Research Design - - Topic 19
Exploratory Factor Analysis

General Rationale and Purpose

An Example Using SPSS Factor

Constructs and Terms

Principal Components Analysis
    Basic Mathematics

SPSS Factor with Principal Components and Oblique rotation

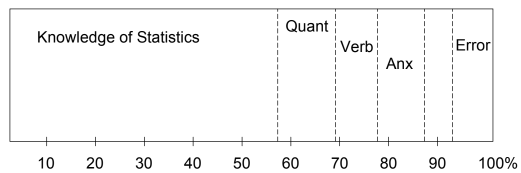
Principal Axis Factor Analysis

Factor Scores

General Rationale and Purpose

The logic underlying Factor Analysis is that measures are influenced by a number of underlying constructs, and that the dimensions in factor analysis (principal components) help to uncover them. This is done by decomposing the correlations (or covariances, etc.) among them.

For example, scores on a statistics test might tap a number of dimensions, such as knowledge of the course material, quantitative ability, verbal ability, test anxiety, other possible elements, and error. It might be diagrammed as follows:



The purpose of factor analysis is to understand the relationships among variables. This is achieved by identifying a number of dimensions and seeing how the variables relate to them. **These dimensions are simply weighted aggregates of the variables.** They are also referred to as factors, components, or latent variables.

When conducting a factor analysis, one seeks to determine the number of dimensions necessary to explain the bulk of the relationships among the variables, and then to interpret them by investigating the correlations of the variables with the dimensions.

One of the earliest applications of factor analysis identified dimensions underlying scores on measures of mental ability (Thurstone, 1938). Another application (McCrae & Costa, 1987) identified the Big 5 dimensions of personality as comprising:

- Openness
- Conscientiousness
- Extraversion
- Agreeableness
- Neuroticism

Steps in conducting a factor analysis

There are three stages in conducting a factor analysis:

1. Calculate a matrix of associations. Most times, this is a correlation matrix but it could be a covariance matrix, a cross product matrix, etc.
2. Calculate an initial factor matrix. There are many types of solutions, principal components, principal axis, unweighted least squares, maximum likelihood, alpha analysis, etc…
3. Produce a rotated solution. Sometimes, this isn't done, but if it is, there are a number of alternatives, varimax, quartimax, equamax, oblimin, promax, etc…

Concepts and Terms

Fundamental Theorem. The correlation matrix can be reproduced from the factor matrices as follows:

$$R = A A^T \quad \text{for an orthogonal factor matrix}$$

$$R = A \phi A^T \quad \text{for a non-orthogonal factor matrix}$$

where A is the structure matrix, A^T is its transpose, and Φ is the matrix of correlations among the factors.

Pattern Matrix. A matrix of the weights used to define the factors.

Structure Matrix. A matrix of the correlations of each variable with the factors.

Eigenvalue. The variance of a principal component

Eigenvalue 1 criterion. Retain all factors with eigenvalues > 1.

Scree test (Cattell, 1966). A plot of the eigenvalues to determine the number of factors to retain.

5

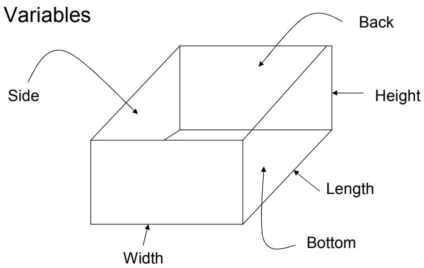
An Example Using SPSS Factor:

One of the initial demonstrations of the logic underlying factor analysis was Thurstone's Box problem. Our version factor analyzes correlations of the length, width, and height of each box, the areas of the 3 sides, the diagonals of the 3 sides, the perimeters of the 3 sides, and the volume. The data are on the webpage (Box Problem Data), and the Syntax file is as follows:

```
FACTOR
 /VARIABLES LENGTH WIDTH HEIGHT BOTA BACKA SIDEA BOTP
BACKP SIDEP BOTD BACKD SIDED VOLUME
 /MISSING LISTWISE
 /ANALYSIS LENGTH WIDTH HEIGHT BOTA  BACKA SIDEA BOTP
BACKP SIDEP BOTD BACKD SIDED VOLUME
 /PRINT INITIAL CORRELATION EXTRACTION ROTATION
 /PLOT EIGEN
 /CRITERIA MINEIGEN(1) ITERATE(25)
 /EXTRACTION PC
 /CRITERIA ITERATE(25)
 /ROTATION VARIMAX
 /METHOD=CORRELATION .
```

