An introduction to R: a short course

Multivariate analysis

The zelig website

ZELIG: EVERYONE'S STATISTICAL SOFTWARE

Kosuke Imai, Gary King, Olivia Lau

Version:3.4-5 (Download) (What's new)

Zelig is named after a Woody Allen movie about a man who had the strange ability to become the physical and psychological reflection of anyone he met and thus to fit perfectly in any situation.

Zelig is a single, easy-to-use program that can estimate, help interpret, and present the results of a large range of statistical methods. It literally *is* "everyone's statistical software" because Zelig uses (R) code from many researchers. We also hope it will *become* "everyone's statistical software" for applications, and we have designed it so that anyone can use it or add their methods to it. Zelig comes with detailed, self-contained documentation that minimizes startup costs for Zelig and R, automates graphics and summaries for all models, and, with only three simple commands required, generally makes the power of R accessible for all users. Zelig also works well for teaching, and is designed so that scholars can use the same program with students that they use for their research.

Zelig adds considerable infrastructure to improve the use of existing methods. It generalizes the program Clarify (for Stata), which translates hard-to-interpret coefficients into quantities of interest; combines multiply imputed data sets (such as output from Amelia) to deal with missing data; automates bootstrapping for all models; uses sophisticated nonparametric matching commands which improve parametric procedures (via MatchIt); allows one-line commands to run analyses in all designated strata; automates the creation of replication data files so that you (or, if you wish, anyone else) can replicate the results of your analyses (hence satisfying the replication standard); makes it easy to evaluate counterfactuals (via WhatIf); and allows conditional population and superpopulation inferences. Zelig includes many specific methods, based on likelihood, frequentist, Bayesian, robust Bayesian, and nonparametric theories of inference.

- Documentation and Installation
- Installation, Frequently Asked Questions
- Please send **All** questions, bugs and requests: Zelig Mailing List, [Un]Subscribe, or Browse/Search Archives
- A paper that describes the advances underlying Zelig software: Kosuke Imai, Gary King, and Olivia Lau. **"Toward A Common Framework for Statistical Analysis and Development"** *Journal of Computational and Graphical Statistics*, Vol. 17, No. 4 (December), pp. 892-913 (**Abstract:** HTML | **Paper:** PDF)
- Slides used to introduce Zelig
- Be notified of Zelig changes

http://gking.harvard.edu/zelig/

Multivariate analysis

I. Dimension reduction through factor analysis, principal components analysis, cluster analysis, and multidimensional scaling

II. Multiple measures of reliability

III. Practical use of R for scoring inventories

Dimension Reduction

I. The problem: How best to summarize and think about many variables some of which are moderately correlated

II. The solutions: rank reduction through FA, PCA, CA, MDS

III. Examples will be tests and then items

The Thurstone data set

```
> data(bifactor)
> colnames(Thurstone) <-
c("Sentences","Vocab","S.comp","F.letter","4.letter","Suffix",
     "Series","Pedi","letters")



> round(Thurstone,2)
                Sentences Vocab S.comp F.letter 4.letter Suffix Series Pedi letters
Sentences            1.00  0.83   0.78     0.44     0.43   0.45   0.45 0.54    0.38
Vocabulary           0.83  1.00   0.78     0.49     0.46   0.49   0.43 0.54    0.36
Sent.Completion      0.78  0.78   1.00     0.46     0.42   0.44   0.40 0.53    0.36
First.Letters        0.44  0.49   0.46     1.00     0.67   0.59   0.38 0.35    0.42
4.Letter.Words       0.43  0.46   0.42     0.67     1.00   0.54   0.40 0.37    0.45
Suffixes             0.45  0.49   0.44     0.59     0.54   1.00   0.29 0.32    0.32
Letter.Series        0.45  0.43   0.40     0.38     0.40   0.29   1.00 0.56    0.60
Pedigrees            0.54  0.54   0.53     0.35     0.37   0.32   0.56 1.00    0.45
Letter.Group         0.38  0.36   0.36     0.42     0.45   0.32   0.60 0.45    1.00
```

How many dimensions

I. Chi Square test

II. Scree plot

III. Parallel analysis of random data

IV. Minimum Average Partial correlation

V. Very Simple Structure

VI. Do not use eigen value > 1 rule!

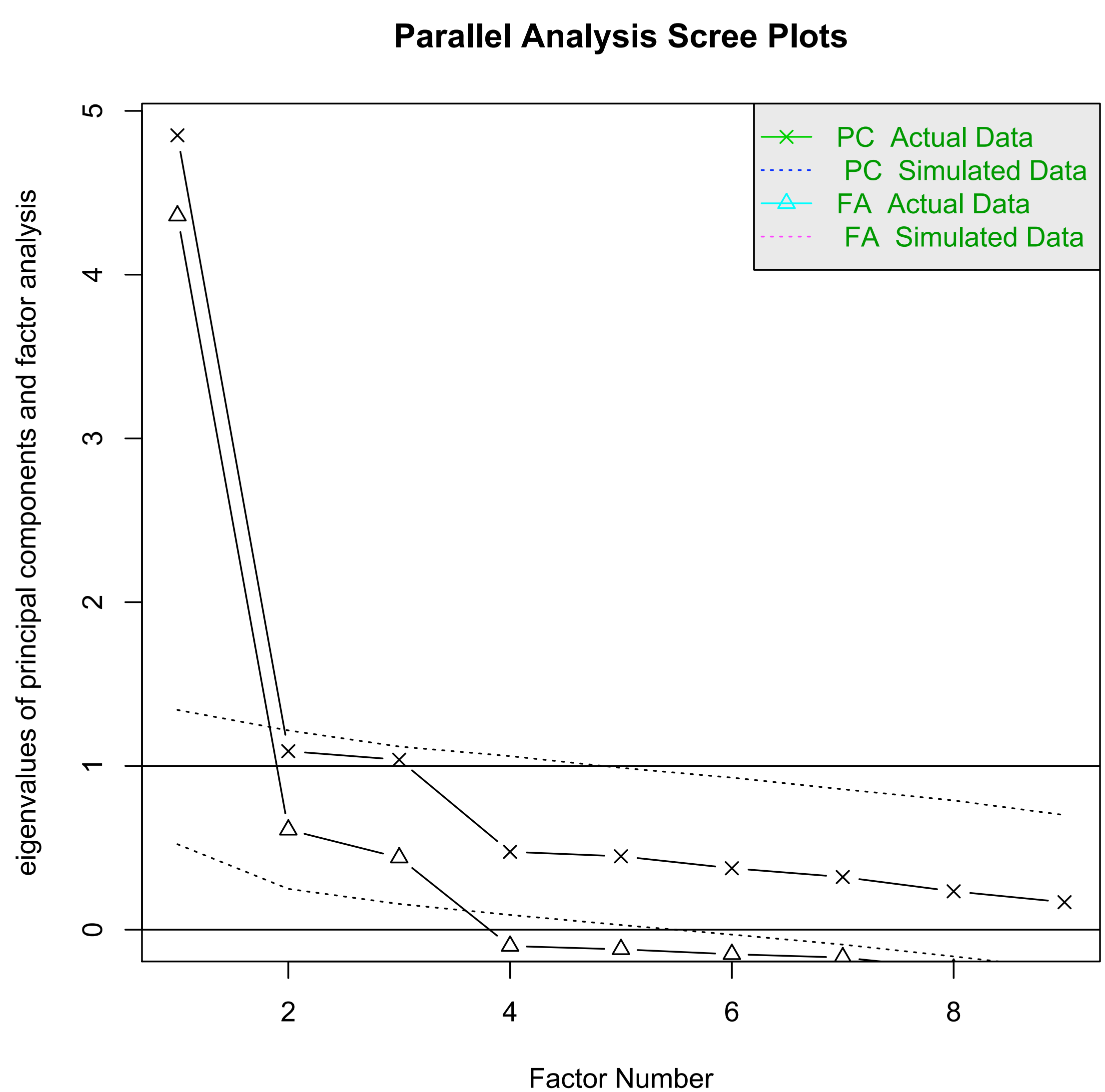
```
> factanal(covmat=Thurstone,factors=3,n.obs =213)
Call:
factanal(factors = 3, covmat = Thurstone, n.obs = 213)
Uniquenesses:
Sentences     Vocab    S.comp  F.letter  4.letter    Suffix    Series
Pedi   letters
    0.175     0.165     0.268     0.268     0.372     0.504     0.282
0.496     0.473
Loadings:
               Factor1 Factor2 Factor3
Sentences       0.834   0.244   0.264
Vocabulary      0.827   0.318   0.223
Sent.Completion 0.775   0.284   0.227
First.Letters   0.228   0.792   0.230
4.Letter.Words  0.213   0.706   0.291
Suffixes        0.314   0.616   0.134
Letter.Series   0.232   0.179   0.795
Pedigrees       0.446   0.166   0.527
Letter.Group    0.154   0.311   0.638
               Factor1 Factor2 Factor3
SS loadings      2.454   1.902   1.642
Proportion Var   0.273   0.211   0.182
Cumulative Var   0.273   0.484   0.666
Test of the hypothesis that 3 factors are sufficient.
The chi square statistic is 2.82 on 12 degrees of freedom.
The p-value is 0.997
```

MLE factor analyis: factanal

```
Call:
factanal(factors = 2, covmat = Thurstone, n.obs = 213)
Uniquenesses:
Sentences      Vocab     S.comp  F.letter  4.letter    Suffix    Series
Pedi   letters
    0.168      0.178      0.269      0.322      0.344      0.537      0.680
0.612     0.677
Loadings:
                Factor1 Factor2
Sentences        0.866   0.287
Vocabulary       0.839   0.343
Sent.Completion  0.795   0.314
First.Letters    0.255   0.783
4.Letter.Words   0.235   0.775
Suffixes         0.317   0.602
Letter.Series    0.372   0.426
Pedigrees        0.524   0.336
Letter.Group     0.269   0.501
               Factor1 Factor2
SS loadings      2.793   2.420
Proportion Var   0.310   0.269
Cumulative Var   0.310   0.579
Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 82.84 on 19 degrees of freedom.
The p-value is 5.99e-10
```

