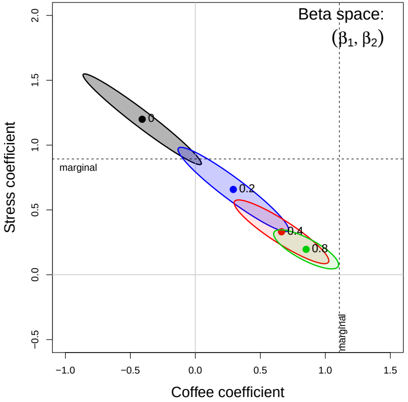
- β_1 decreases with increasing error
- the intercept, β_0 increases
- The increasing size of confidence ellipses shows decreased precision of the estimates

Measurement error: Example

Now, consider a multiple regression model, with coffee as an additional predictor

$$\text{Heart} = \beta_0 + \beta_1 \text{Stress} + \beta_2 \text{Coffee}$$

What is the effect of measurement error in Stress on both coefficients, (β_1, β_2)



- The coefficient β_1 for Stress goes towards 0, as before
- The coefficient β_2 for Coffee decreases towards its marginal value (Stress not included in the model)
- Thus, measurement error in even one x variable has effects throughout the model

Latent variables

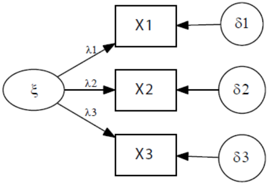
In EFA, CFA & SEM, measurement error in observed variables is handled by positing an underlying latent variable ("factor") responsible for producing the observed score x

$$x_i = \lambda \xi_i + \delta_i$$

- ξ ("ksi" or "xi") is the true latent variable measured by x
- λ is the regression coefficient ("factor loading") of x on ξ
- δ is the error of measurement
- x is called an indicator of the latent variable ξ

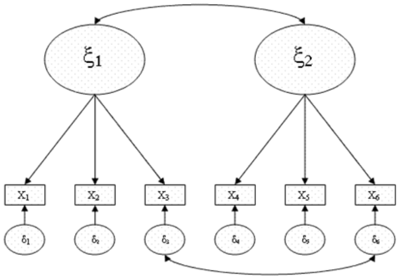
There there are usually multiple observed indicators, $x_1, x_2, \ldots$ measuring a given (latent) construct

$$\begin{aligned} x_{1i} &= \lambda_1 \xi_i + \delta_{1i} \\ x_{2i} &= \lambda_2 \xi_i + \delta_{2i} \\ x_{3i} &= \lambda_3 \xi_i + \delta_{3i} \end{aligned}$$



Latent variables

- The observed variables can also be considered as measures of two (or more) latent variables
- The latent variables (factors) can be correlated
- There can also be correlations among the error terms



$$
\begin{aligned}
x_1 &= \lambda_{11}\xi_1 + \lambda_{12}\xi_1 + \delta_1 \\
x_2 &= \lambda_{21}\xi_1 + \lambda_{22}\xi_2 + \delta_2 \\
x_3 &= \lambda_{31}\xi_1 + \lambda_{32}\xi_2 + \delta_3 \\
\vdots &= \vdots
\end{aligned}
$$

The General CFA model

The general CFA measurement model is

$$\boldsymbol{x} = \boldsymbol{\Lambda}\boldsymbol{\xi} + \boldsymbol{\delta}$$

where

- $\boldsymbol{x}$ is the $q \times 1$ vector of observed or measured variables
- $\boldsymbol{\Lambda}$ is the $q \times k$ matrix of factor loadings
- $\boldsymbol{\xi}$ is the vector of latent variables
- i.e., λ_{ij} is the partial regression coefficient for x_i on ξ_j in the regression of x_i on $\xi_1, \xi_2, \ldots, \xi_k$
- $\boldsymbol{\delta}$ is the vector of errors of measurement or disturbance terms

This model, together with assumptions implies that the covariance matrix of $\boldsymbol{x}$ is

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}^{\mathsf{T}} + \boldsymbol{\Theta}$$

where $\boldsymbol{\Phi}$ is the covariance matrix of the factors, $\boldsymbol{\xi}$, and $\boldsymbol{\Theta}$ is the covariance matrix of the errors, $\boldsymbol{\delta}$

Testing Equivalence of Measures with CFA

Test theory is concerned with ideas of reliability, validity and equivalence of measures.

- The same ideas apply to other constructs (e.g., anxiety scales or experimental measures of conservation).
- Test theory defines several degrees of “equivalence”.
- Each kind may be specified as a confirmatory factor model with a single common factor.
- The CFA approach allows a more nuanced approach to these issues.

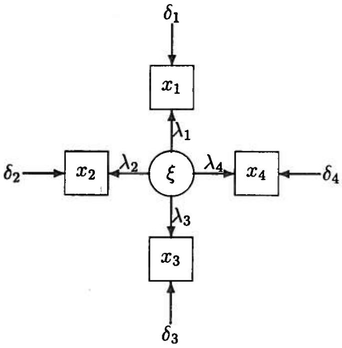
$$\boldsymbol{\Sigma} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} \begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \end{pmatrix} + \begin{bmatrix} \theta_{11} & & & \\ & \theta_{22} & & \\ & & \theta_{33} & \\ & & & \theta_{44} \end{bmatrix}$$

Testing Equivalence of Measures with CFA

One-factor model:

$$\boldsymbol{\Sigma} = \boldsymbol{\lambda}\boldsymbol{\lambda}^{\mathsf{T}} + \boldsymbol{\Theta} = \begin{bmatrix} \lambda_1^2 + \theta_{11} & & & \\ \lambda_2\lambda_1 & \lambda_2^2 + \theta_{22} & & \\ \lambda_3\lambda_1 & \lambda_3\lambda_2 & \lambda_3^2 + \theta_{33} & \\ \lambda_4\lambda_1 & \lambda_4\lambda_2 & \lambda_4\lambda_3 & \lambda_4^2 + \theta_{44} \end{bmatrix}$$

Path diagram:



Congeneric measurement model

- The single factor model is called the congeneric measurement model
- It implies that the true scores, $\tau_i = \lambda_i \xi$ are perfectly correlated
- The true score variance in x_i is λ_i^2 — also called comunality in EFA lingo
- The reliability of x_i is

$$\rho_i = \frac{\lambda_i^2}{\mathrm{var}(x_i)} = \frac{\lambda_i^2}{\lambda_i^2 + \theta_{ii}} = 1 - \frac{\theta_{ii}}{\lambda_i^2 + \theta_{ii}}$$

- Strictly speaking, the error term δ_i ("unique factor") is considered to be the sum of two uncorrelated components

$$\begin{aligned} \delta_i &= s_i + e_i \\ \text{unique} &= \text{specific} + \text{error} \end{aligned}$$

- ρ_i is a lower bound on true reliability

Kinds of equivalence

- ***Parallel tests***: Measure the same thing with equal precision. The strongest form of “equivalence”.
- ***Tau-equivalent tests***: Have equal true score variances (λ_i^2), but may differ in error variance (θ_{ii}).
  Like parallel tests, this requires tests of the same length & time limits.
  E.g., short forms cannot be τ-equivalent.
- ***Congeneric tests***: The weakest form of equivalence: All tests measure a single common factor, but the loadings & error variances may vary.

