Factor Analysis Using SPSS

For an overview of the theory of factor analysis please read Field (2000) Chapter 11 or refer to your lecture.

Factor analysis is frequently used to develop questionnaires: after all if you want to measure an ability or trait, you need to ensure that the questions asked relate to the construct that you intend to measure. I have noticed that a lot of students become very stressed about SPSS. Therefore I wanted to design a questionnaire to measure a trait that I termed 'SPSS anxiety'. I decided to devise a questionnaire to measure various aspects of students' anxiety towards learning SPSS. I generated questions based on interviews with anxious and non-anxious students and came up with 23 possible questions to include. Each question was a statement followed by a five-point Likert scale ranging from 'strongly disagree' through 'neither agree or disagree' to 'strongly agree'. The questionnaire is printed in Field (2000, p. 442).

The questionnaire was designed to predict how anxious a given individual would be about learning how to use SPSS. What's more, I wanted to know whether anxiety about SPSS could be broken down into specific forms of anxiety. So, in other words, are there other traits that might contribute to anxiety about SPSS? With a little help from a few lecturer friends I collected 2571 completed questionnaires (at this point it should become apparent that this example is fictitious!). The data are stored in the file **SAQ.sav**.

Initial Considerations

Sample Size

Correlation coefficients fluctuate from sample to sample, much more so in small samples than in large. Therefore, the reliability of factor analysis is also dependent on sample size. Field (2000) reviews many suggestions about the sample size necessary for factor analysis and concludes that it depends on many things. In general over 300 cases is probably adequate but communalities after extraction should probably be above 0.5 (see Field, 2000).

Data Screening

SPSS will nearly always find a factor solution to a set of variables. However, the solution is unlikely to have any real meaning if the variables analysed are not sensible. The first thing to do when conducting a factor analysis is to look at the inter-correlation between variables. If our test questions measure the same underlying dimension (or dimensions) then we would expect them to correlate with each other (because they are measuring the same thing). If we find any variables that do not correlate with any other variables (or very few) then you should consider excluding these variables before the factor analysis is run. The correlations between variables can be checked using the *correlate* procedure (see Chapter 3) to create a correlation matrix of all variables. This matrix can also be created as part of the main factor analysis.

The opposite problem is when variables correlate too highly. Although mild multicollinearity is not a problem for factor analysis it is important to avoid extreme multicollinearity (i.e. variables that are very highly correlated) and *singularity* (variables that are perfectly correlated). As with regression, singularity causes problems in factor analysis because it becomes impossible to determine the unique contribution to a factor of the variables that are highly correlated (as was the case for multiple regression). Therefore, at this early stage we look to eliminate any variables that don't correlate with any other variables or that correlate very highly with other variables ($R < 0.9$). Multicollinearity can be detected by looking at the determinant of the R-matrix (see next section).

As well as looking for interrelations, you should ensure that variables have roughly normal distributions and are measured at an interval level (which Likert scales are, perhaps wrongly,

assumed to be!). The assumption of normality is important only if you wish to generalize the results of your analysis beyond the sample collected.

Running the Analysis

Access the main dialog box (Figure 1) by using the **Analyze⇒Data Reduction⇒Factor ...** menu path. Simply select the variables you want to include in the analysis (remember to exclude any variables that were identified as problematic during the data screening) and transfer them to the box labelled *Variables* by clicking on ▶.

