

F5 - LISREL FOR DUMMIES

A five-step approach

V0.2 [September 12, 2012]

Harry BG Ganzeboom

David Nikoloski

Department of Sociology

VU University Amsterdam

Table of Contents

1.	Preparation: The conceptual model.....	3
1.1.	Latent variables	3
1.2.	Measurement model, causal structure and factor correlation.....	3
2.	A five-step LISREL guide.....	5
2.1.	Step 1: Feeding LISREL data.....	5
2.1.1.	Setting up example data	5
2.1.2.	Importing the correlation matrix from SPSS	6
2.1.3.	Constructing the correlation matrix using the Prelis-extension of LISREL.....	7
2.2.	Step 2: Initializing the model.....	9
2.2.1.	Selecting and re-ordering variables.....	9
2.2.2.	Model declaration	9
2.2.3.	Example	10
2.3.	Step 3: Model Specification	11
2.3.1.	Estimating parameters.....	11
2.3.2.	Setting measurement rules.....	11
2.3.3.	Example	12
2.4.	Step 4: Getting LISREL to run	13
2.4.1.	Guiding LISREL towards model convergence	13
2.4.2.	Output specification	13
2.4.3.	Example	14
2.5.	Step 5: F5.....	14
3.	Reading Output.....	15
3.1.	Reading the path diagram.....	15
3.2.	Reading the output file.....	16
3.2.1.	The FIT statistics	16
3.2.2.	Fitted covariance matrix and residuals.....	16
3.2.3.	Standardized Solutions.....	17
4.	Model adjustment	18
4.1.	Equality constraints	18
4.2.	Modification indices	19

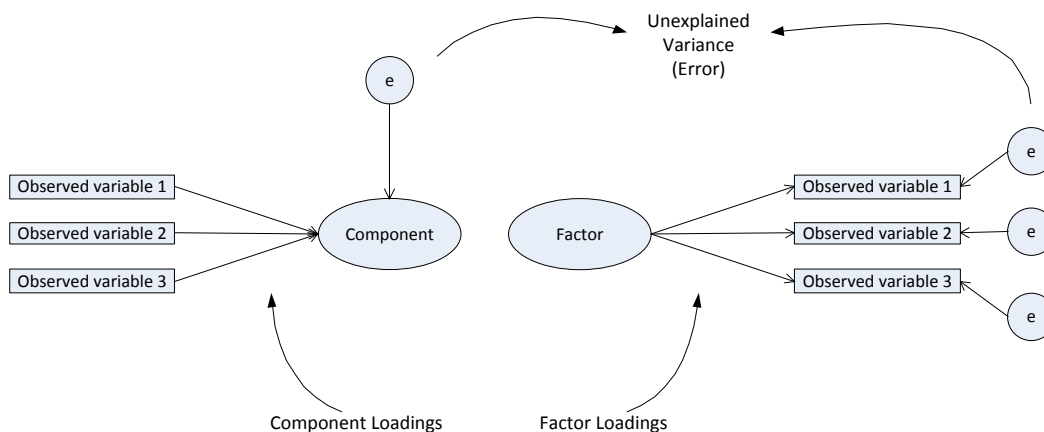
1. Preparation: The conceptual model

The basis of all LISREL modeling is the creation of a causal diagram that specifies how variables causally interact with each other, like in any regression model. In contrast to normal regression models, LISREL estimates parameters using simultaneous equations. This makes it possible to estimate many parameters in complex structures of interaction. Perhaps the most major advantage of this is that one can distinguish between (1) latent variables and (2) observed variables. This idea is based on the acknowledgement of the inevitability of measurement error. Because measurement is inherently flawed, the observed properties of the world should always be regarded as imperfect representations of the properties the researcher is actually interested in. The effect is that many sources of variance can possibly distort the observed causal variance the researcher intends to measure.

Because LISREL models are theory driven, the specification of the theoretical (starting) model is probably the most vital step in the process. One has to specify both the construction of ‘models of variables’ (called latent variables), and the theorized interaction between these variable-models.

1.1. Latent variables

The fundamental idea of factor- or component-analysis is based on the same premises, or the idea that one or more observed variables cause, or are caused by, one or more unobserved ‘latent’ variables, expressed as factors or components. If observed variables ‘cause’ a latent variable, we call this latent variable a *component*. If the observed variables ‘are caused by’ the latent variable, we call this latent variable a *factor*.



