Confirmatory Factor Analysis

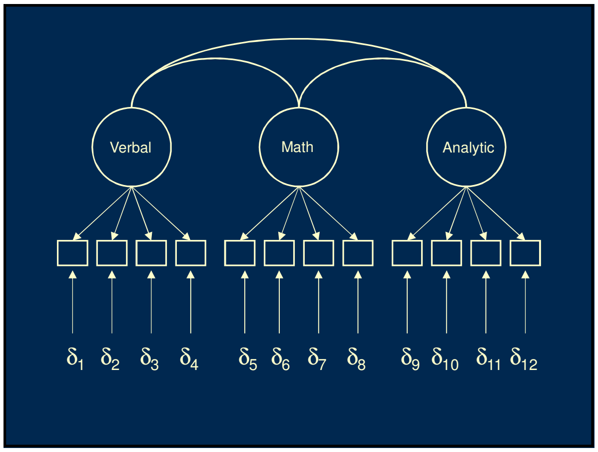
Additional issues in confirmatory factor analysis:

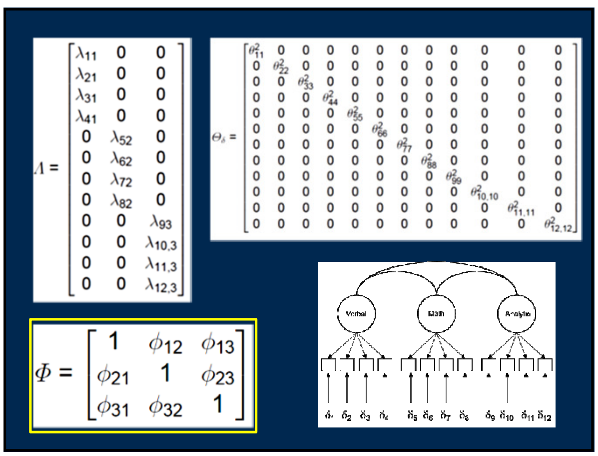
- Scaling of latent variables
- Correlation versus covariance matrices
- Model modification ("exploratory" confirmatory factor analysis)
- Relation to measurement theory
- Measurement invariance (cross-validation)

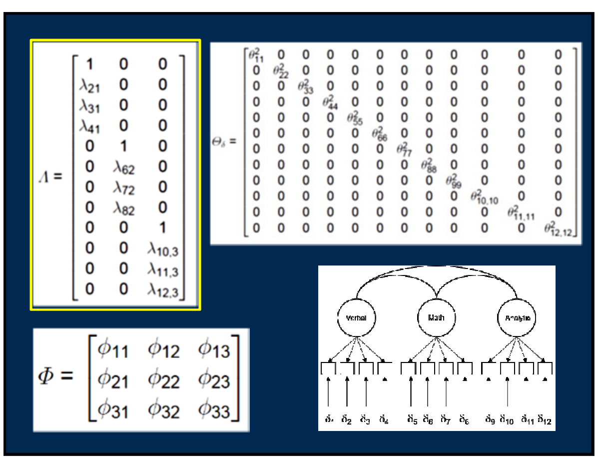
Scaling

For a model to be identified, the latent variables must be given a scale or metric.

This can be done by either standardizing the latent variables (setting the diagonals of the Φ matrix to 1.00) or by setting one λ for each latent variable to 1.00 to give that latent variable the same scale as the observed variable.


