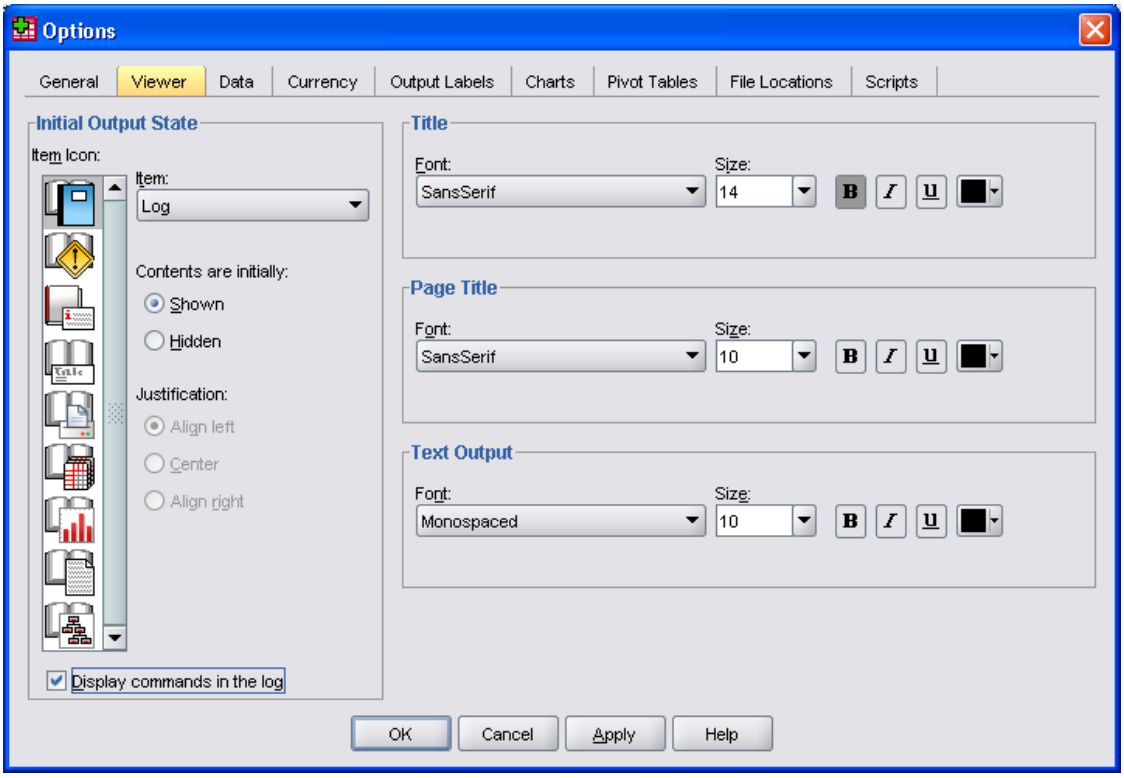
A Brief Introduction to SPSS Factor Analysis

SPSS has a procedure that conducts exploratory factor analysis. Before launching into a step by step example of how to use this procedure, it is recommended that the analyst save the syntax for each factor analysis run for bookkeeping purposes. A quick way to have the syntax nested within the output file is to select commands to be displayed in the log.

This is done by going to Edit → Options → Viewer. In the Viewer tab, shown below, check the box on the bottom left corner Display commands in the log. Hit Apply and then OK to close out the window. Once that option is checked, all syntax generated from will be included in the output.



Reading in Data

There are many ways to read in data into SPSS, and the text import wizard works well for many formats. It may be accessed by going to File → Open → Data. From there, locate the file name and type, and click away. However, this wizard is set up to only read case wise data. To read in covariance or correlation matrices, syntax should be used. To create a syntax file, go to File → New → Syntax. A syntax dialogue box will open and be ready for commands to be typed in.

In SPSS, commands end with a period (.) instead of the semicolon (;) in SAS or Mplus. Additionally, SPSS will not be able to handle comments nested within code. However, comments can be made on separate lines beginning with * or comment and ending with a period (.).

Correlation Matrix:

```
comment read in correlation matrix.
matrix data variables=rowtype_ mechanics vectors algebra analysis statistics.
begin data
N 88 88 88 88 88
corr 1.00000
corr 0.55341 1.00000
corr 0.54675 0.60964 1.00000
corr 0.40939 0.48508 0.71081 1.00000
corr 0.38910 0.43645 0.66474 0.60717 1.00000
end data.
```

The first line in the syntax is a comment. The second line defines the type of data as well as the labels used for each column of data. The third command uses 8 lines where the number of observations for each column is entered followed by the triangular correlation matrix. Note that each line of data begins with corr.

