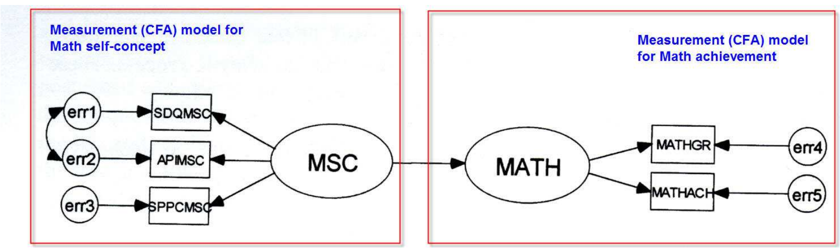
# Exploratory and Confirmatory Factor Analysis

Michael Friendly

Feb. 25, Mar. 3, 10, 2008
SCS Short Course
http://www.math.yorku.ca/SCS/Courses/factor/




## Part 3: CFA Outline

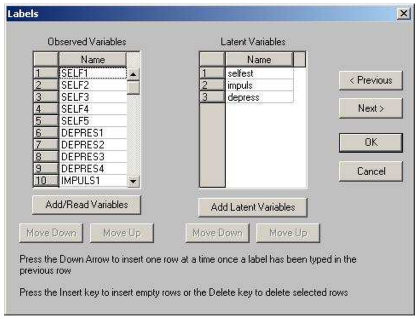
1. Indeterminacy of the Common Factor Model
2. Restricted maximum likelihood FA
   - Example: Ability and Aspiration
   - Using PROC CALIS
3. Higher-order factor analysis: ACOVS model
4. LISREL model: CFA and SEM
   - Testing equivalence of measures with CFA
   - Several Sets of Congeneric Tests
   - Example: Lord's data
   - Example: Speeded & unspeeded tests
   - Simplex models for ordered variables
5. Factorial invariance
   - Example: Academic and Non-academic boys
6. Identifiability in CFA models
7. Power and sample size for EFA and CFA

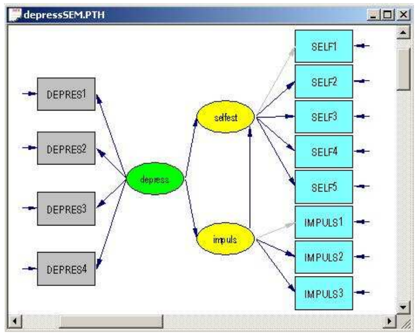
## Confirmatory Factor Analysis

- Development of CFA models
  - Restricted maximum likelihood factor analysis (RIMLFA model)
  - Analysis of covariance structures (ACOVS model)
  - Structural equations, mean structures (LISREL model)
- Applications of CFA
  - Higher-order factor analysis
  - Test theory models of "equivalence"
  - Sets of congeneric tests
  - Inter-rater reliability
  - Multi-trait, Multi-method data
  - Simplex models for ordered latent variables
  - Factorial invariance
  - Power and sample size for EFA and CFA models

## Prelude: CFA software

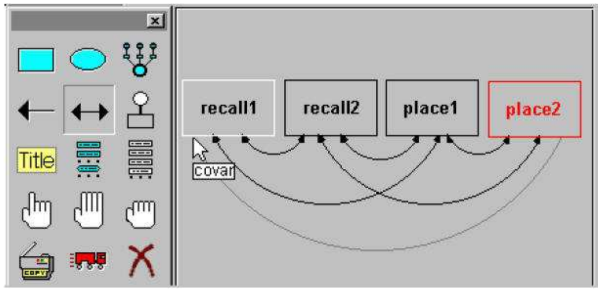
- LISREL (http://www.ssicentral.com/)
  - Originally designed as stand-alone program with matrix syntax
  - LISREL 8.5+ for Windows/Mac: Includes
    - interactive, menu-driven version;
    - PRELIS (pre-processing, correlations and models for categorical variables);
    - SIMPLIS (simplified, linear equation syntax)
    - path diagrams from the fitted model



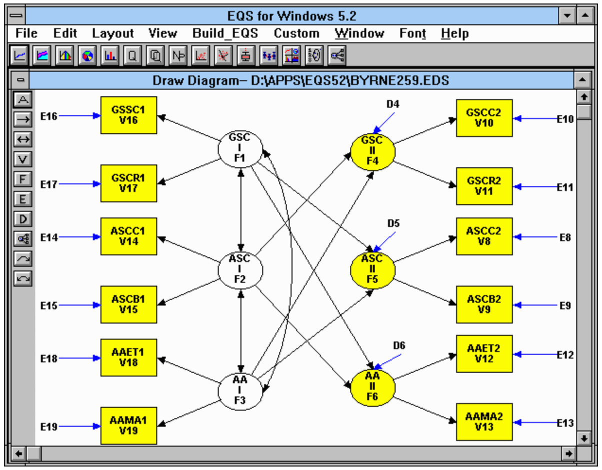
## Prelude: CFA software

- Amos (http://www.spss.com/amos/): Linear equation syntax + path diagram model description
  - import data from SPSS, Excel, etc; works well with SPSS
  - Create the model by drawing a path diagram
  - simple facilities for multi-sample analyses
  - nice comparative displays of multiple models




## Prelude: CFA software

- EQS (http://www.mvsoft.com/): Uses linear equation syntax
  - spreadsheet data editor; cut/paste from other apps
  - now has Diagrammer à la Amos
  - polyserial and polychoric correlations for categorical data

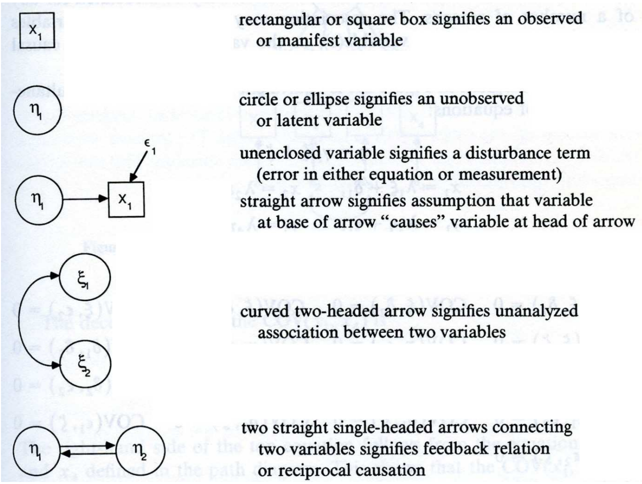



## Prelude: CFA software

- SAS: PROC CALIS
  - MATRIX (à la LISREL), LINEQS (à la EQS), RAM, ... syntax
  - Does not handle multi-sample analyses
- SAS macros `http://www.math.yorku.ca/SCS/sasmac/`:
  - `calisgfi` macro: more readable display of PROC CALIS fit statistics
  - `caliscmp` macro: compare model fits from PROC CALIS à la Amos
  - `csmpower` macro: power estimation for covariance structure models
  - `ram2dot` macro: path diagram from PROC CALIS RAM data set
  - `eqs2ram` macro: translate linear equations to RAM data set
- R: `sem` package (John Fox): uses RAM model syntax
  - path diagrams using `graphviz`

