Factor Analysis - SPSS©

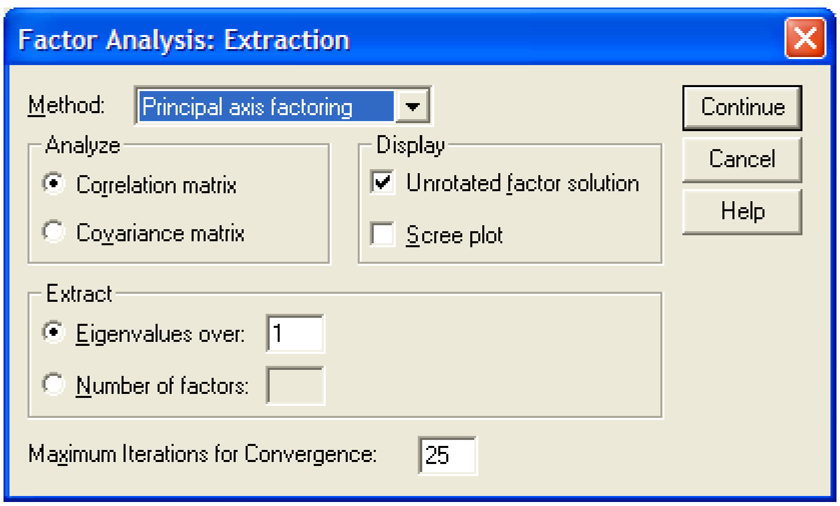
- First Read Principal Components Analysis.

The methods we have employed so far attempt to repackage all of the variance in the *p* variables into principal components. We may wish to restrict our analysis to variance that is common among variables. That is, when repackaging the variables' variance we may wish not to redistribute variance that is unique to any one variable. This is **Common Factor Analysis**. A common factor is an abstraction, a hypothetical dimension that affects at least two of the variables. We assume that there is also one unique factor for each variable, a factor that affects that variable but does not affect any other variables. We assume that the *p* unique factors are uncorrelated with one another and with the common factors. It is the variance due to these unique factors that we shall exclude from our FA.

Iterated Principal Factors Analysis

The most common sort of FA is **principal axis FA**, also known as **principal factor analysis**. This analysis proceeds very much like that for a PCA. We eliminate the variance due to unique factors by replacing the 1's on the main diagonal of the correlation matrix with estimates of the variables' communalities. Recall that a variable's communality, its SSL across components or factors, is the amount of the variable's variance that is accounted for by the components or factors. Since our factors will be common factors, a variable's communality will be the amount of its variance that is common rather than unique. The R^2 between a variable and all other variables is most often used initially to estimate a variable's communality.

Using the beer data, change the extraction method to principal axis:



Take a look at the initial communalities (for each variable, this is the R^2 for predicting that variable from an optimally weighted linear combination of the remaining variables). Recall that they were all 1's for the principal components analysis we did earlier, but now each is less than 1. If we sum these communalities we get 5.675. We started with 7 units of standardized variance and we have now reduced that to 5.675 units of standardized variance (by eliminating unique variance).

Communalities

	Initial	Extraction
COST	.738	.745
SIZE	.912	.914
ALCOHOL	.866	.866
REPUTAT	.499	.385
COLOR	.922	.892
AROMA	.857	.896
TASTE	.881	.902

Extraction Method: Principal Axis Factoring.

For an iterated principal axis solution SPSS first estimates communalities, with R^2 's, and then conducts the analysis. It then takes the communalities from that first analysis and inserts them into the main diagonal of the correlation matrix in place of the R^2 's, and does the analysis again. The variables' SSL's from this second solution are then inserted into the main diagonal replacing the communalities from the previous iteration, etc. etc., until the change from one iteration to the next iteration is trivial.

Look at the communalities after this iterative process and for a two-factor solution. They now sum to 5.60. That is, 5.6/7 = 80% of the variance is common variance and 20% is unique. Here you can see how we have packaged that common variance into two factors, both before and after a varimax rotation:

Total Variance Explained

	Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.123	44.620	44.620	2.879	41.131	41.131
2	2.478	35.396	80.016	2.722	38.885	80.016

Extraction Method: Principal Axis Factoring.