6

Definition of Variables



Measures

Length	Perimeter of Bottom	Volume
Width	Perimeter of Back	
Height	Perimeter of Side	
Area of Bottom	Diagonal of Bottom	
Area of Back	Diagonal of Back	
Area of Side	Diagonal of Side	

7

Summary Statistics of the Basic Data

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
LENGTH	15.0300	2.98389	200
WIDTH	10.1100	2.19133	200
HEIGHT	6.1200	1.48885	200
BOTA	151.0500	41.89083	200
BACKA	61.6500	19.19766	200
SIDEA	92.1950	28.14065	200
BOTP	50.2650	6.95894	200
BACKP	32.5100	5.10492	200
SIDEP	42.3750	6.47798	200
BOTD	18.3050	2.60691	200
BACKD	11.9050	1.98638	200
SIDED	16.3400	2.73713	200
VOLUME	922.7950	334.47256	200

8

Correlation Matrix

Correlation Matrix

		LENGTH	WIDTH	HEIGHT	BOTA	BACKA	SIDEA	BOTP	BACKP	SIDEP	BOTD	BACKD	SIDED	VOLUME
Correlation	LENGTH	1.000	-.116	-.074	.612	-.102	.593	.765	-.149	.880	.860	-.169	.965	.455
	WIDTH	-.116	1.000	-.083	.684	.599	-.113	.539	.804	-.147	.377	.895	-.136	.466
	HEIGHT	-.074	-.083	1.000	-.076	.710	.731	-.100	.500	.390	-.109	.310	.144	.591
	BOTA	.612	.684	-.076	1.000	.426	.363	.964	.546	.530	.890	.595	.585	.733
	BACKA	-.102	.599	.710	.426	1.000	.521	.315	.944	.244	.200	.841	.056	.820
	SIDEA	.593	-.113	.731	.363	.521	1.000	.443	.334	.893	.494	.172	.746	.805
	BOTP	.765	.539	-.100	.964	.315	.443	1.000	.406	.660	.972	.445	.733	.699
	BACKP	-.149	.804	.500	.546	.944	.334	.406	1.000	.103	.259	.962	-.037	.753
	SIDEP	.880	-.147	.390	.530	.244	.893	.660	.103	1.000	.742	-.008	.959	.699
	BOTD	.860	.377	-.109	.890	.200	.494	.972	.259	.742	1.000	.286	.826	.640
	BACKD	-.169	.895	.310	.595	.841	.172	.445	.962	-.008	.286	1.000	-.102	.654
	SIDED	.965	-.136	.144	.585	.056	.746	.733	-.037	.959	.826	-.102	1.000	.576
	VOLUME	.455	.466	.591	.733	.820	.805	.699	.753	.699	.640	.654	.576	1.000

The correlation matrix is subjected to an analysis that extracts factors accounting for decreasing amounts of variance. This produces a table of the variance extracted by each factor (see next slide). These variances are referred to as eigenvalues. Given 13 variables, that table presents the variances associated with each of 13 factors extracted.

The first section gives the eigenvalues for each of the 13 dimensions, the second repeats this for the 3 factors retained, and the third section gives the sum of squared loadings for the 3 factors after rotation. Note:

1. The percentage is the eigenvalue divided by 13.
2. The sum of the 13 eigenvalues = 13.
3. The rotation sums of squared loadings attempts to even out the values.