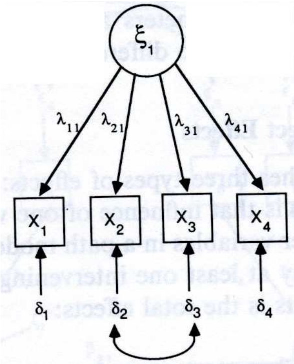
## Prelude: Path diagrams

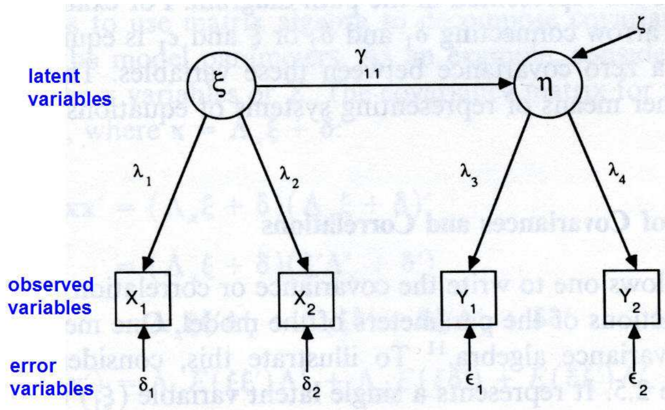
- Visual representation of a set of simultaneous equations for EFA, CFA, SEM models



# Prelude: Path diagrams

Examples:




CFA, 1-factor model SEM, two latent variables



# Indeterminacy of the Common Factor Model

- The general Factor Analysis model allows the factors to be correlated. Let $\mathbf{\Phi}_{k \times k}$ be the variance-covariance matrix of the common factors. Then the model is

$$\mathbf{\Sigma} = \mathbf{\Lambda} \mathbf{\Phi} \mathbf{\Lambda}^{\mathsf{T}} + \mathbf{\Psi} \quad (6)$$

- However, model (6) is not ***identified***, since it has more parameters than there are correlations (or variances and covariances).
- That is, any rotation of $\mathbf{\Lambda}$ by a non-singular transformation matrix, $\boldsymbol{T}_{k \times k}$ will fit equally well:

$$F(\mathbf{\Lambda}, \mathbf{\Phi}) = F(\mathbf{\Lambda}\boldsymbol{T},\ \boldsymbol{T}^{-1}\mathbf{\Phi}\boldsymbol{T}^{-1})$$

- The transformation matrix, $\boldsymbol{T}_{(k \times k)}$ corresponds to the fact that $k^2$ restrictions need to be imposed in $\mathbf{\Phi}$ and/or $\mathbf{\Lambda}$ to obtain a unique solution.
- Setting $\mathbf{\Phi} = I$ (uncorrelated factors) gives $k(k+1)/2$ restrictions; all methods of estimating factor impose an additional $k(k-1)/2$ restrictions on $\mathbf{\Lambda}$.



# Indeterminacy of the Common Factor Model

- Therefore, the number of effective unknown parameters is:

$$\overbrace{pk}^{\mathbf{\Lambda}} + \overbrace{k(k+1)/2}^{\mathbf{\Phi}} + \overbrace{p}^{\mathbf{\Psi}} - \overbrace{k^2}^{\boldsymbol{T}} = pk + p - k(k-1)/2$$

  so the number of ***degrees of freedom*** for the model is:

| | |
|---|---|
| Sample moments (***S***) | $p(p+1)/2$ |
| - Parameters estimated | $pk + p - k(k-1)/2$ |
| = Degrees of freedom | $[(p-k)^2 - (p+k)]/2$ |

- E.g., with $p = 6$ tests, $k = 3$ factors will always fit perfectly

| k | 1 | 2 | 3 |
|---|---|---|---|
| df | 9 | 4 | 0 |



# Restricted Maximum Likelihood FA

The essential ideas of CFA can be introduced most simply as follows:

- Jöreskog (1969) proposed that a factor hypothesis could be tested by imposing restrictions on the factor model, in the form of **fixed** elements in $\Lambda$ and $\Phi$ (usually 0).
- The maximum likelihood solution is then found for the remaining **free** parameters in $\Lambda$ and $\Phi$.
- The $\chi^2$ for the restricted solution provides a test of how well the hypothesized factor structure fits.

# Restricted Maximum Likelihood FA

For example, the pattern below specifies two non-overlapping oblique factors, where the *x*'s are the only free parameters.

$$\Lambda = \begin{bmatrix} x & 0 \\ x & 0 \\ x & 0 \\ 0 & x \\ 0 & x \\ 0 & x \end{bmatrix} \qquad \Phi = \begin{bmatrix} 1 & \\ x & 1 \end{bmatrix}$$

- A $k = 2$-factor EFA model would have all parameters free and $df = 15 - 11 = 4$ degrees of freedom.
- This model has only 7 free parameters and $df = 15 - 7 = 8$.
- If this restricted model fits (has a small $\chi^2/df$), it is strong evidence for two non-overlapping oblique factors.



# Restricted vs. Unrestricted solutions

**Unrestricted solution** Factor solutions with $m = k^2$ restrictions are mathematically equivalent:
- same communalities,
- same goodness of fit $\chi^2$.
- Any unrestricted solution can be rotated to any other.

**Restricted solution** Solutions with $m > k^2$ restrictions
- have different communalities,
- do not reflect the same common factor space, and
- cannot be rotated to one another.



# 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:

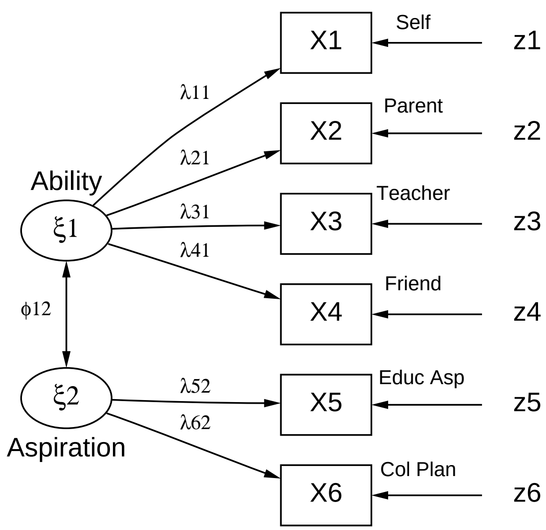
| | S-C | Par | Tch | Frnd | Educ | Col |
|---|---|---|---|---|---|---|
| S-C Abil | 1.00 | | | | | |
| Par Eval | 0.73 | 1.00 | | | | |
| Tch Eval | 0.70 | 0.68 | 1.00 | | | |
| FrndEval | 0.58 | 0.61 | 0.57 | 1.00 | | |
| Educ Asp | 0.46 | 0.43 | 0.40 | 0.37 | 1.00 | |
| Col Plan | 0.56 | 0.52 | 0.48 | 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.
- If so, are the two factors correlated?

