Exploratory Factor Analysis Example

Note: The SPSS analysis does not match the R or SAS analyses requesting the same options, so caution in using this software and these settings is warranted. The promax rotation may be the issue, as the oblimin rotation is somewhat closer between programs.[1] Because the results in R match SAS more closely, I've added SAS code below the R output. All three specifications should produce very similar results, but the results from SPSS differ substantially from the results from R and SAS and there are fairly minor differences between SAS and R.

```
get file='c:\jason\amos\semclass\se1.sav'.

*EFA example with self-esteem scale from care recipient study; principle axis factoring with promax
oblique rotation.
FACTOR VAR=rfelpos rnotprdr ramable ramfailr rnumqal rnotworr
   /method=covariance
   /analysis=rfelpos rnotprdr ramable ramfailr rnumqal rnotworr
   /print=initial extraction rotation correlation sig
   /plot=eigen
   /diagonal=default
   /CRITERIA=FACTORS(2)
   /EXTRACTION=paf
   /rotation=promax.
```

I chose two factors extract, based on my examination of the scree plot. I ran this syntax twice—once to get the eigenvalues and communalities (PCA), and then to extract the number of factors based on examination of the scree plot. I obtained the bivariate (zero-order) correlations first to examine the associations among all of the variables.

Correlation Matrix

		rfelpos feel positively	rnotprdr	ramable I am able to do things	ramfailr	rnumqal I have a number of good qualities	rnotworr
Correlation	rfelpos feel positively	1.000	.260	.037	.246	.174	.134
	rnotprdr	.260	1.000	.097	.543	.518	.406
	ramable I am able to do things	.037	.097	1.000	.158	.108	.022
	ramfailr	.246	.543	.158	1.000	.289	.470
	rnumqal I have a number of good qualities	.174	.518	.108	.289	1.000	.071
	rnotworr	.134	.406	.022	.470	.071	1.000
Sig. (1-tailed)	rfelpos feel positively		.002	.344	.004	.030	.073
	rnotprdr	.002		.147	.000	.000	.000
	ramable I am able to do things	.344	.147		.044	.122	.407
	ramfailr	.004	.000	.044		.001	.000
	rnumqal I have a number of good qualities	.030	.000	.122	.001		.222
	rnotworr	.073	.000	.407	.000	.222	

The table below gives the eigenvalues for each factor (raw score portion of the total variance of the variables accounted for by each of the possible number of factors) and the percent (each eigenvalue divided the total number of items). Then the total variance accounted for by each extracted factor (I extracted two).

[1] I've used promax in the past, because at some point I learned that it was a somewhat preferable oblique rotation method to other options (see Schmitt & Sass, 2011). Kaiser normalization is used before the promax rotation in SPSS and this process may differ between packages. However, this can be turned off with SPSS (use `nokaiser` on the `/criteria` subcommand) but does not appear to resolve the discrepancies at all. The number of iterations used also can make a difference although changing this setting does not resolve the discrepancies either (I restricted iterations in R, in fact, to make those results match SAS better). After considerable trial and error with options and searching for information on the subject, I cannot resolve the discrepancies. I am inclined to trust the SAS and R outputs more than the SPSS output in this case.