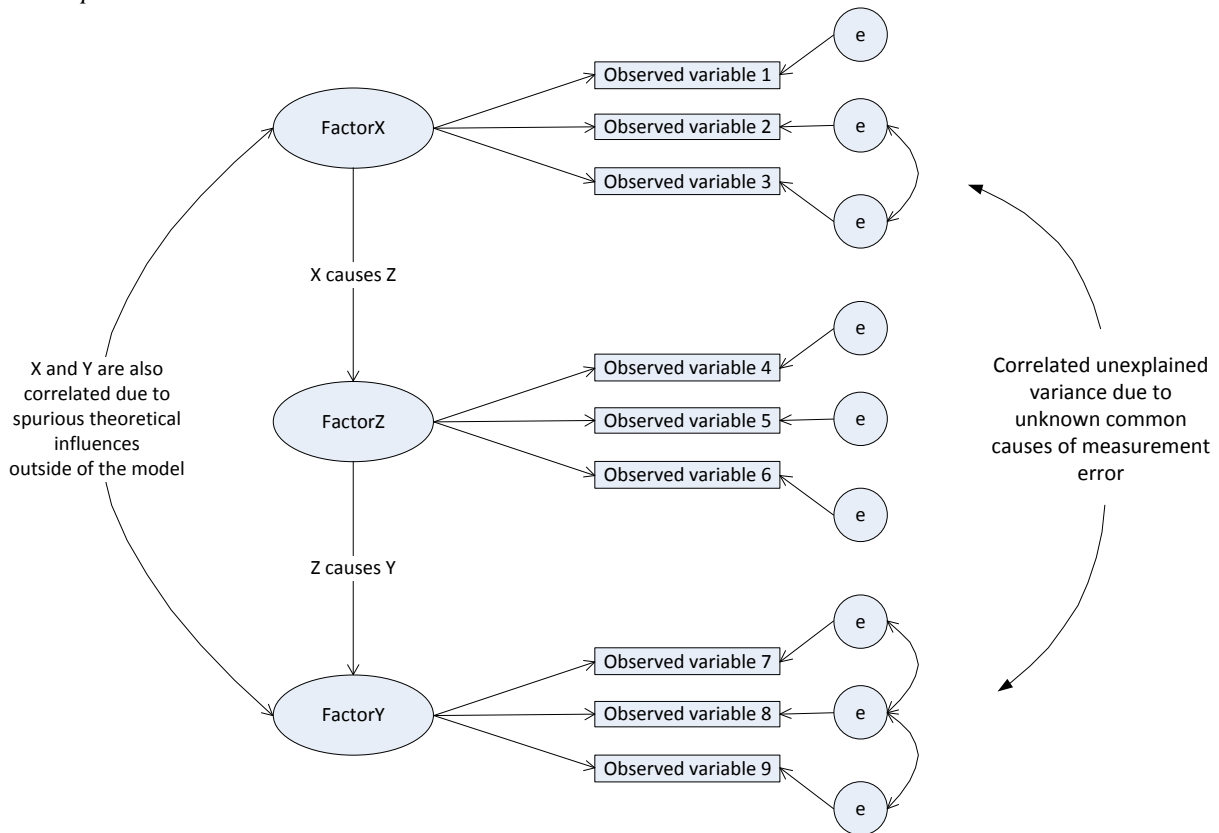
Note that component analysis is a descriptive method, seeking linear combinations of weighted observed variables such that maximum variance is found among these variables. It treats the observed variables as being perfect, and the latent variables as descriptions of shared variance amongst the observed variables. Factor analysis is a causal method, seeking the latent common cause of combined actual observations, treating the observed variables as imperfect effects of a perfect common cause.

Key insights here are twofold. The first is that this means that factors are meaningful constructs, while components are constructs without an understandable meaning. The second is then that, in a structural equation model for estimation of causal structures and measurement error simultaneously, we are looking for a model describing causal interaction between factors, not components.

1.2. Measurement model, causal structure and factor correlation

Based on the insights above, a factor model can be seen as a measurement model of the observed world, and the interaction between multiple factor models can be seen as the causal structure of the actual world (within the boundaries of the scope of the model). A full LISREL model describes both simultaneously. The left side

(interaction between factors) we call the *structural model*, and the right side (factor loadings and error terms) we call the *pattern model or measurement model*.



In the complete LISREL model, four different types of model parameters (arrows in the model diagram) can be specified, making them either (1) free for estimation, (2) fixed to a certain specified value, or (3) free for estimation, but constrained to be equal to other specified parameters. These types of model parameters are often referred to by characters of the Greek alphabet.

- Causal coefficients (Beta)
- Factor loadings (Lambda)
- Factor correlations (Psi)
- Error correlations (Theta)

Effectively, the model consists of a matrix of parameter estimations and specifications for each of the types of model parameters. The researcher has to have a general idea which of the parameters are theoretically predetermined and which of the parameters should be estimated. In order to think about all of these types of relations simultaneously, the researcher has to ask at least the following questions:

- In which causal (time-) sequence should the latent variables be? Which 'earlier' variable (cause) influences which 'later' variable (effect).
- Which observed variables are caused by which latent variables?
- Which latent variables correlate because they are (partially) caused by other causes, external to the model (real-world spuriousness)?
- Which portions of unexplained variance in the observed variables are caused by a common measurement bias external to the model (measurement-world spuriousness)?