# Specifying the model

The model can be shown as a path diagram:






# Specifying the model

This can be cast as the restricted factor analysis model:

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

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 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 $\xi_1$, and College Plans is the most reliable measure of $\xi_2$.
- 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

```
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 $\Lambda$ (MATRIX _F_) and $\Phi$(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

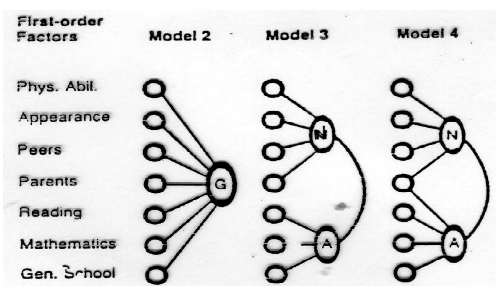
- With the LINEQS statement, specify linear equations for the observed variables, using F1, F2, ... for common factors and E1, E2, ... for unique factors.
- STD statement specifies variances of the factors
- COV statement specifies covariances

```
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;
```



# 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.






# Analysis of Covariance Structures (ACOVS)

Jöreskog (1970, 1974) proposed a generalization of the common factor model which allows for second-order factors.

$$\begin{aligned} \Sigma &= \boldsymbol{B}(\Lambda \Phi \Lambda^{\mathsf{T}} + \Psi^2)\boldsymbol{B}^{\mathsf{T}} + \Theta^2 \\ &= \boldsymbol{B}\Gamma\boldsymbol{B}^{\mathsf{T}} + \Theta^2 \end{aligned}$$

where:

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

# Analysis of Covariance Structures (ACOVS)

In applications of ACOVS, any parameters in $\boldsymbol{B}, \boldsymbol{\Lambda}, \boldsymbol{\Phi}, \boldsymbol{\Psi}$, or $\boldsymbol{\Theta}$ may be

- **free** to be estimated,
- **fixed constants** by hypothesis, or
- **constrained** to be equal to other parameters.

The maximum likelihood solution minimizes:

$$F(\boldsymbol{B}, \boldsymbol{\Lambda}, \boldsymbol{\Phi}, \boldsymbol{\Psi}, \boldsymbol{\Theta}) = \text{tr}(S\widehat{\boldsymbol{\Sigma}}^{-1}) + \log|\widehat{\boldsymbol{\Sigma}}| - \log|S| - p$$

with respect to the independent free parameters. At the minimum, $(N-1)F_{min} \sim \chi^2$ with degrees of freedom $= p(p+1)/2$ - (number of free parameters in model).



# 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).

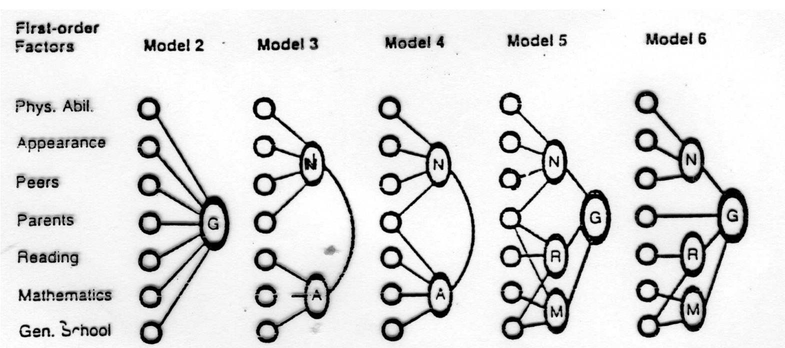



*Figure 1.* Higher order fac?r structures for Models 2 to 6.



# LISREL/SEM Model

- Jöreskog (1973) further generalized the ACOVS model to include structural equation models along with CFA.
- Two parts:

  Measurement model - How the latent variables are measured in terms of the observed variables; measurement properties (reliability, validity) of observed variables. [Traditional factor analysis models]

  Structural equation model - Specifies causal relations among observed and latent variables.
  - Endogenous variables - determined within the model ($y$s)
  - Exogenous variables - determined outside the model ($x$s)

Structural eqn. for latent variables

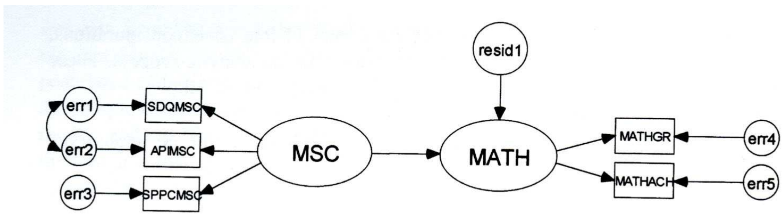
$$\boldsymbol{\eta} = \boldsymbol{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}$$

Measurement models for observed variables

$$\boldsymbol{x} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}$$
$$\boldsymbol{y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\epsilon}$$

# LISREL/SEM Model

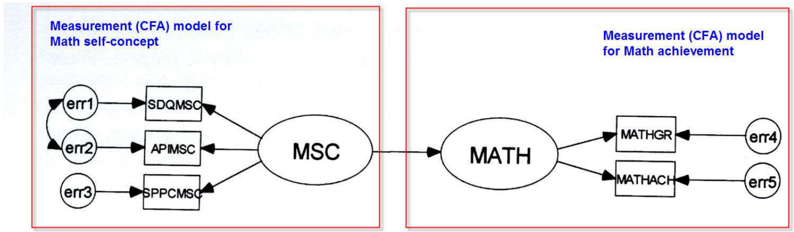
SEM model for measures of math self-concept and math achievement:



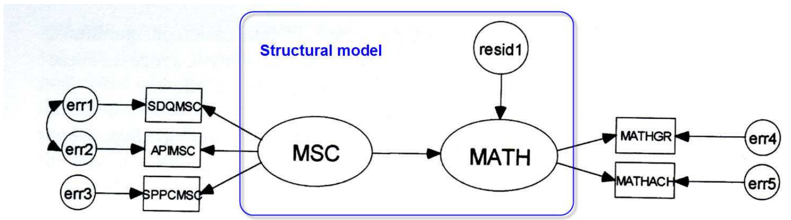



# LISREL/SEM Model

Measurement sub-models for ***x*** and ***y***




Structural model, relating $\xi$ to $\eta$





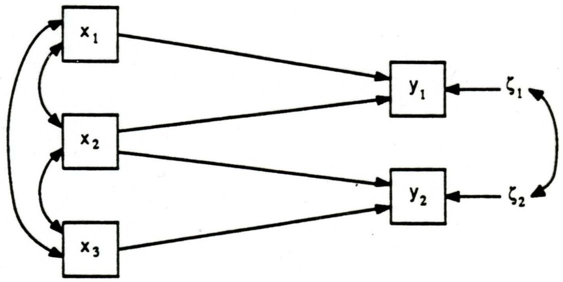

## LISREL submodels

- NX = # of observed eXdogenous $x$ variables
- NY = # of observed endogenous $y$ variables
- NKsi = # of latent eXdogenous $\xi$ variables
- NEta = # of latent endogenous $\eta$ variables
- ***Structural equations, GLM*** $[NY, NX > 0, NK = NE = 0]$

$$\boldsymbol{y} = \boldsymbol{B}\boldsymbol{y} + \boldsymbol{\Gamma}\boldsymbol{x} + \boldsymbol{\zeta}$$

Path analysis, structural equation causal model for directly observed variables. Ordinary regression models and GLM are the special case:

$$\boldsymbol{B} = 0 \Rightarrow \boldsymbol{y} = \boldsymbol{\Gamma}\boldsymbol{x} + \boldsymbol{\zeta}$$






- ***Factor analysis model*** $[NX, NK > 0, NY = NE = 0]$

$$\begin{aligned} \boldsymbol{x} &= \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta} \\ \Rightarrow \boldsymbol{\Sigma}_{xx} &= \boldsymbol{\Lambda}_x \boldsymbol{\Phi} \boldsymbol{\Lambda}_x^{\mathsf{T}} + \boldsymbol{\Theta}_\delta \end{aligned}$$

- ***Second-order factor analysis model*** $[NY, NE, NK > 0]$

$$\begin{aligned} \boldsymbol{\eta} &= \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \\ \boldsymbol{y} &= \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\epsilon} \\ \Rightarrow \boldsymbol{\Sigma}_{yy} &= \boldsymbol{\Lambda}_y (\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Gamma}^{\mathsf{T}} + \boldsymbol{\Theta}_\delta) \boldsymbol{\Lambda}_y^{\mathsf{T}} + \boldsymbol{\Theta}_\epsilon \end{aligned}$$

# Using LISREL: Ability & Aspiration

