

# Admin

- Assignment 2:
  - Due now.
- Final Exam
  - 11 June 2:30pm, Room HEATH, UnionComplex.
  - An exam guide and practice questions will be provided next week.
- Small Group Presentations

# Kaleidoscope eyes: Anomalous visual experiences in synaesthesia

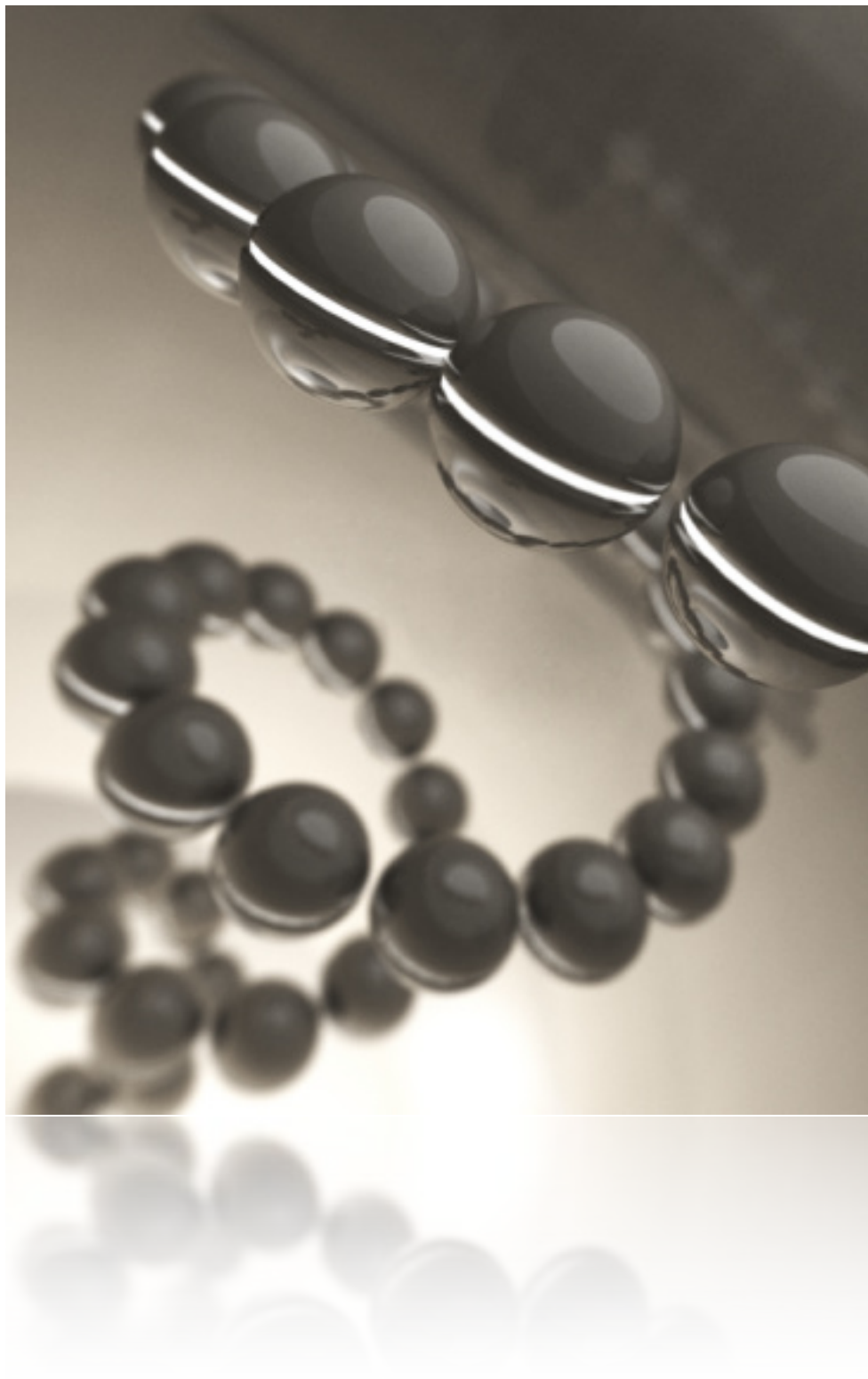
Professor Jason Mattingley

Wednesday, 13 May 2008 from 12-1pm  
McElwain Building, Room 317

Synaesthesia is an unusual phenomenon in which a stimulus in one sensory modality elicits a vivid and involuntary perception in another modality. Thus, for example, the sound of the letter "A" may induce a sensation of redness, or the taste of roast chicken may feel jagged and angular. The phenomenon has intrigued philosophers, cognitive scientists and neurologists for over a century, yet little is known about its underlying neural and cognitive bases. In this talk I will review the heterogeneous manifestations of synaesthesia, and provide examples of subjective reports given by individuals with these unusual perceptual experiences. I will then describe the results of a recent series of laboratory experiments, in which we have identified a reliable cognitive marker for the colour-graphemic form of synaesthesia. Our data show for the first time that synaesthesia arises automatically, and that it cannot be suppressed even when it is detrimental to task performance. I will conclude by providing a tentative framework within which to understand the neural and cognitive bases of synaesthesia, and offer some suggestions for future research.

# Factor Analysis via PCA

- Overview
- The 'number of factors' problem
- The rotation problem
- Modelling data in factor analysis
- Schematic representation of factor analysis via PCA
- **SPSS**





# Consider an investigation into the nature of intelligence.



ability to recite song lyrics from memory



ability to hold two conversations at once



speed at completing crosswords



ability to assemble something from IKEA



ability to use a street directory



speed at completing jigsaw puzzles

What might be the 'underlying factors'?

**X** =

Participant



1	-0.14	-2.33	-2.75	0.71	1.87	-0.32
2	0.42	-0.35	0.38	-0.16	-0.62	0.52
3	-1.18	-0.53	-2.74	0.87	-0.61	0.51
4	0.67	1.44	-1.02	-1.11	-1.87	-0.54
5	0.27	0.26	1.66	0.81	0.58	1.29
6	0.11	0.25	-0.85	1.26	1.17	2.52
7	1.11	0.97	-1.66	1.30	0.79	2.60
8	0.54	0.43	-1.10	-1.51	-0.46	-1.96
9	0.50	1.72	2.25	-1.02	-0.97	-0.08
10	0.86	2.07	-0.40	-0.02	-0.30	-0.34
11	0.64	0.40	-0.56	1.42	1.31	1.95
12	-0.51	0.25	-0.05	0.39	0.25	-0.28
13	0.48	0.92	-0.95	-0.59	0.31	-2.62
14	0.54	1.13	0.11	-0.07	-0.74	0.12
15	-0.39	-0.92	-1.41	1.54	1.34	2.61
16	0.87	2.16	0.70	-0.83	-2.75	-1.41
17	-0.62	-1.88	-0.18	-0.32	-0.80	-1.00
18	1.80	0.44	1.68	0.90	1.38	0.08
19	0.65	-1.01	1.70	2.55	2.75	4.25
20	-1.42	-0.22	-2.00	-0.11	0.69	-1.00
⋮	⋮	⋮	⋮	⋮	⋮	⋮
100	0.29	0.87	-1.99	0.24	-0.37	-0.28

$X =$

intelligence.sav [DataSet1] - SPSS Data Editor

1 : lyrics      -0.14      Visible: 6 of 6 Variables

	lyrics	converse	crossword	ikea	directory	jigsaw
1	-0.1400	-2.3300	-2.7500	0.7100	1.8700	-0.3200
2	0.4200	-0.3500	0.3800	-0.1600	-0.6200	0.5200
3	-1.1800	-0.5300	-2.7400	0.8700	-0.6100	0.5100
4	0.6700	1.4400	-1.0200	-1.1100	-1.8700	-0.5400
5	0.2700	0.2600	1.6600	0.8100	0.5800	1.2900
6	0.1100	0.2500	-0.8500	1.2600	1.1700	2.5200
7	1.1100	0.9700	-1.6600	1.3000	0.7900	2.6000
8	0.5400	0.4300	-1.1000	-1.5100	-0.4600	-1.9600
9	0.5000	1.7200	2.2500	-1.0200	-0.9700	-0.0800
10	0.8600	2.0700	-0.4000	-0.0200	-0.3000	-0.3400
11	0.6400	0.4000	-0.5600	1.4200	1.3100	1.9500
12	-0.5100	0.2500	-0.0500	0.3900	0.2500	-0.2800
13	0.4800	0.9200	-0.9500	-0.5900	0.3100	-2.6200
14	0.5400	1.1300	0.1100	-0.0700	-0.7400	0.1200
15	-0.3900	-0.9200	-1.4100	1.5400	1.3400	2.6100
16	0.8700	2.1600	0.7000	-0.8300	-2.7500	-1.4100
17	-0.6200	-1.8800	-0.1800	-0.3200	-0.8000	-1.0000
18	1.8000	0.4400	1.6800	0.9000	1.3800	0.0800
19	0.6500	-1.0100	1.7000	2.5500	2.7500	4.2500
20	-1.4200	-0.2200	-2.0000	-0.1100	0.6900	-1.0000
21	-0.3400	1.1000	-0.6900	-0.6200	-0.0100	-0.4600
22	0.3600	1.5500	-0.3200	1.5900	3.5100	2.7400
23	1.0400	1.2700	1.3100	-0.4200	0.9900	0.2600
24	0.6200	0.7000	1.0800	-0.3400	1.3100	-1.6500
25	-0.3500	-1.2500	-0.4100	0.9000	0.5500	2.0200
26	0.4100	-0.0600	-1.1800	-1.6600	-3.1900	-2.7500
27	-2.0600	-2.7000	-3.3000	-0.6500	-2.2800	-0.7600
28	-0.5600	0.8400	0.2100	-1.1900	-1.3400	-3.2000
29	0.2600	-1.2400	0.9200	-0.7000	-0.6400	-0.0500
30	-1.3000	0.2400	-0.2900	-1.3800	-0.7100	-2.5000
31	-0.5100	-0.9700	0.1700	-0.3200	0.0200	-0.4000
32	-0.4700	-3.3000	-0.8600	0.5700	1.3400	-0.9500
33	0.0700	0.2000	2.1500	0.2700	-0.3700	-1.0400
34	-0.8800	-1.8500	-1.2400	-0.4300	0.8900	-0.6600
35	1.7400	2.1700	2.2800	0.6400	1.6400	0.7200
36	-0.0600	0.8700	0.4300	-0.3200	-1.1500	1.0800
37	-0.7100	0.0100	-1.3100	0.6900	-1.2000	1.1800
38	0.5600	2.1900	2.3600	0.2800	0.5100	1.2600
39	-0.7100	0.3400	-0.4500	-0.1600	0.2000	2.3600
40	-0.1900	-0.7900	-1.4600	-0.5600	-1.1900	0.0900
41	1.5600	2.1000	2.9800	-2.5600	-1.6800	-1.2300
42	0.4100	1.8600	0.0200	0.3000	-0.0600	1.1800
43	-1.0400	-1.7800	-1.9700	0.7800	-0.2200	1.2200
44	0.7700	0.2000	3.0300	-0.0500	-1.8300	-1.1000
45	-1.4000	-1.6100	-1.2800	-0.3100	-0.0300	-0.3700
46	0.0300	0.9200	0.2600	-0.6300	-0.2700	0.7700
47	-0.2500	-0.0900	1.2100	-0.7100	-2.9200	0.1100
48	-0.8700	0.3700	1.4100	-1.7200	-2.0700	-3.4400

Data View    Variable View

SPSS Processor is ready

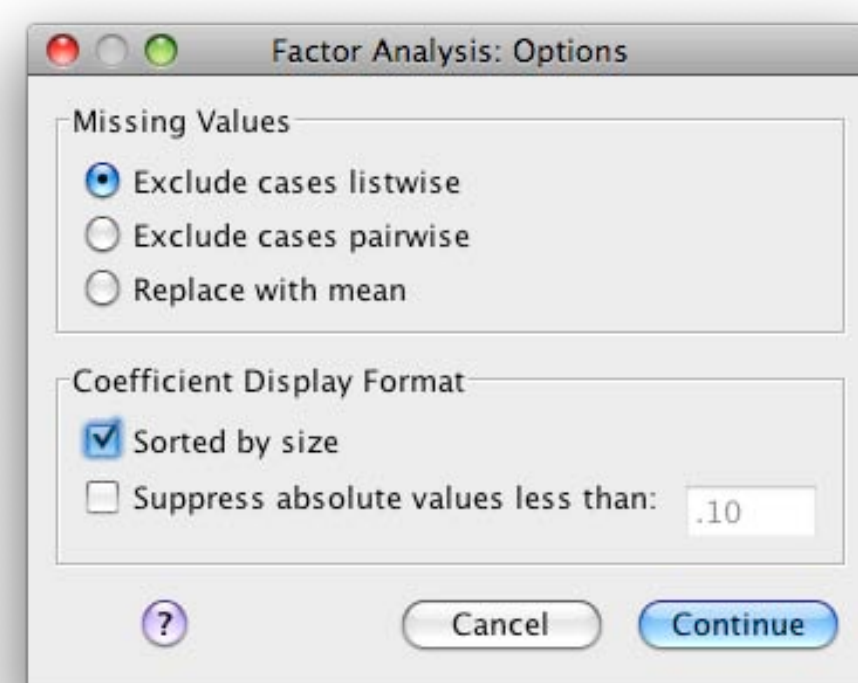
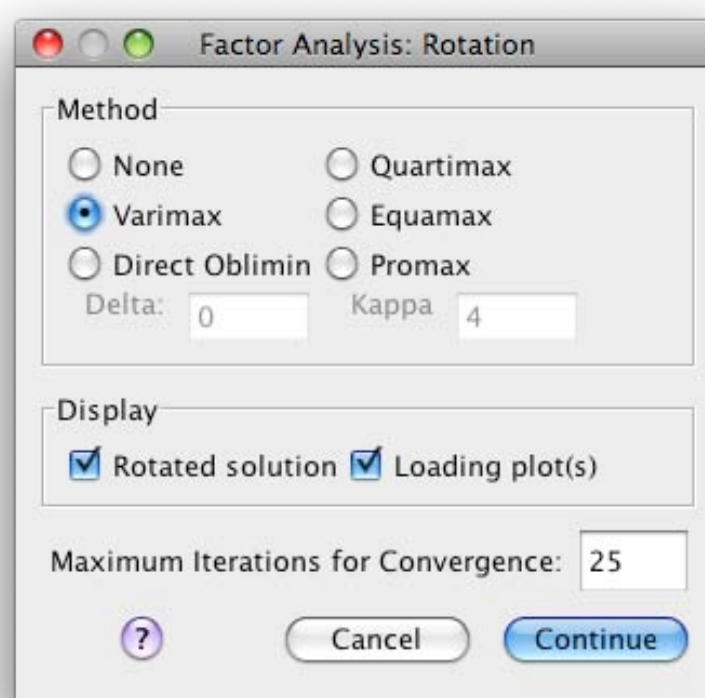
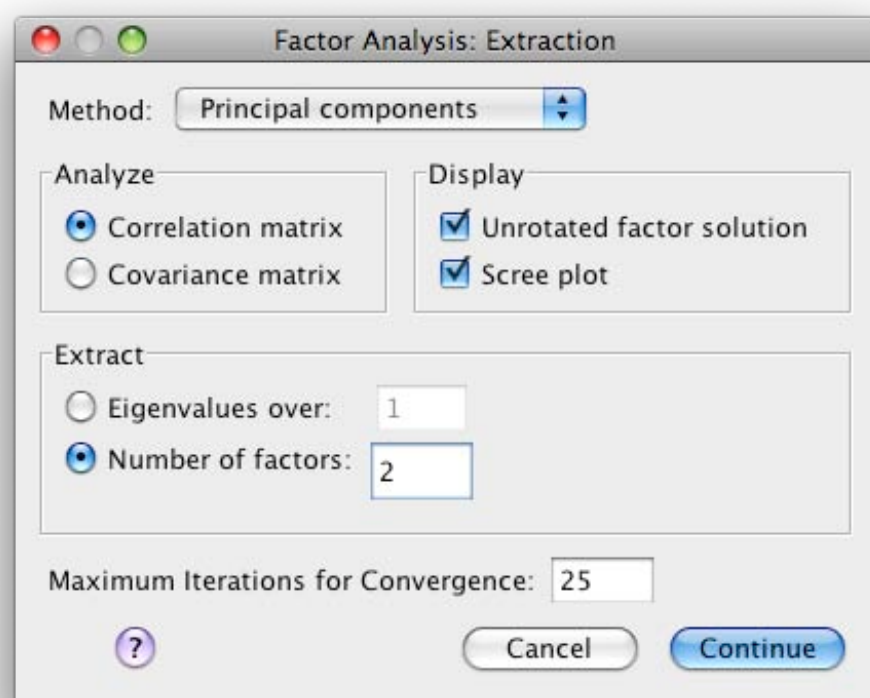
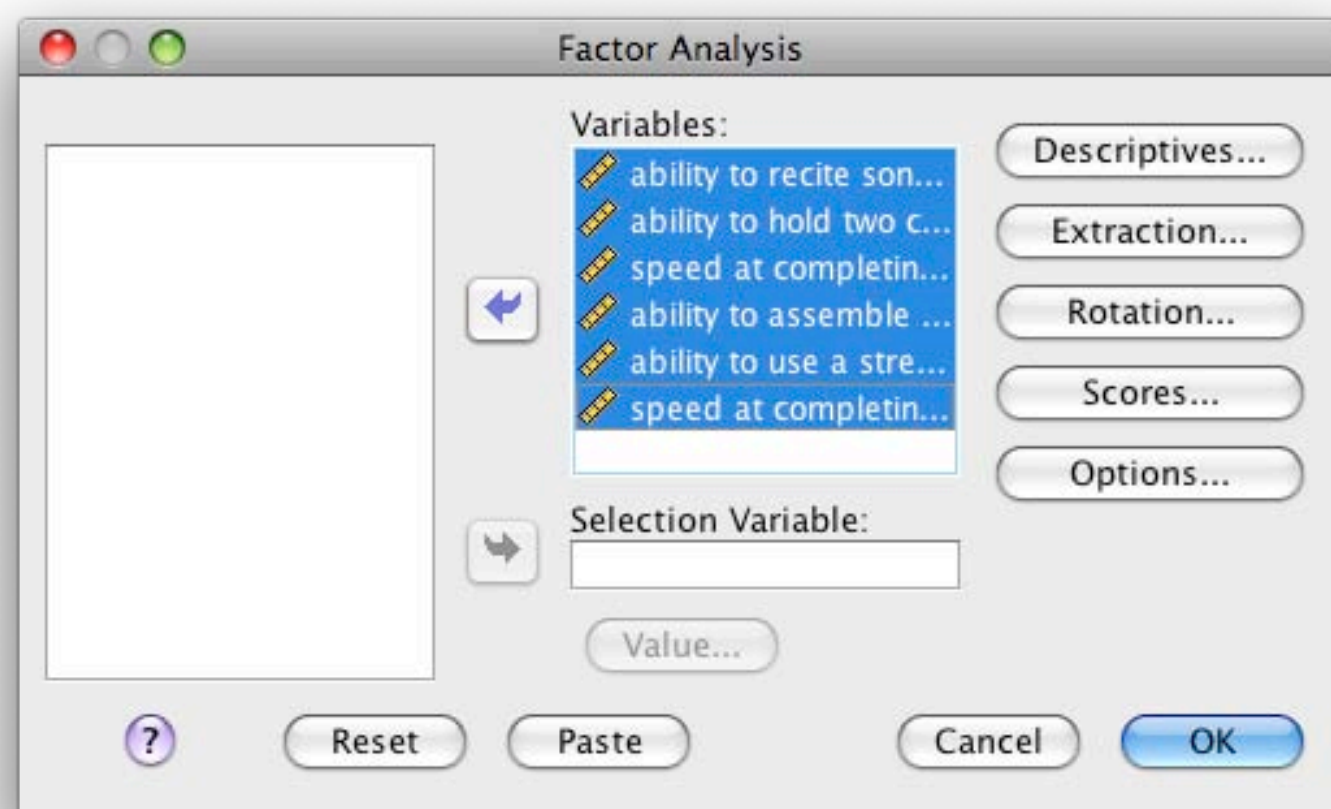
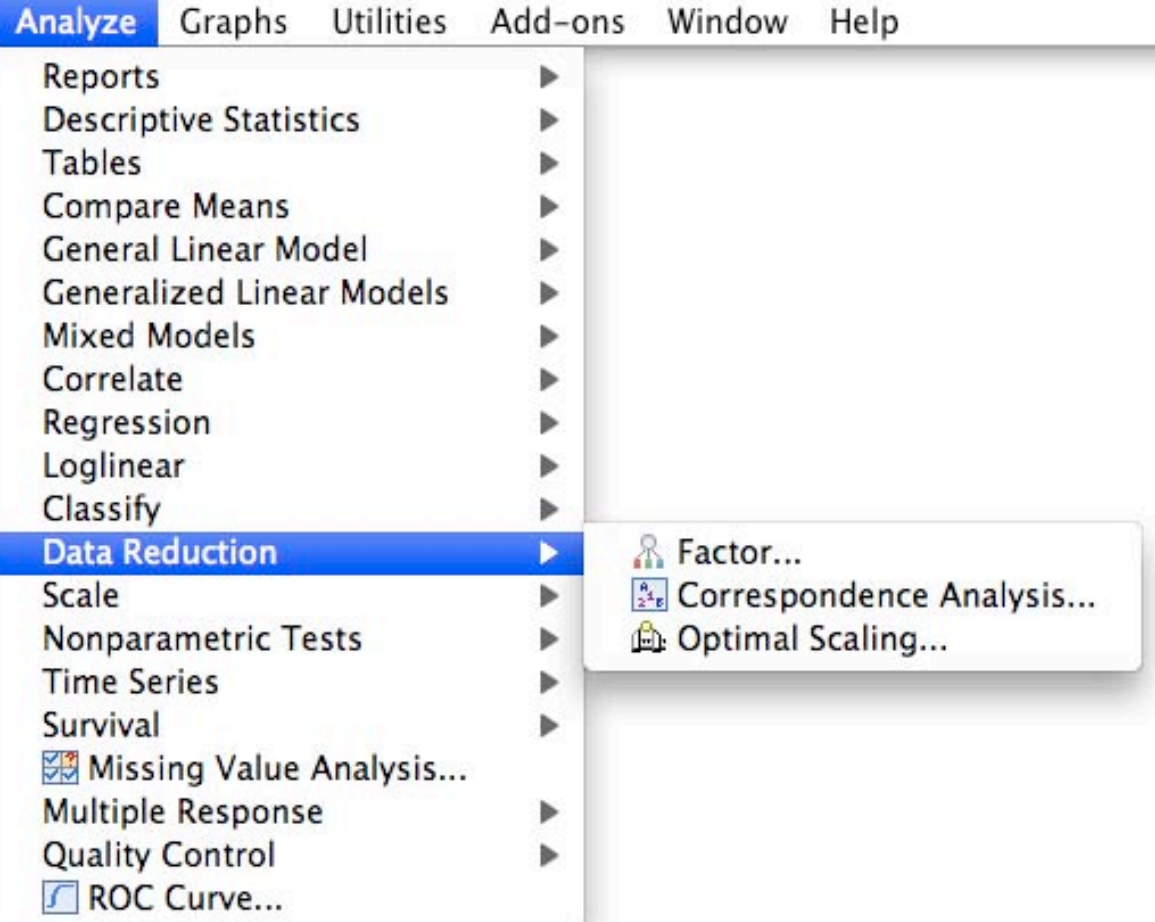
# Limitations