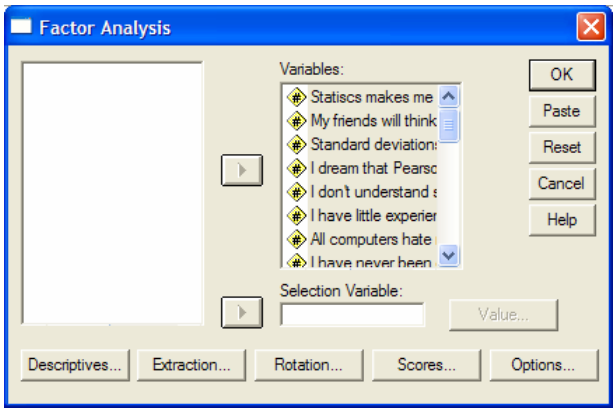


Figure 1: Main dialog box for factor analysis

There are several options available, the first of which can be accessed by clicking on Descriptives... to access the dialog box in Figure 2. The *Coefficients* option produces the *R*-matrix, and the *Significance levels* option will produce a matrix indicating the significance value of each correlation in the *R*-matrix. You can also ask for the *Determinant* of this matrix and this option is vital for testing for multicollinearity or singularity. The determinant of the *R*-matrix should be greater than 0.00001; if it is less than this value then look through the correlation matrix for variables that correlate very highly ($R > 0.8$) and consider eliminating one of the variables (or more depending on the extent of the problem) before proceeding. The choice of which of the two variables to eliminate will be fairly arbitrary and finding multicollinearity in the data should raise questions about the choice of items within your questionnaire.

KMO and Bartlett's test of sphericity produces the Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's test (see Field, 2000, Chapters 10 & 11). The value of KMO should be greater than 0.5 if the sample is adequate.

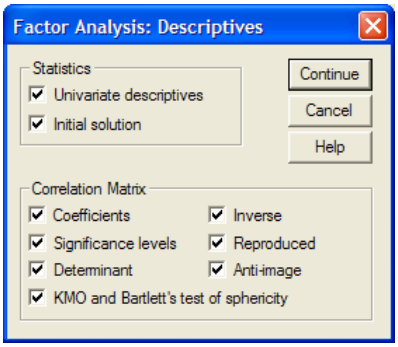


Figure 2: Descriptives in factor analysis

Factor Extraction on SPSS

To access the *extraction* dialog box (Figure 3), click on Extraction... in the main dialog box. There are a number of ways of conducting a factor analysis and when and where you use the various

methods depend on numerous things. For our purposes we will use *principal component* analysis, which strictly speaking isn't factor analysis; however, the two procedures usually yield identical results (see Field, 2000, section 11.2.2). The method chosen will depend on what you hope to do with the analysis (see Field, 2000 for details).

The *Display* box has two options within it: to display the *Unrotated factor solution* and a *Scree plot*. The scree plot was described earlier and is a useful way of establishing how many factors should be retained in an analysis. The unrotated factor solution is useful in assessing the improvement of interpretation due to rotation. If the rotated solution is little better than the unrotated solution then it is possible that an inappropriate (or less optimal) rotation method has been used.

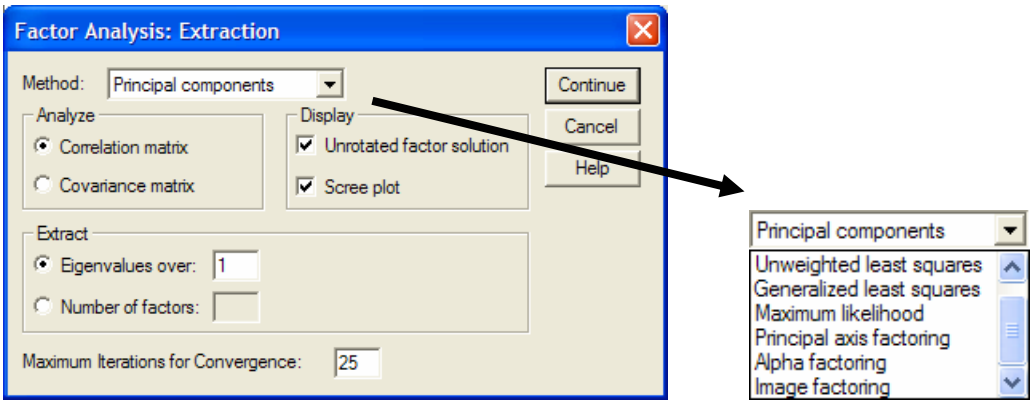


Figure 3: Dialog box for factor extraction

The *Extract* box provides options pertaining to the retention of factors. You have the choice of either selecting factors with eigenvalues greater than a user-specified value or retaining a fixed number of factors. For the *Eigenvalues over* option the default is Kaiser's recommendation of eigenvalues over 1. It is probably best to run a primary analysis with the *Eigenvalues over* 1 option selected, select a scree plot, and compare the results. If looking at the scree plot and the eigenvalues over 1 lead you to retain the same number of factors then continue with the analysis and be happy. If the two criteria give different results then examine the communalities and decide for yourself which of the two criteria to believe. If you decide to use the scree plot then you may want to redo the analysis specifying the number of factors to extract. The number of factors to be extracted can be specified by selecting *Number of factors* and then typing the appropriate number in the space provided (e.g. 4).