These hypotheses may be tested with CFA/SEM by testing equality of the factor loadings (λ_i) and unique variances (θ_{ii}).

$$\underbrace{\overbrace{\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4}^{\tau \text{ equivalent}} \qquad \theta_{11} = \theta_{22} = \theta_{33} = \theta_{44}}_{\text{Parallel}}$$

Example: Reliability in essay scoring

- Essay exams present a challenge for standardized testing (SAT, LSAT, etc.)
- An early study by Votaw (1948) analyzed scores for N=126 examinees given a 3-part English composition test
  - x_1: score on an original copy of the part 1 essay
  - x_2: score on a hand-written copy of the part 1 essay
  - x_3: score on a carbon-copy of the hand-written part 1 essay
  - x_4: score on an original copy of the part 2 essay
- Questions:
  - Can these scores be used interchangeably– as strictly parallel or τ-equivalent tests?
  - If not, are the scores on original copies more reliable than those on copies?
  - Are the scores for part 1 and part 2 originals equally reliable?

Example: Reliability in essay scoring

Read the covariance matrix:

```
library(sem)
votaw <- readMoments(diag=TRUE,
    names=c('orig1', 'hcpy1', 'ccpy1', 'orig2'), text="
25.0704
12.4363  28.2021
11.7257   9.2281  22.7390
20.7510  11.9732  12.0692  21.8707
")
```

Fit the congeneric model:

```
votaw.mod1 <- specifyEquations(text="
orig1 = lam1 * Ability
hcpy1 = lam2 * Ability
ccpy1 = lam3 * Ability
orig2 = lam4 * Ability
V(Ability) = 1
")
```

Other models

More restrictive models are specified simply by using the same parameter names for equal parameters.

τ-equivalent model

```
votaw.mod2 <- specifyEquations(
  text="
orig1 = lam * Ability
hcpy1 = lam * Ability
ccpy1 = lam * Ability
orig2 = lam * Ability
V(Ability) = 1
")
```

parallel model

```
votaw.mod3 <- specifyEquations(
  text="
orig1 = lam * Ability
hcpy1 = lam * Ability
ccpy1 = lam * Ability
orig2 = lam * Ability
V(Ability) = 1
V(orig1) = error
V(hcpy1) = error
V(ccpy1) = error
V(orig2) = error
")
```

An intermediate “semi-parallel” model specified two sets of equal loadings λ_1 for `orig1` and `orig2`, λ_1 for `hcpy1` and `ccpy1`

Example: Reliability in essay scoring

Summary of analyses:

Model	Hypothesis	df	χ^2	p
1	congeneric	2	2.28	0.32
2	tau-equivalent	5	40.42	0.00
3	parallel	8	109.12	0.00
4	semi-parallel	6	8.99	0.17

Results for congeneric model:

Variable	$\widehat{\lambda}_i$	s.e.($\widehat{\lambda}_i$)	$\widehat{\rho}_i$
orig1	4.57	0.36	0.83
hcpy1	2.68	0.45	0.25
ccpy1	2.65	0.40	0.31
orig2	4.54	0.33	0.94

However, semi-parallel model is simpler, and fits well.

Several Sets of Congeneric Tests

For several sets of measures, the test theory ideas of congeneric tests can be extended to test the equivalence of each set.
If each set is congeneric, the estimated correlations among the latent factors measure the strength of relations among the underlying "true scores".

Example: Correcting for Unreliability

- Given two measures, x and y, the correlation between them is limited by the reliability of each.
- CFA can be used to estimate the correlation between the true scores, τ_x, τ_y, or to test the hypothesis that the true scores are perfectly correlated:

$$H_0 : \rho(\tau_x, \tau_y) = 1$$

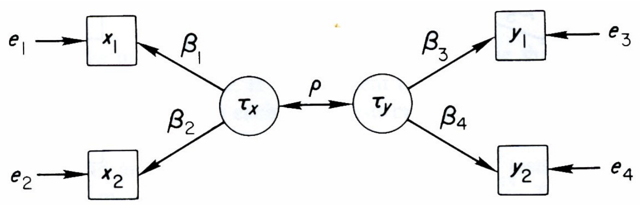
- The estimated true-score correlation, $\hat{\rho}(\tau_x, \tau_y)$ is called the correlation of x, y corrected for attenuation.

Several Sets of Congeneric Tests

The analysis requires two "parallel" forms of each test, x_1, x_2, y_1, y_2. Tests are carried out with the model:

$$\begin{bmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \beta_1 & 0 \\ \beta_2 & 0 \\ 0 & \beta_3 \\ 0 & \beta_4 \end{bmatrix} \begin{bmatrix} \tau_x \\ \tau_y \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix} = \boldsymbol{\Lambda \tau} + \boldsymbol{e}$$

with $corr(\boldsymbol{\tau}) = \rho$, and $var(\boldsymbol{e}) = \mathrm{diag}\{\theta_1^2, \theta_2^2, \theta_3^2, \theta_4^2\}$. The model is shown in this path diagram:



Several Sets of Congeneric Tests

Hypotheses

The following hypotheses can be tested. The difference in χ^2 for H_1 vs. H_2, or H_3 vs. H_4 provides a test of the hypothesis that $\rho = 1$.

$$
\begin{aligned}
H_1 &: \quad \rho = 1 \text{ and } H_2 \\
H_2 &: \quad \left\{ \begin{array}{ll} \beta_1 = \beta_2 & \theta_1^2 = \theta_2^2 \\ \beta_3 = \beta_4 & \theta_3^2 = \theta_4^2 \end{array} \right. \\
H_3 &: \quad \rho = 1, \text{ all other parameters free} \\
H_4 &: \quad \text{all parameters free}
\end{aligned}
$$

H_1 and H_2 assume the measures x_1, x_2 and y_1, y_2 are parallel. H_3 and H_4 assume they are merely congeneric.

Several Sets of Congeneric Tests

These four hypotheses actually form a 2×2 factorial

- parallel vs. congeneric: H_1 and H_2 vs. H_3 and H_4 and
- $\rho = 1$ vs. $\rho \neq 1$.

For nested models, model comparisons can be done by testing the difference in χ^2, or by comparing other fit statistics (AIC, BIC, RMSEA, etc.)