Verbal
Math
Analytic
δ_1 δ_2 δ_3 δ_4 δ_5 δ_6 δ_7 δ_8 δ_9 δ_{10} δ_{11} δ_{12}


$$\Lambda = \begin{bmatrix} \lambda_{11} & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ \lambda_{31} & 0 & 0 \\ \lambda_{41} & 0 & 0 \\ 0 & \lambda_{52} & 0 \\ 0 & \lambda_{62} & 0 \\ 0 & \lambda_{72} & 0 \\ 0 & \lambda_{82} & 0 \\ 0 & 0 & \lambda_{93} \\ 0 & 0 & \lambda_{10,3} \\ 0 & 0 & \lambda_{11,3} \\ 0 & 0 & \lambda_{12,3} \end{bmatrix}$$
$$\Theta_\delta = \begin{bmatrix} \theta^2_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta^2_{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta^2_{33} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta^2_{44} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta^2_{55} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \theta^2_{66} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \theta^2_{77} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta^2_{88} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta^2_{99} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta^2_{10,10} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta^2_{11,11} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta^2_{12,12} \end{bmatrix}$$
$$\Phi = \begin{bmatrix} 1 & \phi_{12} & \phi_{13} \\ \phi_{21} & 1 & \phi_{23} \\ \phi_{31} & \phi_{32} & 1 \end{bmatrix}$$
Verbal
Math
Analytic


$$\Lambda = \begin{bmatrix} 1 & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ \lambda_{31} & 0 & 0 \\ \lambda_{41} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda_{62} & 0 \\ 0 & \lambda_{72} & 0 \\ 0 & \lambda_{82} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & \lambda_{10,3} \\ 0 & 0 & \lambda_{11,3} \\ 0 & 0 & \lambda_{12,3} \end{bmatrix}$$
$$\Theta_\delta = \begin{bmatrix} \theta^2_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta^2_{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta^2_{33} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta^2_{44} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta^2_{55} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \theta^2_{66} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \theta^2_{77} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta^2_{88} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta^2_{99} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta^2_{10,10} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta^2_{11,11} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta^2_{12,12} \end{bmatrix}$$
$$\Phi = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix}$$
Verbal
Math
Analytic

The same number of parameters are estimated, but the particular parameters that are estimated changes.

$$\Phi = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix}$$

	Estimate	Std.err	Z-value	P(>|z|)	Std.lv	Std.all
Covariances:						
Verbal ~~						
Math	0.199	0.029	6.782	0.000	0.417	0.417
Reasoning	0.215	0.029	7.442	0.000	0.461	0.461
Math ~~						
Reasoning	0.176	0.026	6.759	0.000	0.405	0.405

$$\Phi = \begin{bmatrix} 1 & \phi_{12} & \phi_{13} \\ \phi_{21} & 1 & \phi_{23} \\ \phi_{31} & \phi_{32} & 1 \end{bmatrix}$$

	Estimate	Std.err	Z-value	P(>|z|)	Std.lv	Std.all
Covariances:						
Verbal ~~						
Math	0.417	0.046	9.108	0.000	0.417	0.417
Reasoning	0.461	0.048	9.528	0.000	0.461	0.461
Math ~~						
Reasoning	0.405	0.048	8.412	0.000	0.405	0.405

The standardized solutions are identical.

The same number of parameters are estimated, but the particular parameters that are estimated changes.

$$\Phi = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix}$$

	Estimate	Std.err	Z-value	P(>|z|)	Std.lv	Std.all
Latent variables:						
Verbal =~						
Grammar	1.000				0.716	0.734
Prgrph_Cmprhn	1.065	0.080	13.256	0.000	0.762	0.718
Vocabulary	0.960	0.070	13.772	0.000	0.687	0.689
Sentnc_Cmpltn	1.108	0.079	13.940	0.000	0.793	0.731

$$\Phi = \begin{bmatrix} 1 & \phi_{12} & \phi_{13} \\ \phi_{21} & 1 & \phi_{23} \\ \phi_{31} & \phi_{32} & 1 \end{bmatrix}$$

	Estimate	Std.err	Z-value	P(>|z|)	Std.lv	Std.all
Latent variables:						
Verbal =~						
Grammar	0.716	0.042	17.040	0.000	0.716	0.734
Prgrph_Cmprhn	0.762	0.044	17.213	0.000	0.762	0.718
Vocabulary	0.687	0.041	16.598	0.000	0.687	0.689
Sentnc_Cmpltn	0.793	0.046	17.357	0.000	0.793	0.731

The standardized solutions are identical.

