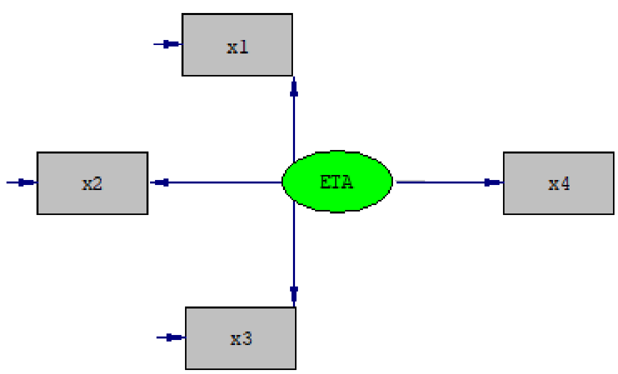
The one-factor congeneric measurement model

The most common type of measurement model is the one-factor congeneric measurement model, see Joreskog (1971b). A path diagram of this model is shown in the figure below.



The corresponding equations are written in matrix form as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} \xi + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix}$$

or

$$\mathbf{x} = \boldsymbol{\lambda}\xi + \boldsymbol{\delta}.$$

The model is empirically not directly verifiable since there are more unobserved variables than observed. However, with the assumption that the latent variable is standardized, the equations imply that the covariance matrix of the observed variables is of the form

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

In this equation, $\boldsymbol{\Theta}$ is a diagonal matrix with elements θ_{ii}, the variances of δ_i (i = 1, 2, 3, 4).

The hypothesis that the population covariance matrix has this form is testable from a random sample of observations. In addition, the following subhypotheses are testable.

The above model is called the *congeneric measurement* model. The measures $x_1, x_2, ..., x_q$ are said to be *congeneric* if their true values $\tau_1, \tau_2, ..., \tau_q$ have all pair-wise correlations equal to unity. This is true of the model, since $\tau_i = \lambda_i\xi = x_i - \delta_i$ (i = 1, 2, 3, 4) and all τ 's are linearly related and hence have unit correlation. The true variance in x_i is λ_i^2 and the reliability of x_i is

$$\rho_{ii} = \frac{\lambda_i^2}{\delta_{ii}} = \frac{\lambda_i^2}{\lambda_i^2 + \theta_{ii}} = 1 - \frac{\theta_{ii}}{\lambda_i^2 + \theta_{ii}}.$$

Strictly speaking, the error δ_i is considered to be sum of two uncorrelated random components s_i and e_i is the true measurement error. However, unless there are several replicate measures x_i with the same s_i, one cannot distinguish between these two components or separately estimate their variances. In consequence, ρ_{ii} is a lower bound for the true reliability.

Parallel measures have equal true score variances and equal error variances, i.e.,

$$\lambda_1^2 = \cdots = \lambda_4^2 \qquad \theta_{11} = \cdots = \theta_{44}.$$

Tau-equivalent measures have equal true score variances, but possibly different error variances.

Tests of parallelism and tau-equivalence are demonstrated in the following example.

In an experiment (Votaw, 1948) to establish methods of obtaining reader reliability in essay scoring, 126 examinees were given a three-part English Composition examination. Each part required the examinee to write an essay, and for each examinee, scores were obtained on the following:

1. the original part 1 essay
2. a handwritten copy of the original part 1 essay

3. a carbon copy of the handwritten copy in (2), and
4. the original part 2 essay.

Scores were assigned by a group of readers using procedures designed to counterbalance certain experimental conditions. The investigator would like to know whether, on the basis of this sample of size 126, the four scores can be used interchangeably or whether scores on the copies (2) and (3) are less reliable than the originals (1) and (4).

The covariance matrix of the four measurements is given in the command file below. The hypotheses to be tested are that the measurements are

1. parallel,
2. tau-equivalent, and
3. congeneric, respectively.

All analyses use the ML fit function.

The LISREL command file for this analysis is (**EX31A.LIS** in the **LISREL Examples** folder):

```
Analysis of Reader Reliability in Essay Scoring Votaw's Data
Congeneric model estimated by ML
DA NI=4 NO=126
LA
ORIGPRT1 WRITCOPY CARBCOPY ORIGPRT2
CM
25.0704
12.4363 28.2021
11.7257  9.2281 22.7390
20.7510 11.9732 12.0692 21.8707
MO NX=4 NK=1 LX=FR PH=ST
LK
Esayabil
OU SE ND=2
```

The DA command specifies four observed variables and a sample size of 126; the MA default is assumed, so the covariance matrix will be analyzed.

Labels for the input variables follow the LA command.

The CM command indicates that a covariance matrix is to be input. Because an external file is not specified, the matrix follows in the command file. A format statement does not appear, so the input is in free format.

The MO command specifies four *x*-variables and one latent variable; the elements of $\boldsymbol{\lambda}$ are all free (LX = FR), and the latent variable is standardized (PH = ST).

A label for the latent variable follows the LK command.

The OU command requests only standard errors (SE) as additional output, and the number of decimal places is set to two.

To obtain the input for the hypothesis of tau-equivalence, insert the command

```
EQUAL LX(1) – LX(4)
```

before the OU command. This specifies that the elements of $\boldsymbol{\lambda}$ should be equal.

The hypothesis of parallel measurements is specified by adding one more EQ command:

EQUAL TD(10 – TD(4)

See **EX31B.LIS** and **EX31C.LIS** for the other two models for which results are given in the tables below.

Table: Essay scoring data: summary of analyses

Hypothesis	*Df*	χ^2	*P*
(1) Parallel	8	109.12	0.000
(2) Tau-equivalent	5	40.42	0.000
(3) Congeneric	2	2.28	0.320

Table: Essay scoring data: results for congeneric model

I	$\hat{\lambda}_i$	$se\left(\hat{\lambda}_i\right)$	$\hat{\rho}_{ii}$
1	4.57	0.36	0.83
2	2.68	0.45	0.25
3	2.65	0.40	0.31
4	4.54	0.33	0.94

In the results of this analysis, as summarized in the first table, it is seen that the hypotheses (1) and (2) are untenable, but the hypothesis (3) is acceptable.

The results under the hypothesis (3) are given in the second table. The three columns of this table can be read off directly from the output for the ML solution. The reliabilities in column 3 appear where the output says “squared multiple correlations for *x*-variables”.

Inspecting the different λ ’s, it is evident that these are different even taking their respective standard errors of estimate into account. Comparing the reliabilities in the last column, one sees that they are high for scores

(1) and (4) and low for scores (2) and (3). Thus, it appears that scores obtained from originals are more reliable than scores based on copies.