chi square rejects a 2 factor solution

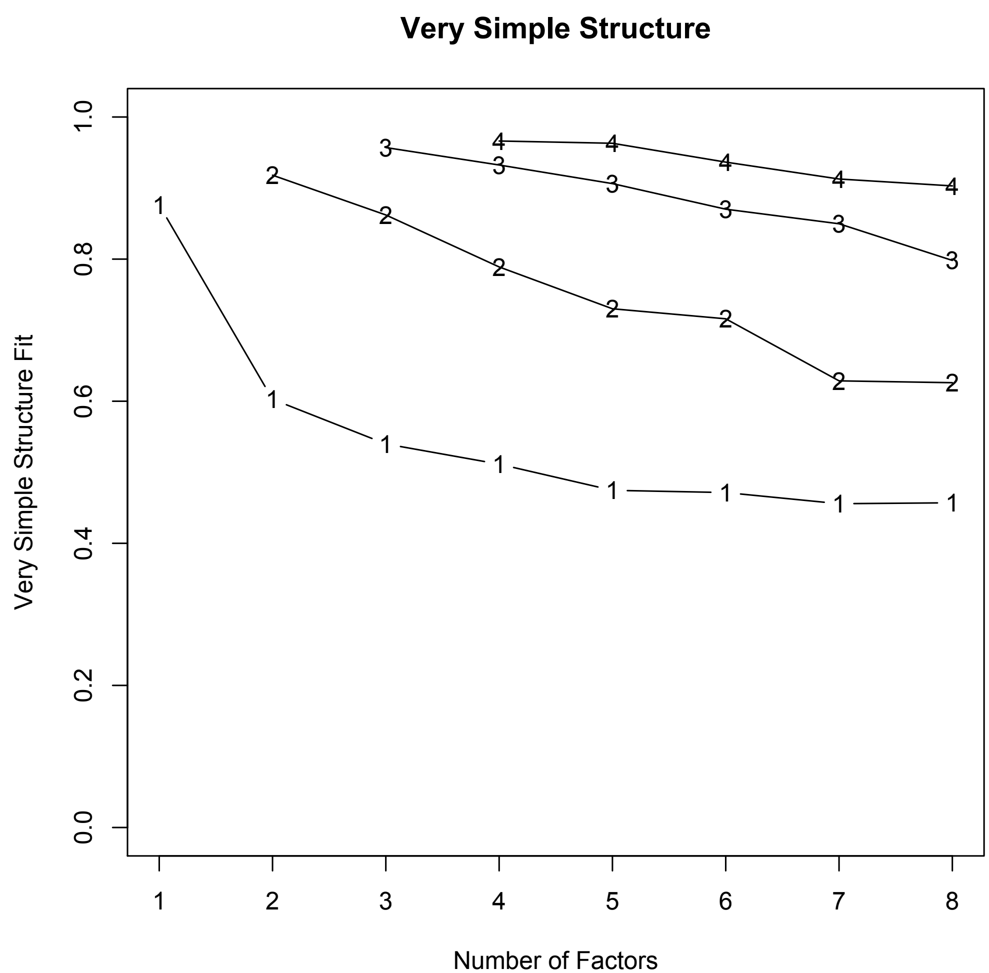
Parallel analysis



Very Simple Structure and MAP

```
> vss <- VSS(Thurstone,n.obs=213,SMC=FALSE)
> vss

Very Simple Structure
Call: VSS(x = Thurstone, n.obs = 213, SMC = FALSE)
VSS complexity 1 achieves a maximimum of 0.88  with  1  factors
VSS complexity 2 achieves a maximimum of 0.92  with  2  factors

The Velicer MAP criterion achieves a minimum of 1  with  3  factors

Velicer MAP
[1] 0.07 0.07 0.07 0.11 0.20 0.31 0.59 1.00

Very Simple Structure Complexity 1
[1] 0.88 0.60 0.54 0.51 0.47 0.47 0.46 0.46

Very Simple Structure Complexity 2
[1] 0.00 0.92 0.86 0.79 0.73 0.72 0.63 0.63
```

Very Simple Structure



Principal Axis FA

```
> pa3 <- factor.pa(Thurstone,nfactors=3,n.obs=213)
> pa3
          V  PA1  PA2  PA3
Sentences 1 0.83
Vocab     2 0.83 0.32
S.comp    3 0.78
F.letter  4      0.79
4.letter  5      0.71
Suffix    6 0.31 0.62
Series    7           0.79
Pedi      8 0.45      0.53
letters   9      0.31 0.64

                PA1  PA2  PA3
SS loadings    2.46 1.91 1.64
Proportion Var 0.27 0.21 0.18
Cumulative Var 0.27 0.49 0.67

Test of the hypothesis that 3 factors are sufficient.

The degrees of freedom for the model is 12  and the fit was  0.01
The number of observations was  213  with Chi Square =  2.97  with
prob <  1
```

Principal Components

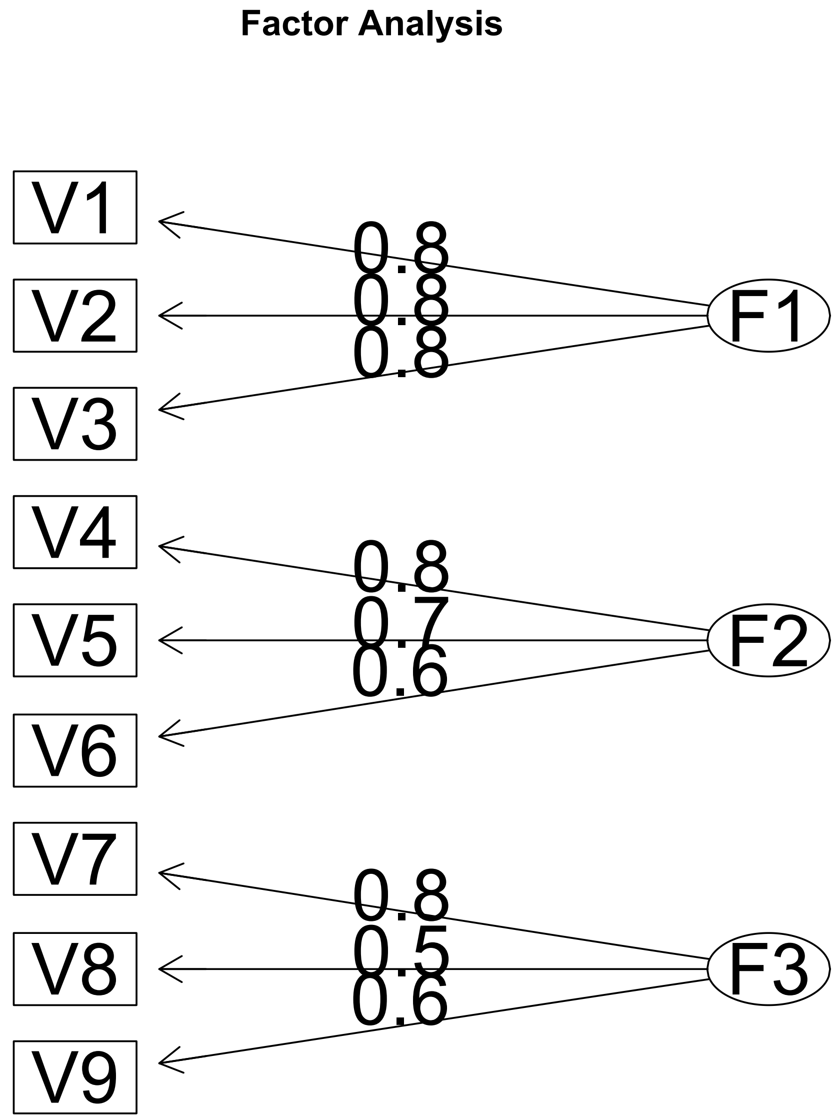
```
> pc3
                V   PC1   PC2   PC3
Sentences       1 0.863
Vocabulary      2 0.854  0.31
Sent.Completion 3 0.849
First.Letters   4        0.82
4.Letter.Words  5        0.79 0.301
Suffixes        6 0.314  0.77
Letter.Series   7             0.834
Pedigrees       8 0.534       0.613
Letter.Group    9        0.31 0.805

                PC1  PC2  PC3
SS loadings    2.73 2.25 1.99
Proportion Var 0.30 0.25 0.22
Cumulative Var 0.30 0.55 0.78

Test of the hypothesis that 3 factors are sufficient.

The degrees of freedom for the model is 12  and the fit was  0.62
The number of observations was  213  with Chi Square =  127.9  with
prob <  1.6e-21
```

Comparing solutions: factor congruence

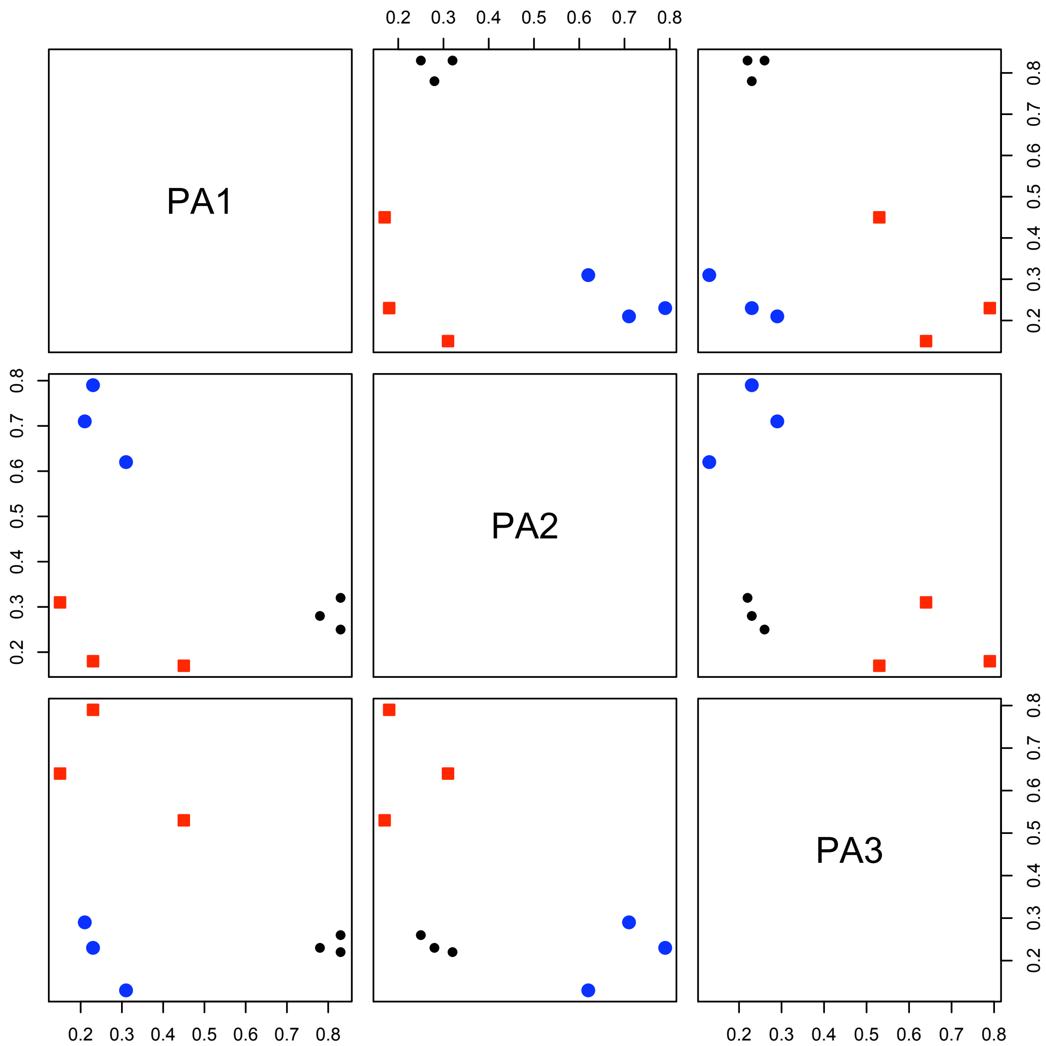
```
> factor.congruence(list(f3,pa3,pc3))
```