- LISREL can fit multiple models, but you have to do the model comparison tests "by hand."
- AMOS can fit multiple models, and does the model comparisons for you.
- With PROC CALIS, the CALISCMP macro provides a flexible summary of multiple-model comparisons.
- `sem()` provides an `anova()` method

Example: Lord's data

Lord's vocabulary test data:

- x_1, x_2: two 15-item tests, liberal time limits
- y_1, y_2: two 75-item tests, highly speeded

Analyses of these data give the following results:

Hypothesis	Free Parameters	df	χ^2	p-value	AIC
H_1: par, $\rho = 1$	4	6	37.33	0.00	25.34
H_2: par	5	5	1.93	0.86	-8.07
H_3: cong, $\rho = 1$	8	2	36.21	0.00	32.27
H_4: cong	9	1	0.70	0.70	-1.30

- Models H2 and H4 are acceptable, by χ^2 tests
- Model H2 is "best" by AIC

Lord's data

The tests of $\rho = 1$ can be obtained by taking the differences in χ^2,

	Parallel		Congeneric	
	χ^2	df	χ^2	df
$\rho = 1$	37.33	6	36.21	2
$\rho \neq 1$	1.93	5	0.70	1
	35.40	1	35.51	1

- Both tests reject the hypothesis that $\rho = 1$,
- Under model H2, the ML estimate is $\widehat{\rho} = 0.889$.
- $\Rightarrow$ speeded and unspeeded vocab. tests do not measure *exactly* the same thing.
- NB: The CFA/SEM approach is far more rigorous than usually applied to social measurements like anxiety, depression, etc.
  SAS example: `datavis.ca/courses/factor/sas/calis1c.sas`

Lord's data: PROC CALIS

```
data lord(type=cov);
   input _type_ $ _name_ $  x1 x2 y1 y2;
datalines;
n    .  649     649     649     649
cov  x1 86.3937   .       .       .
cov  x2 57.7751 86.2632   .       .
cov  y1 56.8651 59.3177 97.2850   .
cov  y2 58.8986 59.6683 73.8201 97.8192
mean .   0       0       0       0
;
```

Model H4: $\beta_1, \beta_2, \beta_3, \beta_4 \ldots \rho$=free

```
title "Lord's data: H4- unconstrained two-factor model";
proc calis data=lord
     cov
     summary   outram=M4;
   lineqs  x1  = beta1 F1  + e1,
           x2  = beta2 F1  + e2,
           y1  = beta3 F2  + e3,
           y2  = beta4 F2  + e4;
   std  F1 F2 = 1 1,
        e1 e2 e3 e4 = ve1 ve2 ve3 ve4;
   cov  F1 F2 = rho;
run;
```

Lord's data: PROC CALIS

The SUMMARY output contains many fit indices:

```
       Lord's data: H4- unconstrained two-factor model

 Covariance Structure Analysis: Maximum Likelihood Estimation

 Fit criterion . . . . . . . . . . . . . . . . . .     0.0011
 Goodness of Fit Index (GFI) . . . . . . . . . . .     0.9995
 GFI Adjusted for Degrees of Freedom (AGFI)  . . .     0.9946
 Root Mean Square Residual (RMR) . . . . . . . . .     0.2715
 Chi-square = 0.7033       df = 1      Prob>chi**2 = 0.4017
 Null Model Chi-square:    df = 6                  1466.5884
 Bentler's Comparative Fit Index . . . . . . . . .     1.0000
 Normal Theory Reweighted LS Chi-square  . . . . .     0.7028
 Akaike's Information Criterion  . . . . . . . . .    -1.2967
 Consistent Information Criterion  . . . . . . . .    -6.7722
 Schwarz's Bayesian Criterion  . . . . . . . . . .    -5.7722
 McDonald's (1989) Centrality. . . . . . . . . . .     1.0002
 Bentler & Bonett's (1980) Non-normed Index. . . .     1.0012
 Bentler & Bonett's (1980) Normed Index. . . . . .     0.9995
 James, Mulaik, & Brett (1982) Parsimonious Index.     0.1666
 ...
```

Lord's data: PROC CALIS

Model H3: H4, with $\rho = 1$

```
title "Lord's data: H3- rho=1, one-congeneric factor";
proc calis data=lord
     cov summary outram=M3;
   lineqs  x1  = beta1 F1 + e1,
           x2  = beta2 F1 + e2,
           y1  = beta3 F2 + e3,
           y2  = beta4 F2 + e4;
   std  F1 F2 = 1 1,
        e1 e2 e3 e4 = ve1 ve2 ve3 ve4;
   cov  F1 F2 = 1;
run;
```

Model H2: $\beta_1 = \beta_2, \beta_3 = \beta_4$ ..., ρ=free

```
title "Lord's data: H2- X1 and X2 parallel, Y1 and Y2 parallel";
proc calis data=lord
     cov summary outram=M2;
   lineqs  x1  = betax F1  + e1,
           x2  = betax F1  + e2,
           y1  = betay F2  + e3,
           y2  = betay F2  + e4;
   std  F1 F2 = 1 1,
        e1 e2 e3 e4 = vex vex vey vey;
   cov  F1 F2 = rho;
run;
```

Lord's data: CALISCMP macro

Model comparisons using CALISCMP macro and the OUTRAM= data sets

```
%caliscmp(ram=M1 M2 M3 M4,
   models=%str(H1 par rho=1/H2 par/H3 con rho=1/H4 con),
   compare=1 2 / 3 4 /1 3/ 2 4);
```

Model Comparison Statistics from 4 RAM data sets

Model	Parameters	df	Chi-Square	P>ChiSq	RMS Residual	GFI	AIC
H1 par rho=1	4	6	37.3412	0.00000	2.53409	0.97048	25.3412
H2 par	5	5	1.9320	0.85847	0.69829	0.99849	-8.0680
H3 con rho=1	8	2	36.2723	0.00000	2.43656	0.97122	32.2723
H4 con	9	1	0.7033	0.40168	0.27150	0.99946	-1.2967

(more fit statistics are compared than shown here.)

Lord's data: CALISCMP macro

```
%caliscmp(ram=M1 M2 M3 M4,
   models=%str(H1 par rho=1/H2 par/H3 con rho=1/H4 con),
   compare=1 2 / 3 4 /1 3/ 2 4);
```

```
          Model Comparison Statistics from 4 RAM data sets
 Model Comparison                              ChiSq   df   p-value
------------------------------------------------------------
H1 par rho=1 vs. H2 par                      35.4092    1   0.00000 ****
H3 con rho=1 vs. H4 con                      35.5690    1   0.00000 ****
H1 par rho=1 vs. H3 con rho=1                 1.0689    4   0.89918
H2 par vs. H4 con                             1.2287    4   0.87335
```

Multi-factor congeneric models

- Multi-factor models are at the heart of CFA
- An important special case is when there are G sets of (assumed) congeneric variables, each of which are indicators of a latent variable
- In EFA lingo, these are called non-overlapping factors
- The measurement models for the variables $\mathbf{x}_g$ in set g are of the form

$$\mathbf{x}_g = \boldsymbol{\lambda}_g \xi_g + \boldsymbol{\delta}_g$$

- Then, the loadings $\boldsymbol{\Lambda}$ for all variables can be represented as