- Tabachnick and Fidell (2007, p. 612; Section 13.3.2) discuss the practical issues with PCA, especially the need to have honest, reliable correlations. Of course, checking the effect of transformations of the data on the interpretations is crucial. In order to check if the **R** matrix is ‘usefully’ factorable, there needs to be several large correlations among the variables. Generally, the percentage of variance accounted for is also a guide.

# PCA using SPSS

- It is important to *not* rely on the default values in the SPSS Factor procedure. There are many choice points and it is crucial that you exercise control. Further, it is unlikely that an appropriate analysis would be completed in a single run of the FACTOR procedure. The instructions below result from knowledge gained from earlier runs. The method of extraction is Principal Components, the number of “factors” is specified to be two (2) and the rotation method is orthogonal using the Varimax method.





## FACTOR

```
/VARIABLES lyrics converse crossword ikea directory jigsaw  
/MISSING LISTWISE  
/ANALYSIS lyrics converse crossword ikea directory jigsaw  
/PRINT INITIAL EXTRACTION ROTATION  
/FORMAT SORT  
/PLOT EIGEN ROTATION  
/CRITERIA FACTORS(2) ITERATE(25)  
/EXTRACTION PC  
/CRITERIA ITERATE(25)  
/ROTATION VARIMAX  
/METHOD=CORRELATION.
```

**Note:** a prior analysis (and the fact that the data were simulated with a known structure) indicated that 2 factors would be rotated.

- /VARIABLES subcommand: specifies the variables to be used in the analysis.
- /FORMAT subcommand: specifies that the loading matrix in the output will be sorted which enables, most times, easier reading of the matrix.
- /PRINT subcommand: specifies the default output, plus the correlation matrix and means and standard deviations.
- /PLOT subcommand: specifies a plot of components 1 and 2 after rotation.
- /EXTRACTION subcommand: specifies that the method of extraction to be PC, principal components.
- /ROTATION subcommand: specifies that the Varimax criterion be used to orthogonally rotate the number of factors specified in the criteria subcommand or using the default that the number of factors to be rotated is the number with eigenvalues greater than 1.

# Deciding the number of “factors” to retain







Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.546	42.442	42.442	2.546	42.442	42.442	2.309	38.478	38.478
2	2.020	33.665	76.107	2.020	33.665	76.107	2.258	37.628	76.107
3	.511	8.509	84.615						
4	.489	8.154	92.769						
5	.253	4.209	96.978						
6	.181	3.022	100.000						

Extraction Method: Principal Component Analysis.

The Initial Eigenvalues for each component are the elements of the diagonal matrix, **L**. From the eigenvalues, a decision about the number of “factors” to retain is made. In this example, it is very clear that two components are sufficient to explain the great majority of the total variance, 76% of the total variance.

**L** =

	1	2	3	4	5	6
	2.55	0	0	0	0	0
	0	2.02	0	0	0	0
	0	0	0.51	0	0	0
	0	0	0	0.49	0	0
	0	0	0	0	0.25	0
	0	0	0	0	0	0.18

$2.546/6 = 42.442\%$

$2.020/6 = 33.665\%$

$.511/6 = 8.509\%$

$.489/6 = 8.154\%$

$.253/6 = 4.209\%$

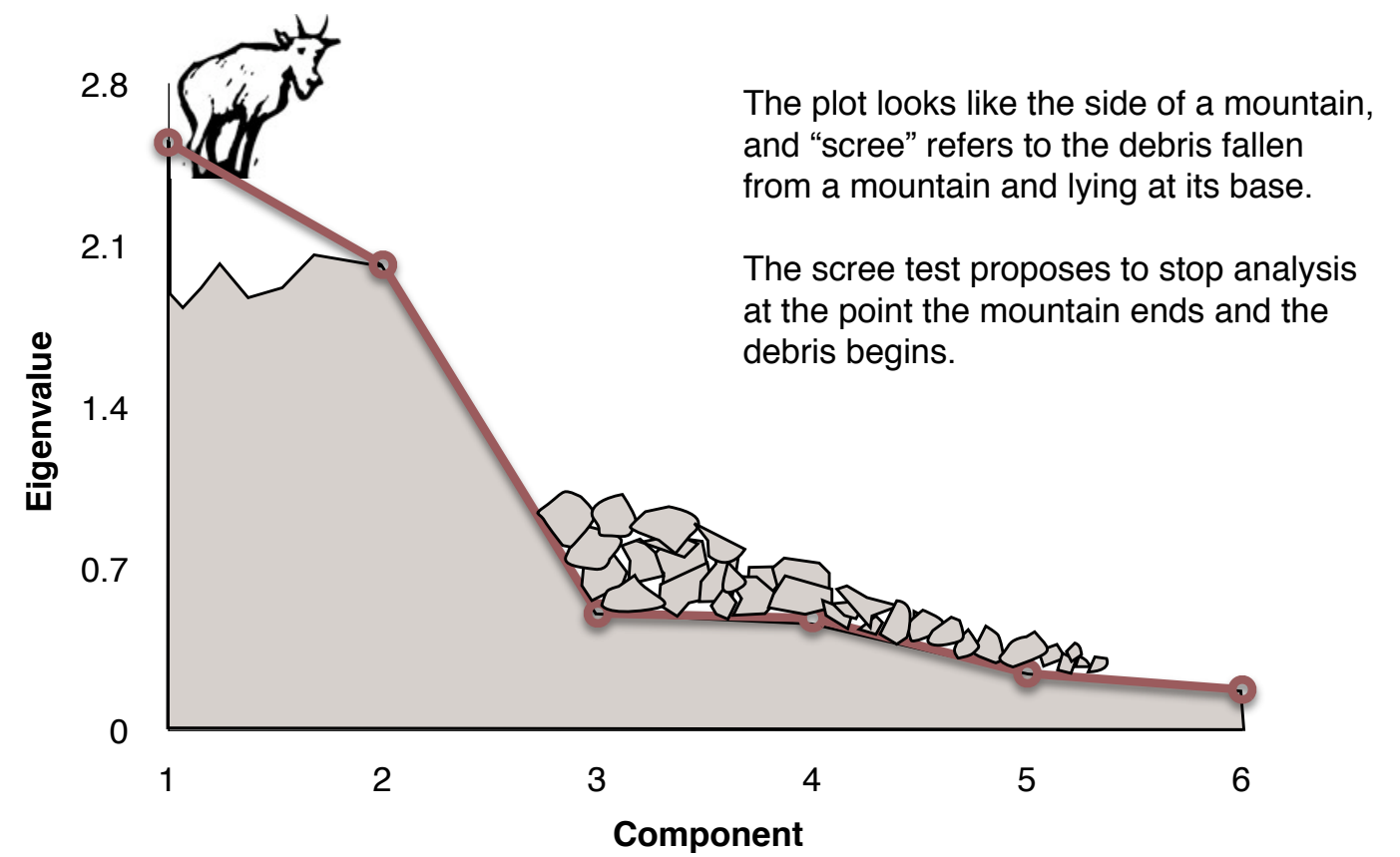
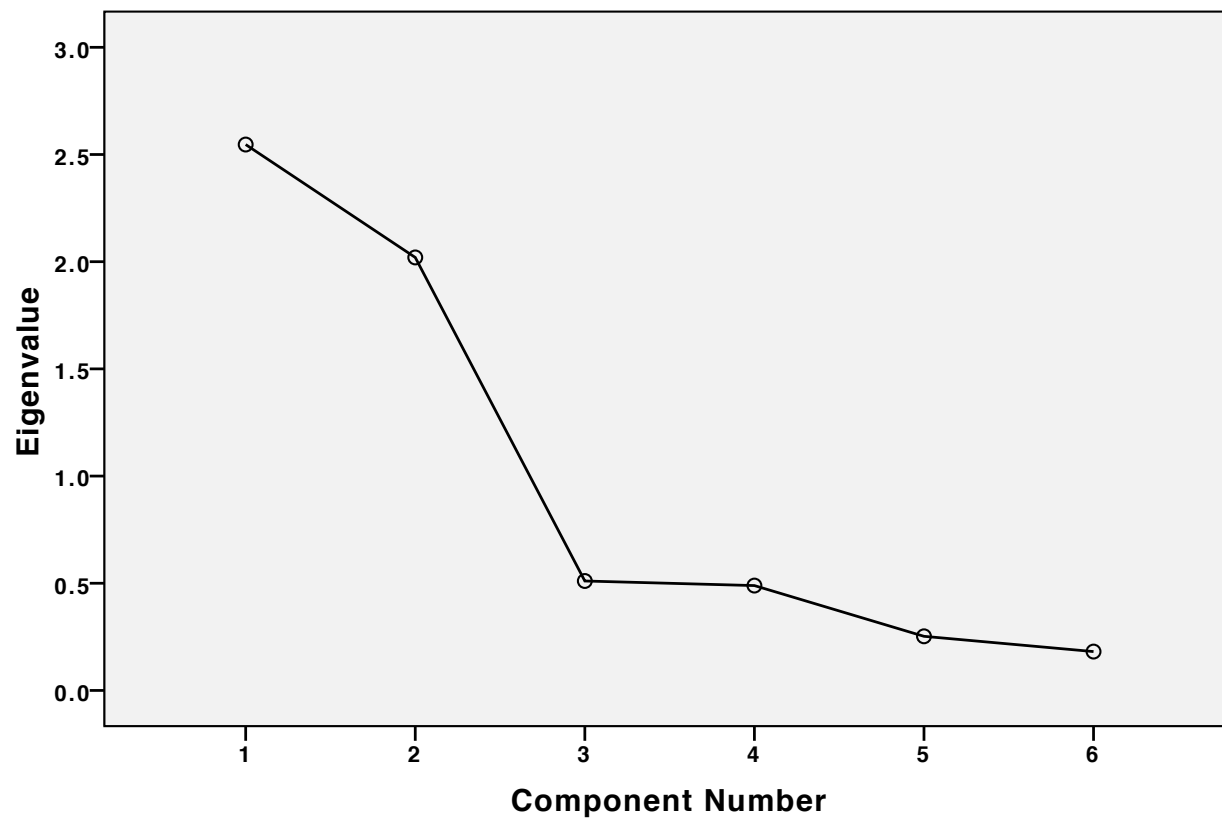
$.181/6 = 3.022\%$

}

76.107%

# Scree Plot

Scree Plot





# The factor loading matrix

Rotated Component Matrix<sup>a</sup>

	Component	
	1	2
ikea	.939	-.071
jigsaw	.835	-.044
directory	.818	.280
lyrics	.229	.885
crossword	-.084	.841
converse	.004	.825

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

a. Rotation converged in 3 iterations.

## $A_{\text{rotated}}$

	1	2
	0.23	0.88
	0.00	0.83
	-0.08	0.84
	0.94	-0.07
	0.82	0.28
	0.84	-0.04

The first Component Matrix is the unrotated matrix,  $A$ . This is not interpreted. Orthogonal rotation using the Varimax method was specified and this Rotated Component Matrix ( $A_{\text{rotated}}$ ) is interpreted. In this clear-cut example, it is easy to see that the **spatial** variables load very highly on Component 1 and the **verbal** variables load highly on Component 2. The Sums of Squared Loadings (SSLs) for the rotated components are given in the table headed 'Rotation Sums of Squared Loadings'.

T&F: If simple structure is present, the columns will have several high and many low values, while the rows will only have one high value. Rows with more than one high correlation correspond to variables that are “complex” because they reflect the influence of more than one factor.







# Rotation Sums of Squared Loadings

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.546	42.442	42.442	2.546	42.442	42.442	2.309	38.478	38.478
2	2.020	33.665	76.107	2.020	33.665	76.107	2.258	37.628	76.107
3	.511	8.509	84.615						
4	.489	8.154	92.769						
5	.253	4.209	96.978						
6	.181	3.022	100.000						

Extraction Method: Principal Component Analysis.