Rotation

We have already seen that the interpretability of factors can be improved through rotation. Rotation maximizes the loading of each variable on one of the extracted factors whilst minimizing the loading on all other factors. Rotation works through changing the absolute values of the variables whilst keeping their differential values constant. Click on Rotation... to access the dialog box in Figure 4.

Varimax, quartimax and equamax are all orthogonal rotations whilst direct oblimin and promax are oblique rotations (see Field 2000 for details). The exact choice of rotation will depend largely on whether or not you think that the underlying factors should be related. If you expect the factors to be independent then you should choose one of the orthogonal rotations (I recommend varimax). If, however, there are theoretical grounds for supposing that your factors might correlate then direct oblimin should be selected.

The dialog box also has options for displaying the *Rotated solution*. The rotated solution is displayed by default and is essential for interpreting the final rotated analysis.

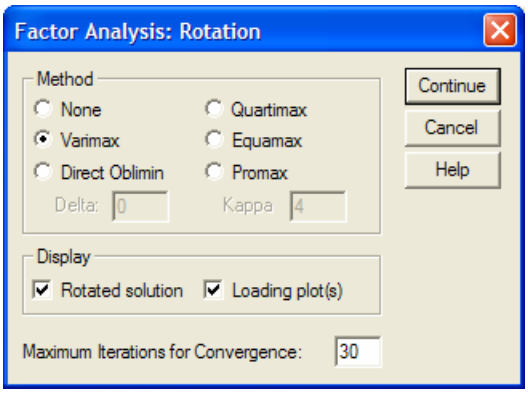


Figure 4: Factor analysis: *rotation* dialog box

Scores

The f*actor scores* dialog box can be accessed by clicking Scores... in the main dialog box. This option allows you to save factor scores for each subject in the data editor. SPSS creates a new column for each factor extracted and then places the factor score for each subject within that column. These scores can then be used for further analysis, or simply to identify groups of subjects who score highly on particular factors. There are three methods of obtaining these scores, all of which were described in sections 11.1.4 and 11.1.4.1 of Field (2000).

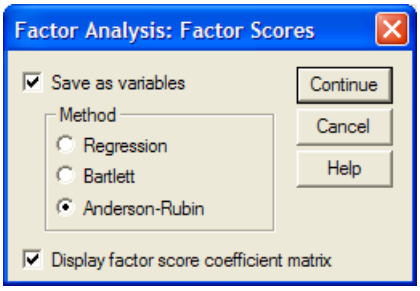


Figure 5: Factor analysis: *factor scores* dialog box

Options

This set of options can be obtained by clicking on Options... in the main dialog box. Two options relate to how coefficients are displayed. By default SPSS will list variables in the order in which they are entered into the data editor. Usually, this format is most convenient. However, when interpreting factors it is sometimes useful to list variables by size. By selecting *Sorted by size*, SPSS will order the variables by their factor loadings. The second option is to *Suppress absolute values less than* a specified value (by default 0.1). This option ensures that factor loadings within ±0.1 are not displayed in the output. Again, this option is useful for assisting in interpretation. The default value is not useful and I recommend changing it either to 0.4 or to a value reflecting the expected value of a significant factor loading given the sample size (see Field section 11.2.5.2). For this example set the value at 0.4.

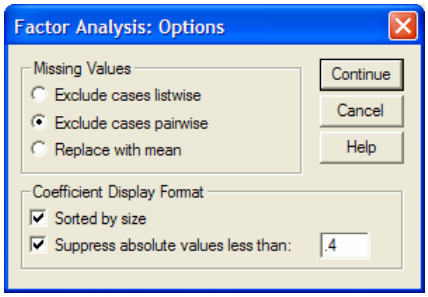


Figure 6: Factor analysis: *options* dialog box

Interpreting Output from SPSS

Select the same options as I have in the screen diagrams and run a factor analysis with orthogonal rotation. To save space each variable is referred to only by its label on the data editor (e.g. Q12). On the output *you* obtain, you should find that the SPSS uses the value label (the question itself) in all of the output. When using the output in this chapter just remember that Q1 represents question 1, Q2 represents question 2 and Q17 represents question 17.