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\lambda}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\lambda}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \boldsymbol{\lambda}_G \end{bmatrix}$$

- The **0**s, of course, are *fixed* parameters. If this model does not fit, some of these can be set free (if there are good reasons!)
- More constrained models can be fit by imposing equality constraints to test stricter parallel or τ-equivalent models

Multi-factor congeneric models

- The covariance matrix $\mathbf{\Sigma}$ of $\boldsymbol{x}$ is again

$$\mathbf{\Sigma} = \mathbf{\Lambda \Phi \Lambda}^{\mathsf{T}} + \mathbf{\Theta}$$

  where $\mathbf{\Phi}$ is the covariance matrix of the factors, $\boldsymbol{\xi}$, and $\mathbf{\Theta}$ is the covariance matrix of the errors, $\boldsymbol{\delta}$
- In congeneric models, errors usually assumed to be uncorrelated: $\mathbf{\Theta}$ = diagonal
- (Some CFA models can allow correlated errors.)
- Model identification: in addition to the t rule,
  - It is necessary to set the scale for the latent ξ variables
  - *Standardized* solution: Set the diagonal entries of $\mathbf{\Phi}$ to 1, so $\mathbf{\Phi}$ is a correlation matrix
  - *Reference variable* solution: Set the loading $\lambda_{ij} = 1$ for one variable i in each column j

Example: Ability and Aspiration

Calsyn & Kenny (1971) studied the relation of perceived ability and educational aspiration in 556 white eigth-grade students. Their measures were:

x_1: self-concept of ability
x_2: perceived parental evaluation
x_3: perceived teacher evaluation
x_4: perceived friend's evaluation
x_5: educational aspiration
x_6: college plans

- Their interest was primarily in estimating the correlation between "true (perceived) ability" and "true apsiration".
- There is also interest in determining which is the most reliable indicator of each latent variable.

The correlation matrix is shown below:

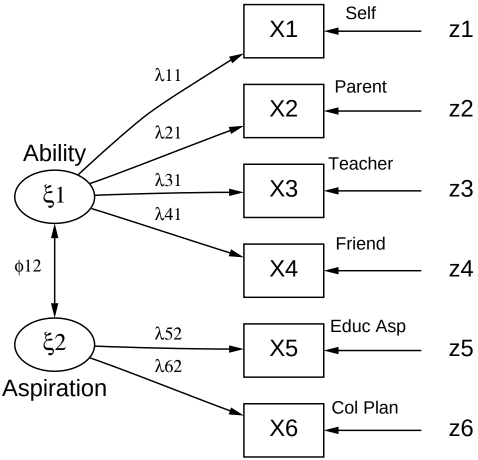
	S-C	Par	Tch	Frnd	Educ	Col
S-C Abil	1.00					
Par Eval	0.73	1.00				
Tch Eval	0.70	0.68	1.00			
FrndEval	0.58	0.61	0.57	1.00		
Educ Asp	0.46	0.43	0.40	0.37	1.00	
Col Plan	0.56	0.52	0.48	0.41	0.72	1.00
	x1	x2	x3	x4	x5	x6

The model to be tested is that

- x_1-x_4 measure only the latent "ability" factor and
- x_5-x_6 measure only the "aspiration" factor.
- i.e., two congeneric factors
- If so, are the two factors correlated?
- i.e., what is the true correlation ϕ_{12} between the latent factors?

Specifying the model

The model can be shown as a path diagram:



Specifying the model

This can be cast as the congeneric CFA model:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ \lambda_{41} & 0 \\ 0 & \lambda_{52} \\ 0 & \lambda_{62} \end{bmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \end{pmatrix}$$

If this model fits, the questions of interest can be answered in terms of the estimated parameters of the model:

- Correlation of latent variables: The estimated value of $\phi_{12} = r(\xi_1, \xi_2)$.
- Reliabilities of indicators: The communality, e.g., $h_i^2 = \lambda_{i1}^2$ is the estimated reliability of each measure.

The solution (found with LISREL and PROC CALIS) has an acceptable fit:

$$\chi^2 = 9.26 \quad \text{df} = 8 \quad (p = 0.321)$$

The estimated parameters (standardized solution) are:

	LAMBDA X		Communality	Uniqueness
	Ability	Aspiratn		
S-C Abil	0.863	0	0.745	0.255
Par Eval	0.849	0	0.721	0.279
Tch Eval	0.805	0	0.648	0.352
FrndEval	0.695	0	0.483	0.517
Educ Asp	0	0.775	0.601	0.399
Col Plan	0	0.929	0.863	0.137

Thus,

- Self-Concept of Ability is the most reliable measure of ξ_1, and College Plans is the most reliable measure of ξ_2.
- The correlation between the latent variables is $\phi_{12} = .67$. Note that this is higher than any of the individual between-set correlations.

Using PROC CALIS

For SAS, a correlation matrix can be input as follows:

```
data calken(TYPE=CORR);
   _TYPE_ = 'CORR'; input _NAME_ $ V1-V6;      % $
   label V1='Self-concept of ability'
         V2='Perceived parental evaluation'
         V3='Perceived teacher evaluation'
         V4='Perceived friends evaluation'
         V5='Educational aspiration'
         V6='College plans';
   datalines;
V1     1.      .       .       .       .       .
V2      .73   1.       .       .       .       .
V3      .70    .68    1.       .       .       .
V4      .58    .61     .57    1.       .       .
V5      .46    .43     .40     .37    1.       .
V6      .56    .52     .48     .41     .72    1.
  ;
```

Using PROC CALIS

The CFA model can be specified in several ways:

- With the FACTOR statement, specify names for the free parameters in Λ (MATRIX _F_) and Φ(MATRIX _P_)

```
proc calis data=calken method=max edf=555 short mod;
 FACTOR n=2;
 MATRIX _F_                                /* loadings */
          [ ,1] = lam1-lam4 ,              /* factor 1 */
          [ ,2] = 4 * 0 lam5 lam6 ;        /* factor 2 */
 MATRIX _P_
          [1,1] = 2 * 1. ,
          [1,2] = COR;     /* factor correlation */
 run;
```

Using PROC CALIS

- With the LINEQS statement, specify linear equations for the observed variables, using `F1`, `F2`, ... for common factors and `E1`, `E2`, ... for unique factors.
- STD statement specifies variances of the factors and errors
- COV statement specifies covariances

```
proc calis data=calken method=max edf=555;
 LINEQS
      V1 = lam1 F1            + E1 ,
      V2 = lam2 F1            + E2 ,
      V3 = lam3 F1            + E3 ,
      V4 = lam4 F1            + E4 ,
      V5 =          lam5 F2 + E5 ,
      V6 =          lam6 F2 + E6 ;
 STD
      E1-E6 = EPS: ,
      F1-F2 = 2 * 1. ;
 COV
      F1 F2 = COR ;
 run;
```

Using `cfa()` in the sem package

In addition to `specifyEquations()`, in the sem package, CFA models are even easier to specify using the `cfa()` function.

```
library(sem)
mod.calken <- cfa()
   F1: v1, v2, v3, v4
   F2: v5, v6

fit.calken <- sem(mod.calken, R.calken, N=556)
```

- Options allow you to specify reference indicators, and to specify covariances among the factors, allowing the factors to be correlated or uncorrelated.
- By default, all factors in CFA models are allowed to be correlated, simplifying model specification.
- sem includes `edit()` and `update()` functions, allowing you to delete, add, replace, fix, or free a path or parameter in a `semmod` object.