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.847	52.669	52.669	6.847	52.669	52.669	5.439	41.839	41.839
2	3.718	28.597	81.265	3.718	28.597	81.265	4.463	34.328	76.167
3	2.246	17.280	98.546	2.246	17.280	98.546	2.909	22.379	98.546
4	.057	.435	98.980						
5	.040	.310	99.291						
6	.031	.236	99.527						
7	.020	.156	99.683						
8	.014	.110	99.793						
9	.010	.079	99.872						
10	.008	.060	99.931						
11	.004	.032	99.964						
12	.004	.031	99.994						
13	.001	.006	100.000						

Extraction Method: Principal Component Analysis.

Communalities

The proportion of variance that variable has in common with all the other variables. They are set at 1.0 in Principal Components; the extraction communalities are the proportions captured by the 3 factors retained.

Note. The communalities are the sum of squares of the factor loadings across the factors in both the initial and rotated factors. That is:
From slide 13,
$.662^2+(-.736)^2+(-.118)^2 = .993$
and from slide 14
$.977^2+(-.190)^2+.058^2 = .993$

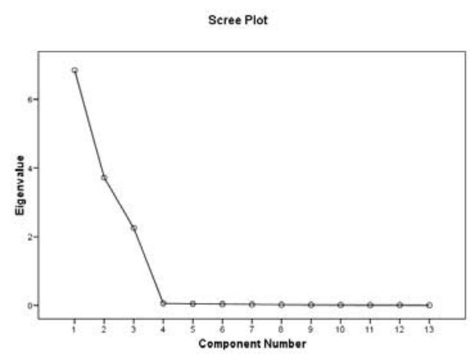
Communalities

	Initial	Extraction
LENGTH	1.000	.993
WIDTH	1.000	.988
HEIGHT	1.000	.982
BOTA	1.000	.977
BACKA	1.000	.979
SIDEA	1.000	.985
BOTP	1.000	.996
BACKP	1.000	.994
SIDEP	1.000	.995
BOTD	1.000	.980
BACKD	1.000	.975
SIDED	1.000	.989
VOLUME	1.000	.978

Extraction Method: Principal Component Analysis.

A Plot of the Eigenvalues associated with each component.

The Scree test is used to distinguish the true factors from those reflecting error. This is identified by the elbow in the plot. This plot indicates three factors.



Initial Factor matrix

The principal component matrix is a structure matrix, consisting of the correlations of each variable with each of the three factors. It is generally not interpreted. Typically, interpretation is done with the rotated matrix.

Note the sum of the squared factor loadings for each factor equals the eigenvalue for that factor. That is:

$.662^2 + \ldots + .943^2 = 6.847$

Component Matrix[a]

	Component		
	1	2	3
LENGTH	.662	-.736	-.118
WIDTH	.487	.683	-.533
HEIGHT	.366	.292	.873
BOTA	.872	.025	-.465
BACKA	.649	.672	.326
SIDEA	.751	-.229	.607
BOTP	.884	-.176	-.429
BACKP	.636	.766	.052
SIDEP	.784	-.540	.299
BOTD	.850	-.346	-.370
BACKD	.584	.781	-.154
SIDED	.734	-.666	.077
VOLUME	.943	.192	.229

Extraction Method: Principal Component Analysis.
a. 3 components extracted.

13

Varimax Rotated Factor Matrix

This is a structure matrix, consisting of the correlations of the variables with the rotated factors. It is typically considered to offer a more parsimonious interpretation than a non-rotated matrix.

Note. The sum of squared factor loadings for each factor equals the value reported in the last part of the table in slide 10. That is:

$.977^2 + \ldots + .545^2 = 5.439$

Rotated Component Matrix[a]

	Component		
	1	2	3
LENGTH	.977	-.190	.058
WIDTH	.082	.963	-.233
HEIGHT	-.105	.154	.973
BOTA	.757	.628	-.099
BACKA	.008	.773	.617
SIDEA	.566	.030	.815
BOTP	.879	.465	-.084
BACKP	.006	.925	.370
SIDEP	.852	-.101	.509
BOTD	.943	.295	-.061
BACKD	.006	.974	.163
SIDED	.943	-.158	.273
VOLUME	.545	.582	.585

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 6 iterations.