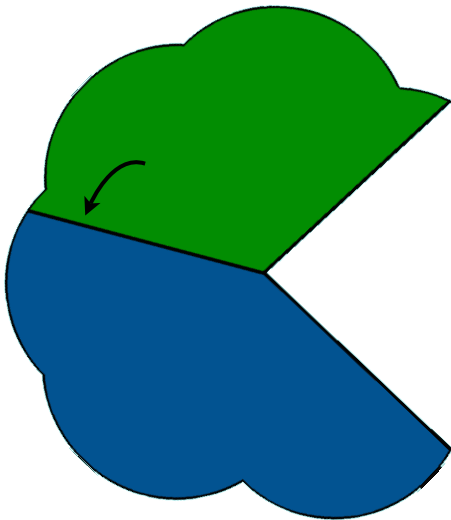
## A<sub>rotated</sub>

	1	2
	0.23	0.88
	0.00	0.83
	-0.08	0.84
	0.94	-0.07
	0.82	0.28
	0.84	-0.04

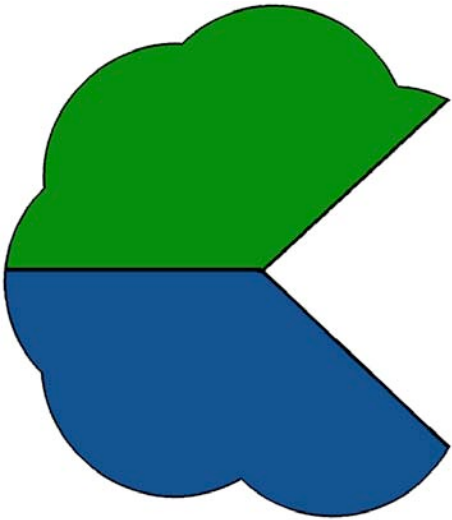
$\lambda = 2.309 \ 2.258$

The Sums of Squared Loadings (SSLs) for the rotated components are given in the table headed ‘Rotation Sums of Squared Loadings’.

$$\left. \begin{aligned} 2.309/6 &= 38.478\% \\ 2.258/6 &= 37.628\% \end{aligned} \right\} 76.107\%$$



Unrotated Solution



Rotated Solution  
(orthogonal)

# Communalities

The table labelled ‘Communalities’ has a column called ‘Extraction’. For an orthogonal solution, these are the sums of squared loadings for each variable. For an oblique solution, communalities are obtained using a regression approach that takes into account the correlations among the components or factors.

The communalities for each variable are the sums of squares of the rows of the component matrix, unrotated or rotated.

Communalities

	Initial	Extraction
lyrics	1.000	.836
converse	1.000	.681
crossword	1.000	.715
ikea	1.000	.887
directory	1.000	.748
jigsaw	1.000	.700

Extraction Method: Principal Component Analysis.

## $A_{\text{rotated}}$

	1	2	$h^2$
	0.23	0.88	<b>0.836</b>
	0.00	0.83	<b>0.681</b>
	-0.08	0.84	<b>0.714</b>
	0.94	-0.07	<b>0.887</b>
	0.82	0.28	<b>0.748</b>
	0.84	-0.04	<b>0.699</b>

$$\sum a^2 = .23^2 + .88^2 = .836$$

That is, 83.6% of the variance in ‘lyric recall’ is accounted for by Factor 1 plus Factor 2.

This gives us an indication of how much ‘lyric recall’ has in common with the two factors.

Sum of Square Loadings  
/6                      2.31    2.26    **4.564**  
                                 38.5% 37.6% **76.11%**

$$\sum a^2 = .23^2 + .00^2 + -.08^2 + .94^2 + .82^2 + .84^2 = 2.31 = \lambda_1$$

# Extraction Sums of Squared Loadings

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.546	42.442	42.442	= 2.546	42.442	42.442	2.309	38.478	38.478
2	2.020	33.665	76.107	= 2.020	33.665	76.107	2.258	37.628	76.107
3	.511	8.509	84.615						
4	.489	8.154	92.769						
5	.253	4.209	96.978						
6	.181	3.022	100.000						

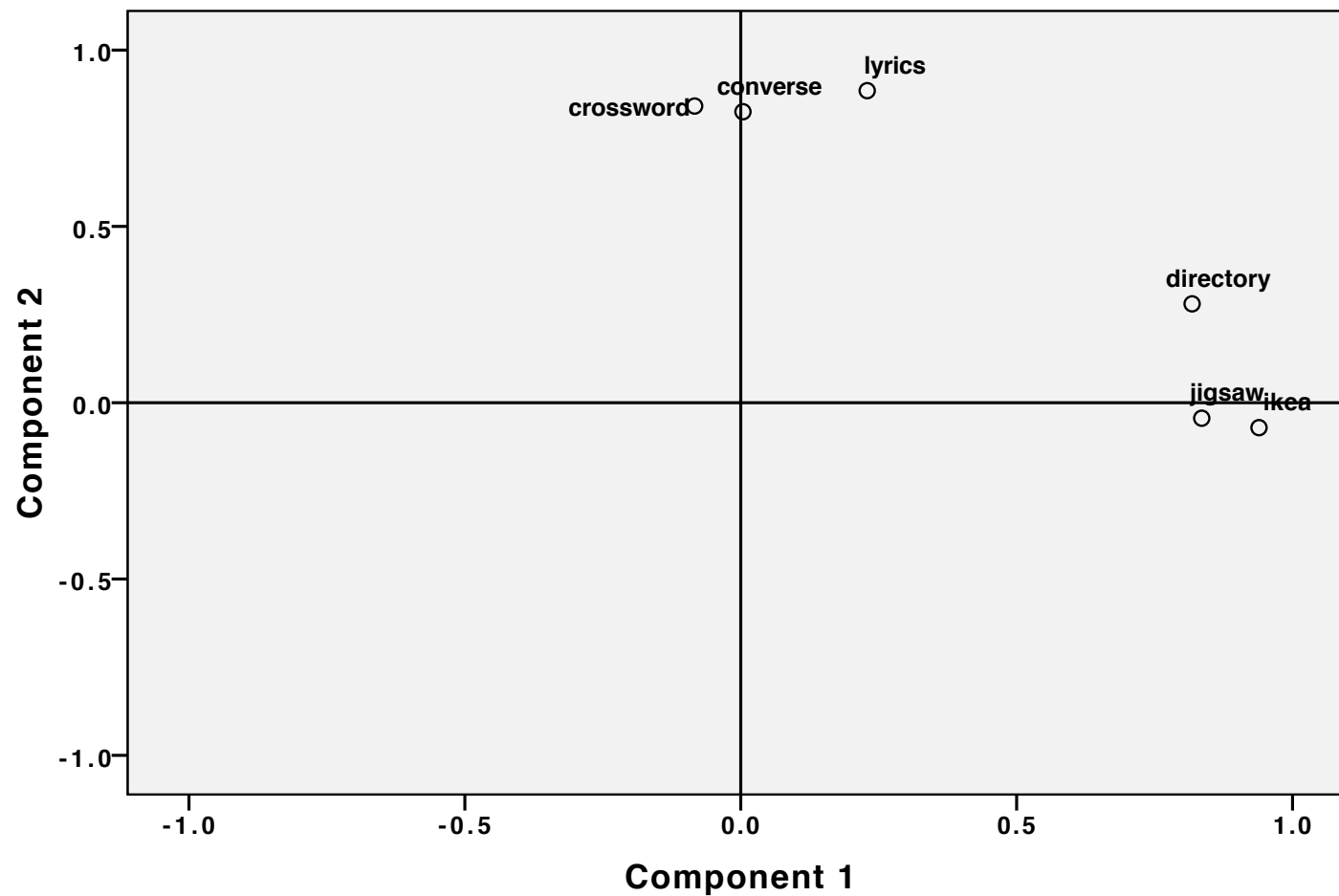
Extraction Method: Principal Component Analysis.

Because a PC model was specified, the eigenvalues and percentage of variance explained in the 'Extraction sums of squared loadings' part of the 'Total Variance Explained Table' are the same as for the initial eigenvalues. However, after rotation, these SSLs change and are called 'Rotation Sums of Squared Loadings'. These are reported in an interpretation.

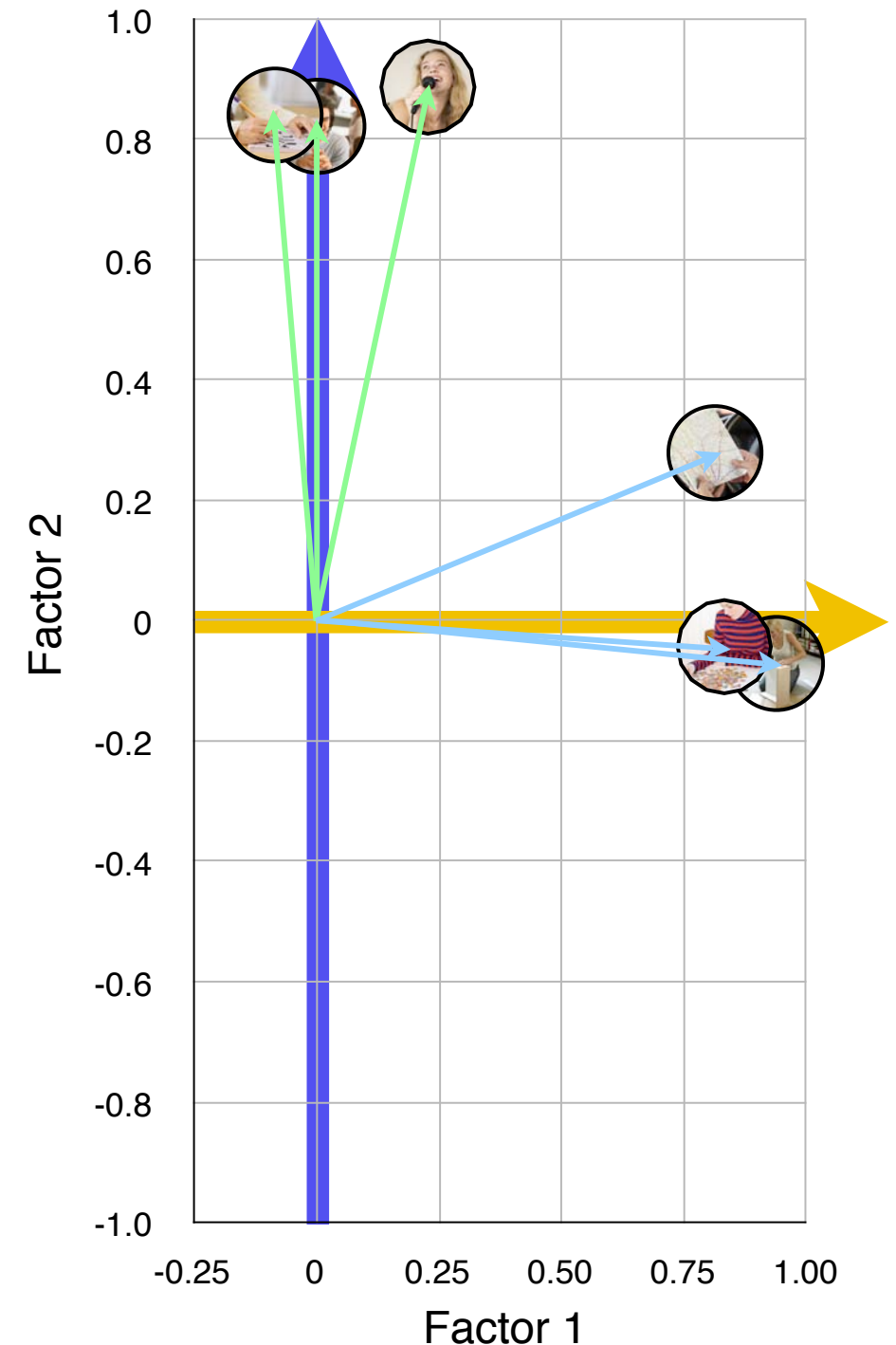


# Loading Plots

Component Plot in Rotated Space



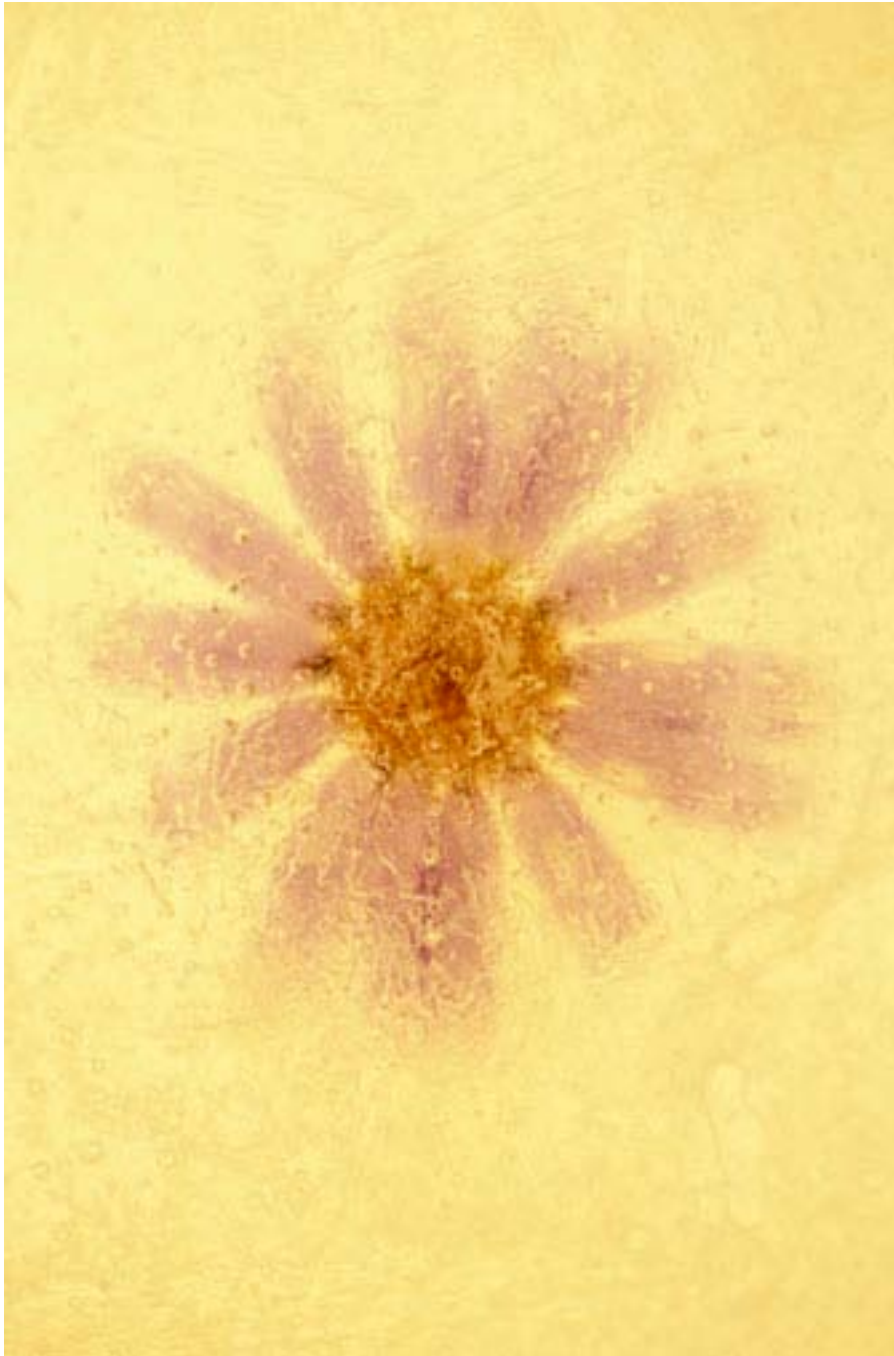
These are useful when two or three factors are rotated. When two components or factors are rotated SPSS produces a two-dimensional plot. For three components/factors, SPSS produces a plot that can be 'spun' to illustrate the patterns in the loadings. It's helpful to add lines as SPSS only plots the labelled points.



Rotated Solution  
(orthogonal)

# Common Factor Analysis

- Types of factor analysis
- The common factor model
- Euler representation
- Modelling the data in factor analysis
- The principal axis factoring method
- Schematic links in factor analysis





# Common Factor Analysis

## Introduction

We have introduced the full principal components analysis in which the original variables are transformed via a linear combination into the same number of uncorrelated components. The aim for a common factor analysis is the same as for factor analysis via PCA. The objective of a factor analysis is to develop a model that makes substantive sense and describes the data to a reasonable extent.

# Types of factor analysis

- Factor Analysis via PCA
  - uses *The Principal Components Model*
- Common Factor Analysis
  - uses *The Common Factor Model*

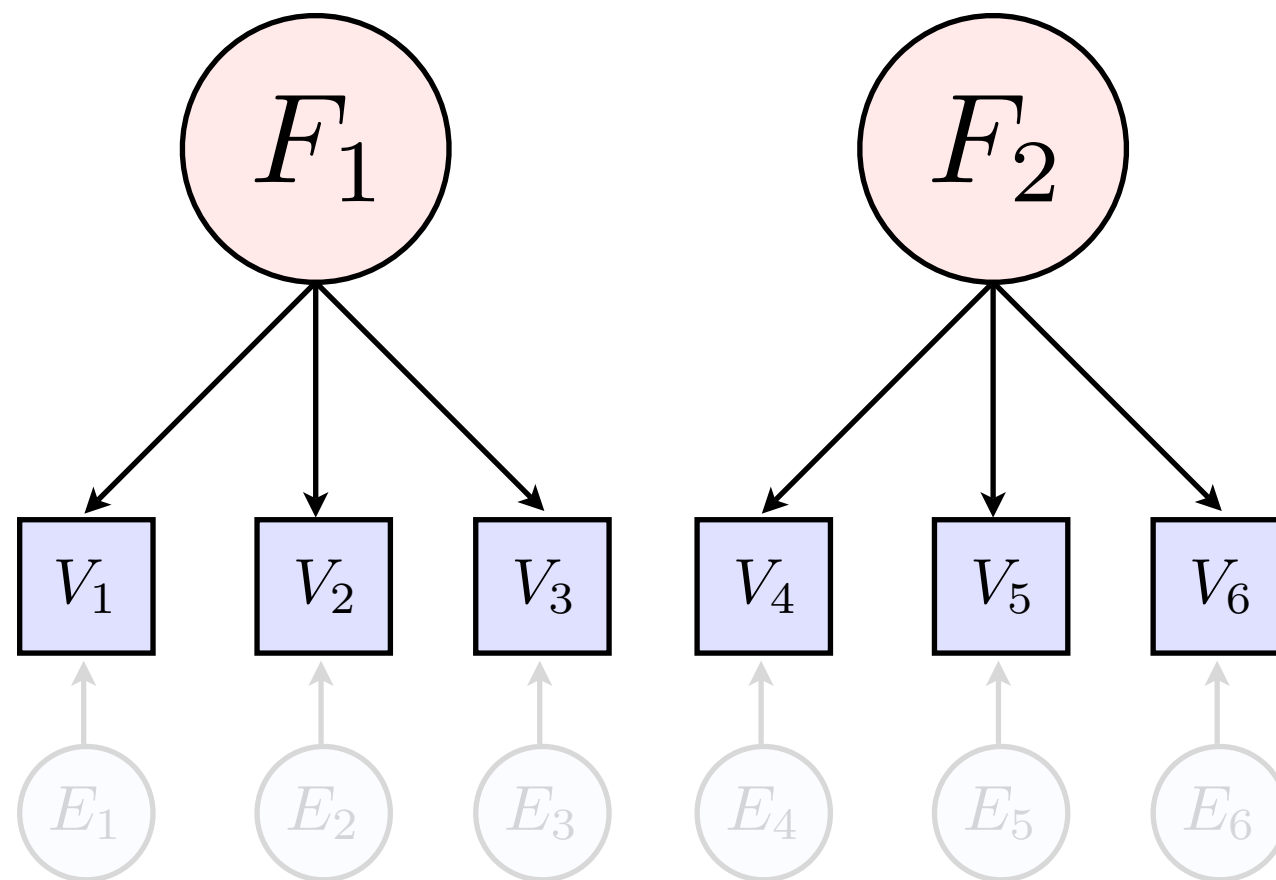
The choice of factor model depends on a researcher's assumptions about the nature of the variance for a variable.