Identical goodness-of-fit statistics are provided by both approaches.

```
Number of observations                          500

Number of missing patterns                        1

Estimator                                        ML      Robust
Minimum Function Test Statistic              55.498      56.546
Degrees of freedom                               51          51
P-value (Chi-square)                          0.309       0.276
Scaling correction factor                                 0.981
  for the Yuan-Bentler correction

User model versus baseline model:

  Comparative Fit Index (CFI)                 0.998       0.997
  Tucker-Lewis Index (TLI)                    0.997       0.996

Root Mean Square Error of Approximation:
  RMSEA                                       0.013       0.015
  90 Percent Confidence Interval       0.000  0.032       0.000  0.033
  P-value RMSEA <= 0.05                       1.000       1.000

Standardized Root Mean Square Residual:

  SRMR                                        0.026       0.026
```

Analyzing Correlation versus Covariance Matrices

Confirmatory factor analysis models are usually based on the decomposition of covariance matrices, not correlation matrices.

The solutions will differ to the extent that variables have different variances (these differences are removed with conversion to a correlation matrix).

Example data: Need for Cognition Scale (18 items) administered to 195 participants.

Latent variables: NC =~	Estimate
nc1	1.000
nc2	1.072
nc3	-0.884
nc4	-1.038
nc5	-1.018
nc6	0.682
nc7	-0.957
nc8	-0.787
nc9	-0.951
nc10	0.801
nc11	0.906
nc12	-0.910
nc13	0.865
nc14	0.800
nc15	0.567
nc16	-0.604
nc17	-0.836
nc18	0.319

C

Latent variables: NC =~	Estimate
nc1	1.000
nc2	1.296
nc3	-1.031
nc4	-1.121
nc5	-1.208
nc6	0.704
nc7	-0.900
nc8	-0.801
nc9	-0.931
nc10	1.046
nc11	1.166
nc12	-1.062
nc13	0.933
nc14	0.905
nc15	0.655
nc16	-0.630
nc17	-0.917
nc18	0.328

R

	[,1]
nc1	1.5344
nc2	1.0491
nc3	1.1274
nc4	1.3138
nc5	1.0898
nc6	1.4429
nc7	1.7357
nc8	1.4778
nc9	1.5986
nc10	0.8989
nc11	0.9268
nc12	1.1247
nc13	1.3192
nc14	1.1994
nc15	1.1477
nc16	1.4141
nc17	1.2740
nc18	1.4587

Variances

The impact of variables in an analysis of correlations will appear to shift as a function of the variances of those variables in relation to the variance-defining indicator.

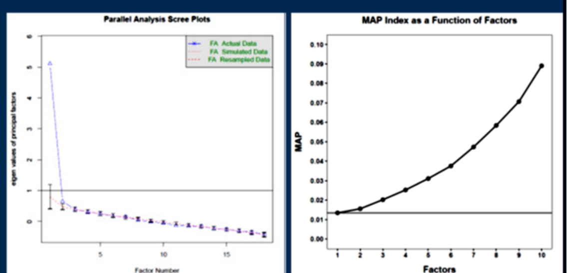
Model Modification

When a presumed model does not fit, changes to the model can be tested for "goodness-of-fit," but the exercise is no longer confirmatory and may capitalize on chance and other biases.

It is surprisingly easy to find a good fitting model by trial and error.

The example data come from a sample of students who completed the 18-item Need for Cognition Scale.

[ ] 1. I would prefer complex to simple problems.
[ ] 2. I like to have the responsibility of handling a situation that involves a lot of thinking.
[ ] 3. Thinking is not my idea of fun.
[ ] 4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.
[ ] 5. I try to anticipate and avoid situations where there is likely a chance I will have to think in depth about something.
[ ] 6. I find joy in deliberating hard and for long hours.
[ ] 7. I only think as hard as I have to.
[ ] 8. I prefer to think about small, daily projects to long-term ones.
[ ] 9. I like tasks that require little thought once I've learned them.
[ ] 10. The idea of relying on thought to make my way to the top appeals to me.
[ ] 11. I really enjoy a task that involves coming up with new solutions to problems.
[ ] 12. Learning new ways to think doesn't excite me very much.
[ ] 13. I prefer my life to filled with puzzles that I must solve.
[ ] 14. The notion of thinking abstractly appeals to me.
[ ] 15. I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.
[ ] 16. I feel relief rather than satisfaction after completing a task that required a lot of attention.
[ ] 17. It's enough for me that something gets the job done; I don't care how or why it works..
[ ] 18. I usually end up deliberating about issues even though they do not affect me personally.