14

Interpreting a Factor Analysis

This analysis of the measurements of the boxes suggested that 3 factors accounted for the relationships among the measures, explaining 98.55% of the total variance. The variance of the first factor was 6.847, accounting for 52.67% of the variance; that for the second factor was 3.718 (28.60%), and that for the third was 2.246 (17.28%). These factors are defined to account for as much of the variance as possible in the matrix.

Generally, the initial matrix is not interpreted. Instead, the factors are rotated to produce a more parsimonious picture, where each variable has a combination of high and low loadings across the factors. Interpretation involves identifying what is common to the variables loading high on a factor and what distinguishes them from the variables not having high loadings on that factor. Consider the rotated factor matrix in slide 14, and attempt to identify the common feature of each factor. The next slide might make the task a bit easier, because there we show the basic elements of each measure. This helps to emphasize that you must know a lot about the measures involved in a factor analysis to be able to make a meaningful interpretation of the factors.

15

Varimax Rotated Factor Matrix

Rotated Component Matrix[a]

		Component		
		1	2	3
L →	LENGTH	.977	-.190	.058
W	WIDTH	.082	.963	-.233
H	HEIGHT	-.105	.154	.973
L*W	BOTA	.757	.628	-.099
W*H	BACKA	.008	.773	.617
L*H	SIDEA	.566	.030	.815
2(L+W)	BOTP	.879	.465	-.084
2(H+W)	BACKP	.006	.925	.370
2(L+H)	SIDEP	.852	-.101	.509
$SQRT(L^2 + W^2)$	BOTD	.943	.295	-.061
$SQRT(H^2 + W^2)$	BACKD	.006	.974	.163
$SQRT(H^2 + L^2)$	SIDED	.943	-.158	.273
L*W*H →	VOLUME	.545	.582	.585

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 6 iterations.

16

Factor Interpretation

Factor I accounts for 41.84% of the variance.[1] It is defined by high loadings (>.30) from Length (L), BOTA (L*W), SIDEA (L*H), BOTP (2(L+W)), SIDEP (2(L+H)), BOTD (sqrt(L^2 + W^2)), SIDED (sqrt(L^2 + H^2)), and Volume (L*W*H). Length is involved in all 8 of these measures so it seems reasonable to define this as a **Length** factor.

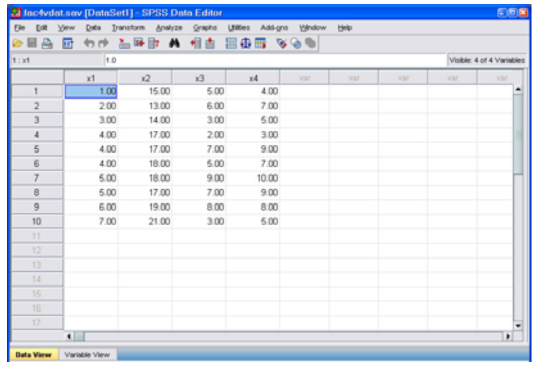
Factor II accounts for 34.33% of the variance. It is defined by high loadings (>.30) from Width (W), BOTA (L*W), BACKA (W*H), BOTP (2(L+W)), BACKP (2(H+W)), BACKD (sqrt(H^2 + W^2)), and Volume (L*W*H), and BOTD (sqrt(L^2 + W^2)),has a loading of .295. Thus, it seems reasonable to define this as a **Width** factor.

Factor III accounts for 22.38% of the variance. It is defined by high loadings (>.30) from Height (H), BACKA (W*H), SIDEA (L*H), BACKP (2(H+W)), SIDEP (2(L+H)) and Volume (L*W*H). Although BACKD (sqrt(H^2 + W^2)) and SIDED (sqrt(L^2 + H^2)) do not contribute substantially to this factor, it seems reasonable nonetheless to define this as a **Height** factor.

[1] See the table in slide 10 in the section under Rotation Sums of Squared Loadings.