2. A five-step LISREL guide

2.1. Step 1: Feeding LISREL data

The estimated parameters in the desired four matrices (Beta, Lambda, Psi and Theta) are estimated based on the total correlation matrix of all observed variables. LISREL needs this correlation matrix as input. One can construct the correlation matrix in two ways. The first is using other programs such as SPSS, and importing the resulting matrix into LISREL. The second is, using the 'Preliis'-extension of LISREL on the original data-set to calculate the correlation matrix inside LISREL.

2.1.1. Setting up example data

To create data that roughly follows the model in the Preparation section, you can use SPSS to generate a dataset as follows:

```
Latents          comp x = uni(1).
                 comp z = x + uni(1).
                 comp y = z + uni(1).
                 execute.

Factor/Error     comp fcorr = uni(1).
Correlations     comp ecorr1 = uni(1).
                 comp ecorr2 = uni(1).
                 comp ecorr3 = uni(1).
                 execute.

Correlated      comp xe = x + fcorr.
latents         comp ye = y + fcorr.
                 execute.

Observed         comp x1 = xe + uni(1).
variables       comp x2 = .8 * xe + ecorr1.
                 comp x3 = .6 * xe + ecorr1.
                 comp z1 = z + uni(.1).
                 comp z2 = z + uni(.3).
                 comp z3 = z + uni(.5).
                 comp y1 = ye + ecorr2.
                 comp y2 = .8 * ye + ecorr2 + ecorr3.
                 comp y3 = .6 * ye + ecorr3.
                 execute.

Standardize     descriptives x1 x2 x3 z1 z2 z3 y1 y2 y3 /save.

Save            save outfile= '[filename.sav]'
                /keep=Zx1 Zx2 Zx3 Zz1 Zz2 Zz3 Zy1 Zy2 Zy3.
```

In the following steps, you do not need this data. The researcher can write the whole syntax based on the picture in the Preparation section. Before going to the final step (F5) however, construct the dataset and choose one of the options available in the current step.

While the usual temptation is to pick the second option right away (and in this example you might as well), in other, less controlled, situations it is generally wise to prepare the data more thoroughly before analysis. Running the data-matrix through a program with more data-manipulation features, such as SPSS, may prevent a number of problems. In order to start the LISREL analysis based on proper data, there are a number of things to worry about before starting, such as missing values, outliers, measurement scales and directions, standardization, etc. These issues can distort (the interpretation of) regression coefficients in the four matrices, so they should be handled with care. LISREL itself offers little possibilities to properly deal with these things.

2.1.2. Importing the correlation matrix from SPSS

2.1.2.1. Preparation

In SPSS it is possible to write the correlation matrix to any of the supported file-types, such as .sav, .txt, or .xls, using the following procedure:

```

1          *****
          Write correlation to a new dataset
          *****
CORRELATIONS
          /variables = Zx1 Zx2 Zx3 Zz1 Zz2 Zz3 Zy1 Zy2 Zy3
          /print=nosig
          /matrix = out('c:\temp\correlations.sav').

2          *****
          Clean the correlation matrix and write to text file
          *****
GET FILE='c:\temp\correlations.sav'.
SELECT if rowtype_='CORR'.
SAVE translate
          /outfile= 'c:\temp\correlations.txt'
          /type=tab
          /cells=values
          /textoptions delimiter=', ' decimal=dot format=plain
          /drop rowtype_ varname_
          /replace.

```

This example writes only the correlation coefficients of the correlation matrix to a tab and comma delimited text file that can be directly imported into LISREL using the following procedure that (1) declares the properties of the data, (2) assigns variable names to the observed variables, and (3) specifies the file from which the data should be read.

2.1.2.2. Basic commands

The DATA statement declares the structure of the correlation matrix, telling LISREL how many observed variables are based on how many observations, how many groups are acknowledged, and whether a correlation or covariance matrix is used.

1	DA NI=kk NO=nnn [NG=gg] [MA=KM]
kk	Number of variables in the correlation matrix.
nnn	Number of observations in the datamatrix the correlation matrix is based on. LISREL does not know this number, because it uses only the correlations as input. Most of the inferential statistics are based on this number! When pairwise deletion is used for missing values, the number of observations becomes arbitrary. In this case, either the minimum or the median N should be chosen.
gg	Number of groups to analyze (usually 1).
MA=KM	This means that the input is a correlation matrix. For more advanced users, the possibility exists to import a covariance matrix (MA=CM), existing of a correlation matrix and standard deviations.

The LA statement labels the observed variables. This is necessary for the output that LISREL produces.

2 **LA**
 varname varname [varname] (etc) /

varname Names of variables are specified in the correct order of the input matrix, separated by a space. Variable labels have a maximum length of eight characters. The variable labels have to start on the next line, after the LA statement.