Example: Speeded and Non-speeded tests

If the measures are cross-classified in two or more ways, it is possible to test equivalence at the level of each way of classification.
Lord (1956) examined the correlations among 15 tests of three types:

- Vocabulary, Figural Intersections, and Arithmetic Reasoning.
- Each test given in two versions: Unspeeded (liberal time limits) and Speeded.

The goal was to identify factors of performance on speeded tests:

- Is speed on cognitive tests a unitary trait?
- If there are several type of speed factors, how are they correlated?
- How highly correlated are speed and power factors on the same test?

Example: Speeded and Non-speeded tests

Hypothesized factor patterns ($\mathbf{\Lambda}$):

(1) 3 congeneric sets

$$\mathbf{\Lambda}_{15\times3} = \begin{array}{c} \begin{matrix} V & I & R \end{matrix} \\ \begin{bmatrix} \beta_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \beta_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \beta_3 \end{bmatrix} \end{array}$$

(2) 3 congeneric sets + speed factor

$$\mathbf{\Lambda}_{15\times4} = \begin{array}{c} \begin{matrix} V & I & R & Sp \end{matrix} \\ \begin{bmatrix} x & & & \\ x & & & \\ x & & & x \\ x & & & x \\ x & & & x \\ & x & & \\ & x & & \\ & x & & x \\ & x & & x \\ & x & & x \\ & & x & \\ & & x & \\ & & x & x \\ & & x & x \\ & & x & x \end{bmatrix} \end{array}$$

Example: Speeded and Non-speeded tests

Hypothesized factor patterns ($\boldsymbol{\Lambda}$): Separate unspeeded and speeded factors

$$
\boldsymbol{\Lambda}_{15\times 6} =
\begin{array}{c}
\begin{array}{cccccc} V & I & R & V & I & R \end{array} \\
\left[\begin{array}{cccccc}
x & & & & & \\
x & & & & & \\
 & & & x & & \\
 & & & x & & \\
 & & & x & & \\
 & x & & & & \\
 & x & & & & \\
 & & & & x & \\
 & & & & x & \\
 & & & & x & \\
 & & x & & & \\
 & & x & & & \\
 & & & & & x \\
 & & & & & x \\
 & & & & & x
\end{array}\right] \\
\textit{Unspeeded} \qquad \textit{Speeded}
\end{array}
$$

Models:

- (3) parallel: equal λ & θ for each factor
- (4) τ-equivalent: equal λ in each col
- (5) congeneric: no equality constraints
- (6) six factors: 3 content, 3 speed

Results:

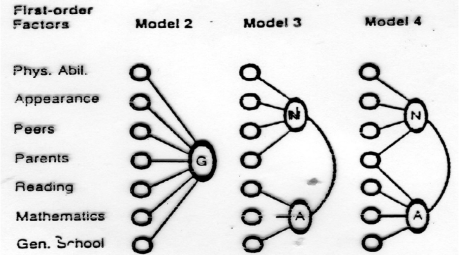
Hypothesis	Parameters	df	χ^2	$\Delta\chi^2$ (df)
1: 3 congeneric sets	33	87	264.35	
2: 3 sets + speed factor	42	78	140.50	123.85 (9)
3: 6 sets, parallel	27	93	210.10	
4: 6 sets, τ-equiv.	36	84	138.72	71.45 (9)
5: 6 sets, congeneric	45	75	120.57	18.15 (9)
6: 6 factors	45	75	108.37	12.20 (0)

Notes:

- Significant improvement from (1) to (2) $\rightarrow$ speeded tests measure something the unspeeded tests do not.
- χ^2 for (2) still large $\rightarrow$ perhaps there are different kinds of speed factors.
- Big improvement from (3) to (4) $\rightarrow$ not parallel

Higher-order factor analysis

- In EFA & CFA, we often have a model that allows the factors to be correlated ($\Phi \neq \boldsymbol{I}$)
- If there are more than a few factors, it sometimes makes sense to consider a 2nd-order model, that describes the correlations among the 1st-order factors.
- In EFA, this was done simply by doing another factor analysis of the estimated factor correlations $\widehat{\Phi}$ from the 1st-order analysis (after an oblique rotation)
- The second stage of development of CFA models was to combine these steps into a single model, and allow different hypotheses to be compared.



Second-order factor analysis: ACOVS model

- Start with a first-order CFA model for the observed variables, $\boldsymbol{y}$ with factors $\boldsymbol{\eta}$

$$\boldsymbol{y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\epsilon}$$

- Now, consider a 2^{nd}-order model for the correlations among the factors $\boldsymbol{\eta}$

$$\boldsymbol{\eta} = \boldsymbol{\Gamma} \boldsymbol{\xi} + \boldsymbol{\zeta}$$

- Combining these equations, we get

$$\boldsymbol{y} = \boldsymbol{\Lambda}_y (\boldsymbol{\Gamma} \boldsymbol{\xi} + \boldsymbol{\zeta}) + \boldsymbol{\epsilon}$$

- This is called the ACOVS model, for “analysis of covariance structures” Jöreskog (1970, 1974)

Second-order factor analysis: ACOVS model

This gives the following model for the covariance matrix $\boldsymbol{\Sigma}$:

$$\begin{aligned}\boldsymbol{\Sigma} &= \boldsymbol{\Lambda}_y(\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Gamma}^\mathsf{T} + \boldsymbol{\Psi})\boldsymbol{\Lambda}_y^\mathsf{T} + \boldsymbol{\Theta}_\epsilon \\ &= \boldsymbol{\Lambda}_y\boldsymbol{\Omega}\boldsymbol{\Lambda}_y^\mathsf{T} + \boldsymbol{\Theta}_\epsilon\end{aligned}$$

where:

- $\boldsymbol{\Lambda}_{y(p\times k)}$ = loadings of observed variables on k 1^{st}-order factors.
- $\boldsymbol{\Omega}_{(k\times k)}$ = correlations among 1^{st}-order factors.
- $\boldsymbol{\Theta}_{(p\times p)}$ = diagonal matrix of unique variances of 1^{st}-order factors.
- $\boldsymbol{\Gamma}_{(k\times r)}$ = loadings of 1^{st}-order factors on r second-order factors.
- $\boldsymbol{\Phi}_{(r\times r)}$ = correlations among 2^{nd}-order factors.
- $\boldsymbol{\Psi}$ = diagonal matrix of unique variances of 2^{nd}-order factors.