# The Common Factor Model

## Common Variance

Factor analysis seeks to explain the correlation among the variables. *The Common Factor Model* assumes that the variation in the scores of individuals on a variable has two sources:



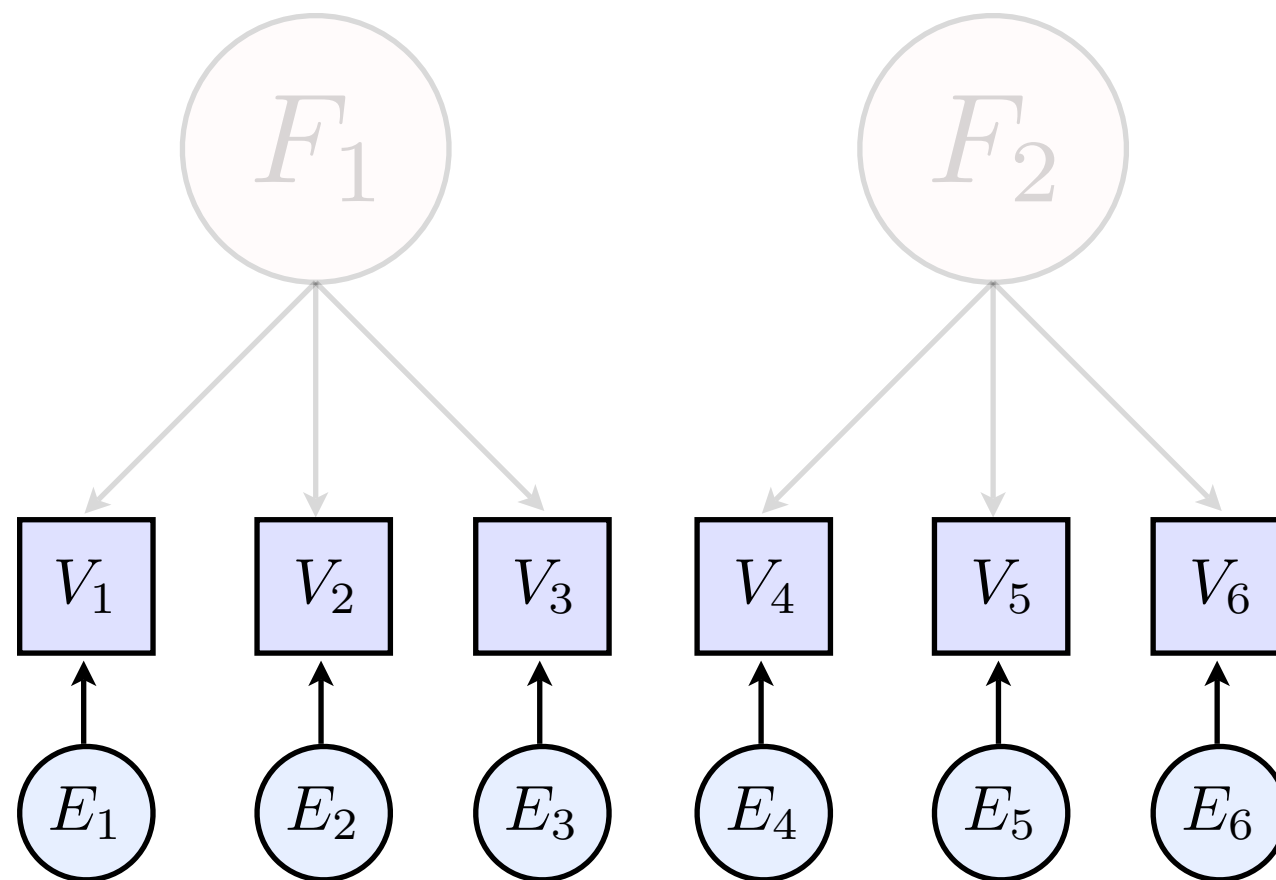
### a. Common Variance

This is part of the variance for a variable due to the assumption that it has something in common with at least one of the other variables in the analysis. Another way of stating this is that the variance is due to common factors,  $F$ 's, which are latent (i.e. unobserved) variables that influence the scores on more than one of the observed variables. This leads to a model for the data called the common factor model.

# The Common Factor Model

## Common Variance

Factor analysis seeks to explain the correlation among the variables. *The Common Factor Model* assumes that the variation in the scores of individuals on a variable has two sources:



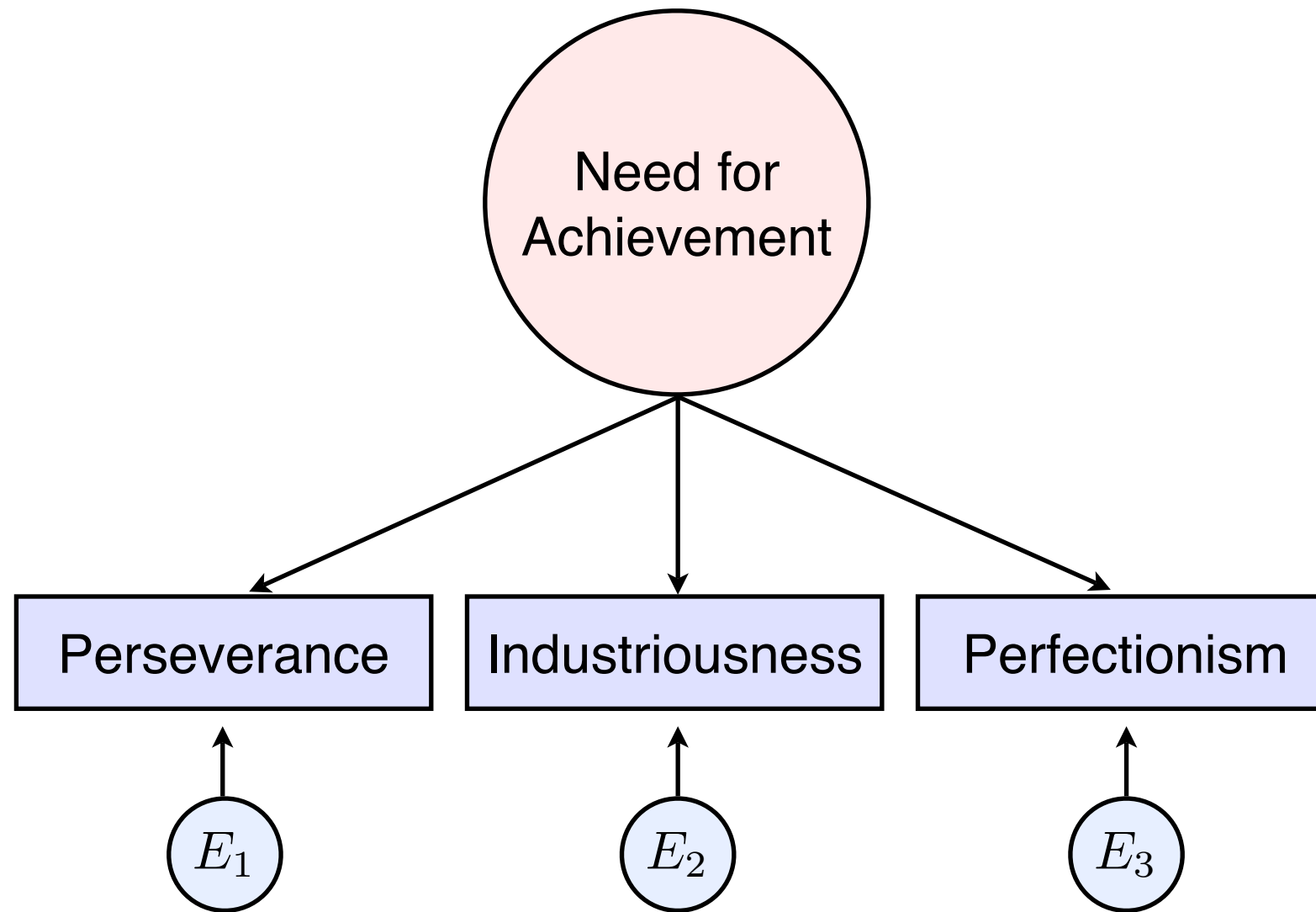
### b. Specific Variance

This is due to influences that are specific to one variable and affect no other variable. These influences include both pure error variance and unique variance that is reliable but specific to a single variable. These latter types cannot be disentangled from each other so they are lumped together into a single latent factor that is unique to just one measure, (the E's).

**The objective of a common factor analysis is to identify the latent factors underlying the common variance.**

# The Common Factor Model

## Common Variance (example)



It's assumed that these three measures of personality correlate is because they are predicted from a "common cause" (need for achievement), the factor in this example, is not a variable that is actually measured. Rather this factor is thought of as the sum of the parts of the variables measured to study it.

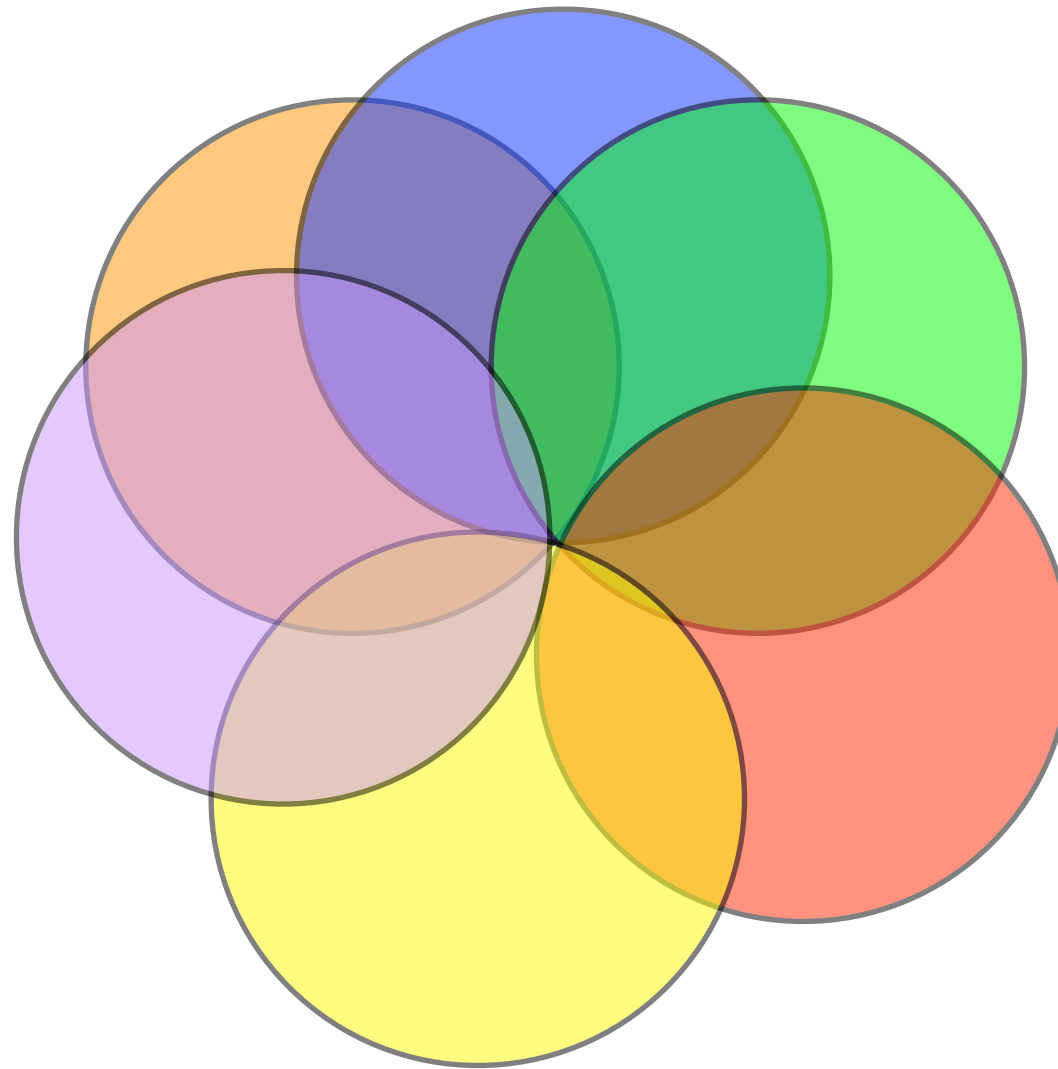
# Hypothesised parts of the variance of a variable





# The Common Factor Model

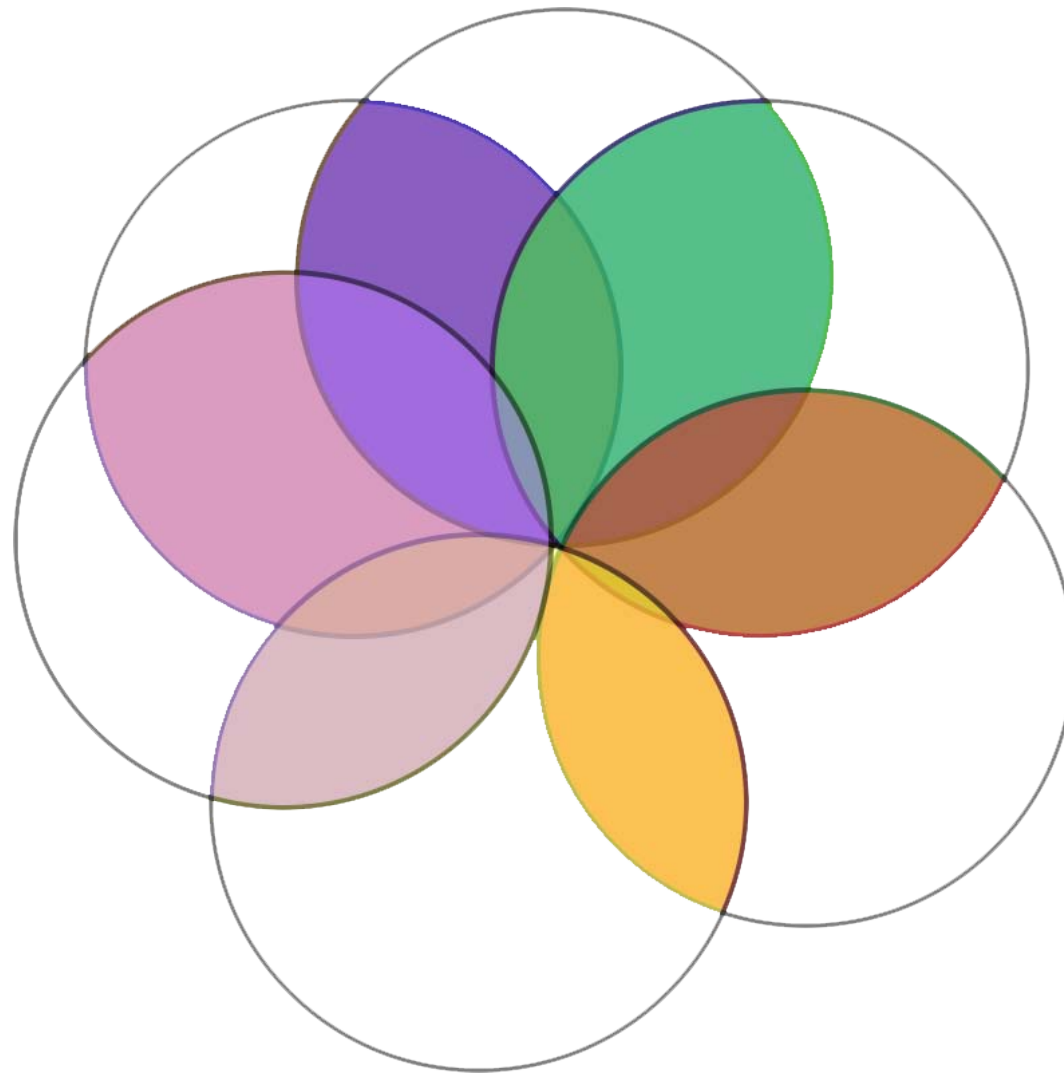
## An Euler representation



Consider six variables that are intercorrelated.