/ The list of variables needs to be terminated by a '/'. This way, LISREL knows when the last observed variable has been declared.

The KM statement specifies the location of the file from which the correlation matrix should be read.

3 **KM FU file=[filename]**

FU This tells LISREL that the correlation matrix is 'full', meaning that both above and below the diagonal, values are provided. The opposite keyword is SY, meaning that only the values below the diagonal are provided, and that the correlation matrix is symmetric.

[filename] The file has to be in the same location as the LISREL syntax. Make sure that the complete path to the data-file does not contain any spaces or special characters, because LISREL cannot deal with them.

2.1.2.3. Example

In the example, used in this tutorial, there are 9 observed variables with an assumed number of observations of 1000, from which we constructed a correlation matrix in SPSS. This leads to the following syntax in LISREL:

```
Step 1        DA NI=9 NO=1000 NG=1 MA=KM
              LA
              Zx1 Zx2 Zx3 Zz1 Zz2 Zz3 Zy1 Zy2 Zy3
              KM FU file= 'c:\temp\correlations.sav'
```

2.1.3. Constructing the correlation matrix using the Prelis-extension of LISREL

The second way to feed data into LISREL is using the Prelis-extension that is provided with the more recent versions of LISREL. Prelis can read data matrices in many formats, such as text/CSV, SPSS, Stata, SAS, XLS, etc, and convert them into a format that LISREL can read directly. Still it is advisable to properly 'prepare' the data matrix in SPSS or another tool before reading it into Prelis, but major advantages here are:

- The number of cases can be fed into LISREL automatically, although a number must still be provided. The output file shows that LISREL ignores the provided number when the actual number of observations is different from what the user declared.
- One can choose more flexible between a correlation matrix and a covariance matrix, because Prelis calculates the input matrix upon execution of the model-syntax.
- Missing values can be handled by Prelis in a number of ways, such as listwise deletion, pairwise deletion and full-information maximum likelihood (FIML).

2.1.3.1. Preparation

The first thing to do is read the data matrix into Prelis, in order to construct a .psf dataset file. In order to do this, in the file menu of LISREL, choose 'Import data'.

After a file is chosen, the program asks the user to specify a Prelis file to store the data in. After saving the .psf file, the researcher can see the data, much like in any spreadsheet program, and use several statistical functions for exploring the data under the Statistics menu. Under the Transformation menu, variables can be computed and recoded. Under the Data Menu, functions are available to rename data variables and variable order.

Most researchers still use SPSS for such data manipulation and preparation, but it is good to know that also in LISREL, some such functions are available. If the data is already prepared in SPSS, nothing more than opening the file and saving it as a .psf file is necessary in order to start LISREL analysis.

2.1.3.2. Basic commands

The basic commands in this version of step 1 are similar to the commands for reading a correlation matrix. The DATA statement still declares the structure of the correlation matrix, telling LISREL how many observed variables are based on how many observations, how many groups are acknowledged, and whether a correlation or covariance matrix should be used. But now the LA and KM statements are obsolete, and replaced by the RA statement, that creates a correlation or covariance matrix as specified by the DATA statement from a .psf datafile.

RA FI= [filename]

filename The file has to be in the same location as the LISREL syntax. Make sure that the complete path to the data-file does not contain any spaces or special characters, because LISREL cannot deal with them.

2.1.3.3. Example

In the example, used in this tutorial, there are 9 observed variables with an assumed number of observations of 1000, in a .psf file. This leads to the following syntax in LISREL:

```
Step 1        DA  NI=9 NO=1000 NG=1 MA=KM  
             RA  = 'c:\temp\data.psf'
```


2.2. Step 2: Initializing the model

2.2.1. Selecting and re-ordering variables

The SELECT statement declares which variables are used in the analysis, and in which order these variables should appear in the output. This can be done by using the variable names and numbers from the Prelis file or, in case of a correlation matrix input, by using the variable names and numbers as specified by the LA statement.

SE	varname/varno varname/varno [varname/varno] (etc) /
varname/ varno	Names or numbers of variables, specified in the correct order the researcher wishes to get in the output, separated by a space. Variable labels have a maximum length of eight characters. The variable labels or numbers have to start on the next line, after the SELECT statement.
/	The list of variables needs to be terminated by a '/'. This way, LISREL knows when the last observed variable has been declared.

LISREL offers the possibility to distinguish between exogenous (X) variables and endogenous (Y) variables, but usually only the Y side is used for creating the complete causal model. The advantage of using the X side apart from the Y side is that the resulting path diagram of the model looks somewhat cleaner when analyzing complex causal structures. In this case, the Y variables must be listed before the X variables in the SELECT statement.