The conceptual model underlying the Need for Cognition Scale assumes a single latent variable. That seems sensible from these preliminary analyses.

```
Number of observations                           195

Number of missing patterns                         1

Estimator                                         ML      Robust
Minimum Function Test Statistic              267.315     228.948
Degrees of freedom                               135         135
P-value (Chi-square)                           0.000       0.000
Scaling correction factor                                  1.168
  for the Yuan-Bentler correction

User model versus baseline model:

  Comparative Fit Index (CFI)                  0.854       0.870
  Tucker-Lewis Index (TLI)                     0.835       0.853

Root Mean Square Error of Approximation:

  RMSEA                                        0.071       0.060
  90 Percent Confidence Interval        0.059  0.084       0.047  0.072
  P-value RMSEA <= 0.05                        0.004       0.095

Standardized Root Mean Square Residual:

  SRMR                                         0.060       0.060
```

A one-factor CFA model is not a particularly good fit.

Most software, lavaan included, provides “modification indices” to help guide follow-up changes to an ill-fitting model. A modification index is the change in the goodness-of-fit χ^2 that would result from setting a parameter free.

```
MI <- modificationIndices(NC_Fit_1)
subset(MI, mi > 10)

##      lhs op  rhs    mi    epc sepc.lv sepc.all sepc.nox
## 105  nc4 ~~  nc5 15.72  0.218   0.218    0.319    0.319
## 157  nc8 ~~ nc11 11.43  0.205   0.205    0.260    0.260
## 158  nc8 ~~ nc12 11.24 -0.234  -0.234   -0.254   -0.254
## 174 nc10 ~~ nc11 11.57  0.149   0.149    0.268    0.268
## 195 nc13 ~~ nc14 13.90  0.258   0.258    0.281    0.281
## 198 nc13 ~~ nc17 10.62  0.232   0.232    0.246    0.246
## 202 nc14 ~~ nc17 12.34  0.240   0.240    0.264    0.264
```

The chi-square would be expected to be smaller by 15.72 if the error variances for Item 4 and Item 5 are allowed to correlate.

```
Number of observations                            195

Number of missing patterns                          1

Estimator                                          ML      Robust
Minimum Function Test Statistic               172.279     147.771
Degrees of freedom                                127         127
P-value (Chi-square)                            0.005       0.100
Scaling correction factor                                   1.166
  for the Yuan-Bentler correction
```

```
User model versus baseline model:

Comparative Fit Index (CFI)                     0.950       0.971
Tucker-Lewis Index (TLI)                        0.940       0.965
```

```
Root Mean Square Error of Approximation:

RMSEA                                           0.043       0.029
90 Percent Confidence Interval          0.025   0.058       0.000  0.046
P-value RMSEA <= 0.05                           0.762       0.982

Standardized Root Mean Square Residual:

SRMR                                            0.050       0.050
```

Successively removing the constraints associated with the largest modification indices will produce a better fit.

It may even produce an excellent fit. After 8 rounds, each time freeing an error covariance, the resulting model fits quite well. But does it make any sense?

```
NC.model.9 <- '
# Latent variable definitions.
NC =~ nc1 + nc2 + nc3 + nc4 + nc5 + nc6 + nc7 + nc8 + nc9 +
nc10 + nc11 + nc12 + nc13 + nc14 + nc15 + nc16 + nc17 + nc18
nc4 ~~ nc5
nc13 ~~ nc14
nc8 ~~ nc11
nc8 ~~ nc12
nc10 ~~ nc11
nc5 ~~ nc12
nc2 ~~ nc4
nc14 ~~ nc17
'
```

There is no apparent pattern to the error variances that were freely estimated . . .