Preliminary Analysis

SPSS Output 1 shows an abridged version of the *R*-matrix. The top half of this table contains the Pearson correlation coefficient between all pairs of questions whereas the bottom half contains the one-tailed significance of these coefficients. We can use this correlation matrix to check the pattern of relationships. First, scan the significance values and look for any variable for which the majority of values are greater than 0.05. Then scan the correlation coefficients themselves and look for any greater than 0.9. If any are found then you should be aware that a problem could arise because of singularity in the data: check the determinant of the correlation matrix and, if necessary, eliminate one of the two variables causing the problem. The determinant is listed at the bottom of the matrix (blink and you'll miss it). For these data its value is 5.271E–04 (which is 0.0005271) which is greater than the necessary value of 0.00001. Therefore, multicollinearity is not a problem for these data. To sum up, all questions in the SAQ correlate fairly well and none of the correlation coefficients are particularly large; therefore, there is no need to consider eliminating any questions at this stage.

Correlation Matrix[a]

		Q01	Q02	Q03	Q04	Q05	Q19	Q20	Q21	Q22	Q23
Correlation	Q01	1.000	-.099	-.337	.436	.402	-.189	.214	.329	-.104	-.004
	Q02	-.099	1.000	.318	-.112	-.119	.203	-.202	-.205	.231	.100
	Q03	-.337	.318	1.000	-.380	-.310	.342	-.325	-.417	.204	.150
	Q04	.436	-.112	-.380	1.000	.401	-.186	.243	.410	-.098	-.034
	Q05	.402	-.119	-.310	.401	1.000	-.165	.200	.335	-.133	-.042
	Q06	.217	-.074	-.227	.278	.257	-.167	.101	.272	-.165	-.069
	Q07	.305	-.159	-.382	.409	.339	-.269	.221	.483	-.168	-.070
	Q08	.331	-.050	-.259	.349	.269	-.159	.175	.296	-.079	-.050
	Q09	-.092	.315	.300	-.125	-.096	.249	-.159	-.136	.257	.171
	Q10	.214	-.084	-.193	.216	.258	-.127	.084	.193	-.131	-.062
	Q11	.357	-.144	-.351	.369	.298	-.200	.255	.346	-.162	-.086
	Q12	.345	-.195	-.410	.442	.347	-.267	.298	.441	-.167	-.046
	Q13	.355	-.143	-.318	.344	.302	-.227	.204	.374	-.195	-.053
	Q14	.338	-.165	-.371	.351	.315	-.254	.226	.399	-.170	-.048
	Q15	.246	-.165	-.312	.334	.261	-.210	.206	.300	-.168	-.062
	Q16	.499	-.168	-.419	.416	.395	-.267	.265	.421	-.156	-.082
	Q17	.371	-.087	-.327	.383	.310	-.163	.205	.363	-.126	-.092
	Q18	.347	-.164	-.375	.382	.322	-.257	.235	.430	-.160	-.080
	Q19	-.189	.203	.342	-.186	-.165	1.000	-.249	-.275	.234	.122
	Q20	.214	-.202	-.325	.243	.200	-.249	1.000	.468	-.100	-.035
	Q21	.329	-.205	-.417	.410	.335	-.275	.468	1.000	-.129	-.068
	Q22	-.104	.231	.204	-.098	-.133	.234	-.100	-.129	1.000	.230
	Q23	-.004	.100	.150	-.034	-.042	.122	-.035	-.068	.230	1.000
Sig. (1-tailed)	Q01		.000	.000	.000	.000	.000	.000	.000	.000	.410
	Q02	.000		.000	.000	.000	.000	.000	.000	.000	.000
	Q03	.000	.000		.000	.000	.000	.000	.000	.000	.000
	Q04	.000	.000	.000		.000	.000	.000	.000	.000	.043
	Q05	.000	.000	.000	.000		.000	.000	.000	.000	.017
	Q06	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Q07	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Q08	.000	.006	.000	.000	.000	.000	.000	.000	.000	.005
	Q09	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Q10	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001
	Q11	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Q12	.000	.000	.000	.000	.000	.000	.000	.000	.000	.009
	Q13	.000	.000	.000	.000	.000	.000	.000	.000	.000	.004
	Q14	.000	.000	.000	.000	.000	.000	.000	.000	.000	.007
	Q15	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001
	Q16	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Q17	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Q18	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Q19	.000	.000	.000	.000	.000		.000	.000	.000	.000
	Q20	.000	.000	.000	.000	.000	.000		.000	.000	.039
	Q21	.000	.000	.000	.000	.000	.000	.000		.000	.000
	Q22	.000	.000	.000	.000	.000	.000	.000	.000		.000
	Q23	.410	.000	.000	.043	.017	.000	.039	.000	.000	