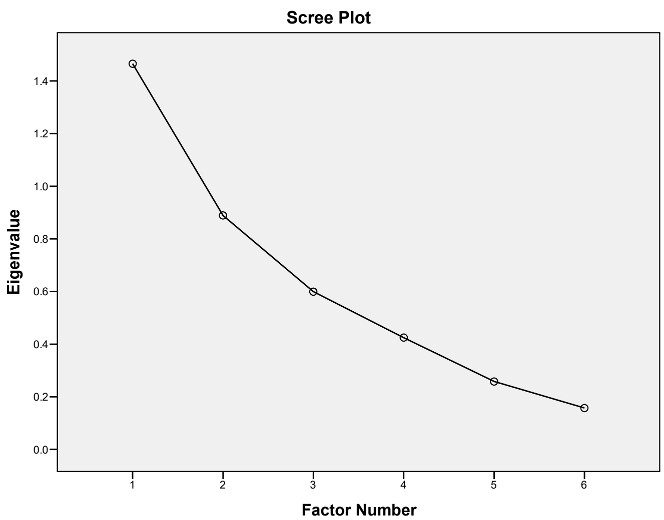
Total Variance Explained

	Factor	Initial Eigenvalues a			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings b
		Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
Raw	1	1.465	38.624	38.624	1.073	28.285	28.285	1.029
	2	.889	23.427	62.051	.764	20.133	48.418	.847
	3	.600	15.802	77.852				
	4	.425	11.198	89.050				
	5	.258	6.802	95.852				
	6	.157	4.148	100.000				
Rescaled	1	1.465	38.624	38.624	1.762	29.361	29.361	1.782
	2	.889	23.427	62.051	.936	15.606	44.967	.983
	3	.600	15.802	77.852				
	4	.425	11.198	89.050				
	5	.258	6.802	95.852				
	6	.157	4.148	100.000				

Extraction Method: Principal Axis Factoring.

a. When analyzing a covariance matrix, the initial eigenvalues are the same across the raw and rescaled solution.
b. When factors are correlated, sums of squared loadings cannot be added to obtain a total variance.

The *Scree Plot* is the plot of the eigenvalue by each factor. Decide on the number of factors to be extracted using this. Note that the common 1.0 Kaiser-Guttman rule does not perform well. Determine where the size of the eigenvalues drops off—the "scree". I guess one or two factors here (?). A more precise determination might be obtained using Horn's (1965) parallel analysis available.[2]



Communalities are the squared multiple correlations for each item predicted by all of the factors.

Communalities

	Raw		Rescaled	
	Initial	Extraction	Initial	Extraction
rfelpos feel positively	.545	.052	1.000	.095
rnotprdr	.487	.291	1.000	.598
ramable I am able to do things	.884	.822	1.000	.930
ramfailr	.521	.329	1.000	.631
rnumqal I have a number of good qualities	.325	.053	1.000	.163
rnotworr	1.032	.290	1.000	.281

Extraction Method: Principal Axis Factoring.

[2] This is not a standard option in SPSS but can be obtained using extensive code from Brian O'Connor (https://oconnor-psych.ok.ubc.ca/nfactors/nfactors.html) and in some R packages (e.g., EFAtools).

The *Factor Matrix* contains the unrotated factor loadings. Raw estimates are unstandardized (covariance metric) and rescaled estimates are standardized (correlation metric).

Factor Matrix a

	Raw		Rescaled	
	Factor		Factor	
	1	2	1	2
rfelpos feel positively	.209	.090	.283	.122
rnotprdr	.495	.214	.709	.307
ramable I am able to do things	.469	-.776	.499	-.825
ramfailr	.540	.193	.748	.267
rnumqal I have a number of good qualities	.224	.055	.392	.097
rnotworr	.471	.260	.464	.256

Extraction Method: Principal Axis Factoring.

a. 2 factors extracted. 15 iterations required.

The *Pattern Matrix* contains the factor loadings from oblique rotated matrix (values most often reported and interpreted). Raw estimates are unstandardized (covariance metric) and rescaled estimates are standardized (correlation metric).

Pattern Matrix a

	Raw		Rescaled	
	Factor		Factor	
	1	2	1	2
rfelpos feel positively	.228	-.005	.309	-.007
rnotprdr	.541	-.012	.776	-.018
ramable I am able to do things	-.001	.907	-.001	.965
ramfailr	.569	.026	.788	.035
rnumqal I have a number of good qualities	.223	.033	.391	.058
rnotworr	.545	-.064	.536	-.063

Extraction Method: Principal Axis Factoring.
Rotation Method: Promax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

The structure matrix gives bivariate correlations between factors and items (not too useful for most researchers). Raw estimates are unstandardized (covariance metric) and rescaled estimates are standardized (correlation metric).

Structure Matrix

	Raw		Rescaled	
	Factor		Factor	
	1	2	1	2
rfelpos feel positively	.227	.031	.308	.042
rnotprdr	.539	.073	.773	.105
ramable I am able to do things	.142	.907	.152	.964
ramfailr	.573	.116	.794	.160
rnumqal I have a number of good qualities	.228	.069	.400	.120
rnotworr	.535	.022	.526	.022

Extraction Method: Principal Axis Factoring.
Rotation Method: Promax with Kaiser Normalization.