The model is thus a nesting of a 2^{nd}-order model for $\boldsymbol{\Gamma}$ within the 1^{st}-order model for $\boldsymbol{\Lambda}_y$.

Example: 2nd Order Analysis of Self-Concept Scales

A theoretical model of self-concept by Shavelson & Bolus (1976) describes facets of an individual's self-concept and presents a hierarchical model of how those facets are arranged.
To test this theory, Marsh & Hocevar (1985) analyzed measures of self-concept obtained from 251 fifth grade children with a Self-Description Questionnaire (SDQ). 28 subscales (consisting of two items each) of the SDQ were determined to tap four non-academic and three academic facets of self-concept:

- physical ability
- physical appearance
- relations with peers
- relations with parents
- reading
- mathematics
- general school

Example: 2nd Order Analysis of Self-Concept Scales

The subscales of the SDQ were determined by a first-order exploratory factor analysis. A second-order analysis was carried out examining the correlations among the first-order factors to examine predictions from the Shavelson model(s).

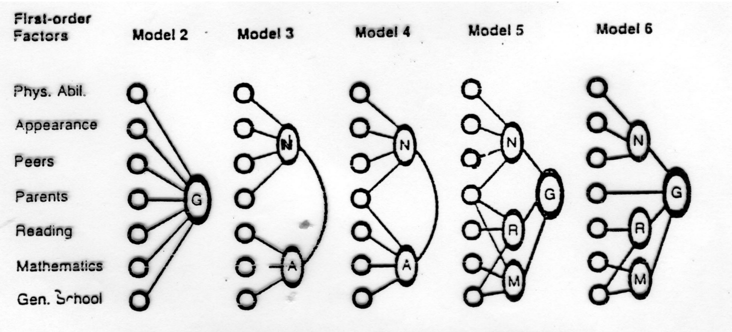


Figure 1. Higher order factor structures for Models 2 to 6.

sem package: Second-order CFA, Thurstone data

Data on 9 ability variables:

```
R.thur <- readMoments(diag=FALSE, names=c(
  'Sentences', 'Vocabulary', 'Sent.Completion',     # verbal
  'First.Letters', '4.Letter.Words','Suffixes',     # fluency
  'Letter.Series','Pedigrees',   'Letter.Group'))   # reasoning
    .828
    .776   .779
    .439   .493    .46
    .432   .464    .425   .674
    .447   .489    .443   .59    .541
    .447   .432    .401   .381   .402   .288
    .541   .537    .534   .35    .367   .32   .555
    .38    .358    .359   .424   .446   .325  .598   .452
```

Thurstone & Thurstone (1941) considered these to measure three factors:

- Verbal Comprehension,
- Word Fluency,
- Reasoning

sem package: Second-order CFA, Thurstone data

Using the `specifyEquations()` syntax:

```
mod.thur.eq <- specifyEquations()
    Sentences         = lam11*F1
    Vocabulary        = lam21*F1
    Sent.Completion   = lam31*F1
    First.Letters     =          lam42*F2
    4.Letter.Words    =          lam52*F2
    Suffixes          =          lam62*F2
    Letter.Series     =                   lam73*F3
    Pedigrees         =                   lam83*F3
    Letter.Group      =                   lam93*F3
    F1 = gam1*F4        # factor correlations
    F2 = gam2*F4
    F3 = gam3*F4
    V(F1) = 1           # factor variances
    V(F2) = 1
    V(F3) = 1
    V(F4) = 1
```

Each line gives a regression equation or the specification of a factor variance (V) or covariance (C)

Fit the model using `sem()`:

```
(fit.thur <- sem(mod.thur.eq, R.thur, 213))

 Model Chisquare =  38.2   Df =  24

 lam11  lam21  lam31  lam41  lam52  lam62  lam73  lam83  lam93   g
0.5151 0.5203 0.4874 0.5211 0.4971 0.4381 0.4524 0.4173 0.4076 1.4
  gam2   gam3    th1    th2    th3    th4    th5    th6    th7
1.2538 1.4066 0.1815 0.1649 0.2671 0.3015 0.3645 0.5064 0.3903 0.4
   th9
0.5051
```

More detailed output is provided by `summary()`:

```
summary(sem.thur)

 Model Chisquare =  38.196   Df =  24 Pr(>Chisq) = 0.033101
 Chisquare (null model) =  1101.9   Df =  36
 Goodness-of-fit index =  0.95957
 Adjusted goodness-of-fit index =  0.9242
 RMSEA index =  0.052822   90% CI: (0.015262, 0.083067)
 Bentler-Bonnett NFI =  0.96534
 Tucker-Lewis NNFI =  0.98002
 Bentler CFI =  0.98668
 SRMR =  0.043595
 BIC =  -90.475
 ...
```

sem package: Second-order CFA, Thurstone data

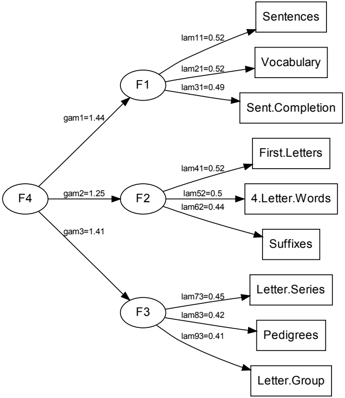
- The same model can be specified using `cfa()`, designed specially for confirmatory factor models
- Each line lists the variables that load on a given factor.

```
mod.thur.cfa <- cfa(reference.indicators=FALSE,
                    covs=c("F1", "F2", "F3", "F4"))
   F1: Sentences, Vocabulary, Sent.Completion
   F2: First.Letters, 4.Letter.Words, Suffixes
   F3: Letter.Series, Pedigrees, Letter.Group
   F4: F1, F2, F3

sem.thur.cfa <- sem(mod.thur.cfa, R.thur, 213)
```

Path diagram:

```
pathDiagram(sem.thur, file="sem-thurstone", edge.labels="both")
Running  dot -Tpdf -o sem-thurstone.pdf  sem-thurstone.dot
```