a. Determinant = 5.271E-04

SPSS Output 1

SPSS Output 2 shows several very important parts of the output: the Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's test of sphericity. The KMO statistic varies

between 0 and 1. A value of 0 indicates that the sum of partial correlations is large relative to the sum of correlations, indicating diffusion in the pattern of correlations (hence, factor analysis is likely to be inappropriate). A value close to 1 indicates that patterns of correlations are relatively compact and so factor analysis should yield distinct and reliable factors. Kaiser (1974) recommends accepting values greater than 0.5 as acceptable (values below this should lead you to either collect more data or rethink which variables to include). Furthermore, values between 0.5 and 0.7 are mediocre, values between 0.7 and 0.8 are good, values between 0.8 and 0.9 are great and values above 0.9 are superb (see Hutcheson and Sofroniou, 1999, pp.224-225 for more detail). For these data the value is 0.93, which falls into the range of being superb: so, we should be confident that factor analysis is appropriate for these data.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.930
Bartlett's Test of Sphericity	Approx. Chi-Square	19334.492
	df	253
	Sig.	.000

SPSS Output 2

Bartlett's measure tests the null hypothesis that the original correlation matrix is an identity matrix. For factor analysis to work we need some relationships between variables and if the *R*-matrix were an identity matrix then all correlation coefficients would be zero. Therefore, we want this test to be *significant* (i.e. have a significance value less than 0.05). A significant test tells us that the *R*-matrix is not an identity matrix; therefore, there are some relationships between the variables we hope to include in the analysis. For these data, Bartlett's test is highly significant ($p < 0.001$), and therefore factor analysis is appropriate.

Factor Extraction

SPSS Output 3 lists the eigenvalues associated with each linear component (factor) before extraction, after extraction and after rotation. Before extraction, SPSS has identified 23 linear components within the data set (we know that there should be as many eigenvectors as there are variables and so there will be as many factors as variables). The eigenvalues associated with each factor represent the variance explained by that particular linear component and SPSS also displays the eigenvalue in terms of the percentage of variance explained (so, factor 1 explains 31.696% of total variance). It should be clear that the first few factors explain relatively large amounts of variance (especially factor 1) whereas subsequent factors explain only small amounts of variance. SPSS then extracts all factors with eigenvalues greater than 1, which leaves us with four factors. The eigenvalues associated with these factors are again displayed (and the percentage of variance explained) in the columns labelled *Extraction Sums of Squared Loadings*. The values in this part of the table are the same as the values before extraction, except that the values for the discarded factors are ignored (hence, the table is blank after the fourth factor). In the final part of the table (labelled *Rotation Sums of Squared Loadings*), the eigenvalues of the factors after rotation are displayed. Rotation has the effect of optimizing the factor structure and one

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	7.290	31.696	31.696	7.290	31.696	31.696	3.730	16.219	16.219
2	1.739	7.560	39.256	1.739	7.560	39.256	3.340	14.523	30.742
3	1.317	5.725	44.981	1.317	5.725	44.981	2.553	11.099	41.842
4	1.227	5.336	50.317	1.227	5.336	50.317	1.949	8.475	50.317
5	.988	4.295	54.612						
6	.895	3.893	58.504						
7	.806	3.502	62.007						
8	.783	3.404	65.410						
9	.751	3.265	68.676						
10	.717	3.117	71.793						
11	.684	2.972	74.765						
12	.670	2.911	77.676						
13	.612	2.661	80.337						
14	.578	2.512	82.849						
15	.549	2.388	85.236						
16	.523	2.275	87.511						
17	.508	2.210	89.721						
18	.456	1.982	91.704						
19	.424	1.843	93.546						
20	.408	1.773	95.319						
21	.379	1.650	96.969						
22	.364	1.583	98.552						
23	.333	1.448	100.000						

Extraction Method: Principal Component Analysis.