	Factor1	Factor2	Factor3	PA1	PA2	PA3	PC1	PC2	PC3
Factor1	1.00	0.64	0.62	1.00	0.64	0.62	1.00	0.59	0.55
Factor2	0.64	1.00	0.62	0.63	1.00	0.62	0.61	0.99	0.57
Factor3	0.62	0.62	1.00	0.62	0.62	1.00	0.61	0.56	0.99
PA1	1.00	0.63	0.62	1.00	0.64	0.62	1.00	0.59	0.55
PA2	0.64	1.00	0.62	0.64	1.00	0.62	0.61	0.99	0.57
PA3	0.62	0.62	1.00	0.62	0.62	1.00	0.61	0.56	0.99
PC1	1.00	0.61	0.61	1.00	0.61	0.61	1.00	0.56	0.54
PC2	0.59	0.99	0.56	0.59	0.99	0.56	0.56	1.00	0.50
PC3	0.55	0.57	0.99	0.55	0.57	0.99	0.54	0.50	1.00

A misleading graph



Plot the loadings: shows some cross loadings




```
plot(pa3)
```

Rotations and tranformations

I. Orthogonal rotations
   A. Varimax, Quartimax
II. Oblique transformations
   A. Promax, Quartimin, biquartimin, ...

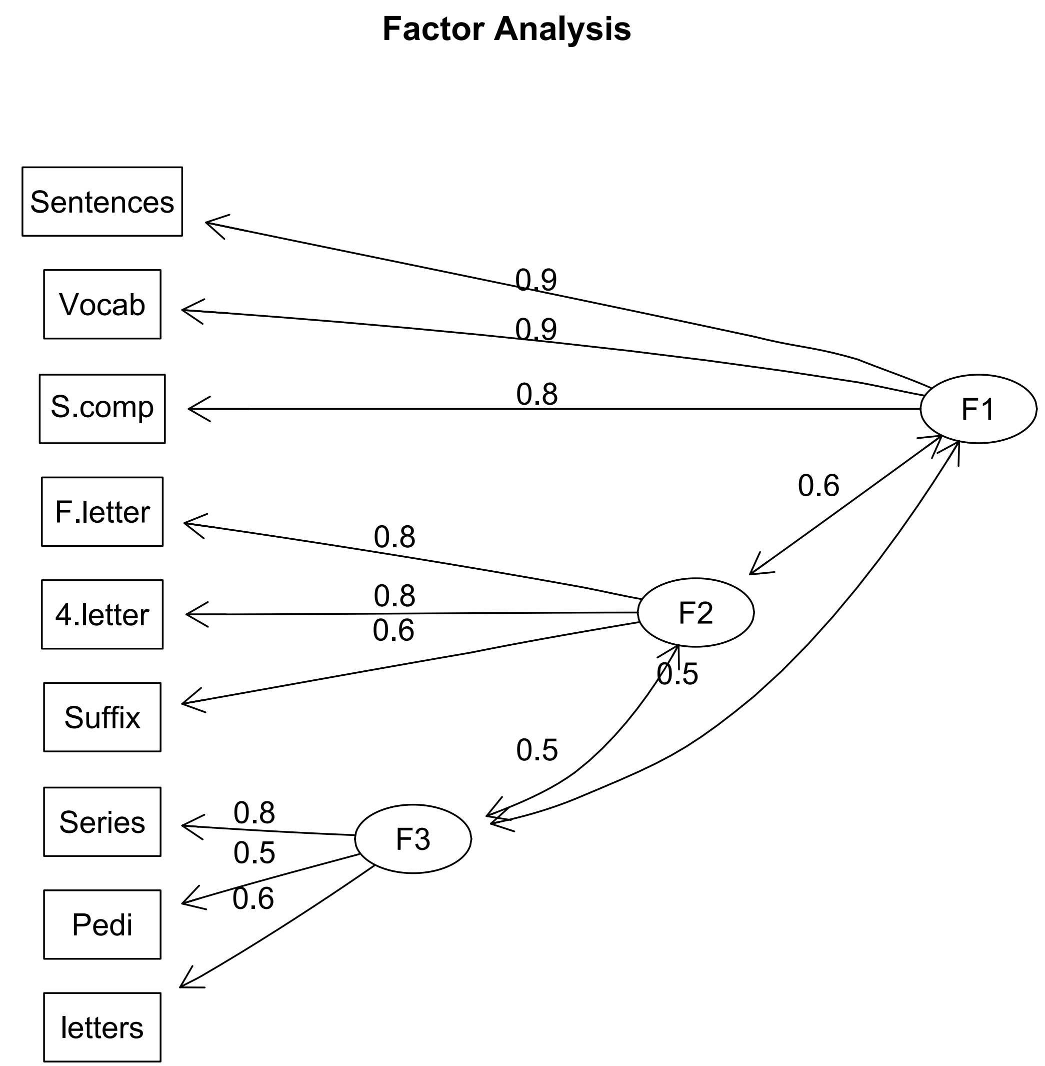
```
> pa3o <- factor.pa(Thurstone,3,rotate="oblimin",n.obs=213)
Loading required package: GPArotation
> pa3o
          V   PA1   PA2   PA3
Sentences 1  0.90
Vocab     2  0.89
S.comp    3  0.84
F.letter  4        0.85
4.letter  5        0.75
Suffix    6        0.63
Series    7              0.83
Pedi      8  0.38        0.47
letters   9              0.64
                PA1  PA2  PA3
SS loadings    2.49 1.74 1.34
Proportion Var 0.28 0.19 0.15
Cumulative Var 0.28 0.47 0.62
 With factor correlations of
     PA1  PA2  PA3
PA1 1.00 0.59 0.54
PA2 0.59 1.00 0.52
PA3 0.54 0.52 1.00
Test of the hypothesis that 3 factors are sufficient.
The degrees of freedom for the model is 12  and the fit was  0.01
The number of observations was  213  with Chi Square =  2.97  with
prob <  1
```

Oblique solution Oblimin

Oblique: Promax

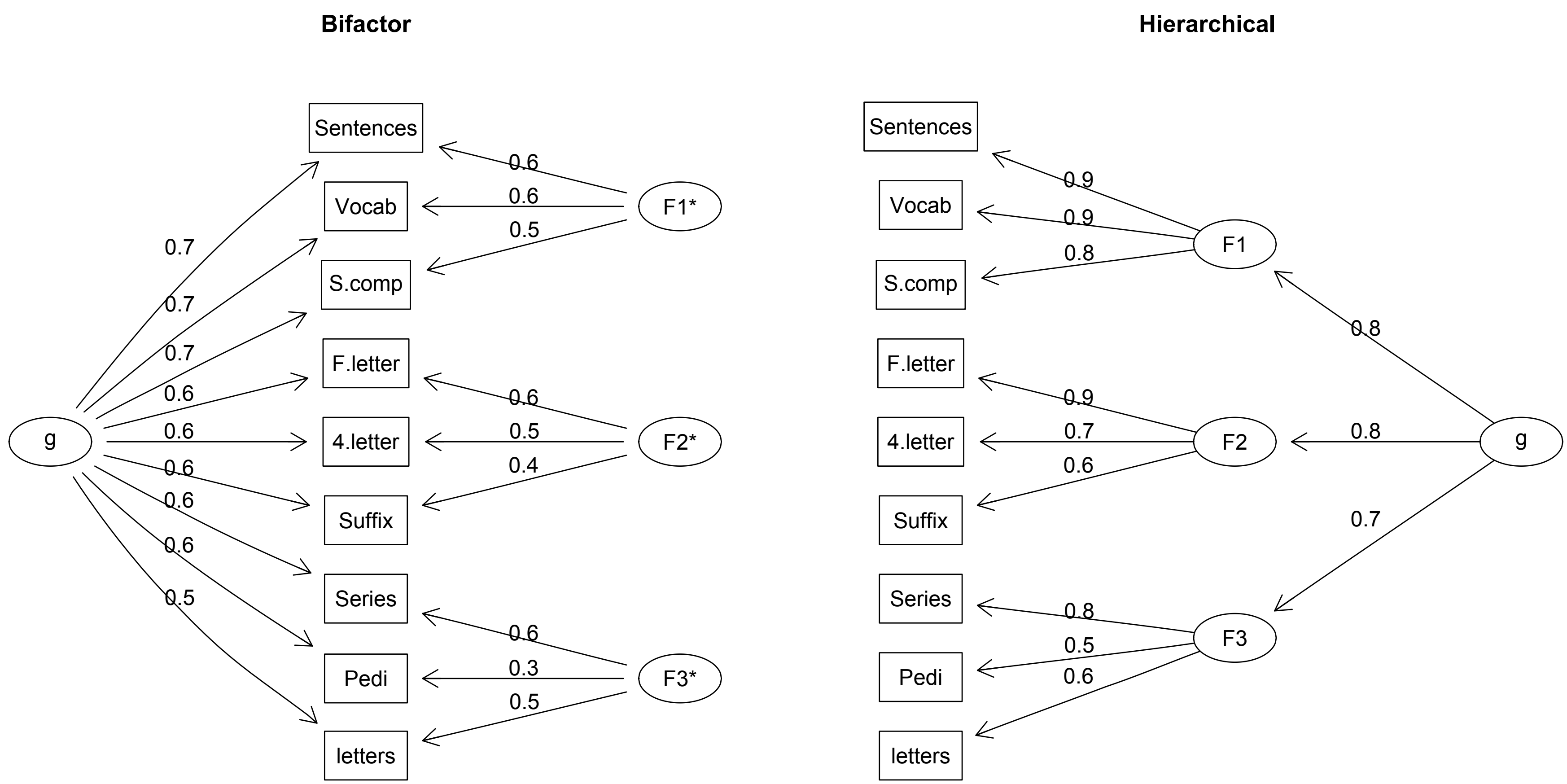
```
> pa3p <- Promax(pa3)
> pa3p
          V    PA1    PA2     PA3
Sentences 1  0.911
Vocab     2  0.904
S.comp    3  0.848
F.letter  4         0.869
4.letter  5         0.759
Suffix    6         0.650
Series    7                0.8865
Pedi      8  0.350         0.4969
letters   9                0.6800

                PA1  PA2  PA3
SS loadings    2.54 1.80 1.52
Proportion Var 0.28 0.20 0.17
Cumulative Var 0.28 0.48 0.65

 With factor correlations of
     PA1  PA2  PA3
PA1 1.00 0.61 0.61
PA2 0.61 1.00 0.58
PA3 0.61 0.58 1.00
```

A more accurate graphic




```
> fa.graph(pa3o,labels=colnames(Thurstone))
```

Hierarchical solutions




```
> omsl <-omega(Thurstone,title="Bifactor")
> omsl <-omega(Thurstone,sl=FALSE,title="Hierarchical")
```