Covariance Matrix:

```
comment read in correlation matrix.
matrix data variables=rowtype_ mechanics vectors algebra analysis statistics.
begin data
N 88 88 88 88 88
cov 305.768
cov 127.223 172.842
cov 101.579 85.157 112.886
cov 106.273 94.673 112.113 220.38
cov 117.405 99.012 121.871 155.536 297.755
end data.
```

The only difference between inputting a correlation matrix versus a covariance matrix is that cov in place of corr is used prior to listing elements of the covariance matrix.

Doing the Factor Analysis

There are two approaches to get SPSS to conduct a factor analysis:

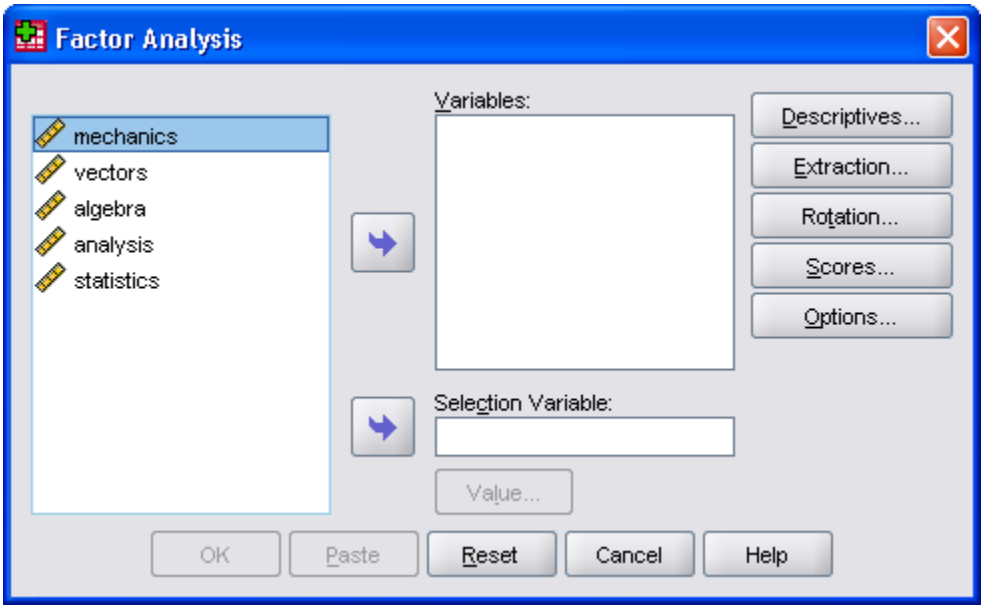
(1) Windows or dialogue boxes

(2) Syntax

Using Dialogue Boxes:

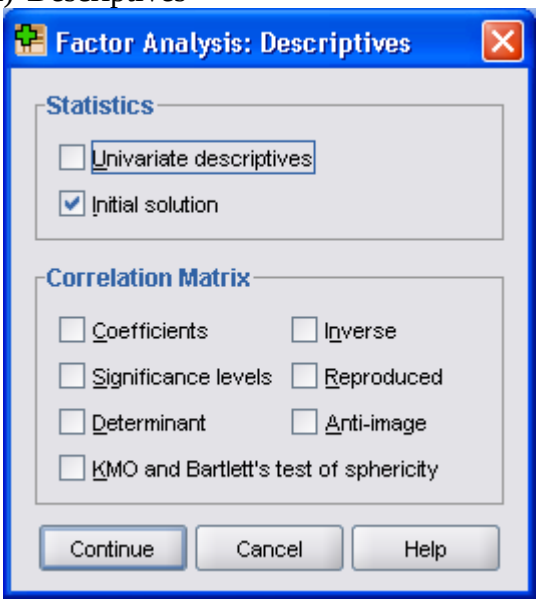
The factor procedure is located in Analyze → Data Reduction → Factor. First, one has to select the variables to be factor analyzed by highlighting the variables on the left box and hitting  to shift them over to the Variables: box. Selection Variable: is for a grouping variable when the factor analysis is to be done to separate groups as identified by the concerned variable. Value... is used to identify the nominal scale numbers to a group name. A snapshot of the Factor Analysis dialogue box is shown below:

Another way to keep a log of syntax used for analyses is to click on the Paste button. That will automatically generate syntax within a syntax file associated with the selected options for the current analysis.



Here is a quick and dirty run down and explanation of the other options in the procedure. Note that the screen shots of the dialogue boxes are the set defaults in SPSS.

(a) Descriptives



Univariate descriptives reports means, standard deviations and sample size of all the selected variables.

Initial solution reports the starting values of communalities used.