SPSS Output 3

consequence for these data is that the relative importance of the four factors is equalized. Before rotation, factor 1 accounted for considerably more variance than the remaining three (31.696% compared to 7.560, 5.725, and 5.336%), however after extraction it accounts for only 16.219% of variance (compared to 14.523, 11.099 and 8.475% respectively).

SPSS Output 4 shows the table of communalities before and after extraction. Principal component analysis works on the initial assumption that all variance is common; therefore, before extraction the communalities are all 1. The communalities in the column labelled *Extraction* reflect the common variance in the data structure. So, for example, we can say that 43.5% of the variance associated with question 1 is common, or shared, variance. Another way to look at these communalities is in terms of the proportion of variance explained by the underlying factors. After extraction some of the factors are discarded and so some information is lost. The amount of variance in each variable that can be explained by the retained factors is represented by the communalities after extraction.

Communalities

	Initial	Extraction
Q01	1.000	.435
Q02	1.000	.414
Q03	1.000	.530
Q04	1.000	.469
Q05	1.000	.343
Q06	1.000	.654
Q07	1.000	.545
Q08	1.000	.739
Q09	1.000	.484
Q10	1.000	.335
Q11	1.000	.690
Q12	1.000	.513
Q13	1.000	.536
Q14	1.000	.488
Q15	1.000	.378
Q16	1.000	.487
Q17	1.000	.683
Q18	1.000	.597
Q19	1.000	.343
Q20	1.000	.484
Q21	1.000	.550
Q22	1.000	.464
Q23	1.000	.412

Extraction Method: Principal Component

Component Matrix[a]

	Component			
	1	2	3	4
Q18	.701			
Q07	.685			
Q16	.679			
Q13	.673			
Q12	.669			
Q21	.658			
Q14	.656			
Q11	.652			-.400
Q17	.643			
Q04	.634			
Q03	-.629			
Q15	.593			
Q01	.586			
Q05	.556			
Q08	.549	.401		-.417
Q10	.437			
Q20	.436		-.404	
Q19	-.427			
Q09		.627		
Q02		.548		
Q22		.465		
Q06	.562		.571	
Q23				.507

Extraction Method: Principal Component Analysis.

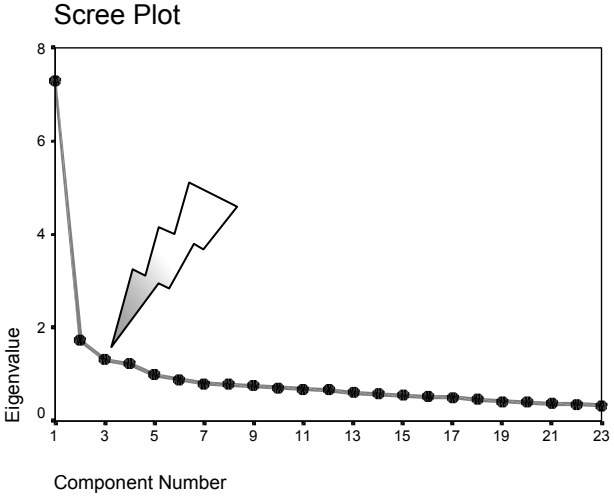
a. 4 components extracted.

SPSS Output 4

This output also shows the component matrix before rotation. This matrix contains the loadings of each variable onto each factor. By default SPSS displays all loadings; however, we requested that all loadings less than 0.4 be suppressed in the output and so there are blank spaces for many of the loadings. This matrix is not particularly important for interpretation.