[ ] 1. I would prefer complex to simple problems.
[g] 2. I like to have the responsibility of handling a situation that involves a lot of thinking.
[ ] 3. Thinking is not my idea of fun.
[a g] 4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.
[a f] 5. I try to anticipate and avoid situations where there is likely a chance I will have to think in depth about something.
[ ] 6. I find joy in deliberating hard and for long hours.
[ ] 7. I only think as hard as I have to.
[c d] 8. I prefer to think about small, daily projects to long-term ones.
[ ] 9. I like tasks that require little thought once I've learned them.
[e] 10. The idea of relying on thought to make my way to the top appeals to me.
[c e] 11. I really enjoy a task that involves coming up with new solutions to problems.
[d f] 12. Learning new ways to think doesn't excite me very much.
[b] 13. I prefer my life to filled with puzzles that I must solve.
[b h] 14. The notion of thinking abstractly appeals to me.
[ ] 15. I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.
[ ] 16. I feel relief rather than satisfaction after completing a task that required a lot of attention.
[h] 17. It's enough for me that something gets the job done; I don't care how or why it works..
[ ] 18. I usually end up deliberating about issues even though they do not affect me personally.

When possible, it is better to consider conceptually or methodologically sensible alternative models that might fit better than the original. One common measurement artifact is direction of scoring.

```
NC_N =~ NA*nc1 + nc2 + nc6 + nc10 + nc11 +
        nc13 + nc14 + nc15 + nc18
NC_P =~ NA*nc3 + nc4 + nc5 + nc7 + nc8 +
        nc9 + nc12 + nc16 + nc17
NC_N ~~ 1*NC_N
NC_P ~~ 1*NC_P
NC_N ~~ NC_P
```

```
Root Mean Square Error of Approximation:

  RMSEA                                  0.061       0.050
  90 Percent Confidence Interval  0.047  0.074       0.035  0.063
  P-value RMSEA <= 0.05                  0.088       0.504
```

This two-factor model is far simpler than the iteratively-modified model and fits reasonably well.

Reliability and Attenuation

Confirmatory factor analysis has a close relation to classical measurement theory. When the latent variable and observed variables are standardized, the λ are the correlations between each variable and the "true score." The square of these λ are the individual item reliabilities (ρ_{ii}). If these are averaged ($\overline{\rho}$), they can be used in the Spearman-Brown formula to provide an estimate of standardized coefficient alpha—an estimate of internal consistency:

$$\alpha_{kk} = \frac{k\overline{\rho}}{1 + (k-1)\overline{\rho}}$$

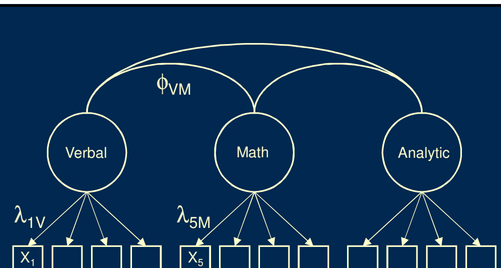
An alternative estimate of reliability uses the loadings and the error variances, resembling the variance ratio form for reliability:

$$\rho^2 = \frac{\left[\sum_{j=1}^{k} \lambda_j\right]^2}{\left[\sum_{j=1}^{k} \lambda_j\right]^2 + \sum_{j=1}^{k} \theta_{jj}^2}$$

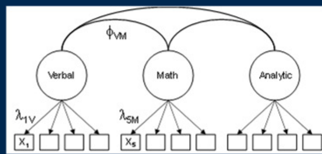
For the Need for Cognition data, the one-factor model parameter estimates yield reliability estimates that are very close to the traditional coefficient alpha:

```
alpha(NC_rescaled)

##
## Reliability analysis
## Call: alpha(x = NC_rescaled)
##
##   raw_alpha std.alpha G6(smc)
##        0.86      0.87    0.88
```

```
# Extract standardized loadings and error variances.
NC_Loadings <- inspect(NC_Fit_11, what = "std")$lambda
NC_error_variances <- diag(inspect(NC_Fit_11, what = "std")$theta)

# Standardized coefficient alpha, version 1.
Mean_Item_Reliability <- mean(NC_Loadings^2)
(18 * Mean_Item_Reliability)/(1 + 17 * Mean_Item_Reliability)

## [1] 0.8771

# Standardized coefficient alpha, version 2.
sum(NC_Loadings)^2/(sum(NC_Loadings)^2 + sum(NC_error_variances))

## [1] 0.8705
```