Principal Components Analysis: Basic Mathematics

A Small Example



	x1	x2	x3	x4
1	1.00	15.00	5.00	4.00
2	2.00	13.00	6.00	7.00
3	3.00	14.00	3.00	5.00
4	4.00	17.00	2.00	3.00
5	4.00	17.00	7.00	9.00
6	4.00	18.00	5.00	7.00
7	5.00	18.00	9.00	10.00
8	5.00	17.00	7.00	9.00
9	6.00	19.00	8.00	8.00
10	7.00	21.00	3.00	5.00

A factor is a weighted aggregate of all the variables in the study. When factor analyzing a correlation matrix, the general form of the aggregate for each factor is:

$$Z_{I_i} = w_1 Z_{1_i} + w_2 Z_{2_i} + w_3 Z_{3_i} + \ldots + w_m Z_{m_i}$$

Given an *mxm* correlation matrix, there will be *m* such factors. For each one, the weights, *w,* and the variance of each factor will differ. The variance of each factor is called the **eigenvalue.**

There are a number of approaches for calculating the eigenvalues. One is to make use of the **determinantal equation**, set it equal to 0, and solve for the eigenvalues. This is written as:

$$|R - \lambda I| = 0$$

R = the correlation matrix, I = an identity matrix, and λ = the eigenvalue

(a) Determinantal Equation $|R - \lambda I| = 0$

Given R (i.e. an *m* x *m* matrix of correlations), this would produce *m* λ's arranged in decreasing order. If all m λ's were obtained,

$$\sum^{m} \lambda = m$$

The λ is called the eigenvalue or the characteristic root, or latent root. It is a measure of the variance of a component.

| R -λ I | = 0

$$\left| \begin{bmatrix} 1.000 & 0.889 & 0.147 & 0.323 \\ 0.889 & 1.000 & 0.091 & 0.172 \\ 0.147 & 0.091 & 1.000 & 0.903 \\ 0.322 & 0.172 & 0.903 & 1.000 \end{bmatrix} - \lambda \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix} \right| = 0$$

$$\begin{vmatrix} 1-\lambda & 0.889 & 0.147 & 0.323 \\ 0.889 & 1-\lambda & 0.091 & 0.172 \\ 0.147 & 0.091 & 1-\lambda & 0.903 \\ 0.323 & 0.172 & 0.903 & 1-\lambda \end{vmatrix} = 0$$

Results in an equation of the type

$$a\lambda^4 + b\lambda^3 + c\lambda^2 + \lambda = 0$$

Yielding 4 values of λ.

The four eigenvalues for this matrix are:

2.269, 1.539, .146, and .047.

Note that they sum to 4, the number of variables.

We can use this information to produce the eigenvector for each factor. The eigenvector is the set of weights for the variables associated with that factor (i.e., a pattern matrix). This is done by solving a set of simultaneous equations such as:

$$[R-\lambda_p I][W]=0$$

$$\begin{bmatrix} 1-\lambda & 0.889 & 0.147 & 0.323 \\ 0.889 & 1-\lambda & 0.091 & 0.172 \\ 0.147 & 0.091 & 1-\lambda & 0.903 \\ 0.323 & 0.172 & 0.903 & 1-\lambda \end{bmatrix}\begin{bmatrix} W_1 \\ W_2 \\ W_3 \\ W_4 \end{bmatrix}=0$$

$$(1-\lambda)w_1+.889w_2+.147w_3+.323w_4=0$$
$$.889w_1+(1-\lambda)w_2+.091w_3+.172w_4=0$$
$$.147w_1+.091w_2+(1-\lambda)w_3+.903w_4=0$$
$$.323w_1+.172w_2+.903w_3+(1-\lambda)w_4=0$$

With 4 equations and 4 unknowns, we can obtain an infinite set of solutions. We obtain a unique one by requiring that the sum of squares of the weights equal 1. In matrix form, this is the side condition that:

$W^TW=1$ which is equivalent to $\sum w^2=1$

For the first factor, this produces the weights
w1 = .521, w2 = .471, w3 = .473, w4 = .533

Thus: $Z_{I_i}=.521Z_{1_i}+.471Z_{2_i}+.473Z_{3_i}+.533Z_{4_i}$

and the variance of this aggregate is the eigenvalue:

That is: $S^2_{Z_I}=\lambda_1=2.269$

For the second factor, this produces the weights
w1 = .467, *w2* = .536, *w3* = -.539, *w4* = -.451

and $S^2_{Z_{II}}=\lambda_2=1.539$

SPSS Factor defines the values in the Component Matrix as:

$a_{j,p}=w_{j,p}\sqrt{\lambda_p}$ where $\sum a^2_{j,p}=\lambda_p\sum w^2_{j,p}=\lambda_p$

It does not output a pattern matrix, though in this case it will be noted that there is a 1 to 1 correspondence between the two. The component matrix is a structure matrix but it differs from the pattern matrix solely in the magnitude of the values so that the sums of squares are λ and 1 respectively.

The loadings on the principal components axes and the varimax rotated axes. These are structure matrices because they are the correlations of the variables with the factors.

Component Matrix[a]

	Component	
	1	2
x1	.784	.580
x2	.709	.665
x3	.712	-.669
x4	.803	-.560

Extraction Method: Principal Component Analysis.
a. 2 components extracted.

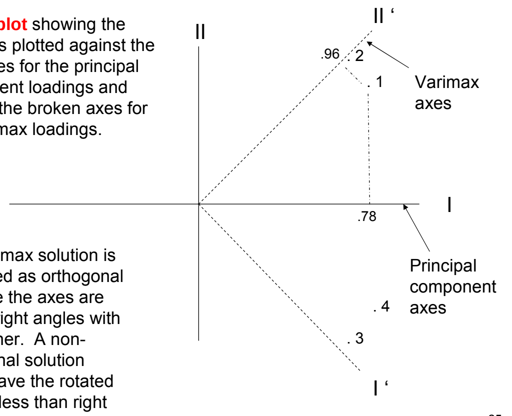
The factor plot for these loadings is shown on the next slide. It will be noted that the points don't move in space. Only the reference axes are rotated.

Rotated Component Matrix[a]

	Component	
	1	2
x1	.151	.963
x2	.037	.971
x3	.977	.025
x4	.965	.166

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 3 iterations.

Factor plot showing the variables plotted against the solid axes for the principal component loadings and against the broken axes for the varimax loadings.

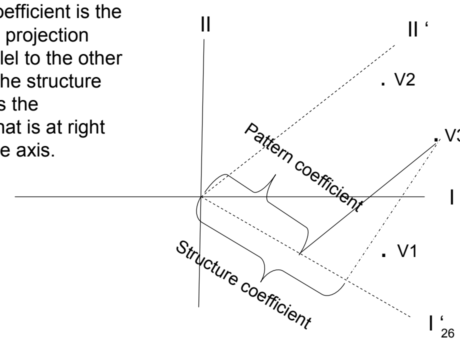


The varimax solution is described as orthogonal because the axes are kept at right angles with each other. A non-orthogonal solution would have the rotated axes at less than right angles.

25

The plot below illustrates the difference between pattern and structure coefficients with non-orthogonal rotation.

A pattern coefficient is the value of the projection that is parallel to the other axis, while the structure coefficient is the projection that is at right angles to the axis.



26

SPSS Analysis of the Box Problem Using Principal Components and an Oblique Rotation

With an oblique rotation, the rotated solution consists of a Pattern Matrix, a Structure Matrix, and a Matrix of Correlations among the Factors.

Pattern Matrix[a]

	Component		
	1	2	3
LENGTH	1.003	-.280	-.017
WIDTH	.040	.966	-.331
HEIGHT	-.155	.167	.974
BOTA	.740	.567	-.229
BACKA	-.060	.782	.548
SIDEA	.541	-.019	.762
BOTP	.873	.391	-.210
BACKP	-.060	.936	.286
SIDEP	.852	-.178	.440
BOTD	.946	.213	-.178
BACKD	-.054	.985	.073
SIDED	.958	-.244	.199
VOLUME	.499	.541	.482

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.
a. Rotation converged in 10 iterations.