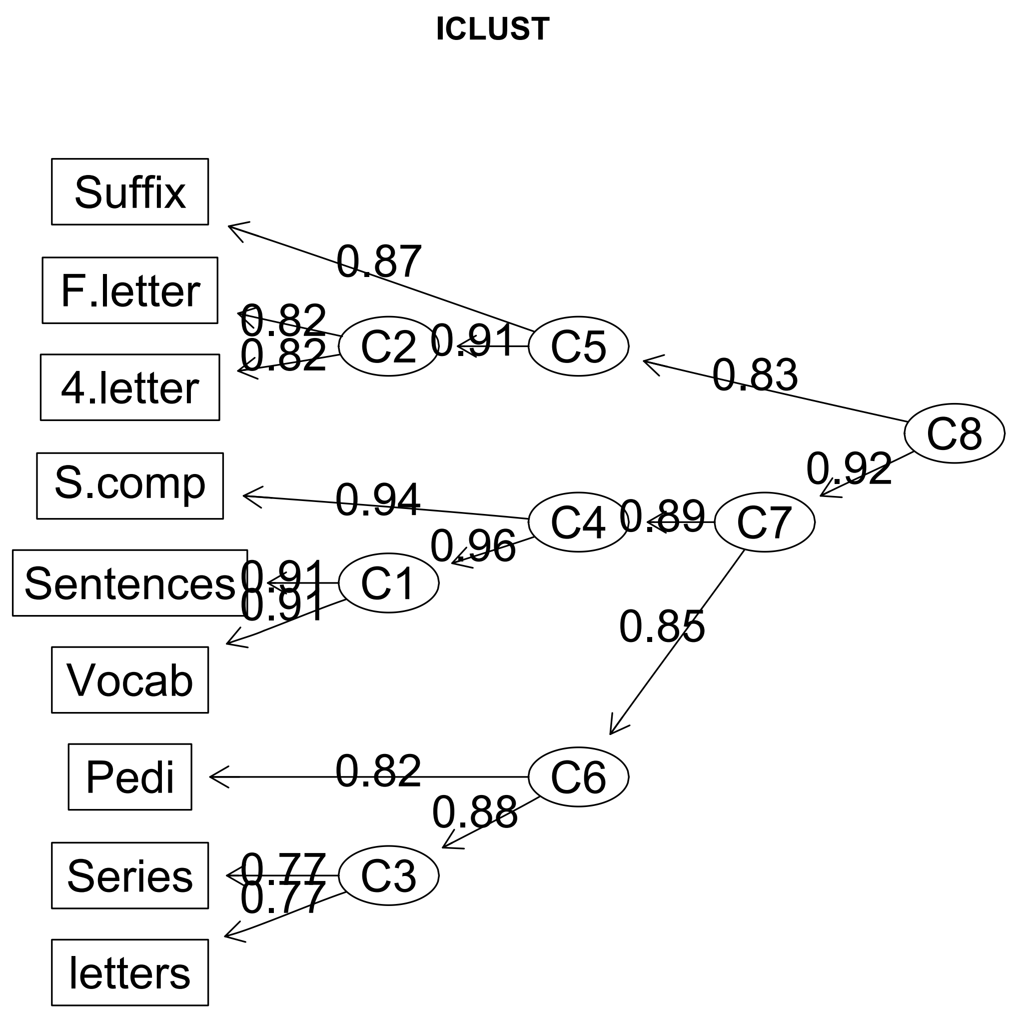
Hierarchical Clustering

I. Find the similarity matrix (correlations)

II. Find the most similar pair of items/tests

III. Combine them and repeat II and III until some criterion (alpha, beta) fails to increase

ICLUST of Thurstone




```
> ic <- ICLUST(Thurstone)
```

Multidimensional scaling

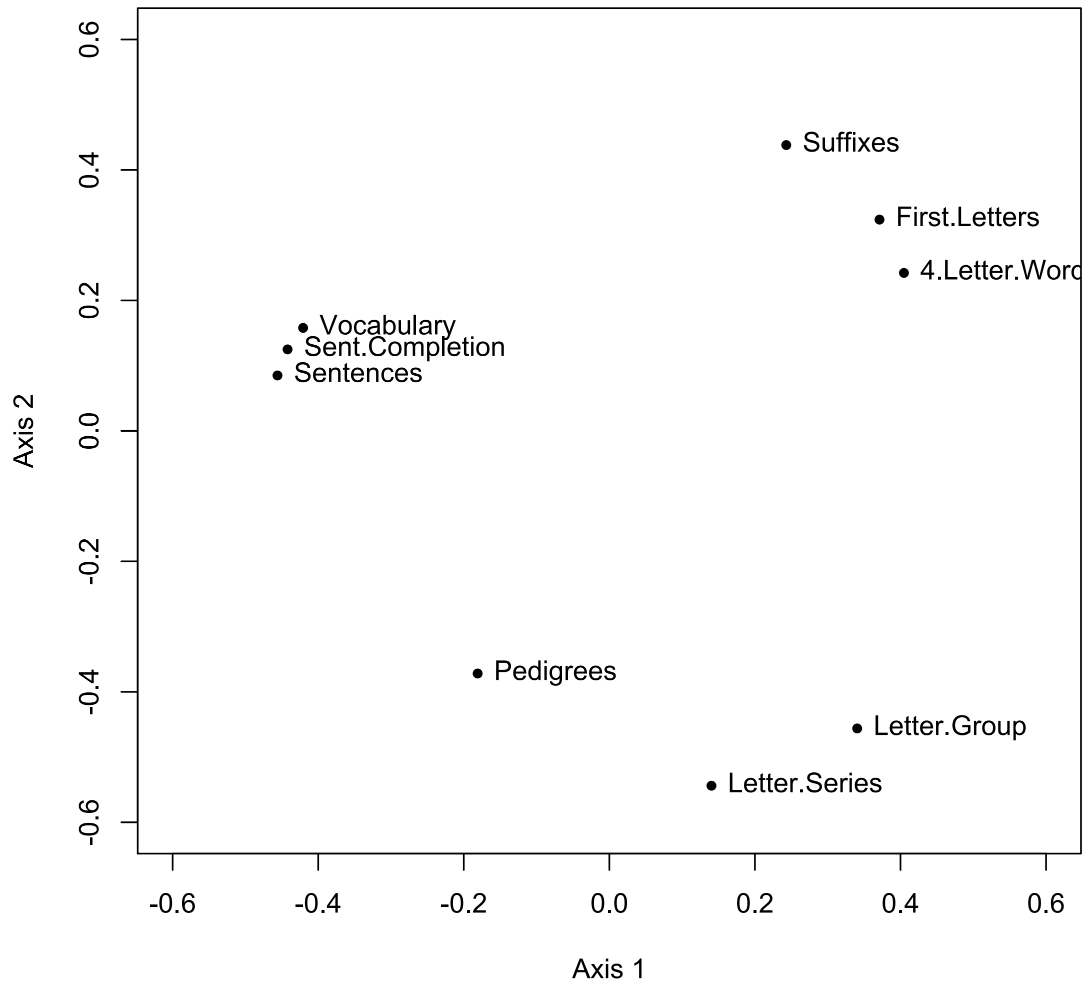
I. Convert correlations to distances

II. Multidimensional scaling will remove a general factor since it considers relative ranking of distances

MDS

```
> Thurs.dist <- sqrt(2*(1-Thurstone))
> mdsT <- cmdscale(Thurs.dist,2)
> round(mdsT,2)
                 [,1]  [,2]
Sentences       -0.46  0.08
Vocabulary      -0.42  0.16
Sent.Completion -0.44  0.12
First.Letters    0.37  0.32
4.Letter.Words   0.40  0.24
Suffixes         0.24  0.44
Letter.Series    0.14 -0.54
Pedigrees       -0.18 -0.37
Letter.Group     0.34 -0.46
```

MDS plot



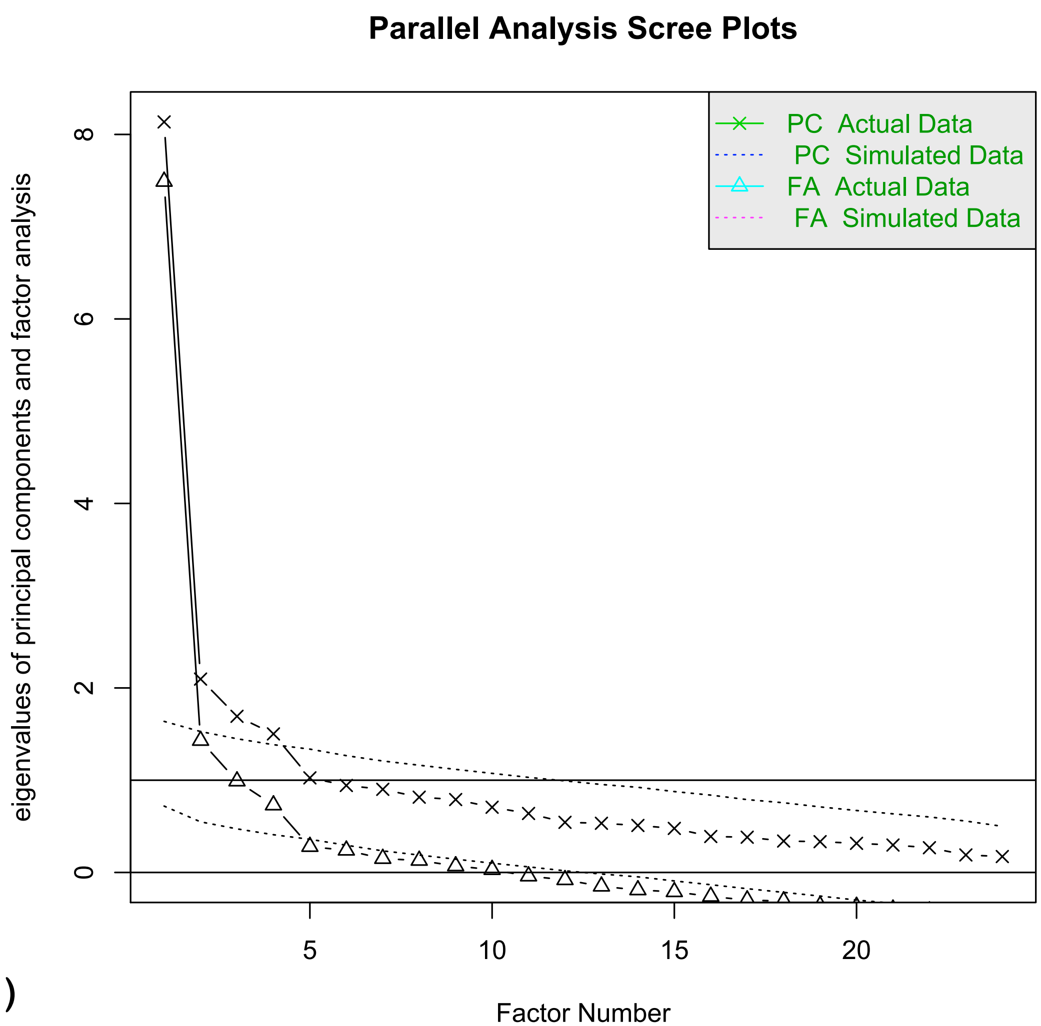

```
position <- rep(4,9)
plot(mdsT,xlim=c(-.6,.6),ylim=c(-.6,.6),ylab="Axis 2",xlab="Axis 1")
text(mdsT,rownames(mdsT),pos=position)
```