2.2.2. Model declaration

The MODEL statement declares the basis from which the model, specified in step 0, will be designed from. First of all, the numbers of variables of each kind need to be specified. Then, the starting configuration, of each of the four effect matrices that have to be estimated, can be set. The four effect matrices are represented in LISREL by two-character codes, as follows:

Parameter type	Greek letter	LISREL code	Matrix description
Causal coefficients	Beta	BE	latent x latent
Factor loadings	Lambda	LY	observed x latent
Factor correlations	Psi	PS	latent x latent
Error correlations	Theta	TH	observed x observed

MO NY=ny NX=nx NE=ne NK=nk [BE=fu,fi] [LY=fu,fi] [PS=sy,fi] [TE=di,fi]	
ny, nx, ne, nk	Respectively, the number of observed Y-variables, observed X-variables, latent Y-variables, and latent X-variables in the correlation matrix
fu, sy, di	'fu' declares that the matrix is full - a saturated matrix is modeled with arrows in both directions. 'sy' declares that the matrix is symmetrical - a saturated matrix is modeled with arrows in one direction. 'di' declares that the matrix consist only of the diagonal - This offers the possibility to free error terms on the PSI-side (latents) or the TE-side (observed), but excluding the possibility for correlations between these error terms
fi, fr	'fi' declares that the matrix initially consists of fixed parameters that do not have to be estimated. Parameters that need to be estimated have to be 'freed' manually later. 'fr' declares that the matrix initially consists of free parameters that have to be estimated. This option usually is only used for the factor correlations (PS). When the factor loadings (LY) or error correlations (TE) matrix is saturated, model identification becomes immediately problematic. The researcher needs to manually 'fix' some parameters later in order to achieve model identification.

The latent variables can be labeled, using the LE statement for the endogenous side (Y) and the LK statement for the exogenous side (X).

LE/LK	
varname [varname] [varname] (etc) /	
varname	Names latent Y variables, specified in the correct order the researcher wishes to get in the output, separated by a space. Variable labels have a maximum length of eight characters. The variable labels or numbers have to start on the next line, after the SELECT statement.
/	The list of variables needs to be terminated by a '/'. This way, LISREL knows when the last observed variable has been declared.

2.2.3. Example

Based on our earlier example, we now expand the model with step 2.

```
Step 1    DA  NI=9 NO=1000 NG=1 MA=KM
          RA = 'c:\temp\data.psf'
Step 2    SE
          Zx1 Zx2 Zx3 Zz1 Zz2 Zz3 Zy1 Zy2 Zy3 /
          MO NY=9 NE=3 BE=fu,fi LY=fu,fi PS=sy,fr TE=sy,fi
          LE
          FactorX FactorZ FactorY /
```

2.3. Step 3: Model Specification

2.3.1. *Estimating parameters*

Let's suppose we initialized all four effect matrices as fixed in the MODEL statement, meaning that no parameters are estimated without manual specification. Parameters can be estimated by freeing the corresponding matrix cells. This can equally be viewed as drawing arrows in the theorized path diagram from step 1. Arrows are 'freed' by using the FREE command. Alternatively, when the effect matrix was initialized as free, arrows can be 'fixed' by using the FIXED command.

FR/FI matrix(var, var)

matrix	The matrix identifier (BE, LA, PS, or TE)
var, var	The identification of the matrix cell, in the form (row, column). In terms of arrows in the LY and BE parts of the path diagram, (var, var) means (end, start). For the PS and TE parts of the diagram, arrows are bidirectional and the matrix is symmetrical around the diagonal, so the order is arbitrary.

2.3.2. *Setting measurement rules*

When a latent variable causes multiple indicators in the model, the factor loadings (LY) and residual variances (TE) should be estimated. This can be done in two ways: through standardization or through the use of a measurement standard. The easiest form is through standardization, freeing all theorized loadings. In this case, LISREL assumes that the latent variable has a standardized measurement scale. This default specification sometimes causes numerical problems in LISREL.

The other way is to use one of the factor loadings of each factor as a reference effect. In this case, these parameters should stay fixed and get a value of 1.