The final rotated loadings are:

Rotated Factor Matrix[a]

	Factor	
	1	2
TASTE	.950	-2.17E-02
AROMA	.946	2.106E-02
COLOR	.942	6.771E-02
SIZE	7.337E-02	.953
ALCOHOL	2.974E-02	.930
COST	-4.64E-02	.862
REPUTAT	-.431	-.447

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

a. Rotation converged in 3 iterations.

These loadings are very similar to those we obtained previously with a principal components analysis.

Reproduced and Residual Correlation Matrices

Having extracted common factors, one can turn right around and try to reproduce the correlation matrix from the factor loading matrix. We assume that the correlations between variables result from their sharing common underlying factors. Thus, it makes sense to try to estimate the correlations between variables from the correlations between factors and variables. The reproduced correlation matrix is obtained by multiplying the loading matrix by the transposed loading matrix. This results in calculating each reproduced correlation as the sum across factors (from 1 to *m*) of the products (*r* between factor and the one variable)(*r* between factor and the other variable). For example, for our 2 factor iterative solution the reproduced correlation between COLOR and TASTE = (*r* for color - Factor1)(*r* for taste - Factor 1) + (*r* for color - Factor 2)(*r* for taste - Factor2) = (.95)(.94) + (.06)(-.02) = .89. The original *r* between color and taste was .90, so our two factors did indeed capture the relationship between Color and Taste.

The residual correlation matrix equals the original correlation matrix minus the reproduced correlation matrix. We want these residuals to be small. If you check "Reproduced" under "Descriptive" in the Factor Analysis dialogue box, you will get both of these matrices:

Reproduced Correlations

		COST	SIZE	ALCOHOL	REPUTAT	COLOR	AROMA	TASTE
Reproduced Correlation	COST	.745[b]	.818	.800	-.365	1.467E-02	-2.57E-02	-6.28E-02
	SIZE	.818	.914[b]	.889	-.458	.134	8.950E-02	4.899E-02
	ALCOHOL	.800	.889	.866[b]	-.428	9.100E-02	4.773E-02	8.064E-03
	REPUTAT	-.365	-.458	-.428	.385[b]	-.436	-.417	-.399
	COLOR	1.467E-02	.134	9.100E-02	-.436	.892[b]	.893	.893
	AROMA	-2.57E-02	8.950E-02	4.773E-02	-.417	.893	.896[b]	.898
	TASTE	-6.28E-02	4.899E-02	8.064E-03	-.399	.893	.898	.902[b]
Residual[a]	COST		1.350E-02	-3.295E-02	-4.02E-02	3.328E-03	-2.05E-02	-1.16E-03
	SIZE	1.350E-02		1.495E-02	6.527E-02	4.528E-02	8.097E-03	-2.32E-02
	ALCOHOL	-3.29E-02	1.495E-02		-3.47E-02	-1.88E-02	-3.54E-03	3.726E-03
	REPUTAT	-4.02E-02	6.527E-02	-3.471E-02		6.415E-02	-2.59E-02	-4.38E-02
	COLOR	3.328E-03	4.528E-02	-1.884E-02	6.415E-02		1.557E-02	1.003E-02
	AROMA	-2.05E-02	8.097E-03	-3.545E-03	-2.59E-02	1.557E-02		-2.81E-02
	TASTE	-1.16E-03	-2.32E-02	3.726E-03	-4.38E-02	1.003E-02	-2.81E-02	

Extraction Method: Principal Axis Factoring.

a. Residuals are computed between observed and reproduced correlations. There are 2 (9.0%) nonredundant residuals with absolute values greater than 0.05.

b. Reproduced communalities

Nonorthogonal (Oblique) Rotation

The data may be better fit with axes that are not perpendicular. This can be done by means of an oblique rotation, but the factors will now be correlated with one another. Also, the factor **loadings** (in the **pattern matrix**) will no longer be equal to the correlation between each factor and each variable. They will still be standardized regression coefficients (Beta weights), the *A*'s in the $X_j = A_{1j}F_1 + A_{2j}F_2 + \ldots + A_{mj}F_m + U_j$ formula presented at the beginning of the handout on principal components analysis. The **correlations** between factors and variables are presented in a factor **structure matrix**.