A more complex example

I. Holzinger-Harman 24 mental ability tests

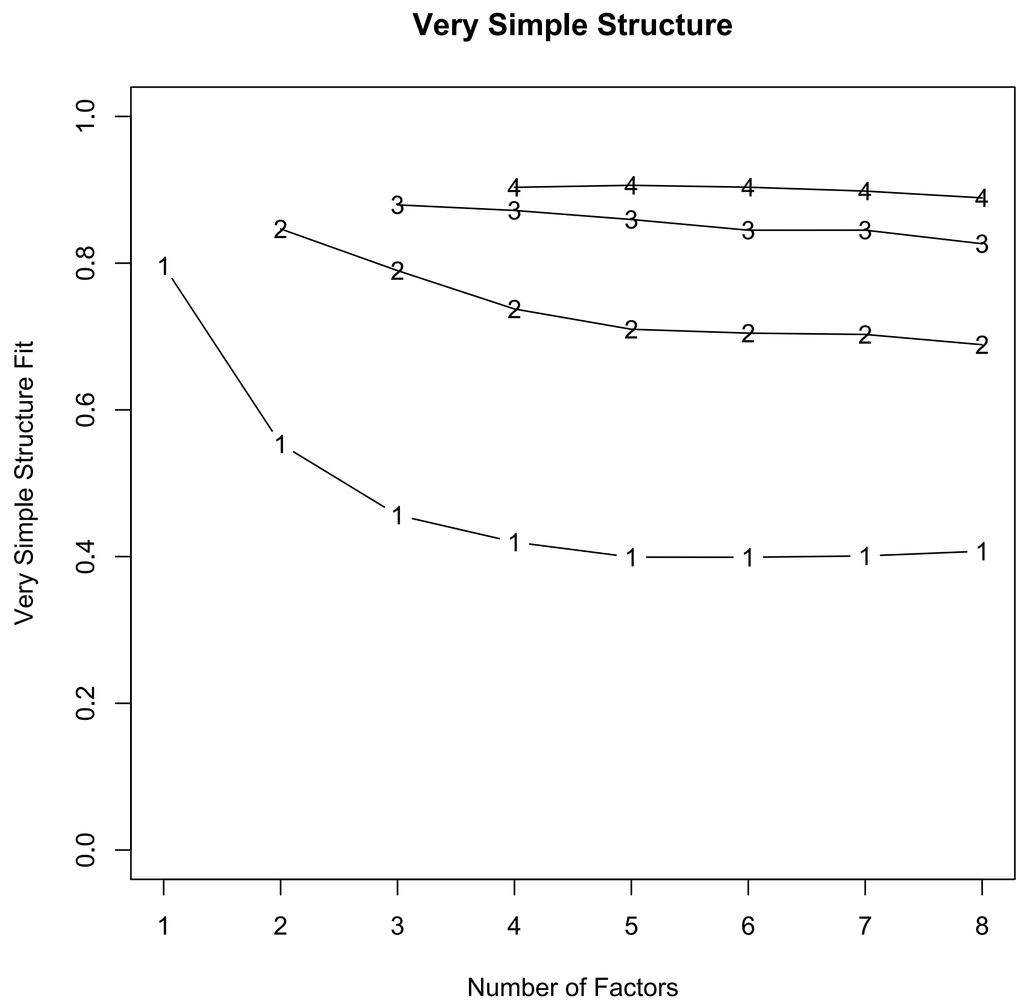
II. Compare FA, CA, MDS

Parallel analysis




```
> hh <- Harman74.cor$cov
> fa.parallel(hh,n.obs=228)
```

VSS says 1 big factor




```
> vss <- VSS(hh,n.obs=228)
```

MAP suggests 4

```
> vss

Very Simple Structure
Call: VSS(x = hh, n.obs = 228)
VSS complexity 1 achieves a maximimum of 0.8  with  1 
factors
VSS complexity 2 achieves a maximimum of 0.85  with  2 
factors

The Velicer MAP criterion achieves a minimum of 0.03  with 
4  factors

Velicer MAP
[1] 0.02 0.02 0.02 0.02 0.02 0.02 0.03 0.03

Very Simple Structure Complexity 1
[1] 0.80 0.55 0.46 0.42 0.40 0.40 0.40 0.41

Very Simple Structure Complexity 2
[1] 0.00 0.85 0.79 0.74 0.71 0.70 0.70 0.69
```

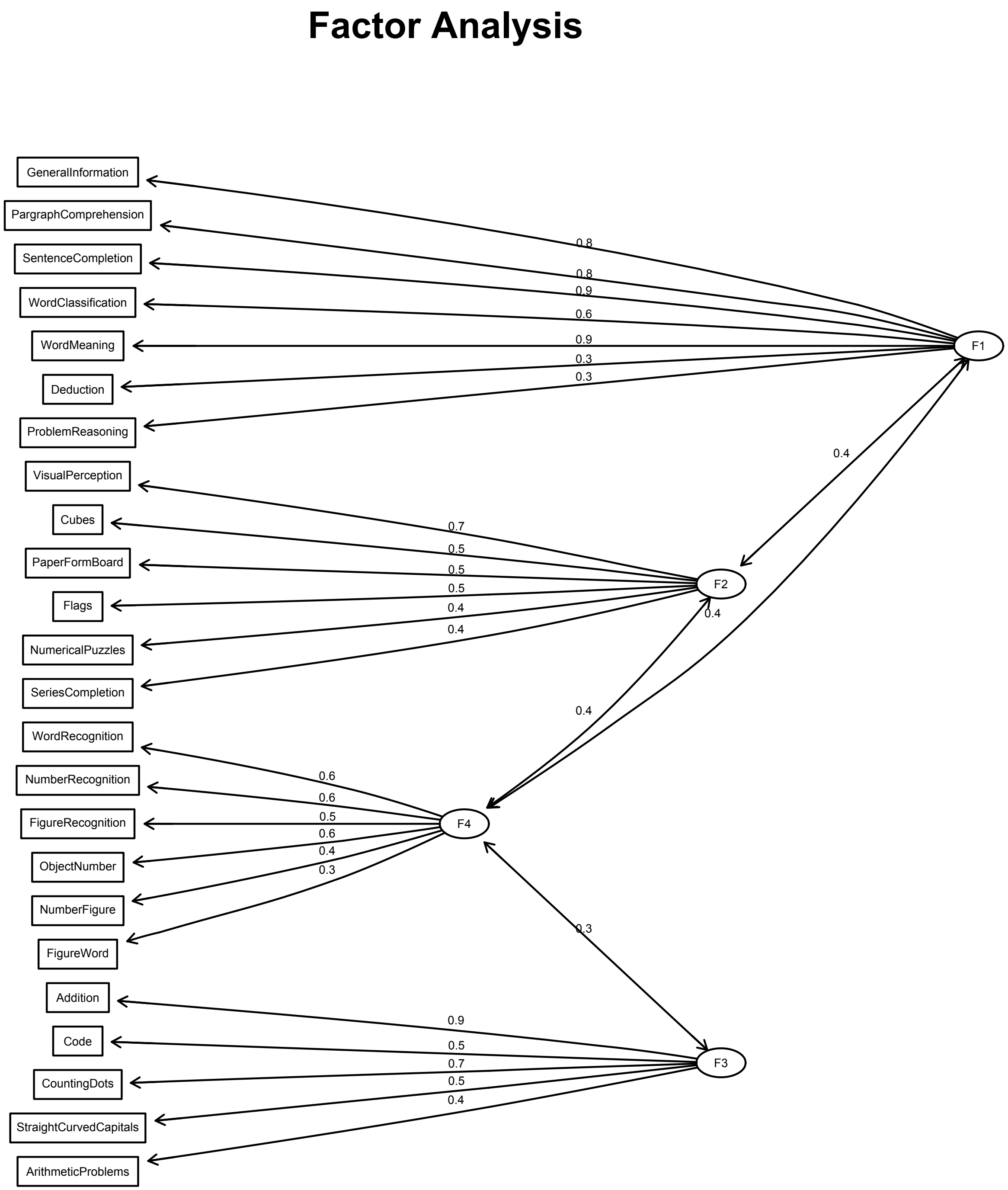
Chi square says > 8

```
> vss$vss.stats[,1:3]
  dof     chisq         prob
1 252 1012.9000 9.259262e-92
2 229  693.9911 2.649458e-48
3 207  485.1213 2.232607e-24
4 186  370.8825 2.174718e-14
5 166  307.7273 1.479600e-10
6 147  266.8324 5.628353e-09
7 129  226.6602 2.396296e-07
8 112  182.6407 2.832707e-05
>
```

4 oblique factors

```
> fa4o <- factor.pa(hh,4,n.obs=228,rotate="oblimin")
> fa4o
                          V   PA1   PA3   PA2   PA4
VisualPerception          1         0.69
Cubes                     2         0.46
PaperFormBoard            3         0.54
Flags                     4         0.52
GeneralInformation        5  0.76
PargraphComprehension     6  0.80
SentenceCompletion        7  0.87
WordClassification        8  0.56
WordMeaning               9  0.86
Addition                 10               0.86
Code                     11               0.49  0.30
CountingDots             12               0.70
StraightCurvedCapitals   13         0.42  0.47
WordRecognition          14                     0.58
NumberRecognition        15                     0.55
FigureRecognition        16         0.33        0.52
ObjectNumber             17                     0.59
NumberFigure             18                     0.43
FigureWord               19                     0.32
Deduction                20  0.33  0.31
NumericalPuzzles         21        0.37  0.33
ProblemReasoning         22  0.31  0.30
SeriesCompletion         23  0.30  0.44
ArithmeticProblems       24              0.41
```