Coefficients – prints out the correlation matrix

Significance values – prints out 1-tailed probabilities of these correlation coefficients

Determinant – prints out the determinant of the correlation matrix

KMO and Barlett's test of sphericity – provides two test statistics. The Kaiser-Meyer-Olkin measure of sampling adequacy tests whether the partial correlations among variables are small. Bartlett's test of sphericity tests whether the correlation matrix is an identity matrix, which would indicate that the factor model is inappropriate

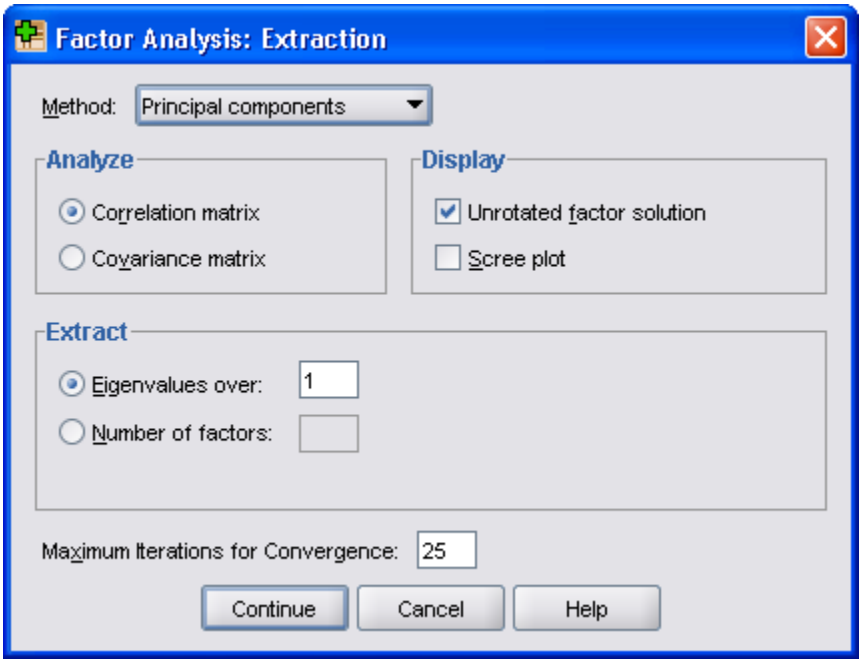
Inverse – prints out the inverse of the correlation matrix

Reproduced – prints out the estimated correlation matrix from the factor solution. The residuals (difference between estimated and observed correlations) are also displayed.

Anti-image – the anti-image correlation matrix contains the negatives of the partial correlation coefficients, and the anti-image covariance matrix contains the negatives of the partial covariances.

(b) Extraction

Method – specifies the method of extraction which are principal components, unweighted least squares, generalized least squares, maximum likelihood, principal axis factoring and image factoring.



Under Analyze, either the Correlation matrix or Covariance matrix may be selected.

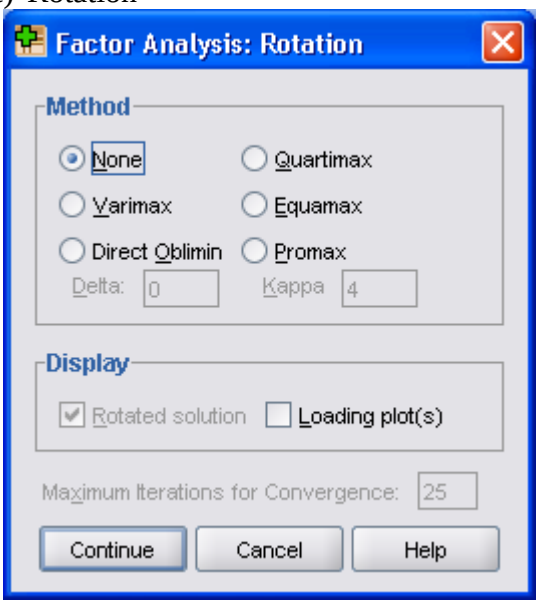
Unrotated factor solution – prints out the unrotated pattern matrix

Scree plot – plots the eigenvalues in descending order.

The number of factors extracted may be based on either Eigenvalues over a set number or a specified Number of Factors.

The Maximum Iterations of Convergence may also be changed.

(c) Rotation



There available rotations are None, Varimax, Quartimax, Equamax, Direct Oblimin and Promax.