I am slowly becoming comfortable with oblique rotations. For this lesson I tried a Promax rotation (a varimax rotation is first applied and then the resulting axes rotated to oblique positions):

Beta Weights	Correlations

Pattern Matrix[a]

	Factor	
	1	2
TASTE	.955	-7.14E-02
AROMA	.949	-2.83E-02
COLOR	.943	1.877E-02
SIZE	2.200E-02	.953
ALCOHOL	-2.05E-02	.932
COST	-9.33E-02	.868
REPUTAT	-.408	-.426

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

a. Rotation converged in 3 iterations.

Structure Matrix

	Factor	
	1	2
TASTE	.947	.030
AROMA	.946	.072
COLOR	.945	.118
SIZE	.123	.956
ALCOHOL	.078	.930
COST	-.002	.858
REPUTAT	-.453	-.469

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

There is some controversy regarding whether it is best to present the beta weights or the correlations when reporting the results of a factor analysis with oblique rotation. While one could report both, that would increase production costs, so usually only one will be presented. When there are only two or three factors, one should report the coefficients (betas or correlations) for every factor that was retained. If, however, there are a larger number of factors, one may shrink the table by reporting, for each variable, only the coefficient with the largest absolute value, and identification of the corresponding factor. This, of course, prevents the reader from seeing to what extent a variable may load on more than one factor. Here is an example of such a table.

Factor Correlation Matrix

Factor	1	2
1	1.000	.106
2	.106	1.000

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

Notice that this solution is not much different from the previously obtained varimax solution, so little was gained by allowing the factors to be correlated.

Exact Factor Scores

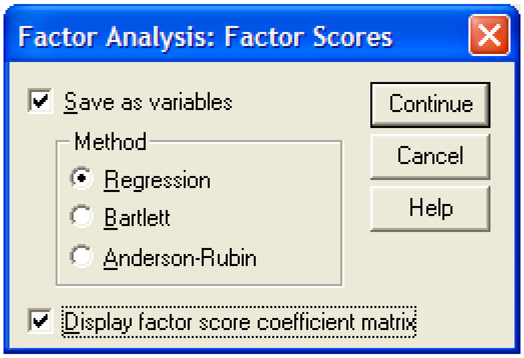
One may wish to define subscales on a test, with each subscale representing one factor. Using an "exact" weighting scheme, each subject's estimated factor score on each factor is a weighted sum of the products of scoring coefficients and the subject's standardized scores on the original variables.

The regression coefficients (standardized scoring coefficients) for converting scores on variables to factor scores are obtained by multiplying the inverse of the original simple correlation matrix by the factor loading matrix. To obtain a subject’s factor scores you multiply e’s standardized scores (*Z*’s) on the variables by these standardized scoring coefficients. For example, subject # 1’s Factor scores are:

Factor 1: (-.294)(.41) + (.955)(.40) + (-.036)(.22) + (1.057)(-.07) + (.712)(.04) + (1.219)(.03) + (-1.14)(.01) = 0.23.

Factor 2: (-.294)(.11) + (.955)(.03) + (-.036)(-.20) + (1.057)(.61) + (.712)(.25) + (.16)(1.219) + (-1.14)(-.04) = 1.06

SPSS will not only compute the scoring coefficients for you, it will also output the factor scores of your subjects into your SPSS data set so that you can input them into other procedures. In the Factor Analysis window, click Scores and select Save As Variables, Regression, Display Factor Score Coefficient Matrix.



Here are the scoring coefficients:

Factor Score Coefficient Matrix

	Factor	
	1	2
COST	.026	.157
SIZE	-.066	.610
ALCOHOL	.036	.251
REPUTAT	.011	-.042
COLOR	.225	-.201
AROMA	.398	.026
TASTE	.409	.110

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
Factor Scores Method: Regression.