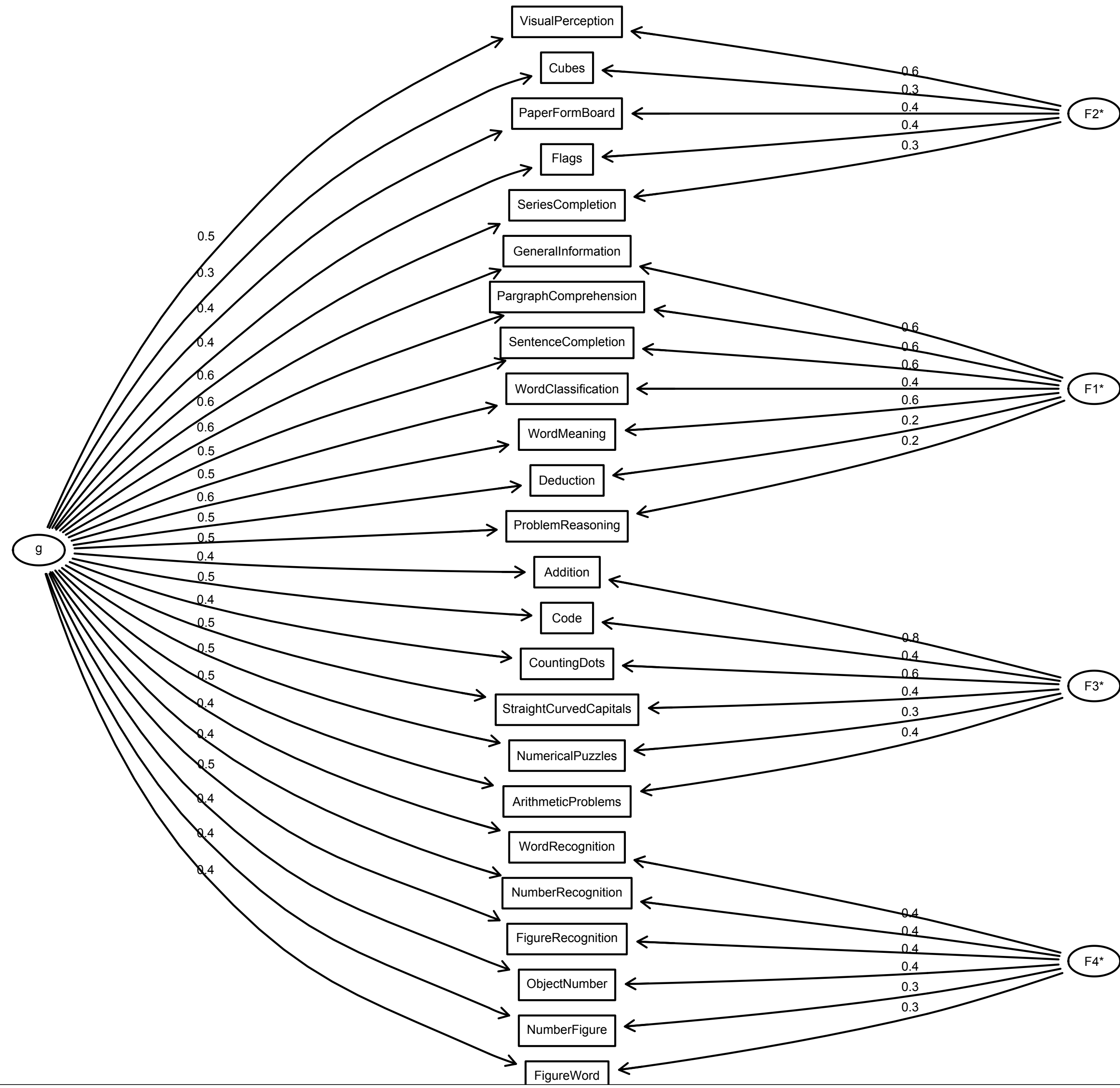
```
 With factor correlations
     PA1  PA3  PA2  PA4
PA1 1.00 0.41 0.30 0.40
PA3 0.41 1.00 0.27 0.38
PA2 0.30 0.27 1.00 0.32
```

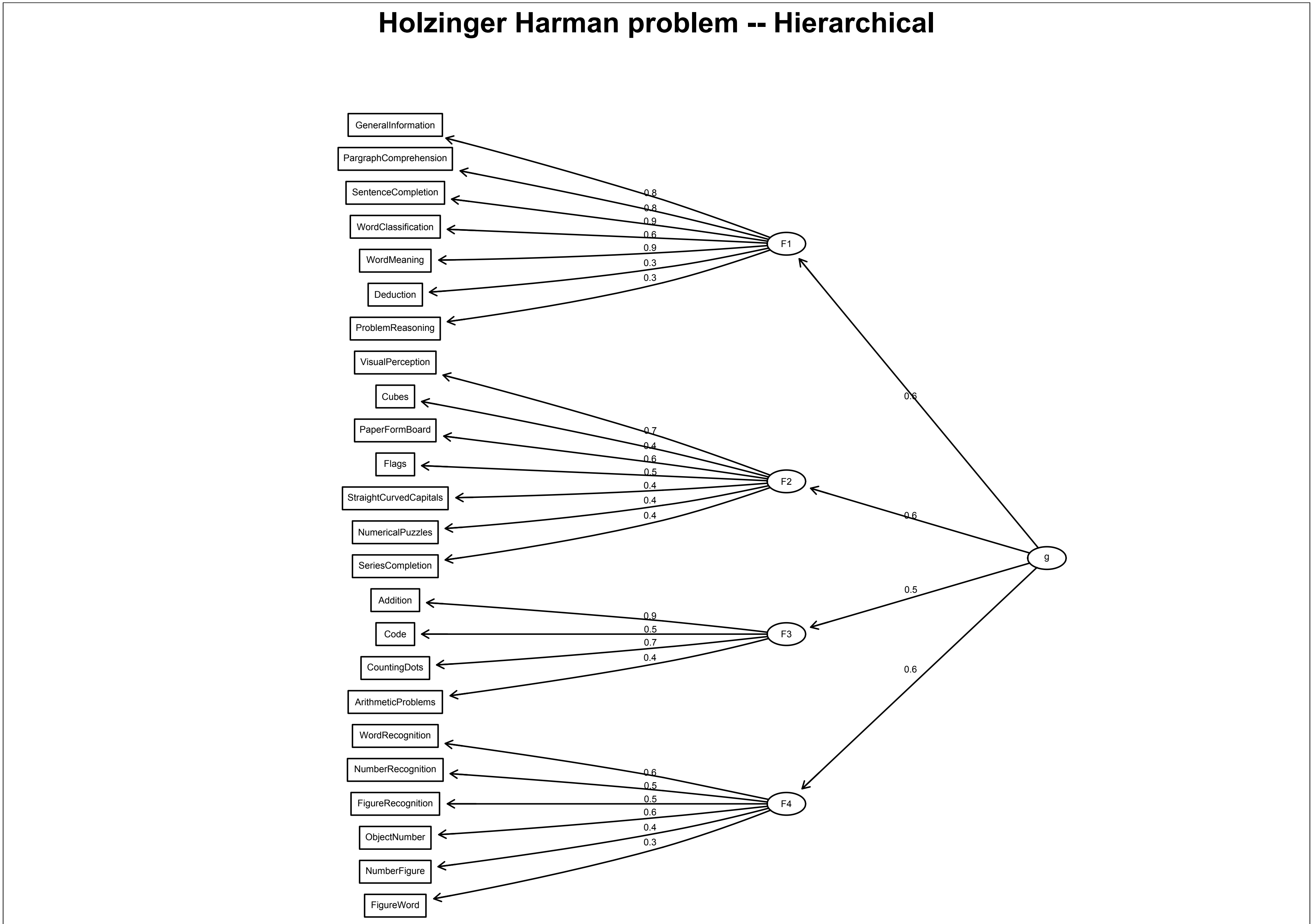


4 correlated factors

Holzinger Harman problem -- Bifactor

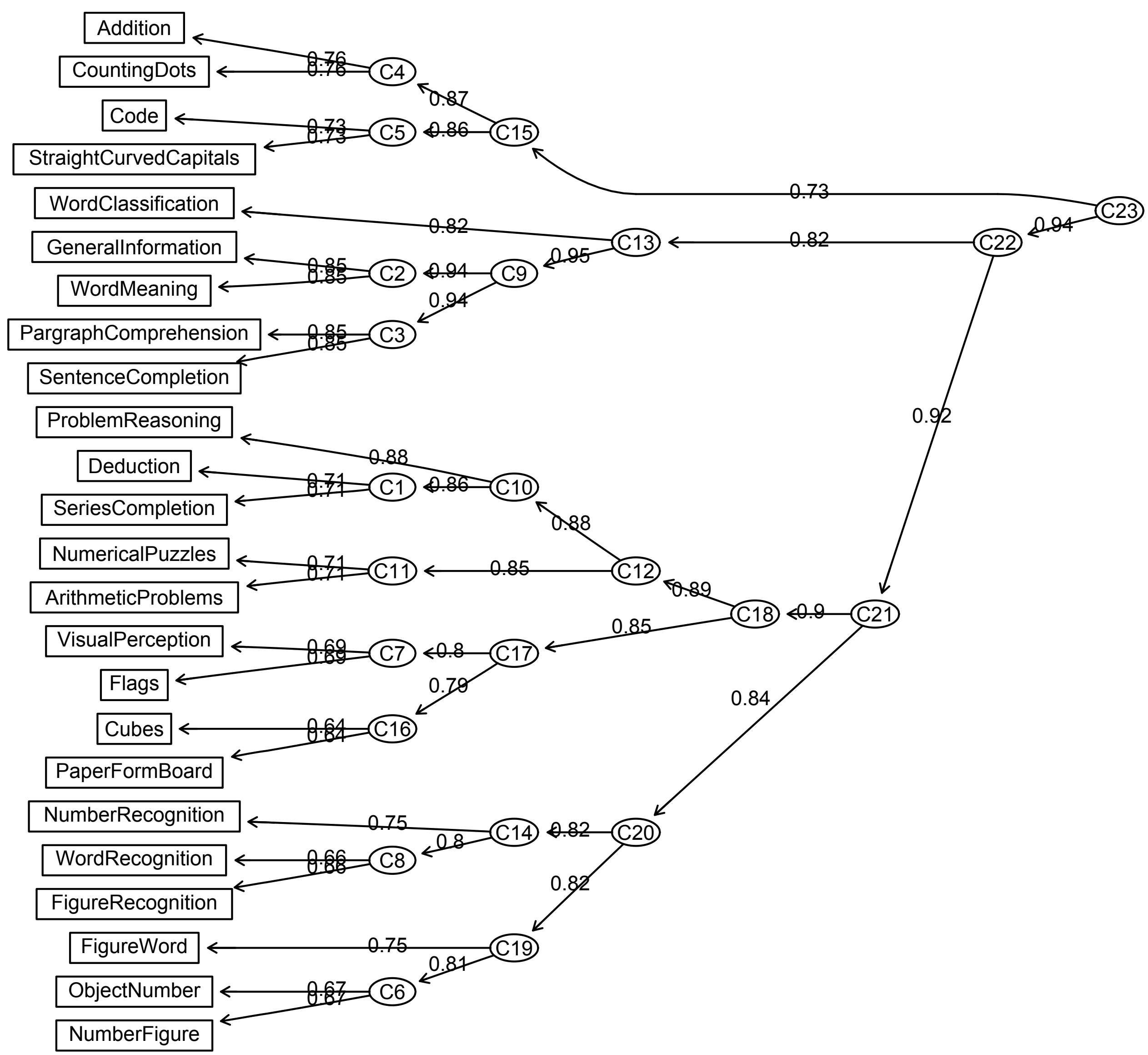
```
om4 <-omega(hh,4,title="Holzinger Harman problem -- Bifactor")
```