```
LISREL1: Ability & Aspiration in 8th grade students

! NOBS gives # of observations on which the correlation matrix is based.
! MATRIX=KMatrix --> a correlation matrix is to be analyzed.
DAta NInputvars=6 NOBS=556 MATRIX=KMATRIX
LAbels
   S_C_Abil Par_Eval Tch_Eval FrndEval Educ_Asp Col_Plan
KMatrix
    1.0
     .73 1.0
     .70  .68 1.0
     .58  .61  .57 1.0
     .46  .43  .40  .37 1.0
     .56  .52  .48  .41  .72 1.0
! MODEL  NKSI=2  --> specified two ksi variables (factors)
!    PHI=STANDARDIZED standardizes the factors to 0 mean and variance 1.
MOdel NX=6 NKSI=2 PHI=STANDARDIZED
```

- Each statement begins with a keyword (or 2-letter abbreviation)
- MOdel statement specifies NX, NY, NKsi, NEta & and matrix forms (symmetric, standardized, etc.)



# Using LISREL: Ability & Aspiration

```
! LKSI specifies names for the factors.
LKsi
 'Ability' 'Aspiratn'
! PATTERN LX specifies the lambda-x (factor loading) parameters
!   to be estimated: 0 indicates a 0 loading; 1 a parameter to
!   be estimated.
PAttern LX
   1 0
   1 0
   1 0
   1 0
   0 1
   0 1
! The FREE LX lines specify exactly the same things; either way will do.
FRee  LX(1 1) LX(2 1) LX(3 1) LX(4 1)
FRee  LX(5 2) LX(6 2)
PDiagram
! OUTPUT: SE requests standard errors; MIndices: "modification indices."
OUtput SE TValues VA MIndices FS TO
```

- PAttern LX or FRee specify free and fixed parameters
- PDiagram: path diagram
- OUtput: specifies output options



# 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.

$$\mathbf{\Sigma} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \end{bmatrix} + \begin{bmatrix} \theta_1^2 & & & \\ & \theta_2^2 & & \\ & & \theta_3^2 & \\ & & & \theta_4^2 \end{bmatrix}$$

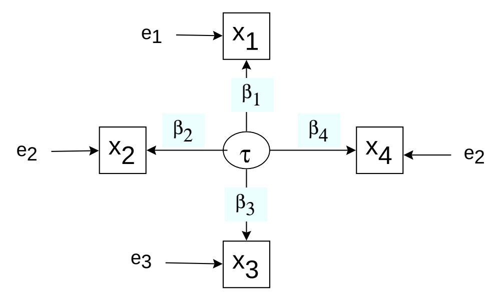


# Testing Equivalence of Measures with CFA

One-factor model:

$$\mathbf{\Sigma} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \end{bmatrix} + \begin{bmatrix} \theta_1^2 & & & \\ & \theta_2^2 & & \\ & & \theta_3^2 & \\ & & & \theta_4^2 \end{bmatrix}$$

Path diagram:



### 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 ($\beta_i^2$), but may differ in error variance ($\theta_i^2$). Like parallel tests, this requires tests of the same length & time limits. E.g., short forms cannot be $\tau$-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 ACOVS/LISREL by testing equality of the factor loadings ($\beta_i$) and unique variances ($\theta_i^2$).

$$\underbrace{\overbrace{\beta_1 = \beta_2 = \beta_3 = \beta_4}^{\tau \text{ equivalent}} \qquad \theta_1^2 = \theta_2^2 = \theta_3^2 = \theta_4^2}_{\text{Parallel}}$$



## 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, $\tau_x$, $\tau_y$, 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."

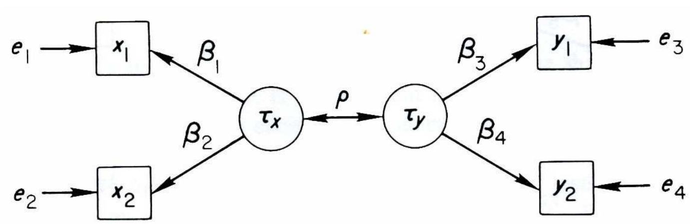


## Several Sets of Congeneric Tests

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

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

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






## Several Sets of Congeneric Tests

### Hypotheses

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

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

$H_1$ 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 $\times$ 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 $\chi^2$, 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.



# 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 | $\chi^2$ | 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 $\chi^2$ tests
- Model H2 is "best" by AIC



# Lord's data

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

| | Parallel | | Congeneric | |
|---|---|---|---|---|
| | $\chi^2$ | df | $\chi^2$ | 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 $\hat{\rho} = 0.889$.
- $\Rightarrow$ speeded and unspeeded vocab. tests do not measure *exactly* the same thing.
  SAS example:
  www.math.yorku.ca/SCS/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: $\beta_1 = \beta_2, \beta_3 = \beta_4$ ..., $\rho$=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 | |

# 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.

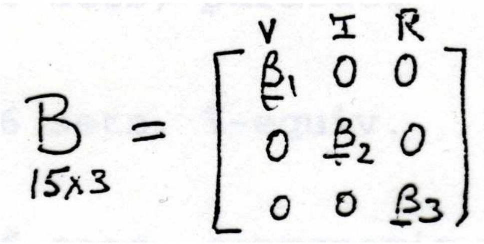
The goal was to identify factors of performance on speeded tests:

- 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?

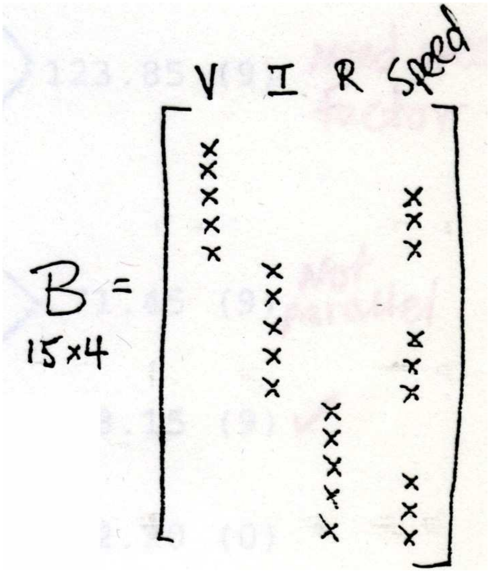


# Example: Speeded and Non-speeded tests

Hypothesized factor patterns (B):




(1) 3 congeneric sets

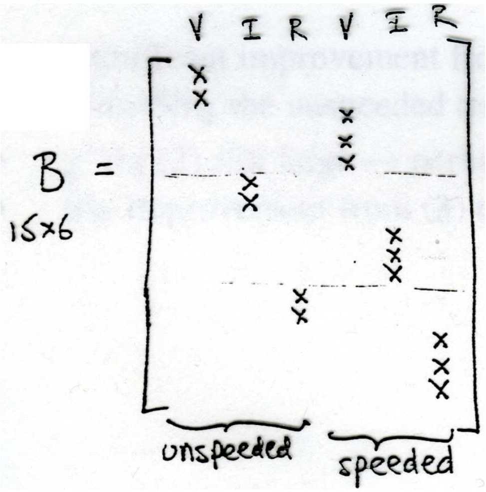



(2) 3 congeneric sets + speed factor



# Example: Speeded and Non-speeded tests

Hypothesized factor patterns (B):




(3)-(6) Six factors

- (3) parallel: equal $\beta$ & $\theta^2$ for each factor
- (4) $\tau$-equivalent: equal $\beta$ in each col
- (5) congeneric: no equality constraints
- (6) six factors: 3 content, 3 speed



Results:

| Hypothesis | Parameters | df | $\chi^2$ | $\Delta\chi^2$ (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, $\tau$-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.
- $\chi^2$ for (2) still large $\rightarrow$ perhaps there are different kinds of speed factors.
- Big improvement from (3) to (4) $\rightarrow$ not parallel

# Simplex Models for Ordered Latent Variables

Guttman (1954; 1957) observed a pattern of correlations, called a ***simplex***, among a set of tests of verbal ability which were *ordered* in complexity:

| | | | | | | |
|---|---|---|---|---|---|---|
| Spelling | 1.00 | | | | | |
| Punctuation | .621 | 1.00 | | | | |
| Grammar | .564 | .742 | 1.00 | | | |
| Vocab | .476 | .503 | .577 | 1.00 | | |
| Literature | .394 | .461 | .472 | .688 | 1.00 | |