Structure Matrix

	Component		
	1	2	3
LENGTH	.957	-.127	.097
WIDTH	.142	.939	-.228
HEIGHT	.008	.241	.968
BOTA	.795	.658	-.067
BACKA	.139	.828	.618
SIDEA	.646	.142	.837
BOTP	.904	.505	-.047
BACKP	.126	.956	.372
SIDEP	.887	-.002	.543
BOTD	.954	.342	-.022
BACKD	.108	.984	.165
SIDED	.949	-.076	.310
VOLUME	.651	.667	.607

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

27

Interpreting an Oblique Solution

Interpretation is based on the pattern matrix, because this describes the unique contributions of the factor to each variable. In the present example, the factors are relatively uncorrelated (see below), so the interpretations are comparable. The structure matrix represents the correlation of each variable with the factor, but in the case of oblique factors this correlation is confounded with the correlations among the factors.

The correlation among the factors is shown in the following table. With a sample size of 200, a correlation greater than .14 is significant at the .05 level, two-tailed.

Component Correlation Matrix

Component	1	2	3
1	1.000	.154	.142
2	.154	1.000	.101
3	.142	.101	1.000

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

28

Other Factor Analytic Solutions

There are many other factor analytic solutions. An important issue in this regard is the distinction between principal components which analyzes the total variance and common factors which analyzes the common variance. There is controversy about which approach is best. Stevens (1996, pp. 383-388) discusses the issues involved and favours principal components analysis. An alternative view is expressed by Fabrigar, Wegener, MacCallum & Strahan (1999).

29

The Principal Axis Solution

This is a common alternative. Rather than analyzing the total variance among the variables, it analyzes the common variance. This is achieved by substituting a measure of common variance into the correlation matrix instead of 1's. A common estimate is the squared multiple correlation of each variable with all the other variables in the matrix.

Principal Axis factoring in SPSS first performs a principal components analysis to determine the number of factors, then estimates the diagonal values (the communalities) using R^2. It then extracts the number of factors initially determined, computes the communalities and compares them with the R^2 values. If they do not agree, the program uses these communalities to extract factors, and continues iterating until stable estimates of the communalities are obtained. It then performs the rotation selected by the user.

30

Factor Scores

Used to estimate the individual's score on factors if they could be measured on the factors. They can be determined either for the unrotated (e.g. Principal Components) matrix or the rotated one. Note, for the Principal Components analysis one possible form is simply the original principal components (e.g. ZI = w1Z1 + w2Z2 + … + wpZp), but as can be seen below they could also be computed using a number of procedures.

These include:

1. Simply summing the variables that define a factor. This procedure is not recommended because the correlations between these scores do not reflect the correlations among the factors.
2. Procedures used by SPSS Factor
   - The Regression Method (see next slide)
   - The Bartlett Method
   - The Anderson-Rubin Method

31

The Regression Method is commonly used. It makes use of multiple regression to estimate scores on each factor for the individuals. Recall that with multiple regression, the β's can be computed as:

$$\beta_{iI} = r_{ii}^{-1} r_{iy} \qquad \text{where}$$

r_{ii}^{-1} is the inverse of the matrix of correlations among the predictors (variables in this case).

r_{iy} are the correlations of the predictors with the criterion (i.e., the factor). In this case, the correlations are the loadings from the appropriate structure matrix.

Note: This procedure results in biased estimates, but as long as the initial solution is principal components, the factor scores will display the same correlations as the factors on which they're based. If the factors are orthogonal the mean of the factor scores = 0, and the standard deviation = 1.0. If the factors are correlated, the means of the factor scores = 0, but the standard deviation may not be 1.0. If the initial solution is principal axis, the factor scores are indeterminate. Among other things, this means that the factor scores do not necessarily reflect the correlations among the factors.

32

References

Cattell, R.B. (1966). The scree test and the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C. & Strahan, E.J. (1999). Evaluating the use of factor analysis in psychological research. *Psychological Methods, 4*, 272-299.

McCrae, R.R. & Costa, P.T. (1987). Validation of the five factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81-90.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences (Third edition).* Mahhaw, NJ: Lawrence Erlbaum.

Thurstone, L.L.(1938). Primary mental abilities. *Psychometric Monographs*, No. 1.

33