Look back at your data sheet. You will find that two columns have been added to the right, one for scores on Factor 1 and another for scores on Factor 2.

SPSS also gives you a Factor Score Covariance Matrix. On the main diagonal of this matrix are, for each factor, the R^2 between the factor and the observed variables. This is treated as an indictor of the internal consistency of the solution. Values below .70 are considered undesirable. These squared multiple correlation coefficients are equal to the variance of the factor scores:

Factor Score Covariance Matrix

Factor	1	2
1	.966	.003
2	.003	.953

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
Factor Scores Method: Regression.

Descriptive Statistics

	N	Mean	Variance
FAC1_1	220	.0000000	.966
FAC2_1	220	.0000000	.953

The input data included two variables (SES and Group) not included in the factor analysis. Just for fun, try conducting a multiple regression predicting subjects' SES from their factor scores and also try using Student's *t* to compare the two groups' means on the factor scores. Do note that the scores for Factor 1 are not correlated with those for Factor 2. Accordingly, in the multiple regression the squared semipartial correlation coefficients are identical to squared zero-order correlation coefficients and the $R^2 = r_{Y1}^2 + r_{Y2}^2$.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.988[a]	.976	.976	.385

a. Predictors: (Constant), FAC2_1, FAC1_1

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1320.821	2	660.410	4453.479	.000[a]
	Residual	32.179	217	.148		
	Total	1353.000	219			

a. Predictors: (Constant), FAC2_1, FAC1_1

b. Dependent Variable: SES

Coefficients[a]

Model		Standardized Coefficients	t	Sig.	Correlations	
		Beta			Zero-order	Part
1	(Constant)		134.810	.000		
	FAC1_1	.681	65.027	.000	.679	.681
	FAC2_1	-.718	-68.581	.000	-.716	-.718

a. Dependent Variable: SES

Group Statistics

	GROUP	N	Mean	Std. Deviation	Std. Error Mean
FAC1_1	1	121	-.4198775	.97383364	.08853033
	2	99	.5131836	.71714232	.07207552
FAC2_1	1	121	.5620465	.88340921	.08030993
	2	99	-.6869457	.55529938	.05580969

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means				
							95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Lower	Upper
FAC1_1	Equal variances assumed	19.264	.000	-7.933	218	.000	-1.16487	-.701253
	Equal variances not assumed			-8.173	215.738	.000	-1.15807	-.708049
FAC2_1	Equal variances assumed	25.883	.000	12.227	218	.000	1.047657	1.450327
	Equal variances not assumed			12.771	205.269	.000	1.056175	1.441809

Unit-Weighted Factor Scores

If one were using FA to define subscales on a test, e might decide to define subscale 1 as being an unweighted sum of all of a subject's scores (for example, 1....7 on a Likert item) on any item loading heavily (>.4) on Factor 1, etc. etc. Such a method is sometimes referred to as a "unit-weighting scheme." For our data, if I answered the Color, Taste, Aroma, Size, Alcohol, Cost, Reputation questions with the responses 80, 100, 40, 30, 75, 60, 10, my Subscale 1 (Aesthetic Quality) subscale score would be 80 + 100 + 40 - 10 = 210 [note that I subtracted the Reputation score since its loading was negative] and my Subscale 2 (Cheap Drunk) score would be 30 + 75 + 60 - 10 = 155.

Certain types of unit-weighting schemes may actually be preferable to exact factor scores in certain circumstances (Grice, 2001; Grice & Harris, 1998). I recommend Grice, J. W. (2001). A comparison of factor scores under conditions of factor obliquity, *Psychological Methods*, *6*, 67-83. Grice reviewed the literature on this topic and presented results of his own Monte Carlo study (with oblique factor analysis). The early studies which suggested that unit-weighted factor scores are preferable to exact factor scores used data for which the structure was simple (each variable loading well on only one factor) and the loadings not highly variable (most loadings either very high or very low). Grice & Harris, 1998, have shown that unit-weighted factor scores based on the loadings perform poorly under conditions of non-simple structure and variable loadings, which are typical of the conditions most often found in actual practice. They developed an alternative unit-weighting scheme which produced factor scores that compared favorably with exact factor scores -- they based the weightings on the factor score coefficients rather than on the loadings.