Inherent in the measurement model is the classical measurement theory notion of attenuation. The correlation between X_1 and X_5 is estimated by $\lambda_{1V}\phi_{VM}\lambda_{5M}$.



$$\sigma_{x_1 x_5} = \sigma_{(\lambda_{1V}\xi_V + \delta_1),(\lambda_{5M}\xi_M + \delta_5)}$$

$$\sigma_{x_1 x_5} = \lambda_{1V}\phi_{VM}\lambda_{5M}$$

$$r_{15} = \lambda_{1V}\phi_{VM}\lambda_{5M}$$

$$\lambda_{1V}^2 = r_{11}$$

$$\lambda_{5M}^2 = r_{55}$$

$$\phi_{VM} = \rho_{15}$$

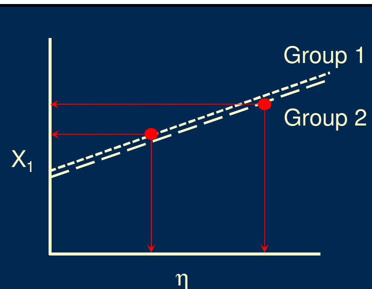
$$r_{15} = \rho_{15}\sqrt{r_{11} r_{55}}$$

Measurement Invariance

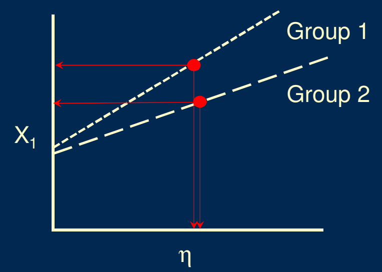
In the context of confirmatory factor analysis, the goals of cross-validation are met through procedures designed to demonstrate measurement invariance.

Measurement invariance is motivated by the desire to infer group differences in latent variable means (and perhaps variances and covariances) that are not an artifact of measurement differences between the groups.

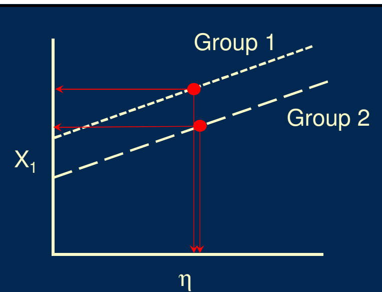
A scale has measurement invariance across groups if identical levels of the latent variable have the same expected raw-score on the measure.



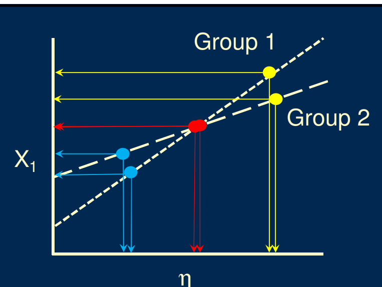
In this graph, the slopes (λ) and intercepts (τ) are the same, allowing group differences on the observed variable (X_1) to inform us about group differences on the latent variable (η).



The slopes (λ) are different in the two groups, making it possible that the groups could be the same on the latent variable (η) even though they exhibit differences on the observed variable (X_1).



The slopes (λ) are the same, but the intercepts (τ) are different in the two groups, making it possible that the groups could be the same on the latent variable (η) even though they exhibit differences on the observed variable (X_1).



If the slopes (λ) <u>and</u> the intercepts (τ) are different in the two groups, considerable error in inference about latent variable (η) differences is possible. Valid inferences about latent variables requires that the measures operate similarly in each group.

To determine if measurement invariance across samples is present, increasingly more stringent models are imposed and tested for equivalence.

The invariance testing process begins with estimation of the same basic factor model in each group. *Configural invariance* is said to exist if the same number of factors and similar patterns of loadings are found in all groups.