At this stage SPSS has extracted four factors. Factor analysis is an exploratory tool and so it should be used to guide the researcher to make various decisions: you shouldn't leave the computer to make them. One important decision is the number of factors to extract. By Kaiser's criterion we should extract four factors and this is what SPSS has done. However, this criterion is accurate when there are less than 30 variables and communalities after extraction are greater than 0.7 or when the sample size exceeds 250 and the average communality is greater than 0.6. The communalities are shown in SPSS Output 4, and none exceed 0.7. The average of the communalities can be found by adding them up and dividing by the number of communalities (11.573/23 = 0.503). So, on both grounds Kaiser's rule may not be accurate. However, you should consider the huge sample that we have, because the research into Kaiser's criterion gives recommendations for much smaller samples. We can also use the scree plot, which we asked SPSS to produce. The scree plot is shown below with a thunderbolt indicating the point of inflexion on the curve. This curve is difficult to interpret because the

curve begins to tail off after three factors, but there is another drop after four factors before a stable plateau is reached. Therefore, we could probably justify retaining either two or four factors. Given the large sample, it is probably safe to assume Kaiser's criterion; however, you could rerun the analysis specifying that SPSS extract only two factors and compare the results.



SPSS Output 5

Factor Rotation

The first analysis I asked you to run was using an orthogonal rotation. SPSS Output 6 shows the rotated component matrix (also called the rotated factor matrix in factor analysis) which is a matrix of the factor loadings for each variable onto each factor. This matrix contains the same information as the component matrix in SPSS Output 4 except that it is calculated *after* rotation. There are several things to consider about the format of this matrix. First, factor loadings less than 0.4 have not been displayed because we asked for these loadings to be suppressed. If you didn't select this option, or didn't adjust the criterion value to 0.4, then your output will differ. Second, the variables are listed in the order of size of their factor loadings because we asked for the output to be *Sorted by size*. If this option was not selected your output will look different. Finally, for all other parts of the output I suppressed the variable labels (for reasons of space) but for this matrix I have allowed the variable labels to be printed to aid interpretation.

Rotated Component Matrix[a]

	Component			
	1	2	3	4
I have little experience of computers	.800			
SPSS always crashes when I try to use it	.684			
I worry that I will cause irreparable damage because of my incompetenece with computers	.647			
All computers hate me	.638			
Computers have minds of their own and deliberately go wrong whenever I use them	.579			
Computers are useful only for playing games	.550			
Computers are out to get me	.459			
I can't sleep for thoughts of eigen vectors		.677		
I wake up under my duvet thinking that I am trapped under a normal distribtion		.661		
Standard deviations excite me		-.567		
People try to tell you that SPSS makes statistics easier to understand but it doesn't	.473	.523		
I dream that Pearson is attacking me with correlation coefficients		.516		
I weep openly at the mention of central tendency		.514		
Statiscs makes me cry		.496		
I don't understand statistics		.429		
I have never been good at mathematics			.833	
I slip into a coma whenever I see an equation			.747	
I did badly at mathematics at school			.747	
My friends are better at statistics than me				.648
My friends are better at SPSS than I am				.645
If I'm good at statistics my friends will think I'm a nerd				.586
My friends will think I'm stupid for not being able to cope with SPSS				.543
Everybody looks at me when I use SPSS				.427

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 9 iterations.

SPSS Output 6

Compare this matrix with the unrotated solution. Before

rotation, most variables loaded highly onto the first factor and the remaining factors didn't really get a look in. However, the rotation of the factor structure has clarified things considerably: there are four factors and variables load very highly onto only one factor (with the exception of one question). The suppression of loadings less than 0.4 and ordering variables by loading size also makes interpretation considerably easier (because you don't have to scan the matrix to identify substantive loadings).

The next step is to look at the content of questions that load onto the same factor to try to identify common themes. If the mathematical factor produced by the analysis represents some real-world construct then common themes among highly loading questions can help us identify what the construct might be. The questions that load highly on factor 1 seem to all relate to using computers or SPSS. Therefore we might label this factor *fear of computers*. The questions that load highly on factor 2 all seem to relate to different aspects of statistics; therefore, we might label this factor *fear of statistics*. The three questions that load highly on factor 3 all seem to relate to mathematics; therefore, we might label this factor *fear of mathematics*. Finally, the questions that load highly on factor 4 all contain some component of social evaluation from friends; therefore, we might label this factor *peer evaluation*. This analysis seems to reveal that the initial questionnaire, in reality, is composed of four sub-scales: fear of computers, fear of statistics, fear of maths, and fear of negative peer evaluation. There are two possibilities here. The first is that the SAQ failed to measure what it set out to (namely SPSS anxiety) but does measure some related constructs. The second is that these four constructs are sub-components of SPSS anxiety; however, the factor analysis does not indicate which of these possibilities is true.