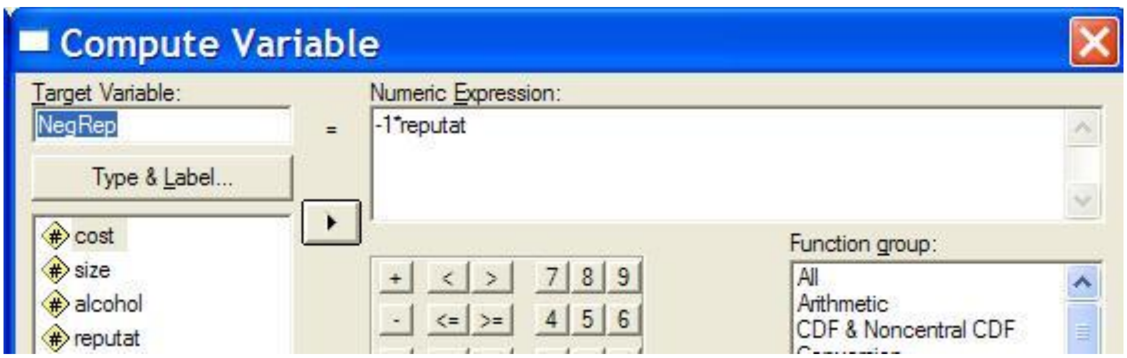
Grice's (2001) article extended the discussion to the case of oblique factor analysis, where one could entertain several different sorts of unit-weighting schemes -- for example, basing them on the pattern matrix (loadings, standardized regression coefficients for predicting), the structure matrix (correlations of variables with factors), or the factor score coefficients. Grice defined a variable as salient on a factor if it had a weighting coefficient whose absolute value was at least 1/3 as large as that of the variable with the largest absolute weighting coefficient on that factor. Salient items' weights were replaced with 1 or –1, and nonsalient variables' weights with 0. The results of his Monte Carlo study indicated that factor scores using this unit-weighting scheme based on scoring coefficients performed better than those using various other unit-weighting schemes and at least as good as exact factor scores (by most criteria and under most conditions). He did note, however, that exact factor scores may be preferred under certain circumstances -- for example, when using factor scores on the same sample as that from which they were derived, especially when sample size is relatively small.

Cronbach's Alpha

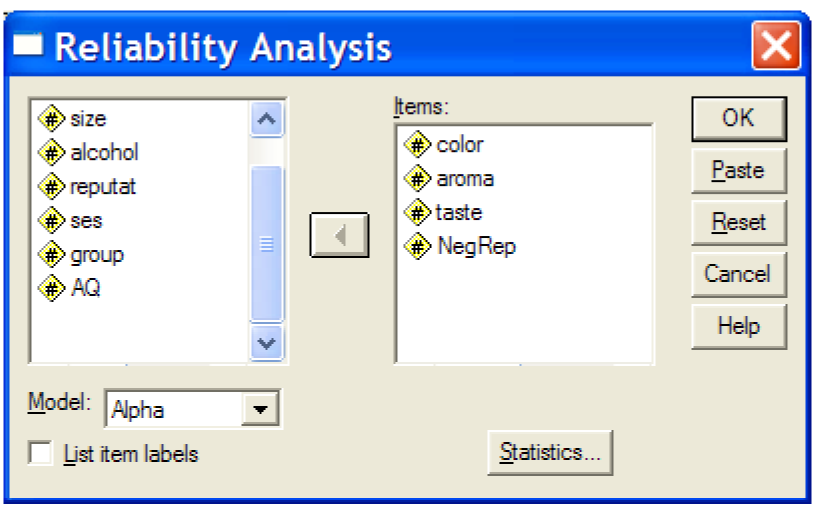
If you have developed subscales such as the Aesthetic Quality and Cheap Drunk subscales above, you should report an estimate of the reliability of each subscale. Test-retest reliability can be employed if you administer the scale to the same persons twice, but usually you will only want to administer it once to each person. Cronbach's alpha is an easy and generally acceptable estimate of reliability.