| Foreign Lit | .389 | .411 | .429 | .548 | .639 | 1.00 |

The pattern is such that the correlations decrease as one moves away from the main diagonal.



# Simplex Models for Ordered Latent Variables

- A simplex pattern is usually found whenever one has a set of latent variables which can be considered ordered on some underlying unidimensional scale.
- This corresponds to saying that the measures closest together on the underlying scale are most highly correlated.
- Some additional examples are:
  - learning data: scores on a set of trials
  - longitudinal data, growth studies
  - developmental studies: tasks ordered by developmental stages

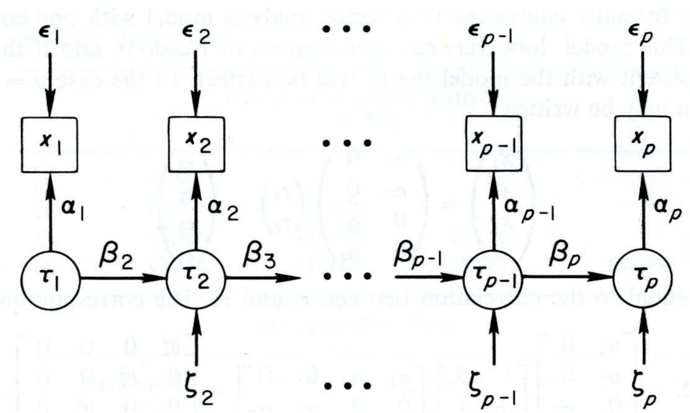


Various simplex models can be tested with LISREL. Consider $p$ fallible variables, $x_1, x_2, \ldots, x_p$. The structural equations for the simplex indicate that each $x_i$ is measured with error, and that the "true scores" are linearly related to the previous one.

$$x_i = \alpha_i \tau_i + \epsilon_i, \quad i = 1, \ldots, p \quad \boxed{\text{measurement model}}$$

$$\tau_i = \beta_i \tau_{i-1} + \xi_i, \quad i = 2, \ldots, p \quad \boxed{\text{structural model}}$$

The path diagram is:




The ***Markov simplex*** defines a set of scale points for the latent variables



# Example: Guttman's Data

Jöreskog (1970) analyzed Guttman's data under several variations of the Markov simplex model.
Solution for Quasi Markov Simplex ($\chi^2 = 43.81$ with 6 df)— $p < .001$, but all residuals are small.

| Test | $\alpha_i$ | $\delta_i$ | $t_i$ |
|---|---|---|---|
| Spelling | .989 | .022 | 0.00 |
| Punctuation | .626 | .212 | 0.69 |
| Grammar | .586 | .212 | 0.83 |
| Vocab | .476 | .216 | 1.47 |
| Literature | .394 | .218 | 1.74 |
| Foreign Lit | .389 | .386 | 1.90 |

The underlying scale ($t_i$) for the latent variables is:

| 0 | .69 | .83 | 1.47 | 1.74 | 1.90 |
|---|---|---|---|---|---|
| \| | \| | \| | \| | \| | \| |
| Spl | Pun | Grm | Voc | Lit | FLit |

# Simplex Models for Ordered Latent Variables

***Notes***:

- Data which fits a simplex model should also fit a simple 1-factor model, but such a model does not really test the assumed ordering.
- The simplex model uses a set of ordered latent variables.
- The estimated spacing of the scale values $t_i$ may be of importance in interpretation.



# 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 ACOVS/LISREL model allows any parameter to be assigned an arbitrary fixed value, or constrained to be equal to some other parameter. We can test any degree of invariance from totally separate factor structures to completely invariant ones.
- Model
  Let $\boldsymbol{x}_g$ 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: Hypotheses

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

- 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?

Both LISREL and AMOS provide convenient ways to do multi-sample analysis. `PROC CALIS` does not.



1. 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

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

which is distributed approx. as $\chi^2$ 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)

2. Same number of factors
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 $\chi^2$'s and degrees of freedom,

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

3. Same factor pattern
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 $\Lambda$ may be either completely unspecified, or be specified to have zeros in certain positions.



To obtain a $\chi^2$ for this hypothesis, estimate $\Lambda$ (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 $H_\Lambda$, *given that the number of factors is the same* in all groups, use

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

4. Same factor pattern and unique variances
A stronger hypothesis is that the unique variances, as well as the factor pattern, are invariant across groups:

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



# 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. $H_{=\Sigma} : \Sigma_1 = \Sigma_2$ | $\begin{cases} \mathbf{\Lambda}_1 = \mathbf{\Lambda}_2 = \mathbf{I}_{(4\times4)} \\ \mathbf{\Psi}_1 = \mathbf{\Psi}_2 = \mathbf{0}_{(4\times4)} \\ \mathbf{\Phi}_1 = \mathbf{\Phi}_2 \text{ constrained, free} \end{cases}$ |
| B. $H_{k=2}$ : $\Sigma_1, \Sigma_2$ 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 | |
|---|---|---|---|---|---|---|---|---|
| | $\chi^2$ | 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: $H_{\Lambda}$ | 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 ($H_{\Lambda}$) in both samples is tenable.
- Unique variances appear to differ in the two samples.
- The factor correlation ($\phi_{12}$) appears to be greater in the Academic sample than in the non-Academic sample.



# Factorial Invariance: LISREL syntax

Model B for Academic group: 2 correlated, non-overlapping factors

```
Ex12: 2 Correlated factors: Hypothesis B (Group A)
DAta NGroup=2 NI=4 NObs=373
LAbels  file=lisrel12.dat; CMatrix file=lisrel12.dat
MOdel NX=4 NKSI=2
! Pattern: 1 free parameter and 1 fixed parameter in each column
FRee    LX(2,1) LX(4,2)
STart 1 LX(1,1) LX(3,2)
OUtput
```

Model B for Non-Academic group: same pattern as Group A

```
Ex12: 2 Correlated factors: Hypothesis B (Group N-A)
DAta NObs=249
LAbels  file=lisrel12.dat; CMatrix file=lisrel12.dat
MOdel LX=PS
PDiagram
OUtput
```

LX=PS: same pattern and starting values as in previous group but loadings are not constrained to be equal



# Factorial Invariance: LISREL syntax

Model C for Academic group: equal $\Lambda_x$— same as Model B