# The Common Factor Model

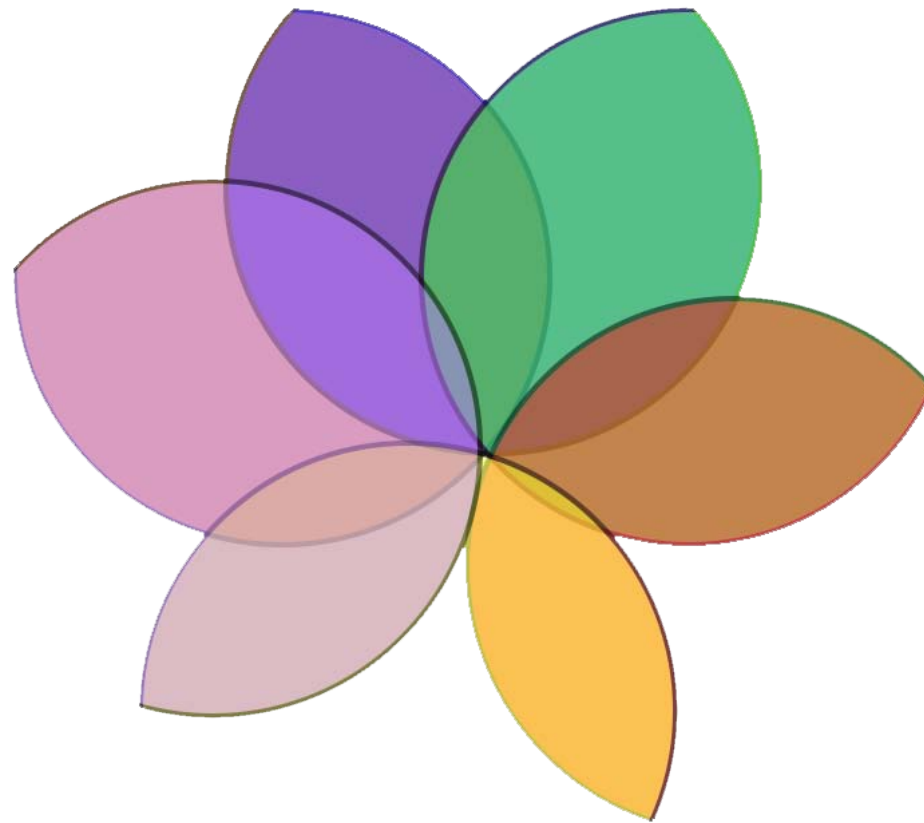
## An Euler representation



Consider the overlap between the six variables

# The Common Factor Model

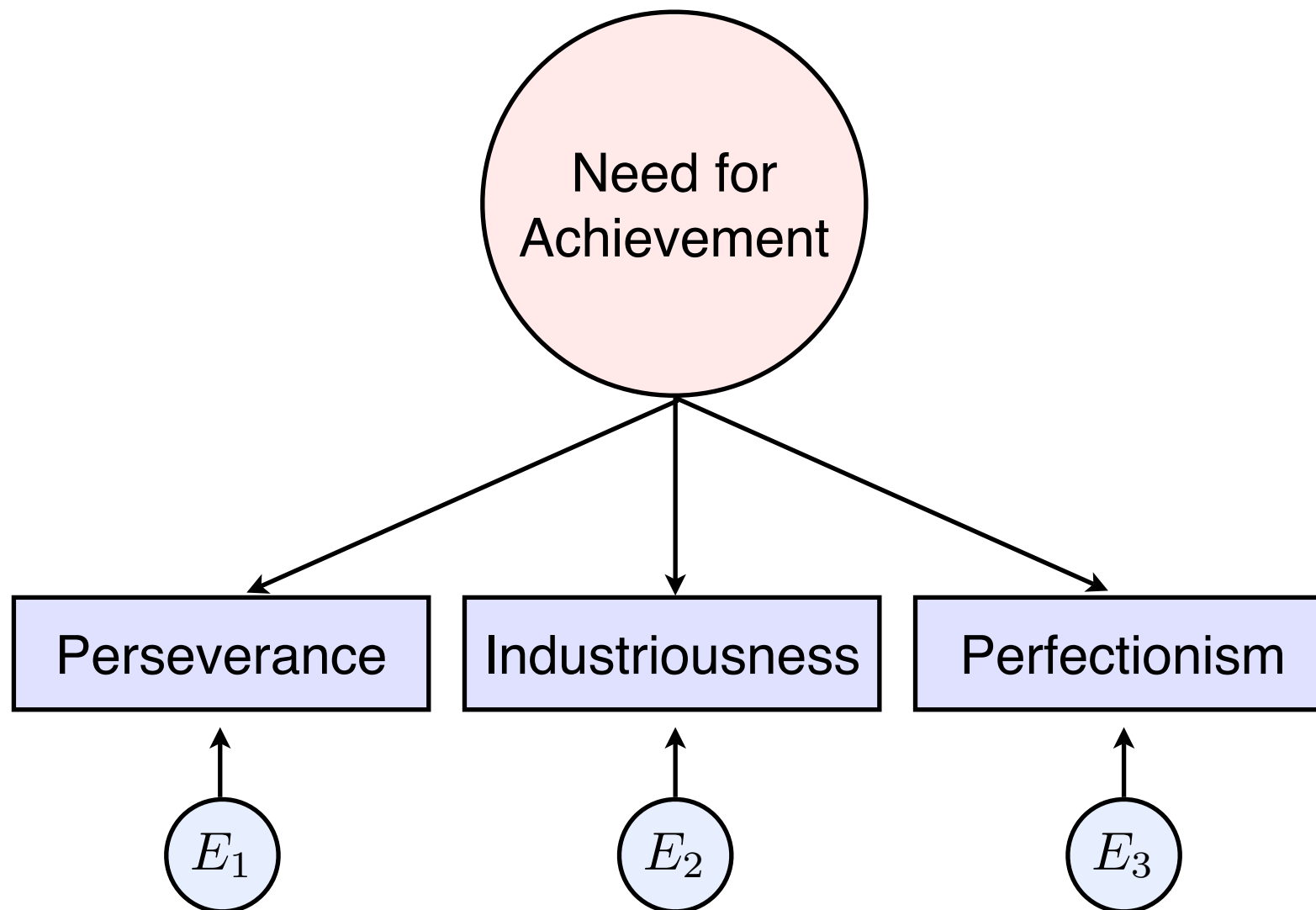
## An Euler representation



Underlying constructs should only be related  
to the common variance

# The Common Factor Model

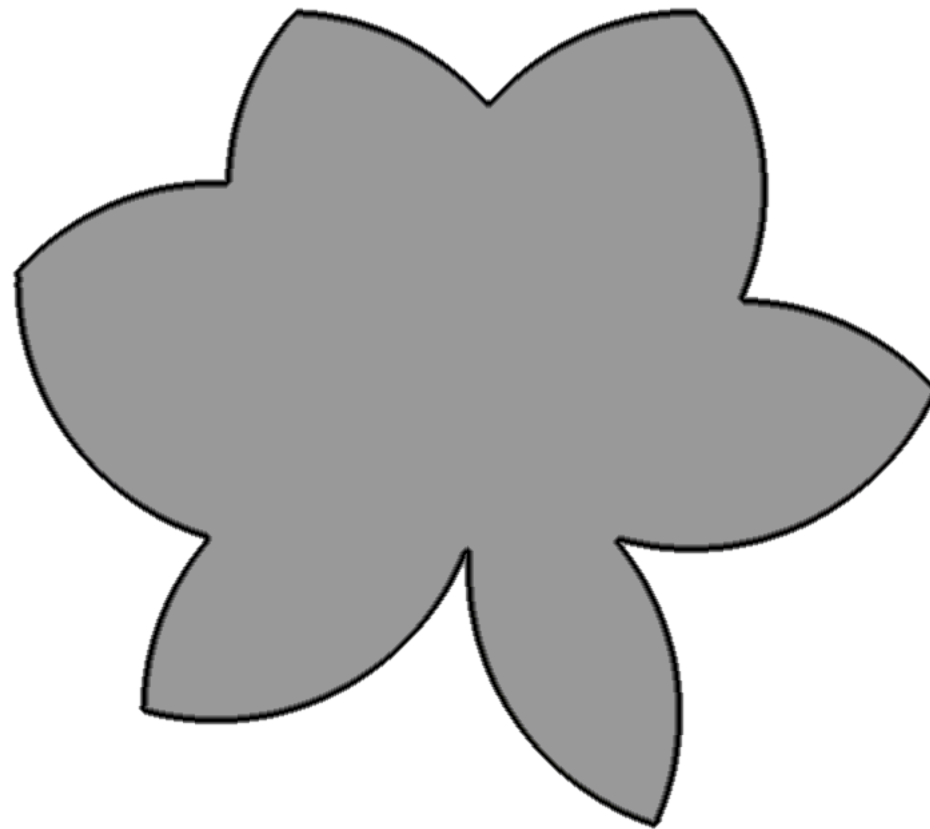
## An Euler representation



For example, “Need for Achievement” should only be related to what’s common among “Perseverance”, “Industriousness”, and “Perfectionism”

# The Common Factor Model

## An Euler representation

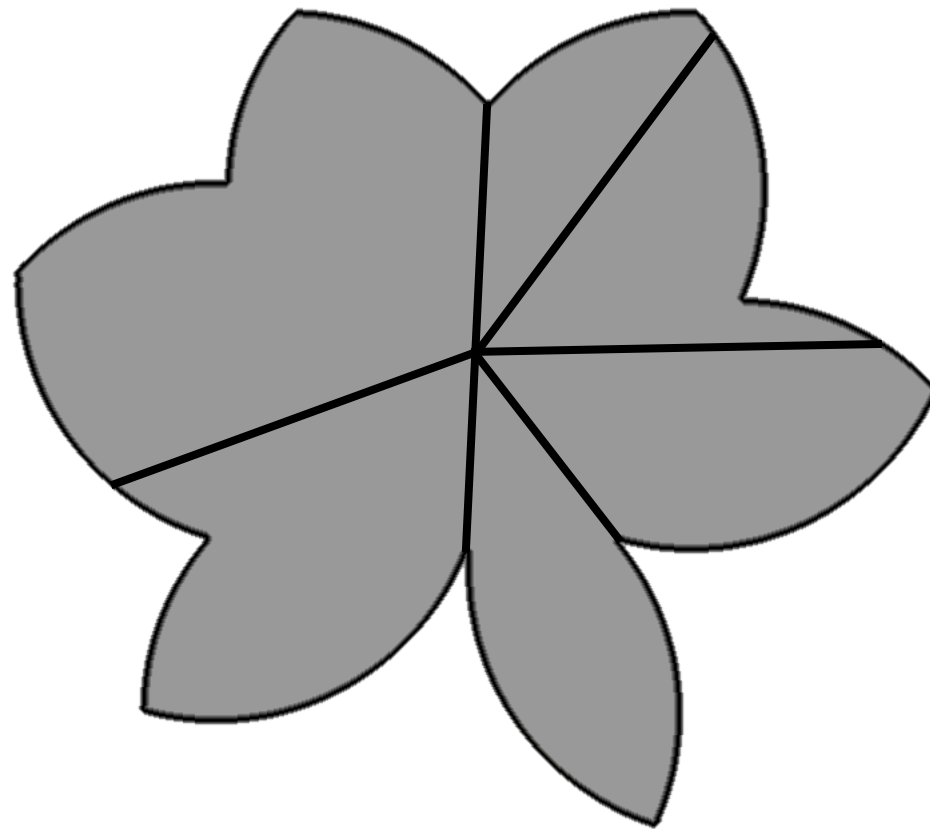


So to find underlying constructs, we're only interested in the common variance.



# The Common Factor Model

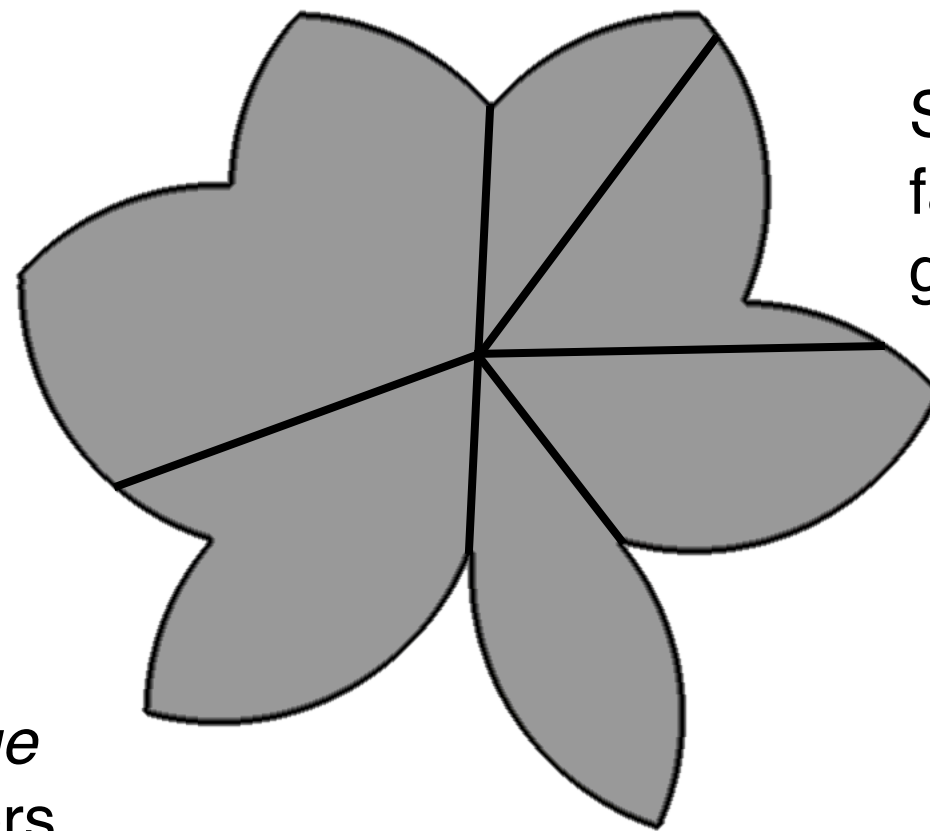
## An Euler representation



So only the common variance is factor analysed

# The Common Factor Model

## An Euler representation



We need to know the *size* of the common variance before we can factor analyse it!

The size of the common variance is given by the communalities.

BUT the communalities *change* depending on how many factors we retain.

So we need to know how many factors we're going to retain to get the communalities.

BUT we can't decide how many factors to retain until we do a factor analysis!

But there's a BIG problem...

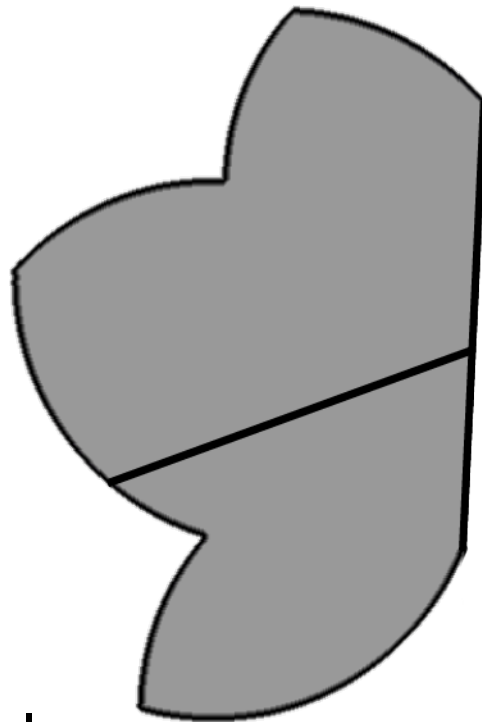
# The Common Factor Model

## An Euler representation

Interpretation is basically the same as for factor analysis via PCA.

The number of factors is decided upon.

The type of rotation (orthogonal or oblique) is decided upon.



Once that problem has been solved...

# Modelling the data in Common Factor Analysis

## Use Regression

In the definition of communalities in factor analysis via principal components analysis each variable was considered as being a linear combination of the factors.

### Communalities

- Consider the previous regression equation:
$$Z_j \leftarrow F_1 F_2 \dots F_m$$
- There's an  $R^2$  value for each variable.
  - How much of the variance of a variable is accounted for by the factors.
  - How much the variable has in common with the factors.
- It is called the *communality* of the variable,  $h^2$ .
- For orthogonal solutions,

$$h_j^2 = a_{j1}^2 + \dots + a_{jm}^2$$



The sum of squared loadings

Recall from last lecture...



# Modelling the data in Common Factor Analysis

## Use Regression

Z							F		
									
Participant							Participant	1	2
1	-0.10	-1.62	-1.78	0.78	1.45	-0.16	1	0.96	-1.25
2	0.50	-0.19	0.31	-0.07	-0.28	0.37	2	0.02	0.20
3	-1.22	-0.32	-1.77	0.93	-0.28	0.37	3	0.51	-1.38
4	0.77	1.10	-0.62	-0.99	-1.15	-0.30	4	-0.87	0.46
5	0.34	0.25	1.16	0.87	0.55	0.86	5	0.79	0.59
6	0.17	0.24	-0.51	1.31	0.96	1.65	6	1.52	-0.17
7	1.24	0.76	-1.05	1.35	0.70	1.70	7	1.55	0.20
8	0.63	0.37	-0.68	-1.38	-0.17	-1.21	8	-1.01	0.29
9	0.59	1.30	1.56	-0.90	-0.53	-0.01	9	-0.71	1.33
10	0.97	1.55	-0.21	0.07	-0.06	-0.18	10	-0.05	0.87
11	0.74	0.35	-0.32	1.47	1.06	1.28	11	1.50	0.18
12	-0.50	0.24	0.02	0.47	0.32	-0.14	12	0.21	-0.10
13	0.56	0.72	-0.58	-0.49	0.36	-1.63	13	-0.63	0.44
14	0.63	0.87	0.13	0.02	-0.37	0.12	14	-0.09	0.58
15	-0.37	-0.60	-0.88	1.58	1.08	1.70	15	1.73	-0.84
16	0.98	1.62	0.52	-0.72	-1.77	-0.86	16	-1.28	1.14
17	-0.62	-1.29	-0.06	-0.22	-0.41	-0.60	17	-0.43	-0.72
18	1.98	0.38	1.18	0.96	1.11	0.09	18	0.82	1.36
19	0.75	-0.67	1.19	2.56	2.06	2.75	19	2.76	0.29
20	-1.47	-0.10	-1.28	-0.02	0.63	-0.60	20	0.01	-1.00
⋮	⋮	⋮	⋮	⋮	⋮	⋮		⋮	⋮
100	0.36	0.69	-1.27	0.32	-0.11	-0.14	100	0.14	-0.12

$$Z_j \leftarrow F_1 F_2 \dots F_m$$

$$Z_{ij} = a_{j1}F_{i1} + \dots + a_{jm}F_{im} + e_{ij}$$

$$Z_{i1} = a_{11}F_{i1} + a_{21}F_{i2} + e_{i1}$$

$$Z_{i2} = a_{12}F_{i1} + a_{22}F_{i2} + e_{i2}$$

$$Z_{i3} = a_{13}F_{i1} + a_{23}F_{i2} + e_{i3}$$

$$Z_{i4} = a_{14}F_{i1} + a_{24}F_{i2} + e_{i4}$$

$$Z_{i5} = a_{15}F_{i1} + a_{25}F_{i2} + e_{i5}$$

$$Z_{i6} = a_{16}F_{i1} + a_{26}F_{i2} + e_{i6}$$



# Modelling the data in Common Factor Analysis

## Modelling the scores on each variable

Scores on a variable can be predicted from scores on the factors.

$$Z \leftarrow F_1 F_2 \dots F_m$$

So for each variable:

$$Z_{ij} = a_{j1}F_{i1} + \dots + a_{jm}F_{im} + e_{ij}$$

Scores = explained by factors + unexplained by factors

DATA = MODEL + RESIDUAL

# Modelling the data in Common Factor Analysis

Modelling the variance of each variable

$$\text{var}(\mathbf{Z}) = \text{var}(\text{regression}) + \text{var}(\text{error})$$

$$1 = h^2 + u^2$$

$$\text{Total variance} = \text{Common variance} + \text{Specific variance (unique + error)}$$

$$\text{DATA} = \text{MODEL} + \text{RESIDUAL}$$

# Correlation matrices used in factor analysis

$R_{full}$

1					
	1				
		1			
			1		
				1	
					1

The full correlation matrix used in PCA

1 = the total variance of each variable

The reproduced correlation matrix


$R_{full}$

=


$R_{reproduced}$

+

$R_{residual}$

					
1	0.64	0.65	0.15	0.40	0.14
	0.64	1	0.49	-0.04	0.19
	0.65	0.49	1	-0.13	0.15
	0.15	-0.04	-0.13	1	0.71
	0.40	0.19	0.15	0.71	1
	0.14	-0.01	-0.04	0.70	0.47

=

					
0.84	0.73	0.73	0.15	0.44	0.15
	0.73	0.68	0.69	-0.05	0.23
	0.73	0.69	0.71	-0.14	0.17
	0.15	-0.05	-0.14	0.89	0.75
	0.44	0.23	0.17	0.75	0.67
	0.15	-0.03	-0.11	0.79	0.67

+

					
0.16	-0.09	-0.07	-0.00	-0.03	-0.02
	-0.09	0.32	-0.20	0.02	-0.04
	-0.07	-0.20	0.29	0.01	-0.02
	-0.00	0.02	0.01	0.11	-0.04
	-0.03	-0.04	-0.02	-0.04	0.25
	-0.02	0.03	0.07	-0.09	-0.20

**Note:** The closer that these reproduced correlations match the original correlations, the better the factor analysis captured the relationship among the variables. This difference comes out in the residual correlations. The goal is to get these as small as possible.