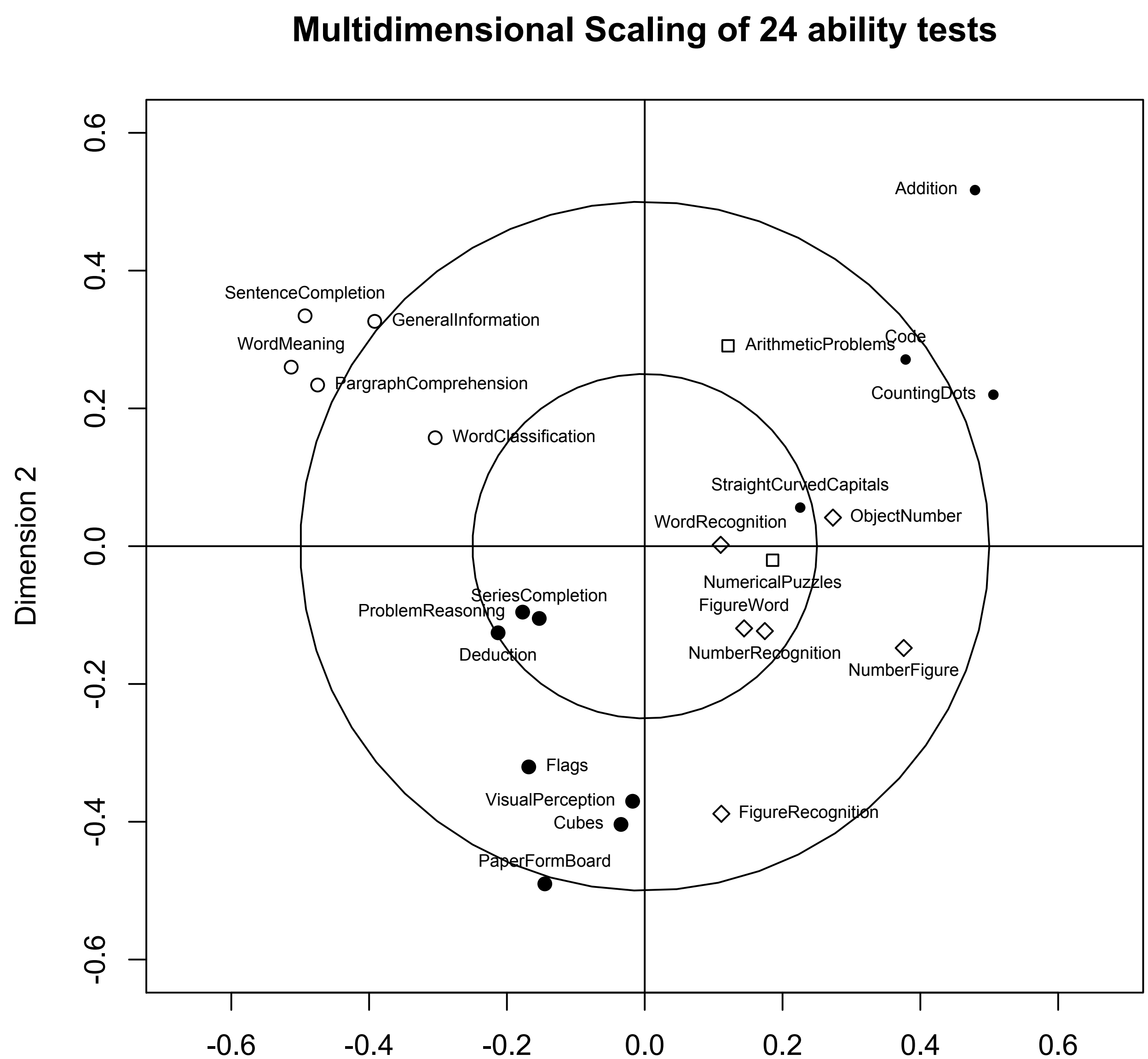

```
> om4 <-omega(hh,4,title="Holzinger Harman problem -- Hierarchical",sl=FALSE)
```

ICLUST of Holzinger Harman 24 mental measurements




```
ic4 <- ICLUST(hh,title="ICLUST of Holzinger Harman 24 mental measurements")
```

MDS of HH problem



code for MDS plot

```
> dis24 <- sqrt(2*(1-Harman74.cor$cov))
>  mds24 <- cmdscale(dis24,2)
>  plot.char <- c( 19, 19, 19, 19,   21, 21,  21, 21,   21,
20, 20, 20,
+     20, 23, 23, 23,   23, 23, 23, 19,   22, 19, 19, 22 )
>  plot(mds24,xlim=c(-.6,.6),ylim=c(-.6,.6),xlab="Dimension
1",ylab="Dimension 2",asp=1,pch=plot.char)
>  position <- c(2,2,3,4,  4,4,3,4,  3,2,3,2, 3,3,1,4,
4,1,3,1,  1,2,3,4)
>   text(mds24,rownames(mds24),cex=.6,pos=position)
>  abline(v=0,h=0)
>  title("Multidimensional Scaling of 24 ability tests")
>  #draw circles at .25 and .50 units away from the center
>   segments = 51
>   angles <- (0:segments) * 2 * pi/segments
>          unit.circle <- cbind(cos(angles), sin(angles))
>     lines(unit.circle*.25)
>     lines(unit.circle*.5)
```

Reliability of scales and raters

I. First consider the reliability of raters using the IntraClass Correlation (see Shrout and Fleiss for the definitive discussion)

II. Types of reliability of ratings (1 or n per target)

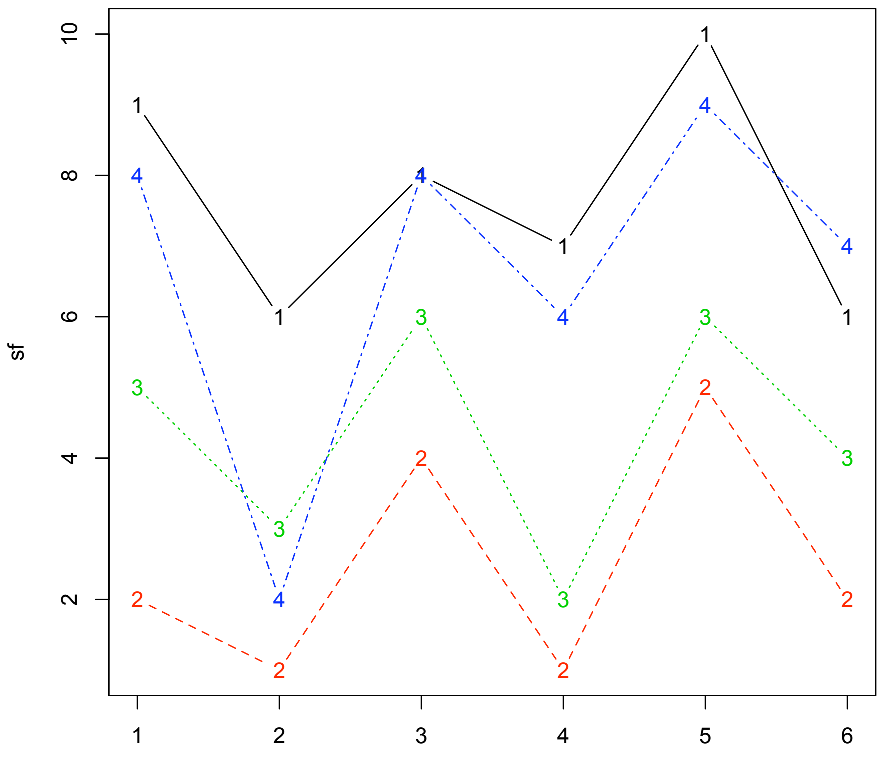
   A. each target rated by a different judge, judges are random

   B. random sample of k judges rate targets

   C. Fixed set of k judges give ratings

4 judges rate 6 subjects

```
> sf
   J1 J2 J3 J4
S1  9  2  5  8
S2  6  1  3  2
S3  8  4  6  8
S4  7  1  2  6
S5 10  5  6  9
S6  6  2  4  7
```



Simple correlations (these will remove means for raters)

```
> round(cor(sf),2)
     J1   J2   J3   J4
J1 1.00 0.75 0.73 0.75
J2 0.75 1.00 0.89 0.73
J3 0.73 0.89 1.00 0.72
J4 0.75 0.73 0.72 1.00
```

ICC

```
> ICC(sf)
                            type  ICC     F df1 df2    p lower bound upper bound
Single_raters_absolute   ICC1 0.17  1.79   5  18 0.16       -0.13        0.72
Single_random_raters     ICC2 0.29 11.03   5  15 0.00        0.02        0.76
Single_fixed_raters      ICC3 0.71 11.03   5  15 0.00        0.34        0.95
Average_raters_absolute ICC1k 0.44  1.79   5  18 0.16       -0.88        0.91
Average_random_raters   ICC2k 0.62 11.03   5  15 0.00        0.07        0.93
Average_fixed_raters    ICC3k 0.91 11.03   5  15 0.00        0.68        0.99
```

Alpha of raters

```
> alpha(sf)

Reliability analysis
Call: alpha(x = sf)

  raw_alpha std.alpha G6(smc) average_r mean  sd
      0.91      0.93    0.92      0.76   21 6.7

 Reliability if an item is dropped:
   raw_alpha std.alpha G6(smc) average_r
J1      0.88      0.91    0.89      0.78
J2      0.87      0.89    0.85      0.73
J3      0.87      0.90    0.85      0.74
J4      0.92      0.92    0.90      0.79
```

```
Item statistics
   n    r r.cor mean  sd
J1 6 0.89  0.83  7.7 1.6
J2 6 0.93  0.92  2.5 1.6
J3 6 0.92  0.91  4.3 1.6
J4 6 0.88  0.82  6.7 2.5
```

Reliability of a single scale

```
> round(cor(bfi[,1:10],use="pairwise"),2)
      A1    A2    A3    A4    A5    C1    C2    C3    C4    C5
A1  1.00 -0.30 -0.23 -0.12 -0.19 -0.03 -0.06  0.01  0.19  0.08
A2 -0.30  1.00  0.39  0.24  0.41  0.06  0.06  0.22 -0.15 -0.12
A3 -0.23  0.39  1.00  0.27  0.45  0.07  0.12  0.16 -0.14 -0.12
A4 -0.12  0.24  0.27  1.00  0.22  0.07  0.17  0.08 -0.15 -0.17
A5 -0.19  0.41  0.45  0.22  1.00  0.10  0.06  0.20 -0.14 -0.10
C1 -0.03  0.06  0.07  0.07  0.10  1.00  0.44  0.41 -0.39 -0.23
C2 -0.06  0.06  0.12  0.17  0.06  0.44  1.00  0.35 -0.36 -0.24
C3  0.01  0.22  0.16  0.08  0.20  0.41  0.35  1.00 -0.37 -0.36
C4  0.19 -0.15 -0.14 -0.15 -0.14 -0.39 -0.36 -0.37  1.00  0.53
C5  0.08 -0.12 -0.12 -0.17 -0.10 -0.23 -0.24 -0.36  0.53  1.00
```