Suppose that we are going to compute AQ (Aesthetic Quality) as color + taste + aroma – reputat and CD as cost + size + alcohol – reputat. How reliable would such subscales be? We conduct an item analysis to evaluate the reliability (and internal consistency) of the each subscale.

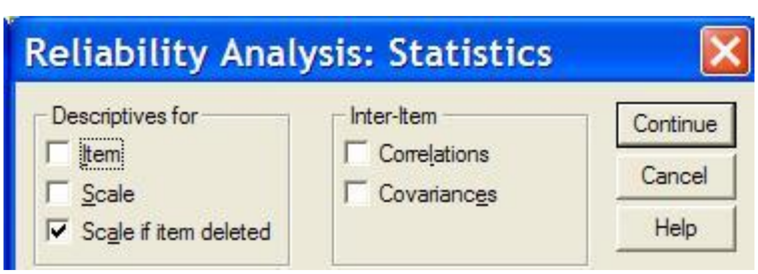
Before conducting the item analysis, we shall need to multiply the Reputation variable by minus 1, since it is negatively weighted in the AQ and CD subscale scores. Transform Compute NegRep = –1 ∗ reputat.



Analyze, Scale, Reliability Analysis. Scoot color, aroma, taste, and NegRep into the items box.



Click Statistics. Check "Scale if item deleted." Continue, OK.



Reliability Statistics

Cronbach's Alpha	N of Items
.886	4

Alpha is .886. Shoot for an alpha of at least .7 for research instruments.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
color	63.7500	3987.814	.859	.810
aroma	70.0000	4060.959	.881	.802
taste	47.5000	4273.174	.879	.807
NegRep	163.0000	5485.936	.433	.961

Notice that NegRep is not as well correlated with the corrected total scores as are the other items and that dropping it from the scale would increase the value of alpha considerably. That might be enough to justify dropping the reputation variable from this subscale.

If you conduct an item analysis on the CD items you will find alpha = .878 and increases to .941 if Reputation is dropped from the scale.

Comparing Two Groups' Factor Structure

Suppose I wish to compare the factor structure in one population with that in another. I first collect data from randomly selected samples from each population. I then extract factors from each group using the same criteria for number of factors etc. in both groups. An **eyeball test** should tell whether or not the two structures are dissimilar. Do I get the same number of well defined factors in the two groups? Is the set of variables loading well on a given factor the same in Group A as in Group B (minor differences would be expected due to sampling error, of course - one can randomly split one sample in half and do separate analyses on each half to get a picture of such error). Consider the following hypothetical loading matrices:

	Group A			Group B		
Variable	Factor1	Factor2	Factor3	Factor1	Factor2	Factor3
X1	.90	.12	.21	.45	.49	-.70
X2	.63	.75	.34	.65	-.15	.22
X3	.15	.67	.24	.27	.80	.04
X4	-.09	-.53	-.16	-.15	.09	.67
X5	-.74	-.14	-.19	.95	-.79	.12
SSL	1.78	1.32	0.28	1.62	1.53	1.00

One may **use a simple Pearson *r* to compare two factors**. Just correlate the loadings on the factor in Group A with the loadings on the maybe similar factor in Group B. For Factor 1 in Group A compared with Factor 1 in Group B, the A(B) data are .90(.45), .63(.65), .15(.27), -.09(-.15), & -.74(.95). The *r* is -0.19, indicating little similarity between the two factors.

The Pearson *r* can detect not only differences in two factors' patterns of loadings, but also differences in the relative magnitudes of those loadings. One should beware that with factors having a large number of small loadings, those small loadings could cause the *r* to be large (if they are similar between factors) even if the factors had dissimilar loadings on the more important variables.