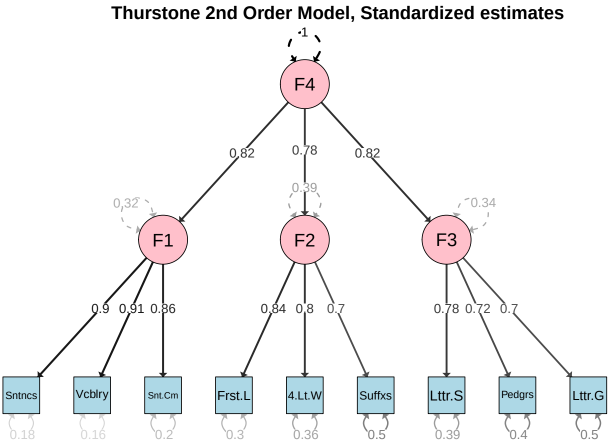


sem package: Other features

- With raw data input, `sem` provides `robust` estimates of standard errors and robust tests
- Can accommodate missing data, via full-information maximum likelihood (FIML)
- **`miSem()`** generates multiple imputations of missing data using the `mi` package
- **`bootSem()`** provides nonparametric bootstrap estimates by independent random sampling
- A given model can be easily modified via **`edit()`** and **`update()`** methods
- Multiple-group analyses and tests of factorial invariance: **`multigroupModel()`**.
- Related: semPlot: lovely, flexible, pub. quality path diagrams

Path diagram from semPlot

```
library(semPlot)
semPaths(sem.thur, what="std", color=list(man="lightblue", lat="pink"),
         nCharNodes=6, sizeMan=6, edge.color="black")
title("Thurstone 2nd Order Model, Standardized estimates", cex=1.5)
```



Factorial Invariance

Multi-sample analyses:

- When a set of measures have been obtained from samples from several populations, we often wish to study the similarities in factor structure across groups.
- The CFA/SEM model allows any parameter to be assigned an arbitrary fixed value, or constrained to be equal to some other parameter. Constraints across groups provide the way to test these models.
- We can test any degree of invariance from totally separate factor structures to completely invariant ones.
- **Model**
  Let $\mathbf{x}_g$ be the vector of tests administered to group g, $g = 1, 2, \ldots, m$, and assume that a factor analysis model holds in each population with some number of common factors, k_g.

$$\Sigma_g = \Lambda_g \Phi_g \Lambda_g^{\mathsf{T}} + \Psi_g$$

Factorial Invariance: Examples

- Arguably among the most important recent development in personality psychology is the idea that individual differences in personality characteristics is organized into five main trait domains: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness
    - One widely used instrument is the 60-item NEO-Five factor inventory (Costa & McCrae, 1992), developed and analyzed for a North American, English-speaking population
    - To what extent does the same factor structure apply across gender?
    - To what extent does the same factor structure applies in other cultural and language goups?
- The emerging field of cross-cultural psychology offers many similar examples.

Factorial Invariance: Hypotheses

We can examine a number of different hypotheses about how “similar” the covariance structure is across groups.

Hypotheses

- Can we simply pool the data over groups?
- If not, can we say that the same number of factors apply in all groups?
- If so, are the factor loadings equal over groups?
- What about factor correlations and unique variances?

Software

- LISREL, AMOS, and M Plus all provide convenient ways to do multi-sample analysis.
- PROC `CALIS` in SAS 9.3 does too.
- In R, the lavaan package provides multi-sample analysis and the **`measurementInvariance()`** function. The sem package includes a **`multigroupModel()`** for such models

Equality of Covariance Matrices

$$H_{=\Sigma} : \Sigma_1 = \Sigma_2 = \cdots = \Sigma_m$$

- If this hypothesis is tenable, there is no need to analyse each group separately or test further for differences among them: Simply pool all the data, and do one analysis!
- If we reject $H_{=\Sigma}$, we may wish to test a less restrictive hypothesis that posits some form of invariance.
- The test statistic for $H_{=\Sigma}$ is Box's test,

$$\chi^2_{=\Sigma} = n \log |S| - \sum_{g=1}^{m} n_g \log |S_g|$$

which is distributed approx. as χ^2 with $d_{=\Sigma} = (m-1)p(p-1)/2$ df.

(This test can be carried out in SAS with PROC `DISCRIM` using the `POOL=TEST` option)

- Same number of factors (Configural invariance)
  The least restrictive form of "invariance" is simply that the number of factors is the same in each population:

$$H_k : k_1 = k_2 = \cdots = k_m = \text{ a specified value, } k$$

  - This can be tested by doing an unrestricted factor analysis for k factors on each group separately, and summing the χ^2's and degrees of freedom,

$$\chi^2_k = \sum_g^m \chi^2_k(g) \qquad d_k = m \times [(p-k)^2 - (p+k)]/2$$

  - If this hypothesis is rejected, there is no sense in testing more restrictive models

- Same factor pattern (Weak invariance)
  If the hypothesis of a common number of factors is tenable, one may proceed to test the hypothesis of an invariant factor pattern:

$$H_\Lambda : \Lambda_1 = \Lambda_2 = \cdots = \Lambda_m$$

  - The common factor pattern Λ may be either completely unspecified, or be specified to have zeros in certain positions.
  - To obtain a χ^2 for this hypothesis, estimate Λ (common to all groups), plus $\Phi_1, \Phi_2, \ldots, \Phi_m$, and $\Psi_1, \Psi_2, \ldots, \Psi_m$, yielding a minimum value of the function, F. Then, $\chi^2_\Lambda = 2 \times F_{min}$.
  - To test the hypothesis H_Λ, given that the number of factors is the *same* in all groups, use

$$\chi^2_{\Lambda|k} = \chi^2_\Lambda - \chi^2_k \text{ with } d_{\Lambda|k} = d_\Lambda - d_k \text{ degrees of freedom}$$

- Same factor pattern and unique variances (Strong invariance)
  A stronger hypothesis is that the unique variances, as well as the factor pattern, are invariant across groups:

$$H_{\Lambda\Psi} : \begin{cases} \Lambda_1 = \Lambda_2 = \cdots = \Lambda_m \\ \Psi_1 = \Psi_2 = \cdots = \Psi_m \end{cases}$$

- Same factor pattern, means and unique variances (Strict invariance)
  The strongest hypothesis is that the factor means are also equal across groups as well as the factor patterns and unique variances:

$$H_{\Lambda\Psi_\mu} : \begin{cases} \Lambda_1 = \Lambda_2 = \cdots = \Lambda_m \\ \Psi_1 = \Psi_2 = \cdots = \Psi_m \\ \mu_1 = \mu_2 = \cdots = \mu_m \end{cases}$$

Example: Academic and Non-Academic Boys

Sorbom (1976) analyzed STEP tests of reading and writing given in grade 5 and grade 7 to samples of boys in Academic and Non-Academic programs.