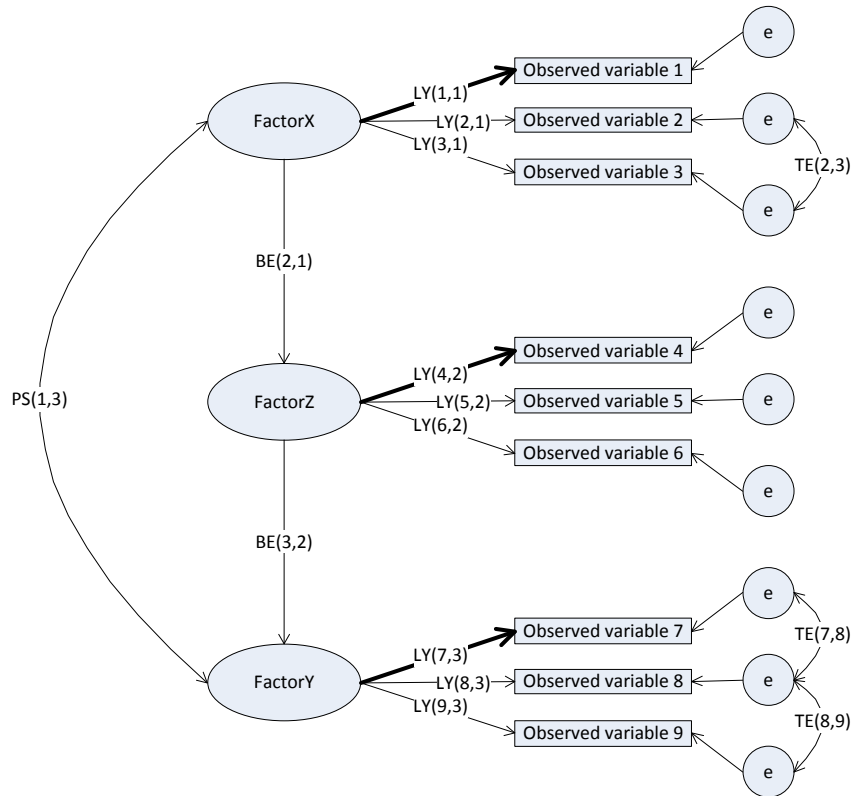
VA value [parameter] [parameter] (etc.)

Value	The value that should be given to a fixed parameter
Parameter	The parameter in the form: matrix(var, var).

The cleanest interpretation of the model is reached when the highest factor loading gets the 1-value. When guessing this incorrectly, this will result in factor loadings greater than 1. The 'measurement fix' can then be changed to the highest factor loading.

2.3.3. Example

Addressing the arrows in the path diagram according to the syntax rules, the path diagram below shows what we wanted. The bold arrows specify the measurement scale of the latent variables.



In our example syntax, we free these matrix cells as follows:

```

Step 1    DA NI=9 NO=1000 NG=1 MA=KM
          RA = 'c:\temp\data.psf'
Step 2    SE
          Zx1 Zx2 Zx3 Zz1 Zz2 Zz3 Zy1 Zy2 Zy3 /
          MO NY=9 NE=3 BE=fu,fi LY=fu,fi PS=sy,fr TE=sy,fi
          LE
Step 3    FactorX FactorZ FactorY /
          VA 1  LY(1,1) LY(4,2) LY(7,3)

          FR  LY(2,1) LY(3,1) LY(5,2) LY(6,2) LY(8,3) LY(9,3)
          FR  BE(2,1) BE(3,2)
          FR  PS(1,3)
          FR  TE(2,3) TE(7,8) TE(8,9)
  
```

2.4. Step 4: Getting LISREL to run

2.4.1. Guiding LISREL towards model convergence

LISREL estimates the model iteratively. This means that it starts with some collection of starting values for all free parameters and assesses model fit. Then, LISREL changes values of parameters in an intelligent way many times, and calculates whether the changes improve the fit each time. It continues to do this until improvements of fit between changes become negligible. We then say the model has converged. LISREL also terminates iterations after the maximum number of iterations is reached and the model has not (yet) converged.

In order to help LISREL in the right direction, smart starting values must be provided, using the START statement.

ST value [parameter] [parameter] (etc.)	
Value	The starting value that should be given to a free parameter
Parameter	The parameter in the form: matrix(var, var).

Usually, setting starting values very specifically is not necessary using modern computers. Setting all starting values to 0.5 should do the trick in most cases.

2.4.2. Output specification

2.4.2.1. *The Path Diagram*

The PD command on the before-last line tells LISREL that it should create the path diagram. If the path diagram, like in example 3.3 does not appear after step 5, something technical is still wrong with the model. We will turn to these kinds of problems in the Problems appendix.

PD
(no options)

2.4.2.2. *The OUTPUT statement*

The LISREL syntax ends with a specification of the output options. While the list of possible options is very long (see Lisrel manual or Kelloway), a good generic choice is to use the following:

OU ML MI SC TV AD=OFF ND=nd RS	
ML	Use maximum likelihood for estimation
MI	Print modification indices
SC	Print standardized solutions
TV	Print standard errors and T-values
AD=off	Continue iterations even if it seems that convergence does not occur
ND=nd	Print [nd] decimals
RS	Print estimated correlation-matrix and residuals

2.4.3. Example

```
Step 1      DA  NI=9 NO=1000 NG=1 MA=KM
           RA = 'c:\temp\data.psf'
Step 2      SE
           Zx1 Zx2 Zx3 Zz1 Zz2 Zz3 Zy1 Zy2 Zy3 /
           MO NY=9 NE=3 BE=fu,fi LY=fu,fi PS=sy,fr TE=sy,fi
           LE
           FactorX FactorZ FactorY /
Step 3      VA 1  LY(1,1) LY(4,2) LY(7,3)
           FR  LY(2,1) LY(3,1) LY(5,2) LY(6,2) LY(8,3) LY(9,3)
           FR  BE(2,1) BE(3,2)
           FR  PS(1,3)
           FR  TE(2,3) TE(7,8) TE(8,9)
Step 4      ST 0.5 ALL
           PD
           OU ML MI SC TV AD=OFF ND=3 RS
```