CC, Tucker's coefficient of congruence. Perform a procrustean rotation of the solution for the one group, which makes it as similar as possible to the solution in the other group. Multiply each loading in the one group by the corresponding loading in the other group. Sum these products and then divide by the square root of (the sum of squared loadings for the one group times the sum of squared loading for the other group). CC of .85 to .94 corresponds to similar factors, and .95 to 1 as essentially identical factors.

See Comparing Two Groups' Factor Structures: Pearson *r* and the Coefficient of Congruence for an example of computing *r* and Tucker's CC.

Cross-correlated factor scores. Compute factor scoring coefficients for Group 1 and, separately, for Group 2. Then for each case compute the factor score using the scoring coefficients from the group in which is it located and also compute it using the scoring coefficients from the other group. Correlate these two sets of factor scores – Same Group and Other Group. A high correlation between these two sets of factor scores should indicate similarity of the two factors between groups. Of course, this method and the other two could be used with random halves of one sample to assess the stability of the solution or with different random samples from the same population at different times to get something like a test-retest measure of stability across samples and times.

RMS, root mean square. For each variable square the difference between the loading in the one group and that in the other group. Find the mean of these differences and then the square root of that mean. If there is a perfect match between the two groups' loadings, RMS = 0. The maximum value of RMS (2) would result when all of the loadings are one or minus one, with those in the one group opposite in sign of those in the other group.

Cattell's salient similarity index (*s*). I have not often seen this method used in recently published articles. If you would like to learn a little about it, see my document Cattell's *s* .

See Factorial Invariance of the Occupational Work Ethic Inventory -- An example of the use of multiple techniques to compare factor structures.

Required Number of Subjects and Variables

"How many subjects do I need for my factor analysis?" If you ask this question of several persons who occasionally employ factor analysis, you are likely to get several different answers. Some may answer "100 or more" or give you some other fixed minimum *N*. Others may say "That depends on how many variables you have, you should have at least 10 times as many subjects as you have variables" (or some other fixed minimum ratio of subjects to variables). Others may say "get as many subjects as you can, the more the better, but you must have at least as many subjects as you have variables." I would say "that depends on several things." So what are those several things? Here I shall summarize the key points of two recent articles that address this issue. The first also addresses the question of how many variables should be included in the analysis. The articles are:

Velicer, W. F., & Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, *3*, 231-251.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*, 84-99.

Velicer & Fava briefly reviewed the literature and then reported the results of their empirical study. The main points were as follows:

- Factors that do not have at least three variables with high loadings should not be interpreted – accordingly, you need to have at least three variables per factor, probably more. Since not all of your variables can be expected to perform as anticipated, you should start out with 6 to 10 variables per factor.
- If the loadings are low, you will need more variables, 10 or 20 per factor may be required. Assuming that some variables will have to be discarded, you should start with 20 to 30 variables per factor.
- The larger the number of subjects, the larger the number of variables per factor, and the larger the loadings, the better are your chances of doing a good job at reproducing the population factor pattern. Strength in one of these areas can compensate for weakness in another – for example, you might compensate for mediocre loadings and variable sample size by having a very large number of subjects. If you can't get a lot of subjects, you can compensate by having many good variables per factor (start with 30 – 60 per factor, but not more variables than subjects).

MacCullum et al. demonstrated, by both theoretical and empirical means that the required sample size or ratio of subjects to variables is not constant across studies, but rather varies as a function of several other aspects of the research. They concluded that rules of thumb about sample size are not valid and not useful. Here are the key points:

- Larger **sample size**, higher **communalities** (low communalities are associated with sampling error due to the presence of unique factors that have nonzero correlations with one another and with the common factors), and high **overdetermination** [each factor having at least three or four high loadings and **simple structure** (few, nonoverlapping factors)] each increase your chances of faithfully reproducing the population factor pattern.
- Strengths in one area can compensate for weaknesses in another area.
- When communalities are high (> .6), you should be in good shape even with *N* well below 100.
- With communalities moderate (about .5) and the factors well-determined, you should have 100 to 200 subjects.
- With communalities low (< .5) but high overdetermination of factors (not many factors, each with 6 or 7 high loadings), you probably need well over 100 subjects.
- With low communalities and only 3 or 4 high loadings on each, you probably need over 300 subjects.
- With low communalities and poorly determined factors, you will need well over 500 subjects.