In factor analysis via principal components analysis the correlation matrix was 'factored'. The total variance was 'repackaged'. The **full correlation matrix** contains 1's on the diagonal. This is the variance of each standardised variable.

Recall from last lecture...



# Correlation matrices used in factor analysis

$R_{full}$

1					
	1				
		1			
			1		
				1	
					1

The full correlation matrix used in PCA

1 = the total variance of each variable

$R_{reduced}$

$h_1^2$					
	$h_2^2$				
		$h_3^2$			
			$h_4^2$		
				$h_5^2$	
					$h_6^2$

The reduced correlation matrix used in PCA

$h^2$  = the communalities of each variable

The reproduced correlation matrix


$R_{full}$

=


$R_{reproduced}$

+

$R_{residual}$

					
1	0.64	0.65	0.15	0.40	0.14
0.64	1	0.49	-0.04	0.19	-0.01
0.65	0.49	1	-0.13	0.15	-0.04
0.15	-0.04	-0.13	1	0.71	0.70
0.40	0.19	0.15	0.71	1	0.47
0.14	-0.01	-0.04	0.70	0.47	1

=

					
0.84	0.73	0.73	0.15	0.44	0.15
0.73	0.68	0.69	-0.05	0.23	-0.03
0.73	0.69	0.71	-0.14	0.17	-0.11
0.15	-0.05	-0.14	0.89	0.75	0.79
0.44	0.23	0.17	0.75	0.75	0.67
0.15	-0.03	-0.11	0.79	0.67	0.70

+

					
0.16	-0.09	-0.07	-0.00	-0.03	-0.02
-0.09	0.32	-0.20	0.02	-0.04	0.03
-0.07	-0.20	0.29	0.01	-0.02	0.07
-0.00	0.02	0.01	0.11	-0.04	-0.09
-0.03	-0.04	-0.02	-0.04	0.25	-0.20
-0.02	0.03	0.07	-0.09	-0.20	0.30

**Note:** The closer that these reproduced correlations match the original correlations, the better the factor analysis captured the relationship among the variables. This difference comes out in the residual correlations. The goal is to get these as small as possible.

In a common factor analysis only the common variance is desired to be factored and it is this common variance that is placed on the diagonal of the correlation matrix. This matrix is called the **reduced correlation matrix** with the communalities in the diagonal. It is factored and the eigenvalues and eigenvectors of this reduced matrix are found and form the basis of the common factor solution.

Recall from last lecture...



# Factor indeterminacy problem

The communality is the variance of the variable that it has in common with the other variables. We need to determine the proportion of each variable's total variance that is *common* variance. The estimate of the communality depends on the number of factors in the model. So we either need to know the communality or the number of factors to be able to get a solution. Unfortunately, these are unknown prior to the factor analysis. This is sometimes called the *factor indeterminacy problem*.

The name for the BIG problem

$A_{\text{rotated}}$

	1	2	$h^2$
	0.23	0.88	0.836
	0.00	0.83	0.681
	-0.08	0.84	0.714
	0.94	-0.07	0.887
	0.82	0.28	0.748
	0.84	-0.04	0.699

$$\sum a^2 = .23^2 + .88^2 = .836$$

That is, 83.6% of the variance in 'lyric recall' is accounted for by Factor 1 plus Factor 2.

This gives us an indication of how much 'lyric recall' has in common with the two factors.

Sum of Square Loadings  
/6

2.31	2.26	4.564
38.5%	37.6%	76.11%

$$\sum a^2 = .23^2 + .00^2 + -.08^2 + .94^2 + .82^2 + .84^2 = 2.31 = \lambda_1$$

One way of solving this problem is called:  
*Principal axis factoring (paf)*...

Recall from last lecture...





# Principal axis factoring (paf)

- This method solves the problem by starting with initial estimates of the communality together with an estimate of the number of factors.
- An initial estimate of the communality for a variable is the square multiple correlation (SMC) of that variable being predicted from the other variables in the analysis:

$$\begin{array}{ll} \mathbf{Z}_1 \leftarrow \mathbf{Z}_2 \dots \mathbf{Z}_p & R_1^2 = \text{initial } \hat{h}_1^2 \\ \vdots & \vdots \\ \mathbf{Z}_p \leftarrow \mathbf{Z}_1 \dots \mathbf{Z}_{p-1} & R_p^2 = \text{initial } \hat{h}_p^2 \end{array}$$

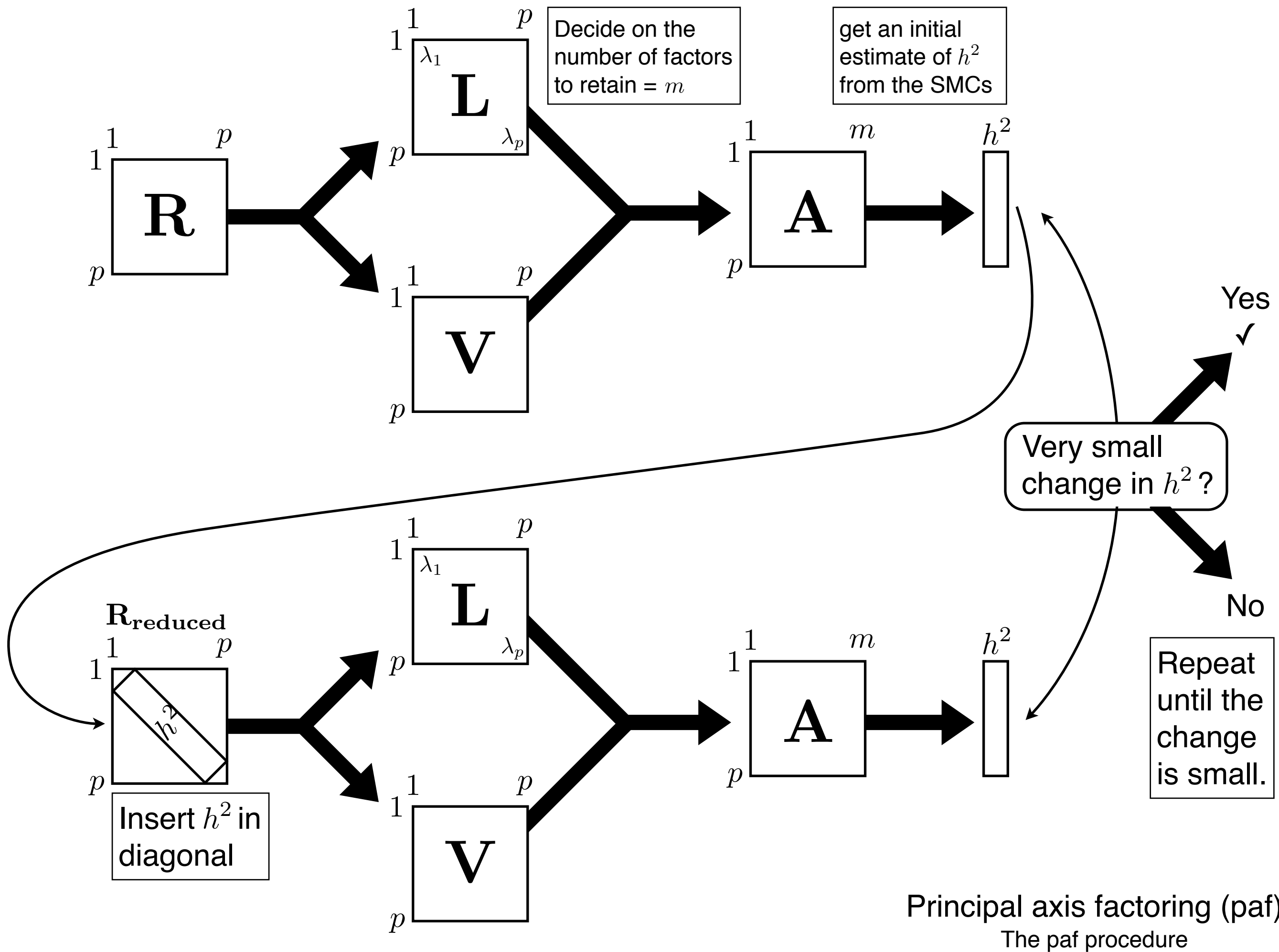
This provides an estimate of how much variance a variable has in common with the other variables.

# Principal axis factoring (paf)

## The paf procedure

1. Estimate the number of factors from the Singular Value Decomposition (SVD) of the  $\mathbf{R}_{full}$  matrix;
2. get an initial estimate of the communalities from the squared multiple correlations;
3. insert these into the diagonal of the  $\mathbf{R}_{full}$  correlation matrix;
4. factor the reduced matrix ( $\mathbf{R}_{reduced}$ );
5. using the estimate of the number of factors, calculate the communalities from this solution using the sum of squared loadings of the unrotated matrix of factor loadings,  $\mathbf{A}$ ;
6. insert the new estimates in the diagonal of the correlation matrix to form a new reduced correlation matrix and factor this new matrix;
7. Repeat this process until there is very little change in the communality estimates.

Once an unrotated factor matrix is found, the procedure is the same as for factor analysis via principal components.



# Using SPSS for Common Factor Analysis

## Research Question and Design

- From T&F (p. 651): During the second year of the panel study described in Appendix B Section B.1, participants completed the Bem Sex Role Inventory (BSRI; Bem, 1974). The sample included 369 middle-class, English-speaking women between the ages of 21 and 60 who were interviewed in person.

Forty-five items from the BSRI were selected for this research, where 20 items measure femininity, 19 masculinity, and 5 social desirability. Respondents attribute traits (e.g., “gentle”, “shy”, “dominant”) to themselves by assigning numbers between 1 (“never or almost never true of me”) and 7 (“always or almost always true of me”) to each of the items. Responses are summed to produce separate masculine and feminine scores. Masculinity and femininity are conceived as orthogonal dimensions of personality with both, one, or neither descriptive of any given individual.

Previous factor analytic work had indicated the presence of between three and five factors underlying the items of the BSRI. Investigation of the factor structure for this sample of women is a goal of this analysis.

# Using SPSS for Common Factor Analysis

## Data Diagnostics

- T&F identified 25 cases that might be considered outliers using a criterion of  $\alpha = .001$  with 44df, critical  $\chi^2 = 78.75$  of the Mahalanobis distance.
- Even though we should run the analysis with and without these outliers and compare the data, we'll eliminate these 25 cases so you can compare these numbers with those obtained by T&F.

factor.sav [DataSet2] - SPSS Data Editor

1 : filter\_5 1.0 Visible: 48 of 48 Variables

	subno	helpful	reliant	defbel	yielding	cheerful	indpt	athlet	shy	assert	strper
1	1	7	7	5	5	7	7	7	1	7	
2	2	5	6	6	6	2	3	3	3	4	
3	3	7	6	4	4	5	5	2	3	4	
4	4	6	6	7	4	6	6	3	4	4	
5	5	6	6	7	4	7	7	7	2	7	
6	7	5	6	7	4	6	6	2	4	4	
7	8	6	4	6	6	6	3	1	3	3	
8	9	7	6	7	5	6	7	5	2	5	
9	10	7	6	6	4	4	5	2	2	5	
10	11	7	4	7	4	7	5	2	1	5	
11	13	7	7	7	4	6	7	1	3	6	
12	14	7	7	5	5	7	1	3	5	6	
13	15	7	7	7	7	7	7	7	1	7	
14	23	5	6	7	5	6	6	1	1	6	
15	25	6	6	7	4	4	7	7	1	5	
16	26	6	6	5	5	7	6	2	2	5	
17	27	7	7	7	6	6	7	1	3	5	
18	28	7	6	7	7	6	5	4	3	3	
19	29	7	5	7	3	4	5	4	4	4	
20	30	6	6	7	4	5	6	4	3	5	
21	31	6	6	4	4	4	4	6	4	4	
22	32	5	5	5	7	5	7	5	4	5	

Data View Variable View

SPSS Processor is ready Filter On



# Using SPSS for Common Factor Analysis

## SPSS Syntax

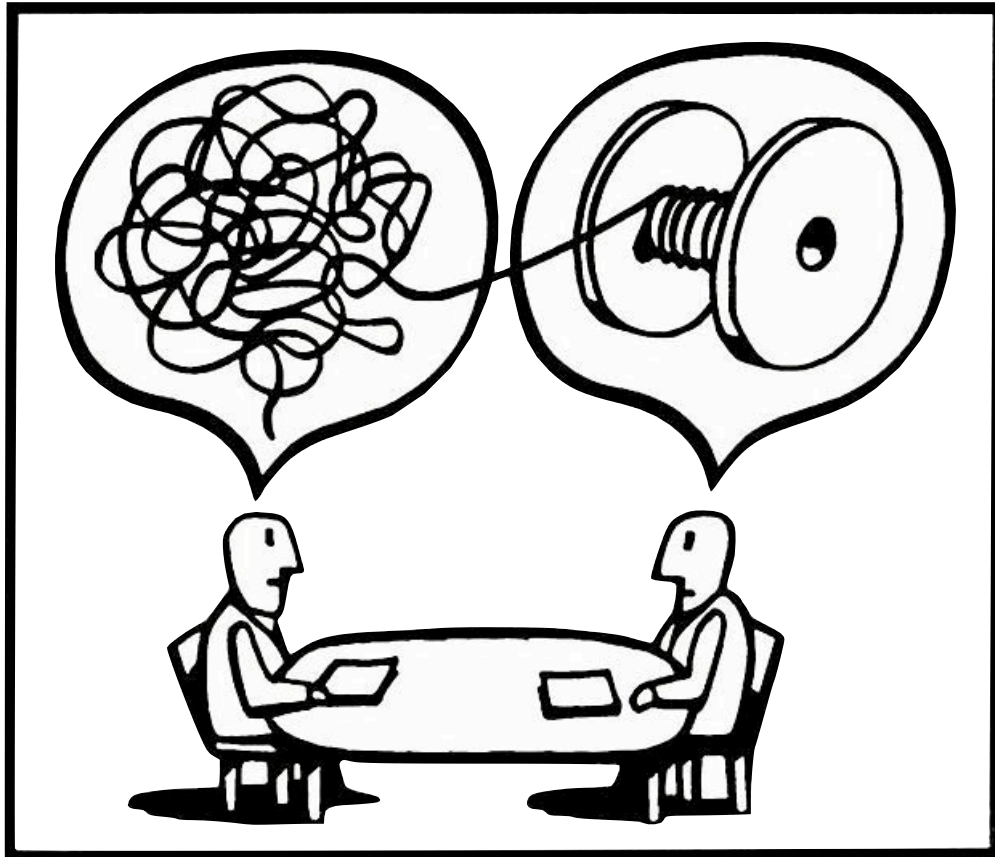
FACTOR

```
/VARIABLES helpful reliant defbel yielding cheerful indpt athlet shy assert  
strpers forceful affect flatter loyal analyt feminine sympathy moody  
sensitiv undstand compass leaderab soothe risk decide selfsuff conscien  
dominant masculin stand happy softspok warm truthful tender gullible leadact  
childlik individ foullang lovchil compete ambitiou gentle  
/MISSING LISTWISE  
/PRINT INITIAL REPR KMO EXTRACTION ROTATION  
/FORMAT SORT  
/PLOT EIGEN ROTATION  
/CRITERIA FACTORS(3)  
/EXTRACTION PAF  
/ROTATION VARIMAX  
/METHOD=CORRELATION .
```

You exert control here

You need another run to get  
an oblique rotation solution

# Interpretation of Common Factor Analysis



- Choice of factor model.
- Factorability and other assumptions.
- Number of factors to retain.
- Type of rotation.
- The interpretation of the factor solution.
- Adequacy of the factor solution.



# Interpretation of Common Factor Analysis

## Choice of factor model

- If your goal is to summarise patterns and/or reduce data.
  - Factor analysis via Principal Components Analysis is appropriate.
- If your goal is to identify underlying constructs.
  - Common factor analysis is appropriate.

For this analysis, it's believed that constructs of masculinity and femininity underlie the data, so a common factor analysis is used.