2.5. **Step 5: F5**

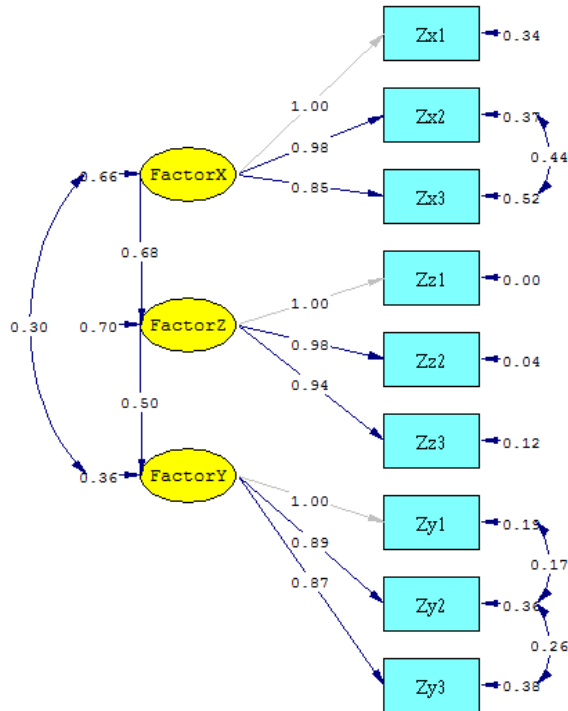
By pressing F5 LISREL is started.

Upon running, LISREL automatically saves (and overwrites!) all the files of the project (usually the .ls8 syntax file and the .pth path diagram file). But before running LISREL, we need data, so we can discuss the results files. If you haven't created the .psf file or correlation matrix from the SPSS-generated data in paragraph 2.1.1, this would be the time to do so.

3. Reading Output

3.1. Reading the path diagram

If the path diagram appears after pressing F5, the model has proven to be able to be estimated. This means that the LISREL syntax works correctly. It does, however, not mean that the estimated values are the correct ones.



Chi-Square=18.84, df=21, P-value=0.59568, RMSEA=0.000

The path diagram shows two goodness-of-fit statistics: the χ^2 with associated p-value and the RMSEA. The χ^2 has to be as low as possible and the RMSEA needs to be close to or lower than .05. Our model has perfect fit statistics, so our model is correct, but this is of course because we generated the data. In practice, such perfect fit statistics are seldom found.

More detailed fit statistics can be found in the output file. The path diagram is produced to compare the picture with the one you drew in the preparation phase. Or in other words, to check if your syntax actually describes what you wanted to describe.

Note that the path diagram in the example looks so neat because all the variables were declared in the right order. It should be clear that if the variables were ordered differently with the same associations, the picture would look like spaghetti.

The PD can also be used by itself to generate LISREL code, based on drawn alterations in the path diagram. It is possible to generate .lpj files. This can save some typing work if certain groups of parameters are (carefully!!) copied from the .lpj file to the .ls8 syntax file. However, in practice the researcher usually ends up with more problems than benefits, so it is highly advisable to just follow the five steps in this document. That way, a researcher always knows exactly what he is doing with regard the matrices that are observed in the output file.

3.2. Reading the output file

The output file contains much more information, based on the options in the output command, specified in step 4. The following sections are present in our example:

- Copy of the syntax
- Overview of the model command
- Covariance matrix
- Detailed specification of all estimated parameters
- Parameter estimates of all used matrices (in our example LY, BE, PS and TE), with t-values and standard errors
- Goodness of fit statistics
- Fitted covariance matrix and residuals
- Standardized solution
- Completely standardized solution

If there are problems, warnings and errors are put in the appropriate sections. The estimates and fitted residuals can provide valuable information about what is wrong. The estimates are usually reported with their standard errors or t-values. Also the goodness-of-fit statistics should be reported always. They ‘prove’ that the estimates are ‘correct’.

3.2.1. *The FIT statistics*

There are many goodness-of-fit statistics and many writers that argue for benefits of the one over the other (see Kelloway). Two fit statistics are really important:

- The χ^2 with connected degrees of freedom: This needs to be as small as possible, and preferably non-significant. Non-significance here means that the empirical correlation matrix does not differ significantly from the fitted (modeled) covariance matrix. Note that the χ^2 by definition increases with the number of cases. On large datasets it is usually impossible to achieve a small and non-significant χ^2 .
- The RMSEA and connected Test of Close Fit: This is a weighed function of the residual correlations and the number of cases, eliminating the problems that the χ^2 suffers from. The norm-value of the RSMEA is 0.5. The test shows if the difference between the fitted covariance matrix and the empirical data-matrix is lower than this norm-value. If this test is non-significant, researchers can accept the model as having a ‘good fit’.