Exercise

Re-run the analysis using oblique rotation (Use Field, 2000 to help you). Compare the results to the current analysis. Also, look over Field (2000) and find out about Factor Scores and how to interpret them.

Example 2:

The University of Sussex is constantly seeking to employ the best people possible as lecturers (no, really, it is). Anyway, they wanted to revise a questionnaire based on Bland's theory of research methods lecturers. This theory predicts that good research methods lecturers should have four characteristics: (1) a profound love of statistics; (2) an enthusiasm for experimental design; (3) a love of teaching; and (4) a complete absence of normal interpersonal skills. These characteristics should be related (i.e. correlated). The 'Teaching Of Statistics for Scientific Experiments' (TOSSE) already existed, but the university revised this questionnaire and it became the 'Teaching Of Statistics for Scientific Experiments — Revised' (TOSSE—R). The gave this questionnaire to 239 research methods lecturers around the world to see if it supported Bland's theory. The questionnaire is in **Error! Reference source not found.**, and the data are in **TOSSE-R.sav**. Conduct a factor analysis (with appropriate rotation) to see the factor structure of the data.

SD = Strongly Disagree, D = Disagree, N = Neither, A = Agree, SA = Strongly Agree						
		SD	**D**	**N**	**A**	**SA**
1	I once woke up in the middle of a vegetable patch hugging a turnip that I'd mistakenly dug up thinking it was Roy's largest root	○	○	○	○	○
2	If I had a big gun I'd shoot all the students I have to teach	○	○	○	○	○
3	I memorize probability values for the F-distribution	○	○	○	○	○
4	I worship at the shrine of Pearson	○	○	○	○	○
5	I still live with my mother and have little personal hygiene	○	○	○	○	○
6	Teaching others makes me want to swallow a large bottle of bleach because the pain of my burning oesophagus would be light relief in comparison	○	○	○	○	○
7	Helping others to understand Sums of Squares is a great feeling	○	○	○	○	○
8	I like control conditions	○	○	○	○	○
9	I calculate 3 ANOVAs in my head before getting out of bed every morning	○	○	○	○	○
10	I could spend all day explaining statistics to people	○	○	○	○	○
11	I like it when people tell me I've helped them to understand factor rotation	○	○	○	○	○
12	People fall asleep as soon as I open my mouth to speak	○	○	○	○	○
13	Designing experiments is fun	○	○	○	○	○
14	I'd rather think about appropriate dependent variables than go to the pub	○	○	○	○	○
15	I soil my pants with excitement at the mere mention of Factor Analysis	○	○	○	○	○
16	Thinking about whether to use repeated or independent measures thrills me	○	○	○	○	○
17	I enjoy sitting in the park contemplating whether to use participant observation in my next experiment	○	○	○	○	○
18	Standing in front of 300 people in no way makes me lose control of my bowels	○	○	○	○	○
19	I like to help students	○	○	○	○	○
20	Passing on knowledge is the greatest gift you can bestow an individual	○	○	○	○	○
21	Thinking about Bonferroni corrections gives me a tingly feeling in my groin	○	○	○	○	○
22	I quiver with excitement when thinking about designing my next experiment	○	○	○	○	○
23	I often spend my spare time talking to the pigeons ... and even they die of boredom	○	○	○	○	○
24	I tried to build myself a time machine so that I could go back to the 1930s and follow Fisher around on my hands and knees licking the floor on which he'd just trodden	○	○	○	○	○
25	I love teaching	○	○	○	○	○

26	I spend lots of time helping students	○	○	○	○	○
27	I love teaching because students have to pretend to like me or they'll get bad marks	○	○	○	○	○
28	My cat is my only friend	○	○	○	○	○

The Teaching of Statistics for Scientific Experiments – Revised (TOSSE-R)