For this demonstration, the Holzinger & Swineford (1939) data will be examined. Nine cognitive abilities tests were administered to students at two schools. The tests are assumed to measure three correlated latent cognitive abilities: visual-spatial processing, text processing, and speed of processing.

```
HS.model <- '
Visual =~ x1+x2+x3
Textual =~ x4+x5+x6
Speed =~ x7+x8+x9
'
config <- cfa(HS.model,data=HS,missing="ML", estimator="MLR",
              likelihood="wishart",representation="LISREL",
              group="school")
summary(config,standardized=TRUE,rsq=TRUE,
        fit.measures = TRUE)
```

```
Group 1 [Pasteur]:

                        Std.all
Latent variables:
  Visual =~
    visual_perc           0.887
    cubes                 0.336
    lozenges              0.515
  Textual =~
    para_comp             0.823
    sent_compl            0.856
    word_meaning          0.838
  Speed =~
    speeded_add           0.547
    speeded_count         0.682
    speeded_disc          0.551

Covariances:
  Visual ~~
    Textual               0.484
    Speed                 0.299
  Textual ~~
    Speed                 0.325
```

```
Group 2 [Grant-Whi

                        Std.all
Latent variables:
  Visual =~
    visual_perc           0.684
    cubes                 0.517
    lozenges              0.695
  Textual =~
    para_comp             0.866
    sent_compl            0.829
    word_meaning          0.826
  Speed =~
    speeded_add           0.660
    speeded_count         0.800
    speeded_disc          0.702

Covariances:
  Visual ~~
    Textual               0.538
    Speed                 0.520
  Textual ~~
    Speed                 0.331
```

The loadings and latent variable correlations show similar patterns in the two groups. Configural invariance appears to hold.

The fit, however, is not terrific, so we would probably want to explore that further before proceeding. We'll forge ahead nonetheless in order to demonstrate the method.

```
User model versus baseline model:

  Comparative Fit Index (CFI)                    0.923       0.914
  Tucker-Lewis Index (TLI)                       0.885       0.871
Root Mean Square Error of Approximation:

  RMSEA                                          0.097       0.101
  90 Percent Confidence Interval          0.075  0.120       0.079  0.125
  P-value RMSEA <= 0.05                          0.001       0.000

Standardized Root Mean Square Residual:

  SRMR                                           0.068       0.068
```

The next constraint requires the factor loadings to be equal across groups. This constraint is called *metric* or *weak invariance* and tests whether respondents across groups attribute the same meaning to the latent constructs.

```
weak <- cfa(HS.model,data=HS,missing="ML",estimator="MLR",
            likelihood="wishart",representation="LISREL",
            group="school",group.equal="loadings")
summary(weak,standardized=TRUE,rsq=TRUE,fit.measures = TRUE)
```

This model is compared to the configural invariance model using a chi-square difference. Invariance holds if the test is not significant. Comparison of CFI is also recommended, with invariance indicated by a difference < .01.

```
## Scaled Chi Square Difference Test (method = "satorra.bentler.2001")
##
##         Df  AIC  BIC Chisq Chisq diff Df diff  Pr(>Chisq)
## config  48 7484 7707   115
## weak    54 7481 7681   123        6.5       6       0.367
```

```
User model versus baseline model:

  Comparative Fit Index (CFI)                    0.924       0.915
```

```
User model versus baseline model:

  Comparative Fit Index (CFI)                    0.921       0.917
```

Weak invariance holds; the latent variables are interpreted the same way in each sample.

Next the intercepts are constrained to be equal. This is called *scalar* or *strong invariance* and implies that the meaning of the constructs (the factor loadings), and the levels of the underlying items (intercepts) are equal in both groups. If this constraint holds, the groups can be compared on their latent variable scores.

```
strong <- cfa(HS.model,data=HS,missing="ML",
              estimator="MLR", likelihood="wishart",
              representation="LISREL",group="school",
              group.equal = c("loadings","intercepts"))
summary(strong,standardized=TRUE,rsq=TRUE,fit.measures = TRUE)
```

Implicit in this test (and the previous one) is the restriction that the latent means are constrained to be equal). Why?