# Interpretation of Common Factor Analysis

## Factorability and other assumptions

Correlations

	helpful	reliant	defbel	yielding	cheerful	indpt	athlet
helpful	1.000	.337	.239	.114	.220	.271	.173
reliant	.337	1.000	.145	-.012	.247	.514	.147
defbel	.239	.145	1.000	-.078	.102	.211	.066
yielding	.114	-.012	-.078	1.000	.181	-.034	.053
cheerful	.220	.247	.102	.181	1.000	.186	.181
indpt	.271	.514	.211	-.034	.186	1.000	.161
athlet	.173	.147	.066	.053	.181	.161	1.000

Looking through the 44 x 44 variable correlation matrix (part shown above), there are several large correlations (many in excess of .3), therefore patterns in responses to the variables are anticipated.

# Interpretation of Common Factor Analysis

## Factorability and other assumptions

### KMO and Bartlett's Test

Kaiser–Meyer–Olkin Measure of Sampling Adequacy.		.852
Bartlett's Test of Sphericity	Approx. Chi-Square	5954.22
	df	946
	Sig.	.000

The KMO measure is also available. It's a measure of the proportion of variance in your variables which is common variance. High values (close to 1) indicate that a factor analysis may be useful. If this value drops below .5, then an FA may not be so useful.

Bartlett's test of sphericity tests the null hypothesis that your correlation matrix is an identity matrix, which would indicate that the variables are completely uncorrelated. Of course you should expect to find some factors because of the conceptual design of the study, so this check is a bit redundant.

Skewness, outliers, linearity, multicollinearity are all potential issues in the sense that they affect the honesty of the correlations. (See T&F's treatment of these in this particular case: p. 652). There is no statistical inference in factor analysis, so normality isn't an issue, but honest correlations are!

# Interpretation of Common Factor Analysis

## General strategy for handling number of factors and rotation

The overall strategy for handling the number of factors to retain and the type of rotation to use is to generate a number of factor solutions from the range of possible values for the number of factors and from the two types of rotation.

The best of these factor solutions is used where “best” is defined as the most interpretable solution that meets the needs of the researcher.

It's highly unlikely that a solution will emerge in a single run of SPSS FACTOR.

Deciding on the number of factors is crucial and no method can give a 'correct' answer. The information for deciding the number of factors is in the initial statistics section of the output.



# Interpretation of Common Factor Analysis

## Number of factors to retain

### a. Eigenvalues > 1 rule

This is the default in SPSS FACTOR and is not necessarily (almost certainly NOT) the best. It generally gives the *maximum* number of interpretable factors.

Total Variance Explained

Factor	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	8.194	18.623	18.623
2	5.154	11.713	30.335
3	2.590	5.887	36.223
4	2.073	4.711	40.934
5	1.648	3.744	44.678
6	1.415	3.216	47.895
7	1.291	2.933	50.828
8	1.221	2.775	53.604
9	1.110	2.522	56.125
10	1.078	2.449	58.575
11	1.032	2.345	60.919
12	.951	2.162	63.081
13	.942	2.140	65.221

Total Variance Explained

Factor	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	8.194	18.623	18.623
2	5.154	11.713	30.335
3	2.590	5.887	36.223
4	2.073	4.711	40.934
5	1.648	3.744	44.678
6	1.415	3.216	47.895
7	1.291	2.933	50.828
8	1.221	2.775	53.604
9	1.110	2.522	56.125
10	1.078	2.449	58.575
11	1.032	2.345	60.919
12	.951	2.162	63.081
13	.942	2.140	65.221
14	.882	2.004	67.225
15	.863	1.962	69.188
16	.828	1.883	71.070
17	.770	1.751	72.821
18	.750	1.704	74.525
19	.737	1.675	76.200
20	.690	1.569	77.769
21	.652	1.481	79.250
22	.646	1.468	80.718
23	.626	1.423	82.141
24	.587	1.335	83.476
25	.571	1.298	84.774
26	.550	1.249	86.023
27	.513	1.166	87.189
28	.501	1.138	88.327
29	.474	1.078	89.405
30	.440	1.001	90.406
31	.435	.990	91.395
32	.392	.890	92.286
33	.368	.835	93.121
34	.366	.831	93.952
35	.350	.795	94.747
36	.330	.751	95.498
37	.319	.725	96.223
38	.302	.685	96.909
39	.287	.653	97.562
40	.260	.590	98.152
41	.250	.568	98.720
42	.230	.524	99.244
43	.207	.470	99.713
44	.126	.287	100.000

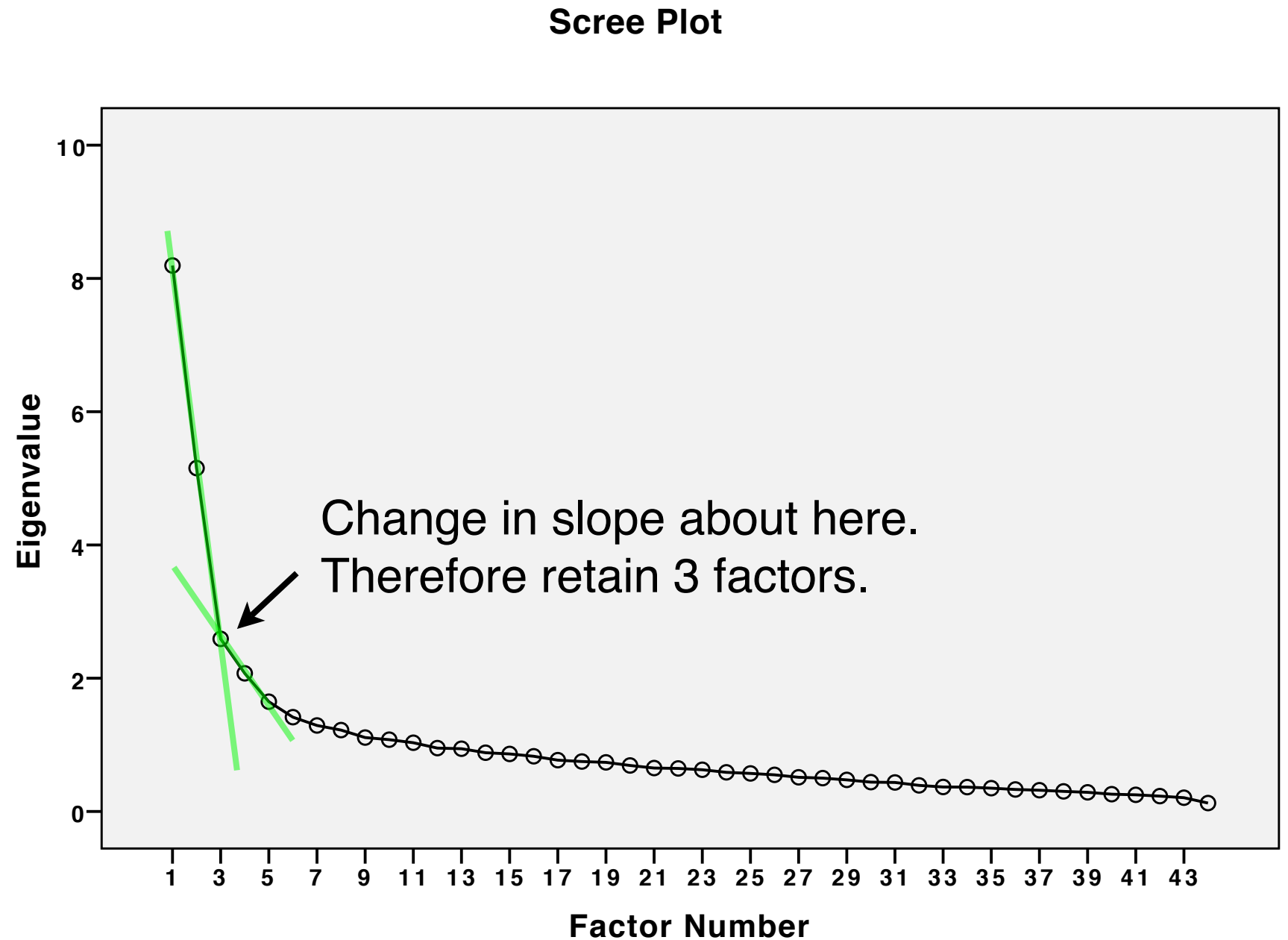
Extraction Method: Principal Axis Factoring.

# Interpretation of Common Factor Analysis

## Number of factors to retain

### b. Scree Test

Discontinuities and changes in slope are used to indicate the range of values for the number of factors.



# Interpretation of Common Factor Analysis

## Number of factors to retain

### c. Parallel analysis test

*Behavior Research Methods, Instruments, & Computers*  
2000, 32 (3), 396-402

## SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test

BRIAN P. O'CONNOR

*Lakehead University, Thunder Bay, Ontario, Canada*

Popular statistical software packages do not have the proper procedures for determining the number of components in factor and principal components analyses. Parallel analysis and Velicer's minimum average partial (MAP) test are validated procedures, recommended widely by statisticians. However, many researchers continue to use alternative, simpler, but flawed procedures, such as the eigenvalues-greater-than-one rule. Use of the proper procedures might be increased if these procedures could be conducted within familiar software environments. This paper describes brief and efficient programs for using SPSS and SAS to conduct parallel analyses and the MAP test.

<http://people.ok.ubc.ca/brioconn/nfactors/nfactors.html>

# Interpretation of Common Factor Analysis

## Number of factors to retain

### c. Parallel analysis test

If our data were random, the size of the eigenvalues would be due to chance alone.

If our factors are meaningful, our observed eigenvalues should be bigger than that expected by chance.

So we can check whether our factors are useful by checking whether they have bigger eigenvalues than factors from random data.

```
* Parallel Analysis program.

set mxloops=9000 printback=off width=80
seed = 1953125.
matrix.

* enter your specifications here.
compute ncases      = 344.
compute nvars       = 44.
compute ndatsets    = 1000.
compute percent     = 95.

* Specify the desired kind of parallel
analysis, where:
    1 = principal components analysis
    2 = principal axis/common factor analysis.
compute kind = 1 .

***** End of user specifications. *****

* principal components analysis.
do if (kind = 1).
compute evals = make(nvars,ndatsets,-9999).
compute nm1 = 1 / (ncases-1).
loop #nds = 1 to ndatsets.
compute x = sqrt(2 * (ln(uniform(ncases,nvars)) * -1) ) &*
           cos(6.283185 * uniform(ncases,nvars) ).
compute vcv = nm1 * (sscp(x) - ((t(csum(x))*csum(x))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute evals(:,#nds) = eval(d * vcv * d).
end loop.
end if.

* principal axis / common factor analysis with SMCs on the diagonal.
do if (kind = 2).
compute evals = make(nvars,ndatsets,-9999).
compute nm1 = 1 / (ncases-1).
loop #nds = 1 to ndatsets.
compute x = sqrt(2 * (ln(uniform(ncases,nvars)) * -1) ) &*
```

# Interpretation of Common Factor Analysis

## Number of factors to retain

### c. Parallel analysis test

Total Variance Explained

Factor	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	8.194	18.623	18.623
2	5.154	11.713	30.335
3	2.590	5.887	36.223
4	2.073	4.711	40.934
5	1.648	3.744	44.678
6	1.415	3.216	47.895
7	1.291	2.933	50.828
8	1.221	2.775	53.604
9	1.110	2.522	56.125
10	1.078	2.449	58.575
11	1.032	2.345	60.919
12	.951	2.162	63.081
13	.942	2.140	65.221

Random Data Eigenvalues

Root	Means	Prcntyle
1	1.762089	1.85146
2	1.6794	1.741623
3	1.613712	1.665987
4	1.561228	1.608289
5	1.512086	1.55255
6	1.46828	1.505712
7	1.427712	1.464369
8	1.389664	1.425624

Random Data Eigenvalues

Root	Means	Prcntyle
1	1.762089	1.85146
2	1.6794	1.741623
3	1.613712	1.665987
4	1.561228	1.608289
5	1.512086	1.55255
6	1.46828	1.505712
7	1.427712	1.464369
8	1.389664	1.425624
9	1.353425	1.387723
10	1.31909	1.352092
11	1.285107	1.316974
12	1.251495	1.283208
13	1.220034	1.250651
14	1.190358	1.21916
15	1.160454	1.188693
16	1.131468	1.159788
17	1.103657	1.129844
18	1.075725	1.101703
19	1.049466	1.074777
20	1.022778	1.047169
21	0.99731	1.021643
22	0.97245	0.996381
23	0.948068	0.972319
24	0.923267	0.947195
25	0.898762	0.922354
26	0.874836	0.899488
27	0.851738	0.874302
28	0.828454	0.85368
29	0.805789	0.829765
30	0.782461	0.806874
31	0.760162	0.783089
32	0.73716	0.760735
33	0.715188	0.73753
34	0.692423	0.715691
35	0.670355	0.693319
36	0.647706	0.670879
37	0.625627	0.649194
38	0.602518	0.626582
39	0.579569	0.603138
40	0.556392	0.581074
41	0.531337	0.554827
42	0.504355	0.53059
43	0.475846	0.503024
44	0.441	0.471835

# Interpretation of Common Factor Analysis

## Number of factors to retain

### c. Parallel analysis test

Total Variance Explained

Factor	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	8.194	18.623	18.623
2	5.154	11.713	30.335
3	2.590	5.887	36.223
4	2.073	4.711	40.934
5	1.648	3.744	44.678
6	1.415	3.216	47.895
7	1.291	2.933	50.828
8	1.221	2.775	53.604
9	1.110	2.522	56.125
10	1.078	2.449	58.575
11	1.032	2.345	60.919
12	.951	2.162	63.081
13	.942	2.140	65.221

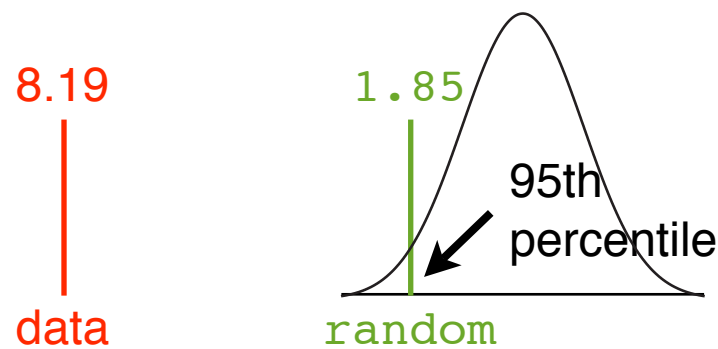
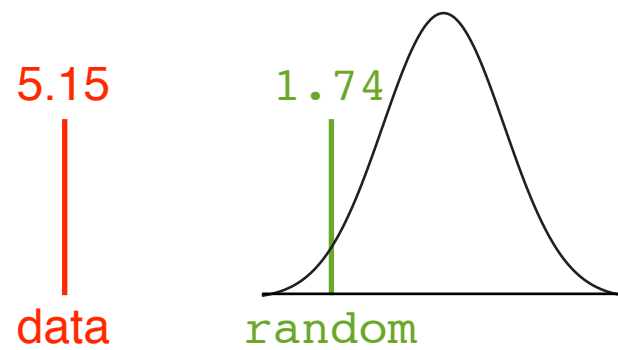
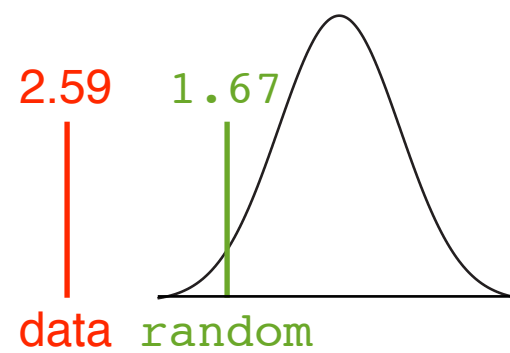
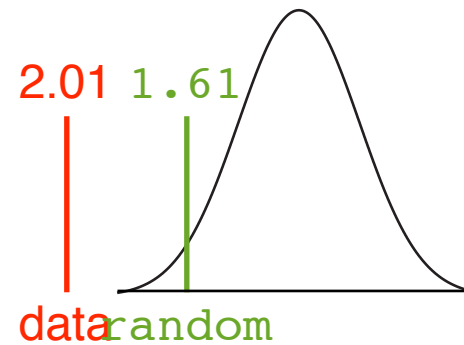
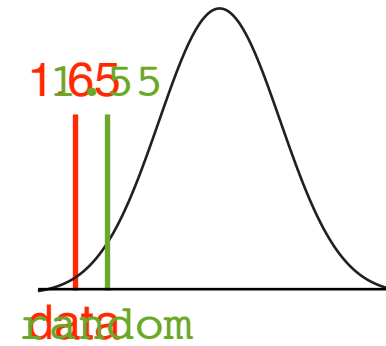
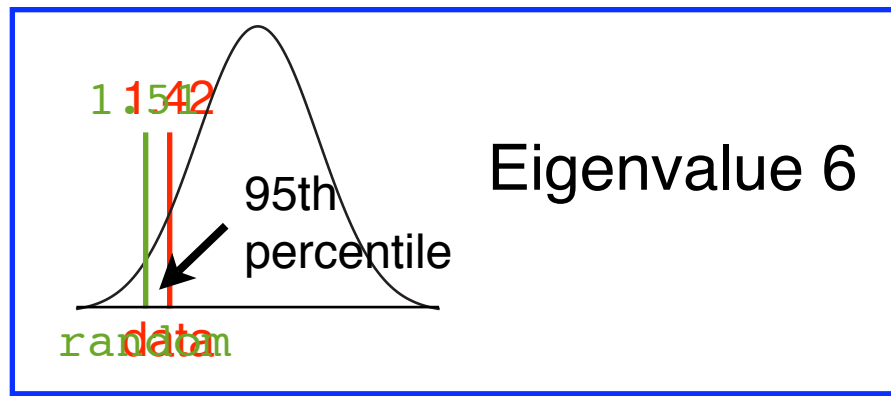
Random Data Eigenvalues

Root	Means	Prcntyle
1	1.762089	1.85146
2	1.6794	1.741623
3	1.613712	1.665987
4	1.561228	1.608289
5	1.512086	1.55255
6	1.46828	1.505712
7	1.427712	1.464369
8	1.389664	1.425624

Continue until the random eigenvalues exceed the eigenvalues obtained from the data.