```
Ex12: Equal Lambda: Hypothesis C (Group A)
DAta NGroup=2 NI=4 NObs=373
LAbels  file=lisrel12.dat; CMatrix file=lisrel12.dat
MOdel NX=4 NKSI=2
! Pattern: 1 free parameter and 1 fixed parameter in each column
FRee    LX(2,1) LX(4,2)
STart 1 LX(1,1) LX(3,2)
OUtput
```

Model C for Non-Academic group: same $\Lambda_x$ as Group A

```
Ex12: 2 Correlated factors: Hypothesis B (Group N-A)
DAta NObs=249
LAbels  file=lisrel12.dat; CMatrix file=lisrel12.dat
MOdel LX=INvariant
PDiagram
OUtput
```

LX=IN: loadings constrained to be equal to those in Group A

Complete example:
www.math.yorku.ca/SCS/Courses/factor/lisrel/lisrel12.ls8



# Factorial Invariance: AMOS Basic syntax

Model B for Academic group:

```
Sub Main
  Dim Sem As New AmosEngine
  With Sem
   .title "Academic and NonAcademic Boys (Sorbom, 1976): "  _
        & "Equality of Factor Structures"

   ' Start out with the least constrained model
     .Model "B: 2 Factors, unconstrained"
     .BeginGroup "invar.xls", "Academic"
     .GroupName "Academic Boys"
     .Structure "Read_Gr5 = ( 1 ) Gr5 + (1) eps1"
     .Structure "Writ_Gr5 = (L1a) Gr5 + (1) eps2"
     .Structure "Read_Gr7 = ( 1 ) Gr7 + (1) eps3"
     .Structure "Writ_Gr7 = (L2a) Gr7 + (1) eps4"
     .Structure "Gr5 <--> Gr7 (phi1)"
     .Structure "eps1 (v1a)"
     .Structure "eps2 (v2a)"
     .Structure "eps3 (v3a)"
     .Structure "eps4 (v4a)"
```

# Factorial Invariance: AMOS Basic syntax

Model B for Non-Academic group:

```
.BeginGroup "invar.xls", "NonAcademic"
.GroupName "NonAcademic Boys"
.Structure "Read_Gr5 = ( 1 ) Gr5 + (1) eps1"
.Structure "Writ_Gr5 = (L1b) Gr5 + (1) eps2"
.Structure "Read_Gr7 = ( 1 ) Gr7 + (1) eps3"
.Structure "Writ_Gr7 = (L2b) Gr7 + (1) eps4"
.Structure "Gr5 <--> Gr7 (phi2)"
.Structure "eps1 (v1b)"
.Structure "eps2 (v2b)"
.Structure "eps3 (v3b)"
.Structure "eps4 (v4b)"
```

Note that the model is the same, but all parameter names are suffixed with 'b', so they are not constrained to be equal



# Factorial Invariance: AMOS Basic syntax

Now, other models can be specified in terms of equality constraints across groups:

```
  ' Fix the loadings in the two groups to be equal
     .Model "C: = loadings", _
      "L1a = L1b; L2a = L2b"

  ' Add contraint that unique variances are equal
     .Model "D: C + = unique var", _
      "L1a=L1b; L2a=L2b; _
        v1a=v1b; v2a=v2b;  v3a=v3b;  v4a=v4b"