Data

	Academic ($N = 373$)				Non-Acad ($N = 249$)			
Read Gr5	281.35				174.48			
Writ Gr5	184.22	182.82			134.47	161.87		
Read Gr7	216.74	171.70	283.29		129.84	118.84	228.45	
Writ Gr7	198.38	153.20	208.84	246.07	102.19	97.77	136.06	180.46

Hypotheses

The following hypotheses were tested:

Hypothesis	Model specifications
A. $H_{=\Sigma}: \Sigma_1 = \Sigma_2$	$\begin{cases} \mathbf{\Lambda}_1 = \mathbf{\Lambda}_2 = \boldsymbol{I}_{(4\times4)} \\ \mathbf{\Psi}_1 = \mathbf{\Psi}_2 = \mathbf{0}_{(4\times4)} \\ \mathbf{\Phi}_1 = \mathbf{\Phi}_2 \text{ constrained, free} \end{cases}$
B. $H_{k=2}$: Σ_1, Σ_2 both fit with $k = 2$ correlated factors	$\begin{cases} \mathbf{\Lambda}_1 = \mathbf{\Lambda}_2 = \begin{bmatrix} x & 0 \\ x & 0 \\ 0 & x \\ 0 & x \end{bmatrix} \\ \mathbf{\Phi}_1, \mathbf{\Phi}_2, \mathbf{\Psi}_1, \mathbf{\Psi}_2 \text{ free} \end{cases}$
C. $H_{\Lambda} : H_{k=2}$ & $\mathbf{\Lambda}_1 = \mathbf{\Lambda}_2$	$\mathbf{\Lambda}_1 = \mathbf{\Lambda}_2$ (constrained)
D. $H_{\Lambda,\Theta} : H_{\Lambda}$ & $\mathbf{\Psi}_1 = \mathbf{\Psi}_2$	$\begin{cases} \mathbf{\Psi}_1 = \mathbf{\Psi}_2 \text{ (constrained)} \\ \mathbf{\Lambda}_1 = \mathbf{\Lambda}_2 \end{cases}$
E. $H_{\Lambda,\Theta,\Phi} : H_{\Lambda,\Theta}$ & $\mathbf{\Phi}_1 = \mathbf{\Phi}_2$	$\begin{cases} \mathbf{\Phi}_1 = \mathbf{\Phi}_2 \text{ (constrained)} \\ \mathbf{\Psi}_1 = \mathbf{\Psi}_2 \\ \mathbf{\Lambda}_1 = \mathbf{\Lambda}_2 \end{cases}$

Analysis

The analysis was carried out with both LISREL and AMOS. AMOS is particularly convenient for multi-sample analysis, and for testing a series of nested hypotheses.

Summary of Hypothesis Tests for Factorial Invariance

Hypothesis	Overall fit				Group A		Group N-A	
	χ^2	df	prob	AIC	GFI	RMSR	GFI	RMSR
A: $H_{=\Sigma}$	38.08	10	.000	55.10	.982	28.17	.958	42.26
B: $H_{k=2}$	1.52	2	.468	37.52	.999	0.73	.999	0.78
C: H_{Λ}	8.77	4	.067	40.65	.996	5.17	.989	7.83
D: $H_{\Lambda,\Psi}$	21.55	8	.006	44.55	.990	7.33	.975	11.06
E: $H_{\Lambda,\Psi,\Phi}$	38.22	11	.000	53.36	.981	28.18	.958	42.26

- The hypothesis of equal factor loadings (H_{Λ}) in both samples is tenable.
- Unique variances appear to differ in the two samples.
- The factor correlation (ϕ_{12}) appears to be greater in the Academic sample than in the non-Academic sample.

lavaan package: Factorial invariance tests

Data

Data for Academic and Non-academic boys:

```
library(sem)
Sorbom.acad <- read.moments(diag=TRUE,
names=c('Read.Gr5', 'Writ.Gr5', 'Read.Gr7', 'Writ.Gr7'))
281.349
184.219 182.821
216.739 171.699 283.289
198.376 153.201 208.837 246.069

Sorbom.nonacad <- read.moments(diag=TRUE,
names=c('Read.Gr5', 'Writ.Gr5', 'Read.Gr7', 'Writ.Gr7'))
174.485
134.468 161.869
129.840 118.836 228.449
102.194  97.767 136.058 180.460

# make the two matrices into a list
Sorbom <- list(acad=Sorbom.acad, nonacad=Sorbom.nonacad)
```

lavaan package: Factorial invariance tests I

Model

Specify `lavaan` model for 2 correlated, non-overlapping factors:

```
library(lavaan)
Sorbom.model <-
'G5 =~ Read.Gr5 + Writ.Gr5
 G7 =~ Read.Gr7 + Writ.Gr7 '
```

Run a cfa model (testing k=2 for each group):

```
(Sorbom.cfa <- cfa(Sorbom.model, sample.cov=Sorbom, sample.nobs=c(373,249)
```

```
Lavaan (0.4-7) converged normally after 240 iterations
  Number of observations per group
  acad                                             373
  nonacad                                          249

  Estimator                                         ML
  Minimum Function Chi-square                    1.525
  Degrees of freedom                                 2
  P-value                                        0.467

Chi-square for each group:
  acad                                           0.863
  nonacad                                        0.662
```

Tests of measurement invariance I

Test all models of measurement invariance:

```
library(semTools)
measurementInvariance(Sorbom.model, sample.cov=Sorbom,
    sample.nobs=c(373,249))
```

```
Measurement invariance tests:

Model 1: configural invariance:
    chisq        df    pvalue       cfi     rmsea        bic
    1.525     2.000     0.467     1.000     0.000 18788.554

Model 2: weak invariance (equal loadings):
    chisq        df    pvalue       cfi     rmsea        bic
    8.806     4.000     0.066     0.997     0.062 18782.970

[Model 1 versus model 2]
  delta.chisq      delta.df delta.p.value     delta.cfi
        7.282         2.000         0.026         0.003
```

Tests of measurement invariance II

```
Model 3: strong invariance (equal loadings + intercepts):
    chisq         df     pvalue        cfi      rmsea        bic
    8.806      6.000      0.185      0.998      0.039 18821.567

[Model 1 versus model 3]
  delta.chisq       delta.df delta.p.value     delta.cfi
        7.282          4.000         0.122         0.002

[Model 2 versus model 3]
  delta.chisq       delta.df delta.p.value     delta.cfi
        0.000          2.000         1.000        -0.001
 ...
```

A fourth model also tests equality of means, but means are not available for this example.

Summary

- ***measurement error*** reduces precision, but worse— introduces bias
- CFA & SEM use latent variables in a measurement model to allow for this

$$\boldsymbol{x} = \boldsymbol{\Lambda} \xi + \delta \quad \Longrightarrow \boldsymbol{\Sigma} = \boldsymbol{\Lambda} \boldsymbol{\Phi} \boldsymbol{\Lambda}^{\mathsf{T}} + \boldsymbol{\Theta}$$

- One-factor models allow for testing various forms of “equivalence” within the SEM framework
  - An essential idea in CFA is allowing for free and fixed parameters and equality contraints
  - These ideas extend directly to more complex models, with multiple factors of possibly different types
- Higher-order CFA models take this a step further, allowing a factor structure for the 1^{st}-order factors
- Multiple-group models allow for testing a variety of measurement invariance models