The *Factor Correlation Matrix* gives the estimated correlation between the two extracted factors. The larger this correlation is the bigger the difference between the factor and pattern matrices. If the correlation is zero, the rotated and unrotated solutions will be the same.

Factor Correlation Matrix

Factor	1	2
1	1.000	.158
2	.158	1.000

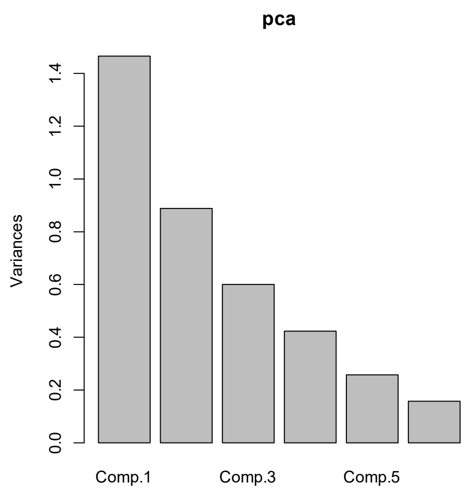
Extraction Method: Principal Axis Factoring.
Rotation Method: Promax with Kaiser Normalization.

R

The results obtained with R differ substantially from those obtained with SPSS, despite code that should produce the same method.

```
library(GPArotation)  #need to load separately for the Promax rotation
library(psych)
#initial analysis to obtain scree plot and decide on number of factors (used correlations to match SPSS)
pca1=princomp(mydata, cor = TRUE)
summary(pca1)
eigens = pca1$sdev*pca1$sdev
eigens

#I decide to extract two factors
efa <- fa(mydata, nfactors=2, rotate="promax",fm="pa", covar=TRUE, max.iter=1)
efa
summary(efa)
```




```
#from the psych package you can do Horn's parallel analysis
#fa.parallel(mydata)


Standardized loadings (pattern matrix)
         item   PA1   PA2    h2   u2
RAMABLE     1  0.00  0.18 0.032 0.97
RAMFAILR    2  0.68  0.07 0.534 0.47
RFELPOS     3  0.18  0.18 0.105 0.89
RNOTPRDR    4  0.48  0.37 0.591 0.41
RNOTWORR    5  0.67 -0.16 0.332 0.67
RNUMQAL     6 -0.12  0.72 0.422 0.58

                 PA1  PA2
SS loadings     1.31 0.48
Proportion Var  0.22 0.08
Cumulative Var  0.22 0.30
Cum. factor Var 0.73 1.00

 With factor correlations of
     PA1  PA2
PA1 1.00 0.64
PA2 0.64 1.00
```

SAS

SAS and R produce very similar final standardized loadings (note the reordering of the items though).

```
proc factor data=one  nfactors = 2 method=principal covariance priors=smc
    rotate=promax norm=kaiser maxiter=25
    outstat=fact_all
    plots=(scree);
    var rfelpos rnotprdr ramable ramfailr rnumqal rnotworr;
```

Inter-Factor Correlations

	Factor1	Factor2
Factor1	1.00000	0.58473
Factor2	0.58473	1.00000

Rotated Factor Pattern (Standardized Regression Coefficients)

		Factor1	Factor2
RFELPOS	feel positively	0.18255	0.18200
RNOTPRDR	RNOTPRDR	0.49038	0.37097
RAMABLE	I am able to do things	0.00660	0.17458
RAMFAILR	RAMFAILR	0.67174	0.09377
RNUMQAL	I have a number of good qualities	-0.07914	0.69294
RNOTWORR	RNOTWORR	0.64396	-0.13373

References

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179-185.
Schmitt, T. A., & Sass, D. A. (2011). Rotation criteria and hypothesis testing for exploratory factor analysis: Implications for factor pattern loadings and interfactor correlations. *Educational and Psychological Measurement*, *71*(1), 95-113.