  ' Add contraint that factor correlations are equal
     .Model "E: D + = factor corr", _
      "L1a=L1b; L2a=L2b; _
        v1a=v1b; v2a=v2b;  v3a=v3b;  v4a=v4b; _
        phi1=phi2"
   End With
End Sub
```



# Identifiability in CFA models

- Because they offer the possibility of fitting hypotheses that are partially specified, care must be take with CFA models to ensure that a unique solution can be found.
- For an unrestricted factor model with $k$ latent factors, we have seen that at least $k^2$ restrictions *must* be imposed.
- It turns out that this is a ***necessary***, but ***not a sufficient*** condition for the model to be identified.



# Identifiability in CFA models

- In addition, it is necessary to specify the *unit of measurement*, or *scale* for each latent variable. This may be done by (arbitrarily) assigning one fixed non-zero loading, typically 1, in each column of the factor matrix.
- For a problem with 6 variables and 2 factors, the loading matrix would look like this:

$$\Lambda_2 = \begin{bmatrix} 1 & 0 \\ x & x \\ x & x \\ 0 & 1 \\ x & x \\ x & x \end{bmatrix} \quad \Phi_2 = \begin{bmatrix} 1 & x \\ x & 1 \end{bmatrix}$$

The fixed 1s correspond to equating the scale of factor 1 to that of variable 1, and factor 2 is equated to variable 4.

# Identifiability condition

- Let $\boldsymbol{\theta}$ be the $t \times 1$ vector containing all unknown parameters in the model, and let

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}^{\mathsf{T}} + \boldsymbol{\Psi}$$

  be the covariance matrix expressed in terms of those parameters.
- Then, the parameters in $\boldsymbol{\theta}$ are identified if you can show that the elements of $\boldsymbol{\theta}$ are uniquely determined functions of the elements of $\boldsymbol{\Sigma}$, that is:

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}_1) = \boldsymbol{\Sigma}(\boldsymbol{\theta}_2) \rightarrow \boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$$



# Identifiability condition

For example, for a 1-factor, 3-variable model:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} [\xi] + \begin{bmatrix} z_1 & & \\ & z_2 & \\ & & z_3 \end{bmatrix}$$

Then, letting $\boldsymbol{\Phi} = var(\xi) = 1$ and $var(z_i) = \psi_i$, the covariance matrix of the observed variables can be expressed as:

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \begin{bmatrix} \lambda_1^2 + \psi_1 & & \\ \lambda_2\lambda_1 & \lambda_2^2 + \psi_2 & \\ \lambda_3\lambda_1 & \lambda_3\lambda_1 & \lambda_3^2 + \psi_3 \end{bmatrix}$$

Each parameter can be solved for, so the model is identified.



# Identifiability rules: t rule

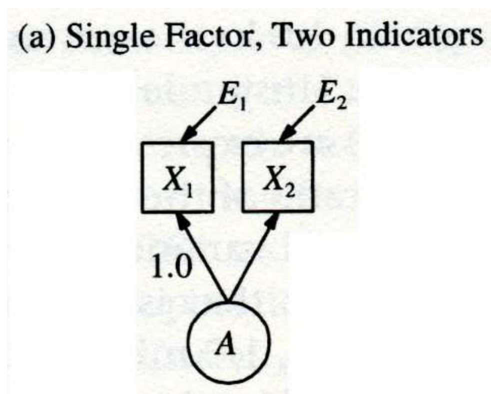
**t-rule:** There cannot be more unknown parameters than there are known elements in the sample covariance matrix. This is a ***necessary***, but ***not sufficient*** condition.

$$t \leq p(p+1)/2$$

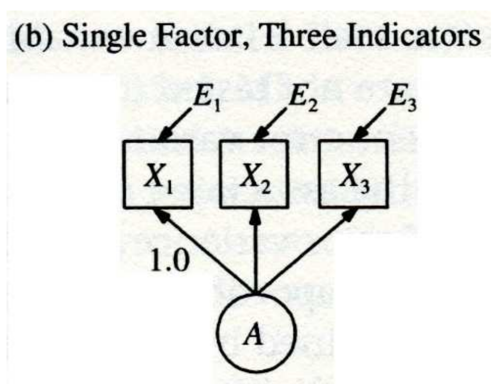
***Example***:
For 6 tests, you can estimate no more than $6 \times 7/2 = 21$ parameters.



# Identifiability rules: t rule




2 variables: 3 var-covariances
model: 4 free parameters
not identified




3 variables: 6 var-covariances
model: 6 free parameters
just identified

# Identifiability rules: 3 variable rules

**3-variable rules:** A factor model is identified when there are:

- three or more variables per factor
- each variable has one and only one non-zero loading
- the unique factors are uncorrelated ($\Psi$ is a diagonal matrix).

There are no restrictions on the factor correlations ($\Phi$). These conditions are jointly ***sufficient***, though ***not necessary***.



# Identifiability rules: 2 variable rules

**2-variable rules** A less restrictive set of rules is:

- two or more variables per factor
- each variable has one and only one non-zero loading
- each factor is scaled by setting one $\lambda_{ij} = 1$ in each column.
- the unique factors are uncorrelated ($\Psi$ is a diagonal matrix).
- there are no fixed zero elements in $\Phi$.

These conditions are also ***sufficient***, though ***not necessary***.

***Example***:

With 4 variables, and 2 latent variables, the model is identified if the parameters are specified as

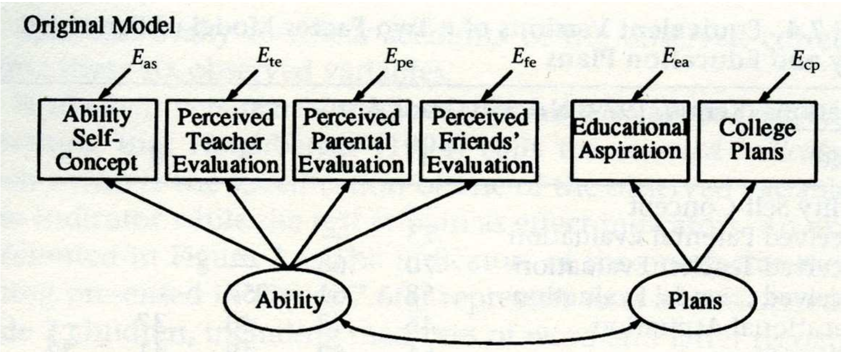
$$\Lambda = \begin{bmatrix} 1 & 0 \\ \lambda_{21} & 0 \\ 0 & 1 \\ 0 & \lambda_{42} \end{bmatrix}, \qquad \Phi = \begin{bmatrix} \phi_{11} & \\ \phi_{21} & \phi_{22} \end{bmatrix} = \text{free}$$



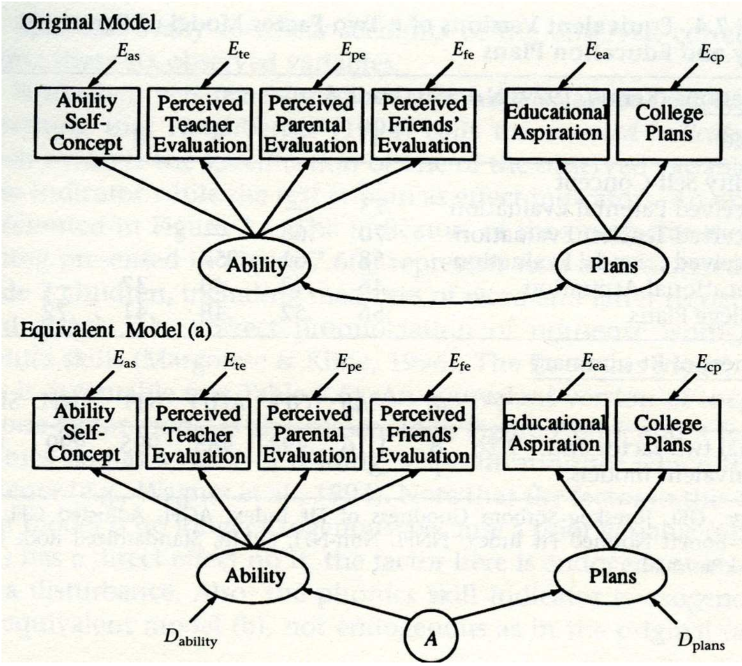
# Equivalent models

- In standard CFA models based on the common factor model (uncorrelated errors, 1st-order model) identified models are unique.
- In more general models, it is possible to have several distinct equivalent models
  - same degrees of freedom, $\chi^2$
  - different substantive interpretation

Recall the model for Ability and aspiration:

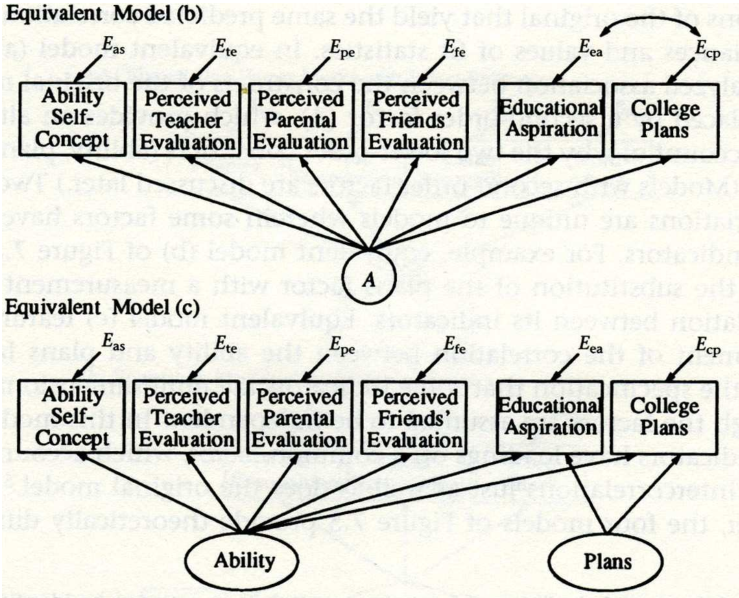





- 2nd-order model, with a general 2nd-order factor

- 1-factor model, with correlated unique variables




- Two overlapping factors



# Power and Sample Size for EFA and CFA

**Bad news** Determining the required sample size, or the power of a hypothesis test are far more complex in EFA and CFA than in other statistical applications (correlation, ANOVA, etc.)

**Good news** There are a few things you *can* do to choose a sample size intelligently.



# Power and Sample Size for EFA and CFA

## Rules of thumb for EFA

For EFA, there is little statistical basis for determining the appropriate sample size, and little basis for determining power (but the overall approach of CFA can be used).
Some traditional "rules of thumb" for EFA:

- ***The more the better!***
  - Reliability and replicability increase directly with $\sqrt{N}$.
  - More reliable factors can be extracted with larger sample sizes.
- Absolute minimum– $N = 5p$, but you should have $N > 100$ for any non-trivial factor analysis. Minimum applies only when communalities are high and $p/k$ is high.
- Most EFA and CFA studies use $N > 200$, some as high as 500-600.
- Safer to use at least $N > 10p$.
- The lower the reliabilities, the larger $N$ should be.



# Using desired standard errors

- An alternative approach for EFA considers the standard errors of correlations, in relation to sample size.
- This usually provides more informed guidance than the rules of thumb. It can be shown that,

$$\sigma(\rho) = \frac{1-\rho^2}{\sqrt{N}} + \mathcal{O}(N^{-1})$$

so, we could determine the sample size to make the standard error of a "typical" correlation smaller than some given value.

$$\sqrt{N} > \frac{1-\rho^2}{\sigma(\rho)}$$

# Using desired standard errors

| | Sample size | | | | |
|---|---|---|---|---|---|
| $\rho$ | 50 | 100 | 200 | 400 | 800 |
| 0.1 | 0.140 | 0.099 | 0.070 | 0.050 | 0.035 |
| 0.3 | 0.129 | 0.091 | 0.064 | 0.046 | 0.032 |
| 0.5 | 0.106 | 0.075 | 0.053 | 0.038 | 0.027 |
| 0.7 | 0.072 | 0.051 | 0.036 | 0.026 | 0.018 |

- Standard error decreases as $|\rho|$ increases.
- So, if you want to keep the standard error less than 0.05, you need $N = 400$ when the "typical" correlation is only 0.1, but $N = 100$ when the "typical" correlation is 0.7.
- In many behavioural and psychology studies, correlations among *different scales* are modest, at best ($0.1 \leq \rho \leq 0.3$).
- For typical scale analysis, one should expect the correlations among items on the *same scale* to be much higher ($0.7 \leq \rho \leq 0.9$), $\Rightarrow$ smaller required sample size for the same standard error.



# Power and Sample size for CFA

- ***Problems:*** The situation in CFA wrt power analysis is typically reversed compared with other forms of hypothesis tests—
  - $\chi^2 = (N-1)F_{min}$, so large $N \Rightarrow$ reject $H_0$.
  - With small specification errors, large sample size will magnify their effects $\Rightarrow$ reject $H_0$.
  - With large specification errors, small sample size will mask their effects $\Rightarrow$ accept $H_0$.
- ***Overall approach:*** MacCallum, Browne and Sugawara (1996) approach allows for testing a null hypothesis of 'not-good-fit', so that a significant result provides support for good fit.
  - Effect size is defined in terms of a null hypothesis and alternative hypothesis value of the root-mean-square error of approximation (RMSEA) index. Typical values for RMSEA:

| | |
|---|---|
| $\leq .05$ | close fit |
| $.05 - .08$ | fair |
| $.08 - .10$ | mediocre |
| $> .10$ | poor |

  - These values, together with the df for the model being fitted, sample size ($N$), and error rate ($\alpha$), allow power to be calculated.



# Power and Sample size for CFA

- The CSMPOWER macro
  - See: http://www.math.yorku.ca/SCS/sasmac/csmpower.html
  - Retrospective power analysis— uses the RMSEA values from the OUTRAM= data set from PROC CALIS for the model fitted.
  - Prospective power analysis— values of RMSEA, DF and N must be provided through the macro arguments.



# Example: Retrospective power analysis

Here, we examine the power for the test of Lord's two-factor model for speeded and unspeeded vocabulary tests, where $N = 649$.