Another index that is often shown in the literature is the AGFI. It needs to be larger than 0.95.

Significance-tests of LISREL models are somewhat dubious. The ask questions of ‘true-or-false’, where essentially questions of ‘how much’ drive these models. The inferential statistics are essentially based on the number of cases in the dataset.

3.2.2. *Fitted covariance matrix and residuals*

The fitted covariance matrix shows correlations between the observed variables that the model predicts, based on its estimates. The residuals show the differences between the estimated correlations in the correlation matrix and the actually observed correlations in the starting matrix. A positive value means that the modeled correlation is overestimated, a negative value that it is underestimated. Both very large residuals and patterns of residuals of the same sign are important bits that can guide model improvement if necessary. However, residuals do not always put the researcher in the right direction. The maximum-likelihood estimation procedure has a tendency to evenly divide lack-of-fit across multiple residuals. Also with regression estimation procedures it is true that the most influential points not always have a high residual.

3.2.3. Standardized Solutions

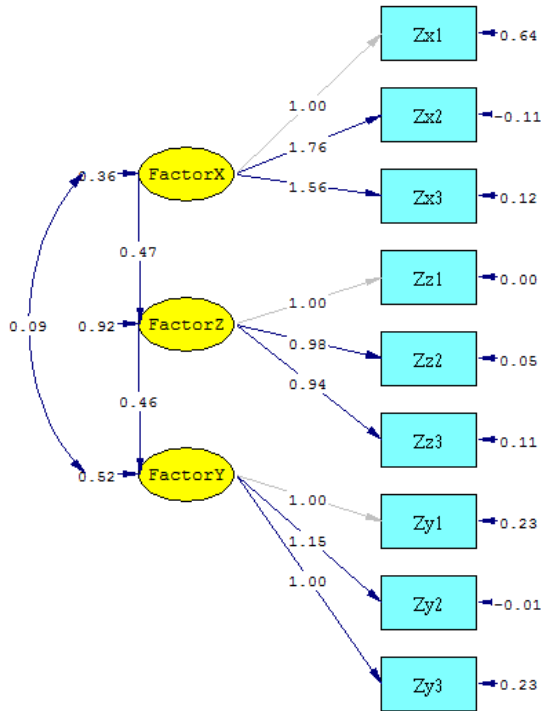
There are two types of standardized solutions:

- Standardized Solution: only the latent variables are standardized.
- Completely Standardized Solution: also the observed variables are standardized.

If MA=KM was specified in step 1, and the measurement scale for latent variables was achieved through standardization in step 3, both of the solutions are equal. Also, when the variables are standardized in SPSS before entering LISREL, like in our example, the solutions are equal.

4. Model adjustment

When the model has proven able to be estimated, the actual modeling usually begins. Often the initial model does not fit well. Suppose that in our model we had overlooked the error correlations, like in the path diagram below. The RSMEA shows a very poor fit, and other parameters change pretty dramatically.



Chi-Square=326.59, df=24, P-value=0.00000, RMSEA=0.147

On real data with many sources of variance, most of the work entails adding and removing parameters, and comparing the χ^2 and RSMEA between models (the critical values of χ^2 for 1 df are 3.84 ($p=.05$) and 6.64 ($P=.01$)). This procedure is basically an iterative repetition of steps 3 through 5. Another tool can be the setting of equality constraints on parameters.

4.1. Equality constraints

With the EQUAL statement, two or more free parameters are forced to be estimated at the same value.

EQ matrix(var, var) matrix(var, var) [matrix(var, var)] (etc.)

matrix	The matrix identifier (BE, LA, PS, or TE)
var, var	The identification of the matrix cell, in the form (row, column). In terms of arrows in the LY and BE parts of the path diagram, (var, var) means (end, start). For the PS and TE parts of the diagram, arrows are bidirectional and the matrix is symmetrical around the diagonal, so the order is arbitrary.

For example, if three parameters are estimated at the same value, this costs only 1 df instead of 3. The possibility for such constrained estimations is one of the major strengths of structural equation based methods over other forms of regression and factor analysis. They add parsimony to models, and when the equality assumption is

correct, they reduce the number of estimated parameters and the standard errors. If the assumption is wrong, they can also reduce model fit pretty dramatically.

4.2. Modification indices

If MI was included in the OU statement (step 4), the output also gives hints about which parameters can be freed that increase model fit. These modification indices are organized in the usual four matrices (LY, BE, PS and TE). Higher values mean a higher positive impact on model fit. This shows in the similar Expected Change section that shows the Chi^2 reduction for these parameters. LISREL also reports the parameter with the highest modification index separately. This can help the researcher along somewhat in the modeling process.