Of course, when planning your research you do not know for sure how good the communalities will be nor how well determined your factors will be, so maybe the best simple advise, for an a priori rule of thumb, is "the more subjects, the better." MacCallum's advise to researchers is to try to keep the number of variables and factors small and select variables (write items) to assure moderate to high communalities.

Missing Data

Unless there is little missing data, best to use multiple imputation. Bruce Weaver has developed easy-to-use code for this. Check it out.

Closing Comments

Please note that this has been an introductory lesson that has not addressed many of the less common techniques available. For example, I have not discussed Alpha Extraction, which extracts factors with the goal of maximizing alpha (reliability) coefficients of the Extracted Factors, or Maximum-Likelihood Extraction, or several other extraction methods.

I should remind you of the necessity of investigating (maybe even deleting) outlying observations. Subjects' factor scores may be inspected to find observations that are outliers with respect to the solution [very large absolute value of a factor score].

References

Cattell, R. B., Balcar, K. R., Horn, J. L., & Nesselroade, J. R. (1969). Factor matching procedures: an improvement of the s index; with tables. *Educational and Psychological Measurement, 29*; 781 – 792. doi 10.1177/001316446902900405

Grice, J. W. (2001). A comparison of factor scores under conditions of factor obliquity. *Psychological Methods*, *6*, 67 - 83.

Grice, J. W., & Harris, R. J. (1998). A comparison of regression and loading weights for the computation of factor scores. *Multivariate Behavioral Research*, *33*, 221 - 247.

Appendix

Table 1*. Oblique Factor Analysis of the WAQ*

Item	Greatest |Beta|	Factor
28 I have difficulty maintaining friendships.	.72	1, Work-Life Conflict
24 My work often seems to interfere with my personal life.	.67	1
29 I have difficulty maintaining intimate relationships.	.66	1
26 I often miss out on important personal activities because of work demands.	.62	1
25 I often put issues in my personal life "on hold" because of work demands.	.62	1
23 I experience conflict with my significant other or with close friends.	.60	1
14 I find myself unable to enjoy other activities because of my thoughts of work.	.47	1
27 I find it difficult to schedule vacation time for myself.	.46	1
6 I constantly feel too tired after work to engage in non-work activities.	.45	1
12 I frequently have work-related insomnia.	.45	1
1 I feel stressed out when dealing with work issues.	.33	1
19 I frequently check over my work many times before I finish it.	.66	2, Work Perfectionism

20 I ask others to check my work often.	.56	2
22 It takes me a long time to finish my work because it must be perfect.	.50	2
21 I frequently feel anxious or nervous about my work.	.50	2
18 I often obsess about goals or achievements at work.	.41	2
8 I prefer to work excessive hours, preferably 60 hours or more per week.	.66	3, Work Addiction
11 I enjoy spending evenings and weekends working.	.63	3
13 I feel very addicted to my work.	.58	3
7 I think about work constantly.	.42	3
9 I have a need for control over my work.	.29	3
17 People would describe me as being impatient and always in a hurry.	.67	4, Unpleasantness
15 I consider myself to be a very aggressive person.	.62	4
16 I get irritated often with others.	.60	4
10 I have a need for control over others.	.39	4
3 I feel anxious when I am not working.	.94	5, Withdrawal Symptoms
2 I feel guilty when I am not working.	.85	5
4 I feel bored or restless when I am not working.	.55	5
5 I am unable to relax at home due to preoccupation at work.	.44	5

- Example of Factor Analysis Reported in APA Style
- Factor Analysis Exercise: Animal Rights, Ethical Ideology, and Misanthropy
- Factor Analysis Exercise: Rating Characteristics of Criminal Defendants
- Polychoric and Tetrachoric Factor Analysis – with data from Likert-type or dichotomous items.
- SAS Version of this Workshop
- Return to Multivariate Analysis with SPSS