```
title "Power analysis: Lord's Vocabulary Data";
title2 "Lord's data: H1- X1 and X2 parallel,
        Y1 and Y2 parallel, rho=1";
proc calis data=lord cov summary outram=ram1;
   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 = 1;
run;

*-- Power analysis from RMSEA statistics in this model;
title 'Retrospective power analysis';
%csmpower(data=ram1);
```

# Example: Retrospective power analysis

Results include:

| Alpha | df | Name of Variable | N | H0 fit value | Ha fit value | Power |
|---|---|---|---|---|---|---|
| 0.05 | 6 | RMSEAEST | 649 | 0.05 | 0.08977 | 0.75385 |
| | | RMSEALOB | 649 | 0.05 | 0.06349 | 0.19282 |
| | | RMSEAUPB | 649 | 0.05 | 0.11839 | 0.99202 |

With this sample size, we have power of 0.75 to distinguish between a fit with RMSEA=0.05 and one with RMSEA=0.09.



## Example: Prospective power analysis

For prospective power analysis, we specify the RMSEA for alternative hypotheses of 'not good fit' with the RMSEAA= parameter (for $H_a$).

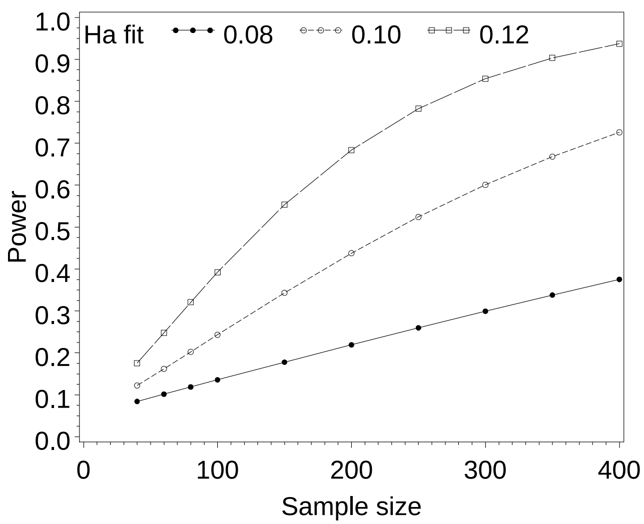
```
*--; title  'Prospective power analysis';
%csmpower(df=6, rmseaa=%str(.08 to .12 by .02),
          plot=%str(power*n =rmseaa));
```

Results include a listing:

| Alpha | df | N | H0 fit value | Ha fit value | Power |
|---|---|---|---|---|---|
| 0.05 | 6 | 40 | 0.05 | 0.08 | 0.08438 |
| | | 40 | 0.05 | 0.10 | 0.12243 |
| | | 40 | 0.05 | 0.12 | 0.17575 |
| | | 60 | 0.05 | 0.08 | 0.10168 |
| | | 60 | 0.05 | 0.10 | 0.16214 |
| | | 60 | 0.05 | 0.12 | 0.24802 |
| | | 80 | 0.05 | 0.08 | 0.11883 |
| | | 80 | 0.05 | 0.10 | 0.20262 |
| | | 80 | 0.05 | 0.12 | 0.32093 |
| | | 100 | 0.05 | 0.08 | 0.13585 |
| | | 100 | 0.05 | 0.10 | 0.24333 |
| | | 100 | 0.05 | 0.12 | 0.39214 |
| | | ... | ... | ... | ... |
| | | 400 | 0.05 | 0.08 | 0.37545 |
| | | 400 | 0.05 | 0.10 | 0.72599 |
| | | 400 | 0.05 | 0.12 | 0.93738 |



Plot of Power by $N$ for each level of RMSEAA:




- For the most stringent test of $H_0$ : RMSEA $=$ 0.05 vs. $H_a$ : RMSEA $=$ 0.08, the largest sample size, $N = 400$ only provides a power of 0.375.
- Good thing they used $N = 649$!



# Individual model specifications

- The overall approach can only evaluate power or required sample size for the whole model.
- It does not distinguish among the *a priori* specifications of free and fixed parameters implied by the model being tested.

Things become more difficult when the focus is on power for deciding on some one or a few specifications (parameters) in a model.

# Individual model specifications

There are some promising results:

- Satorra (1989) found that the modification indices (“Lagrange multipliers” in `PROC CALIS`)— $\Delta\chi^2$ for *fixed parameters* in a model approximate the $\chi^2$ non-centrality parameters required to determine power for a specific fixed parameter.
- Similarly, the Wald tests, $\chi^2_1 = (\text{parm}/s(\text{parm}))^2$ approximate the $\chi^2$ non-centrality parameters required to determine power for *free parameters*.
- These $\chi^2$ values should be studied in relation to the estimated change in the parameter (ECP).
  - A large $\Delta\chi^2$ with a small ECP simply reflects the high power to detect small differences which comes with large $N$.
  - Similarly, a small $\Delta\chi^2$ with a large ECP reflects low power for large differences with small $N$.

  See the paper by Kaplan, “Statistical power in structural equation models”, `www.gsu.edu/~mkteer/power.html` for further discussion and references on these issues.