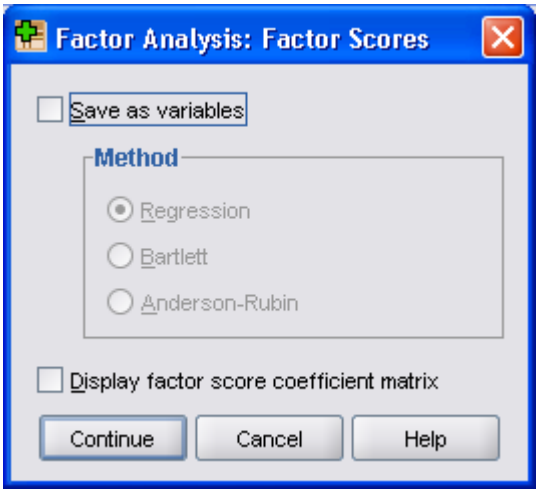
With Direct Oblimin, when Delta equals 0 (the default), solutions are most oblique. As Delta becomes more negative, the factors become less oblique.

With Promax, Kappa specifies the power used in the algorithm. The default for Kappa is 4.

Rotated solution – prints out the rotated pattern matrix and factor transformation matrix for orthogonal rotations. For oblique rotations, the pattern, structure, and factor correlation matrices are displayed.

Loading plot(s) – displays the three-dimensional factor loading plot of the first three factors. For a two-factor solution, a two-dimensional plot is shown. The plot is not displayed if only one factor is extracted. Plots display rotated solutions if rotation is requested.

(d) Scores

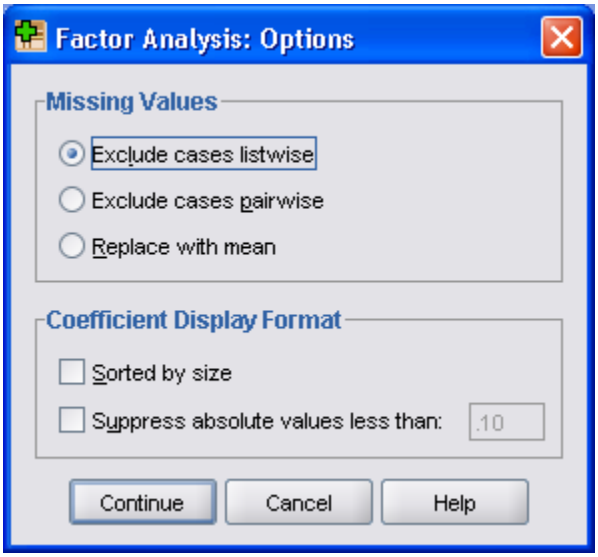


Save as variables – creates a new variable in the working data file with factor scores derived from the chosen method.

The three methods to obtain factor scores are Regression, Bartlett and Anderson-Rubin.

Display factor score coefficient matrix –shows the coefficients by which variables are multiplied to obtain factor scores. Also shows the correlations between factor scores.

(e) Options



Missing Values may be handled by three ways: Exclude cases listwise, Exclude cases pairwise and Replace with mean.

Sorted by size – displays the elements within the output matrices in descending order.

Suppress absolute values less than: - suppresses the display of absolute values that are less than the specified value.

Syntax:

```
FACTOR
  /VARIABLES mechanics vectors algebra analysis statistics
  /MISSING LISTWISE
  /ANALYSIS mechanics vectors algebra analysis statistics
  /PRINT UNIVARIATE
            INITIAL
            CORRELATION
            SIG
            DET
            KMO
            INV
            REPR
            AIC
            EXTRACTION
            ROTATION
            FSCORE
  /FORMAT SORT
  /PLOT EIGEN ROTATION
  /CRITERIA FACTORS(2) ITERATE(25)
  /EXTRACTION ULS
  /CRITERIA ITERATE(25)
```

*variables to be used.
*missing values to be listwise deleted.
*variables to be analyzed.
*print means, *SD*s and sample size.
*print initial communalities.
*print correlation matrix.
*print p-values associated with correlations.
*Print determinant
*conduct KMO test of sphericity
*print inverse matrix
*print reproduced correlations and residuals
*print anti-image correlation matrix
*print pattern matrix, revised communalities
*print rotated and transformation matrix
*print factor score coefficient matrix
*sort factor loadings in descending order
*plot scree plot and variables in factor space
*extract 2 factors, maximum iterations is 25.
*extraction with unweighted least squares
*maximum number of iterations for rotation
*varimax rotation
*save factor scores for the 2 factors, with variable

Here is a short list of keywords available for the following commands:

/MISSING

PAIRWISE = pairwise exclusion of cases or MEANSUB = substituting means for missing values.

/EXTRACTION

PC = principle components, PAF = principle axis factoring, ALPHA = alpha factoring, GLS = generalized least squares and ML = maximum likelihood.

/ROTATION

EQUAMAX = equamax rotation, QUARTIMAX = quartimax rotation, OBLIMIN(n) = direct oblimin rotation where n = delta, PROMAX(n) = promax rotation where n = kappa and NOROTATE = no rotation.