Random Data Eigenvalues		
Root	Means	Prcntyle
1	1.762089	1.85146
2	1.6794	1.741623
3	1.613712	1.665987
4	1.561228	1.608289
5	1.512086	1.55255
6	1.46828	1.505712
7	1.427712	1.464369
8	1.389664	1.425624
9	1.353425	1.387723
10	1.31909	1.352092
11	1.285107	1.316974
12	1.251495	1.283208
13	1.220034	1.250651
14	1.190358	1.21916
15	1.160454	1.188693
16	1.131468	1.159788
17	1.103657	1.129844
18	1.075725	1.101703
19	1.049466	1.074777
20	1.022778	1.047169
21	0.99731	1.021643
22	0.97245	0.996381
23	0.948068	0.972319
24	0.923267	0.947195
25	0.898762	0.922354
26	0.874836	0.899488
27	0.851738	0.874302
28	0.828454	0.85368
29	0.805789	0.829765
30	0.782461	0.806874
31	0.760162	0.783089
32	0.73716	0.760735
33	0.715188	0.73753
34	0.692423	0.715691
35	0.670355	0.693319
36	0.647706	0.670879
37	0.625627	0.649194
38	0.602518	0.626582
39	0.579569	0.603138
40	0.556392	0.581074
41	0.531337	0.554827
42	0.504355	0.53059
43	0.475846	0.503024
44	0.441	0.471835





# Interpretation of Common Factor Analysis

## Number of factors to retain

### a. Eigenvalues $> 1$ rule

- Retain 11 factors

### b. Scree Test

- Retain 3 factors

### c. Parallel analysis test

- Retain 5 factors

Tabachnick & Fidell:

“Previous factor analytic work had indicated the presence of between three and five factors underlying the items of the BSRI. Investigation of the factor structure for this sample of women is a goal of this analysis.”

They retained 4 factors

We'll retain 3 just to make naming easier...

# Interpretation of Common Factor Analysis

## Type of rotation

- Decisions about the methods of rotation.
  - In order to improve interpretability, the initial solution is rotated. A factor solution is easier to interpret if on a factor there are only a few highly loading variables and if a variable loads highly on one factor only. Since we don't know how the factors are related when we start, i.e. the degree of correlation between them, one suggestion is to get both the orthogonal and oblique solutions for each of the number of factors in the estimated range.
- Oblique
  - The axes are rotated and are allowed to become oblique to each other. One procedure is the Oblimin method. The criterion is the same as for orthogonal rotation, that of simple structure. The pattern matrix and the correlations between the factors are interpreted. If the correlations between the factors are low then an orthogonal solution is about the same and is interpreted.
- Orthogonal
  - The axes of the factor space are rotated keeping the axes at right angles (orthogonal). One procedure is a Varimax rotation. The axes are rotated to try and maximise the fit for the conditions. The rotated factor matrix, or structure matrix is interpreted.

# Interpretation of Common Factor Analysis

## Type of rotation

**Factor Correlation Matrix**

Factor	1	2	3
1	1.000	.143	-.106
2	.143	1.000	-.177
3	-.106	-.177	1.000

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

There are low correlations among the factors ( $< .3$ ),  
so there is no real need for an oblique rotation.

So choose an orthogonal solution... but it's a  
judgement call and depends on the purpose.

**Note: T&F used a Promax  
rotation rather than Oblimin**

**Factor Correlation Matrix**

Factor	1	2	3	4
1	1.000	.137	.106	.151
2	.137	1.000	.299	.011
3	.106	.299	1.000	.031
4	.151	.011	.031	1.000

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

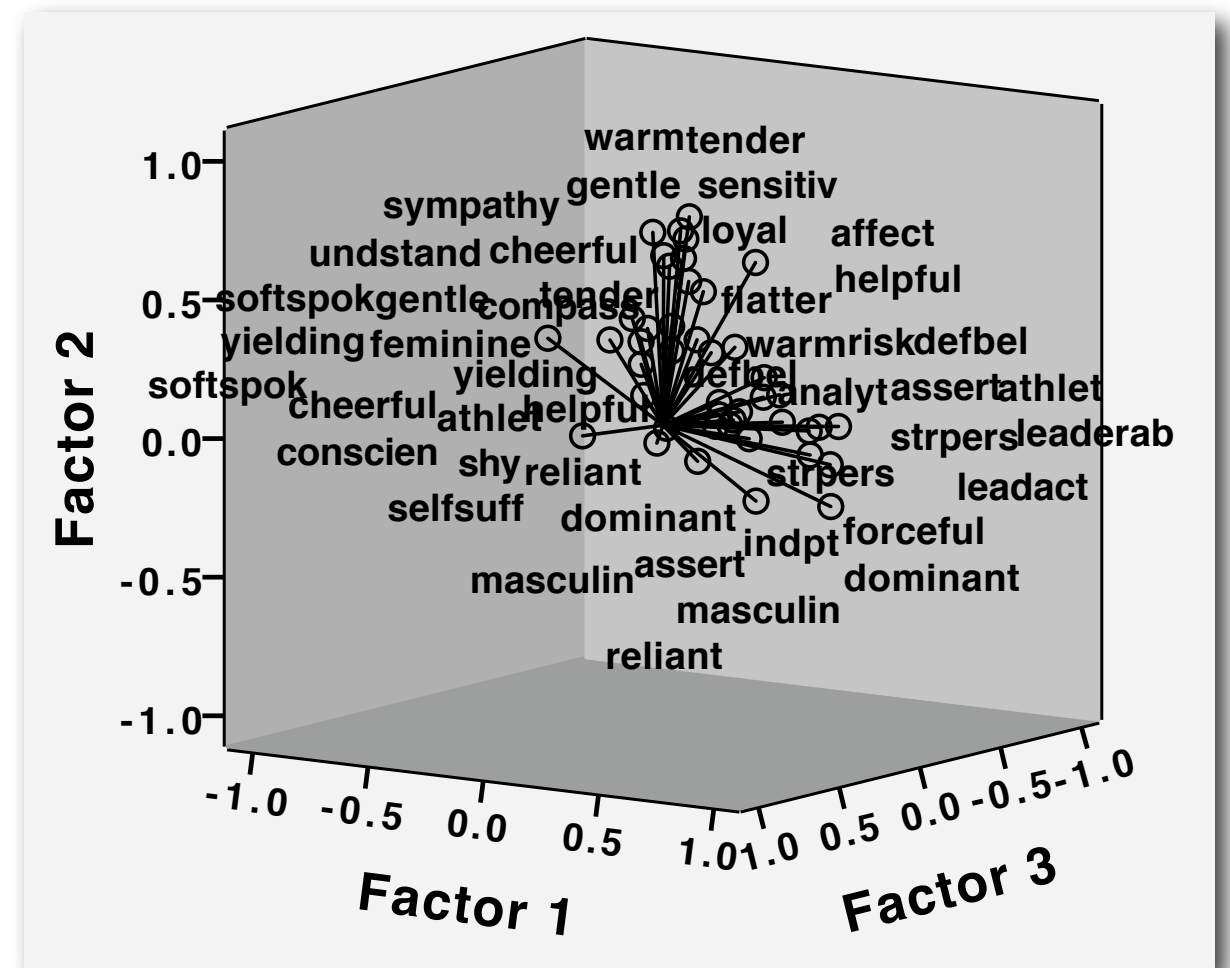
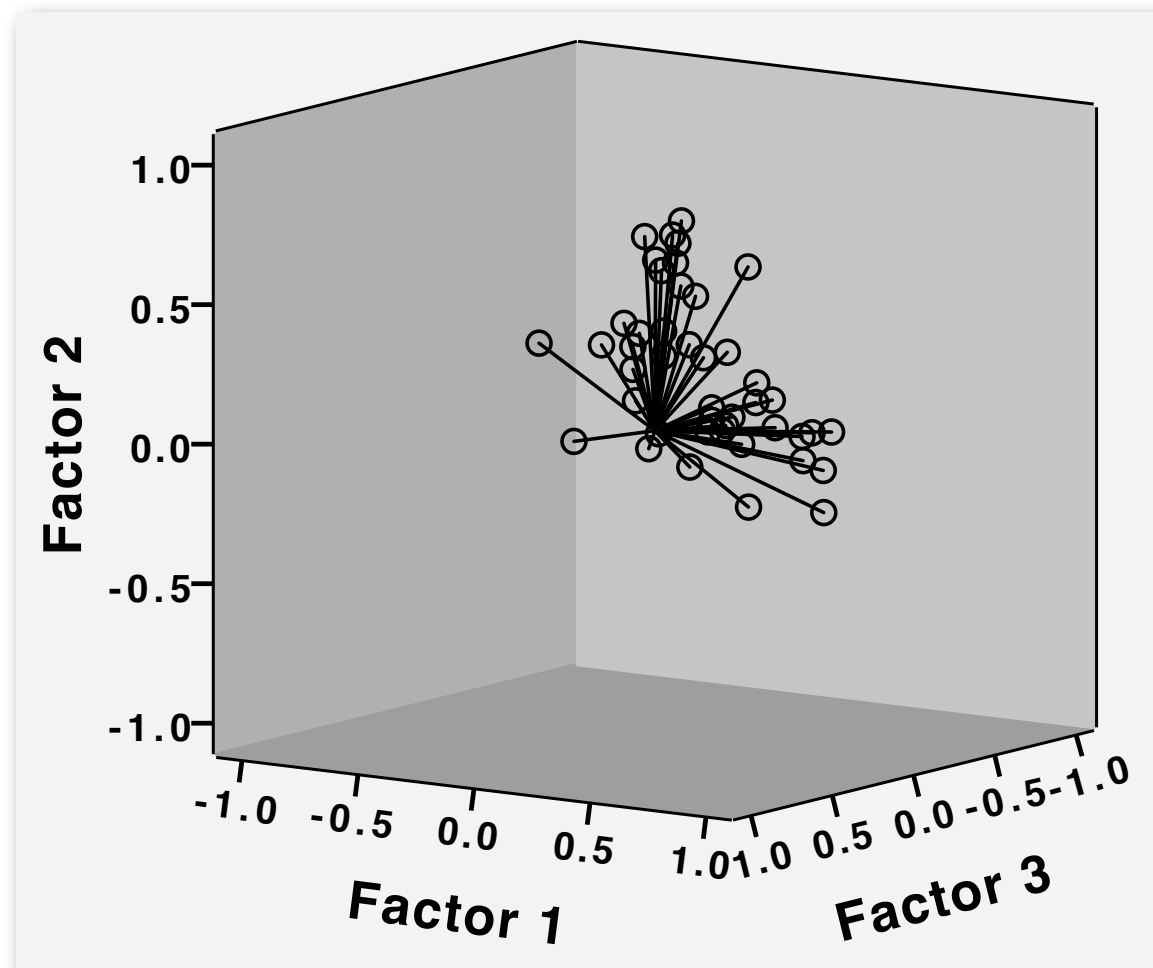
# Interpretation of Common Factor Analysis

## Interpretation of the factor solution

- Factors are named by choosing a phrase or definition that encapsulates the “common thread” amongst the high loading variables.
- The criterion for a high loading must be made explicit ( $> .3$  is commonly used).
  - Reminder: This is a highly subjective process!

# Interpretation of Common Factor Analysis

## Interpretation of the factor solution



- Plotting the factors and making use of Point ID and spinning the factors can be helpful in interpretation.
  - Though it's not terribly useful for presenting data to the reader.



	Factor		
	1	2	3
leaderab	.74	.08	.15
leadact	.72	-.02	.11
strpers	.71	.06	-.08
dominant	.69	-.23	-.05
forceful	.67	-.09	-.08
assert	.66	.06	-.03
stand	.61	.19	.15
compete	.50	.06	-.03
decide	.48	.10	.37
risk	.48	.16	.06
indpt	.48	-.01	.47
ambitiou	.44	.12	.15
individ	.44	.09	.19
defbel	.43	.21	.00
shy	-.38	-.08	-.05
masculin	.32	-.26	-.11
analyt	.30	.13	.09
athlet	.28	.07	.04
warm	.12	.77	.01
tender	.08	.71	.01
gentle	.00	.70	.06
compass	.11	.69	.02
undstand	.08	.63	.11
soothe	.08	.61	.00
affect	.28	.59	-.18
sympathy	.00	.57	-.03
sensitiv	.11	.53	.00
loyal	.20	.51	.03
happy	.07	.43	.29
cheerful	.12	.40	.27
helpful	.31	.37	.23
lovchil	.01	.35	-.04
feminine	.04	.33	.20
softspok	-.31	.32	.27
truthful	.16	.31	.17
yielding	-.16	.30	.09
flatter	.13	.26	-.25
foullang	-.03	.12	.08
selfsuff	.43	.07	.66
reliant	.37	.10	.50
conscien	.22	.31	.44
moody	.06	-.11	-.44
childlik	.00	-.06	-.42
gullible	-.06	.20	-.39

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

a. Rotation converged in 5 iterations.

You need to  
select a cut point  
(T&F chose .45)

Excel File Edit View Insert Format Tools Data Window Help

rfm.xls

New Open Save Print Import Copy Paste Format Undo Redo AutoSum Sort A-Z Sort Z-A Gallery Toolbox

Sheets Charts SmartArt Graphics WordArt

Rotated Factor Matrix<sup>a</sup>

		1	2	3
1	Rotated Factor Matrix <sup>a</sup>			
2	Factor			
3		1	2	3
4	leaderab	0.74	0.08	0.15
5	leadact	0.72	-0.02	0.11
6	strpers	0.71	0.06	-0.08
7	dominant	0.69	-0.23	-0.05
8	forceful	0.67	-0.09	-0.08
9	assert	0.66	0.06	-0.03
10	stand	0.61	0.19	0.15
11	compete	0.50	0.06	-0.03
12	decide	0.48	0.10	0.37
13	risk	0.48	0.16	0.06
14	indpt	0.48	-0.01	0.47
15	ambitiou	0.44	0.12	0.15
16	individ	0.44	0.09	0.19
17	defbel	0.43	0.21	-0.01
18	shy	-0.38	-0.08	-0.05
19	masculin	0.32	-0.26	-0.11
20	analyt	0.30	0.13	0.09
21	athlet	0.28	0.07	0.04
22	warm	0.12	0.77	0.01
23	tender	0.08	0.71	0.01
24	gentle	0.00	0.70	0.06
25	compass	0.11	0.69	0.02
26	undstand	0.08	0.63	0.11
27	soothe	0.08	0.61	-0.01
28	affect	0.28	0.59	-0.18
29	sympathy	0.00	0.57	-0.03
30	sensitiv	0.11	0.53	0.00
31	loyal	0.20	0.51	0.03
32	happy	0.07	0.43	0.29
33	cheerful	0.12	0.40	0.27
34	helpful	0.31	0.37	0.23
35	lovchil	0.01	0.35	-0.04
36	feminine	0.04	0.33	0.20
37	softspok	-0.31	0.32	0.27
38	truthful	0.16	0.31	0.17
39	yielding	-0.16	0.30	0.09
40	flatter	0.13	0.26	-0.25
41	foullang	-0.03	0.12	0.08
42	selfsuff	0.43	0.07	0.66
43	reliant	0.37	0.10	0.50
44	conscien	0.22	0.31	0.44
45	moody	0.06	-0.11	-0.44
46	childlik	-0.01	-0.06	-0.42
47	gullible	-0.06	0.20	-0.39
48	Extraction Method: Principal Axis Factoring. Rotation Method: Varimax with Kaiser Normalization.			
49	a. Rotation converged in 5 iterations.			
50				
51				

Sheet0

Normal View Ready

	Factor		
	1	2	3
leaderab	.74	.08	.15
leadact	.72	-.02	.11
strpers	.71	.06	-.08
dominant	.69	-.23	-.05
forceful	.67	-.09	-.08
assert	.66	.06	-.03
stand	.61	.19	.15
compete	.50	.06	-.03
decide	.48	.10	.37
risk	.48	.16	.06
indpt	.48	-.01	.47
ambitiou	.44	.12	.15
individ	.44	.09	.19
defbel	.43	.21	.00
shy	-.38	-.08	-.05
masculin	.32	-.26	-.11
analyt	.30	.13	.09
athlet	.28	.07	.04
warm	.12	.77	.01
tender	.08	.71	.01
gentle	.00	.70	.06
compass	.11	.69	.02
undstand	.08	.63	.11
soothe	.08	.61	.00
affect	.28	.59	-.18
sympathy	.00	.57	-.03
sensitiv	.11	.53	.00
loyal	.20	.51	.03
happy	.07	.43	.29
cheerful	.12	.40	.27
helpful	.31	.37	.23
lovchil	.01	.35	-.04
feminine	.04	.33	.20
softspok	-.31	.32	.27
truthful	.16	.31	.17
yielding	-.16	.30	.09
flatter	.13	.26	-.25
foullang	-.03	.12	.08
selfsuff	.43	.07	.66
reliant	.37	.10	.50
conscien	.22	.31	.44
moody	.06	-.11	-.44
childlik	.00	-.06	-.42
gullible	-.06	.20	-.39

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

a. Rotation converged in 5 iterations.

You need to  
select a cut point  
(T&F chose .45)

Dominance

Slight factorial complexity

Independence

Empathy

	Factor		
	1	2	3
leadership ability	0.74	0.08	0.15
act as a leader	0.72	-0.02	0.11
strong personality	0.71	0.06	-0.08
dominant	0.69	-0.23	-0.05
forceful	0.67	-0.09	-0.08
assertive	0.66	0.06	-0.03
willing to take a stand	0.61	0.19	0.15
competitive	0.50	0.06	-0.03
makes decisions easily	0.48	0.10	0.37
willing to take risks	0.48	0.16	0.06
independent	0.48	-0.01	0.47
self sufficient	0.43	0.07	0.66
self reliant	0.37	0.10	0.50
affectionate	0.28	0.59	-0.18
loyal	0.20	0.51	0.03
warm	0.12	0.77	0.01
compassionate	0.11	0.69	0.02
sensitive	0.11	0.53	0.00
eager to soothe hurt feelings	0.08	0.61	-0.01
tender	0.08	0.71	0.01
understanding	0.08	0.63	0.11
sympathy	0.00	0.57	-0.03
gentle	0.00	0.70	0.06