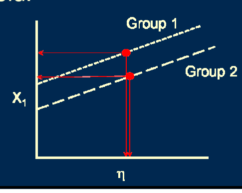
```
## Scaled Chi Square Difference Test (method = "satorra.bentler.2001")
##
weak     54 7481 7681   123        6.5       6       0.367
strong   60 7509 7687   163       40.6       6 0.000000352 ***
```

```
User model versus baseline model:

  Comparative Fit Index (CFI)                    0.924        0.915
```

```
User model versus baseline model:

  Comparative Fit Index (CFI)                    0.883        0.877
```

Strong invariance does not hold.

```
Intercepts:
                     Estimate  Estimate
    x1                  4.941     4.930
    x2                  5.984     6.200
    x3                  2.487     1.996
    x4                  2.823     3.317
    x5                  3.995     4.712
    x6                  1.922     2.469
    x7                  4.432     3.921
    x8                  5.563     5.488
    x9                  5.418     5.327
```



Finally, the residual variances are fixed to be equal across groups. This is called *strict invariance* and means that the explained variance for every item is the same across groups. In other words, the latent construct is measured identically across groups.
If error variances are not equal, groups can still be compared on the latent variable, but this is measured with different amounts of error between groups.

```
strict<- cfa(HS.model,data=HS,missing="ML",estimator="MLR",
             likelihood="wishart", representation="LISREL",
             group="school", group.equal=c("loadings",
             "intercepts","residuals"))
summary(strict,standardized=TRUE,rsq=TRUE,fit.measures = TRUE)
```

```
## Scaled Chi Square Difference Test (method = "satorra.bentler.2001")
##
strong   60 7509 7687   163       40.6       6 0.000000352 ***
strict   69 7508 7653   180       15.6       9       0.076 .
```

```
User model versus baseline model:

  Comparative Fit Index (CFI)                 0.883       0.877
```

```
User model versus baseline model:

  Comparative Fit Index (CFI)                 0.873       0.870
```

The residual variances are similar. Had we passed the previous test, strict invariance would hold.

Additional steps could impose constraints on the latent variable means, variances and covariances.

A convenient function can be used for the basic invariance tests, but it excludes strict invariance and replaces it with latent mean invariance:

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

```
Chi Square Difference Test

                Df  AIC  BIC Chisq Chisq diff Df diff Pr(>Chisq)
fit.configural 48 7484 7707   116
fit.loadings   54 7481 7681   124        8.2       6       0.22
fit.intercepts 60 7509 7687   164       40.1       6    4.4e-07
fit.means      63 7543 7710   205       40.5       3    8.3e-09
```

```
Fit measures:

                 cfi rmsea cfi.delta rmsea.delta
fit.configural 0.923 0.097        NA          NA
fit.loadings   0.921 0.093     0.002       0.004
fit.intercepts 0.882 0.107     0.038       0.015
fit.means      0.840 0.122     0.042       0.015
```

Invariance can be tested with measures collected over time. All of the invariance constraints must be specified explicitly.

```
SE_Model_3 <- '
SE_1 =~ a*r_1_1 + b*r_1_2 + c*r_1_3 + d*r_1_4 + e*r_1_5 +
        f*r_1_6 + g*r_1_7 + h*r_1_8 + i*r_1_9 + j*r_1_10
SE_2 =~ a*r_2_1 + b*r_2_2 + c*r_2_3 + d*r_2_4 + e*r_2_5 +
        f*r_2_6 + g*r_2_7 + h*r_2_8 + i*r_2_9 + j*r_2_10
r_1_1 ~ aa*1
r_2_1 ~ aa*1
r_1_2 ~ bb*1
r_2_2 ~ bb*1
r_1_3 ~ cc*1
r_2_3 ~ cc*1
r_1_4 ~ dd*1
r_2_4 ~ dd*1
r_1_5 ~ ee*1
r_2_5 ~ ee*1
r_1_6 ~ ff*1
r_2_6 ~ ff*1
r_1_7 ~ gg*1
r_2_7 ~ gg*1
r_1_8 ~ hh*1
r_2_8 ~ hh*1
r_1_9 ~ ii*1
r_2_9 ~ ii*1
r_1_10 ~ jj*1
r_2_10 ~ jj*1
SE_1 ~ 0*1
SE_2 ~ 0*1
'
```

Strong invariance for the self-esteem data.

When used wisely, confirmatory factor analysis is a powerful tool that can test well-specified models and compare competing models.

Like any statistical procedure, however, it can be biased when it is used in a more exploratory manner (as in model modification).

In those cases, careful cross-validation is necessary to insure the validity of the best fitting model.

Next time . . .

Multidimensional scaling