Mindless reliability

```
> alpha(bfi[1:10])

Reliability analysis
Call: alpha(x = bfi[1:10])

  raw_alpha std.alpha G6(smc) average_r mean  sd
      0.19      0.25    0.44     0.032   40 4.7

 Reliability if an item is dropped:       Item statistics
   raw_alpha std.alpha G6(smc) average_r         n     r r.cor mean  sd
A1     0.290     0.354    0.51     0.057 A1 1000 0.099 -0.21  2.3 1.3
A2     0.082     0.133    0.35     0.017 A2  994 0.508  0.47  4.8 1.1
A3     0.045     0.101    0.33     0.012 A3  989 0.552  0.54  4.6 1.2
A4     0.108     0.173    0.40     0.023 A4  993 0.448  0.31  4.8 1.4
A5     0.042     0.093    0.32     0.011 A5  988 0.563  0.55  4.6 1.2
C1     0.133     0.190    0.38     0.025 C1  997 0.421  0.34  4.4 1.2
C2     0.123     0.182    0.38     0.024 C2  997 0.434  0.35  4.2 1.3
C3     0.105     0.154    0.36     0.020 C3  995 0.478  0.44  4.3 1.3
C4     0.336     0.392    0.50     0.067 C4  986 0.005 -0.24  2.6 1.4
C5     0.329     0.364    0.49     0.060 C5  997 0.077 -0.17  3.5 1.5
```

somewhat better reliability

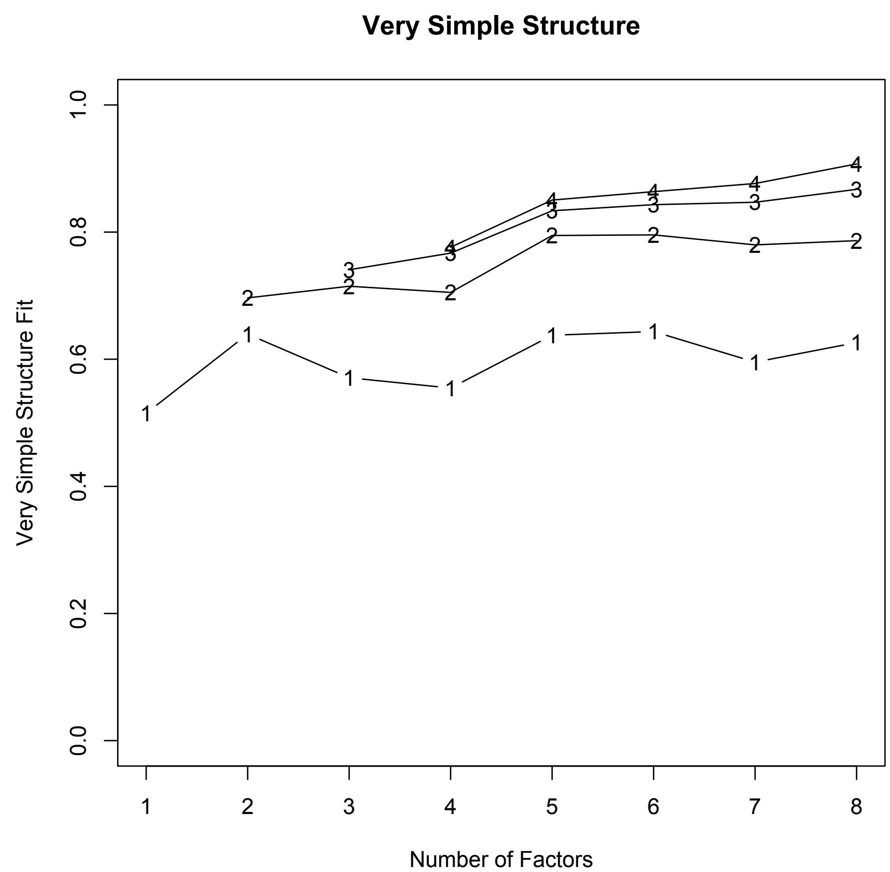
```
> keys <- make.keys(10,list(all=c(-1,2:8,-9,-10)))
> alpha(bfi[1:10],keys)

Reliability analysis
Call: alpha(x = bfi[1:10], keys = keys)

  raw_alpha std.alpha G6(smc) average_r mean  sd
      0.72      0.72    0.75      0.21   40 4.7

 Reliability if an item is dropped:
   raw_alpha std.alpha G6(smc) average_r
A1      0.72      0.72    0.74      0.23
A2      0.70      0.70    0.72      0.20
A3      0.70      0.70    0.72      0.20
A4      0.71      0.71    0.74      0.22
A5      0.70      0.70    0.72      0.21
C1      0.70      0.70    0.72      0.21
C2      0.70      0.70    0.72      0.21
C3      0.69      0.69    0.71      0.20
C4      0.67      0.68    0.70      0.19
C5      0.70      0.70    0.72      0.20

 Item statistics
      n    r r.cor mean  sd
A1 1000 0.41  0.29  2.3 1.3
A2  994 0.55  0.49  4.8 1.1
A3  989 0.55  0.49  4.6 1.2
A4  993 0.46  0.35  4.8 1.4
A5  988 0.53  0.46  4.6 1.2
C1  997 0.52  0.46  4.4 1.2
C2  997 0.54  0.47  4.2 1.3
C3  995 0.59  0.54  4.3 1.3
C4  986 0.64  0.61  2.6 1.4
C5  997 0.55  0.49  3.5 1.5
```

Examine the items

VSS suggests 2 factors!

The items

A1
Am indifferent to the feelings of others.
A2
Inquire about others' well-being.
A3
Know how to comfort others.
A4
Love children.
A5
Make people feel at ease.
C1
Am exacting in my work.
C2
Continue until everything is perfect.
C3
Do things according to a plan.
C4
Do things in a half-way manner.
C5
Waste my time.



Omega reliability

```
> om2 <- omega(bfi[1:10],2)
Warning messages:
1: In schmid(m, nfactors, pc, digits, rotate = rotate, n.obs =
n.obs,  :
  Three factors are required for identification -- general factor
loadings set to be equal. Proceed with caution.
2: In schmid(m, nfactors, pc, digits = digits, n.obs = n.obs, ...) :
  Three factors are required for identification -- general factor
loadings set to be equal. Proceed with caution.
```

```
> om2
Omega
Call: omega(m = bfi[1:10], nfactors = 2)
Alpha:  0.72
G.6:   0.75
Omega Hierarchical:   0.36
Omega Total   0.77

Schmid Leiman Factor loadings greater than  0.2
    g    F1*  F2*  h2   u2
A1- 0.21      0.30      0.87
A2  0.36      0.53 0.41 0.59
A3  0.36      0.55 0.43 0.57
A4  0.24      0.28      0.86
A5  0.36      0.54 0.42 0.58
C1  0.30 0.51      0.36 0.64
C2  0.30 0.48      0.32 0.68
C3  0.37 0.48      0.37 0.63
C4- 0.40 0.58      0.50 0.50
C5- 0.33 0.48      0.33 0.67

With eigenvalues of:
  g F1* F2*
1.1 1.3 1.1
```

Omega h is low

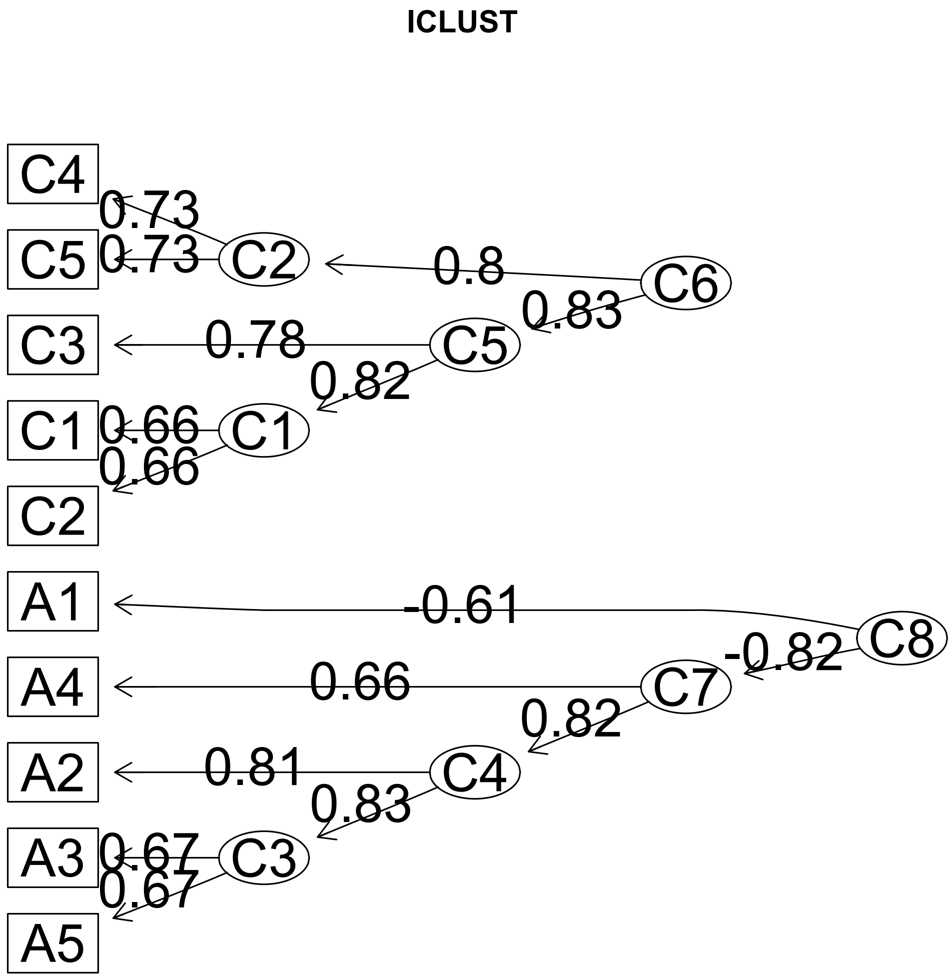
Score 3 scales

```
> keys <-
make.keys(10,list(all=c(-1,2:8,-9,-10),agree=c(-1,2:5),con=c(6:8,-9,-1
0)))
> score.items(keys,bfi[1:10])
Call: score.items(keys = keys, items = bfi[1:10])

(Unstandardized) Alpha:
       all agree  con
alpha 0.72  0.65 0.74

Average item correlation:
          all agree  con
average.r 0.2  0.27 0.36

 Guttman 6* reliability:
          all agree  con
Lambda.6 0.74  0.62 0.72

Scale intercorrelations corrected for attenuation
 raw correlations below the diagonal, alpha on the diagonal
 corrected correlations above the diagonal:
       all agree  con
all   0.72  1.10 1.13
agree 0.75  0.65 0.36
con   0.83  0.25 0.74
```

Item scale correlations

```
Item by scale correlations:
 corrected for item overlap and scale reliability
     all agree   con
A1 -0.28 -0.37 -0.12
A2  0.47  0.63  0.21
A3  0.47  0.63  0.21
A4  0.34  0.37  0.22
A5  0.44  0.58  0.20
C1  0.46  0.13  0.59
C2  0.47  0.19  0.55
C3  0.54  0.25  0.60
C4 -0.62 -0.30 -0.69
C5 -0.50 -0.23 -0.56
>
```

ICLUST shows 2 scales



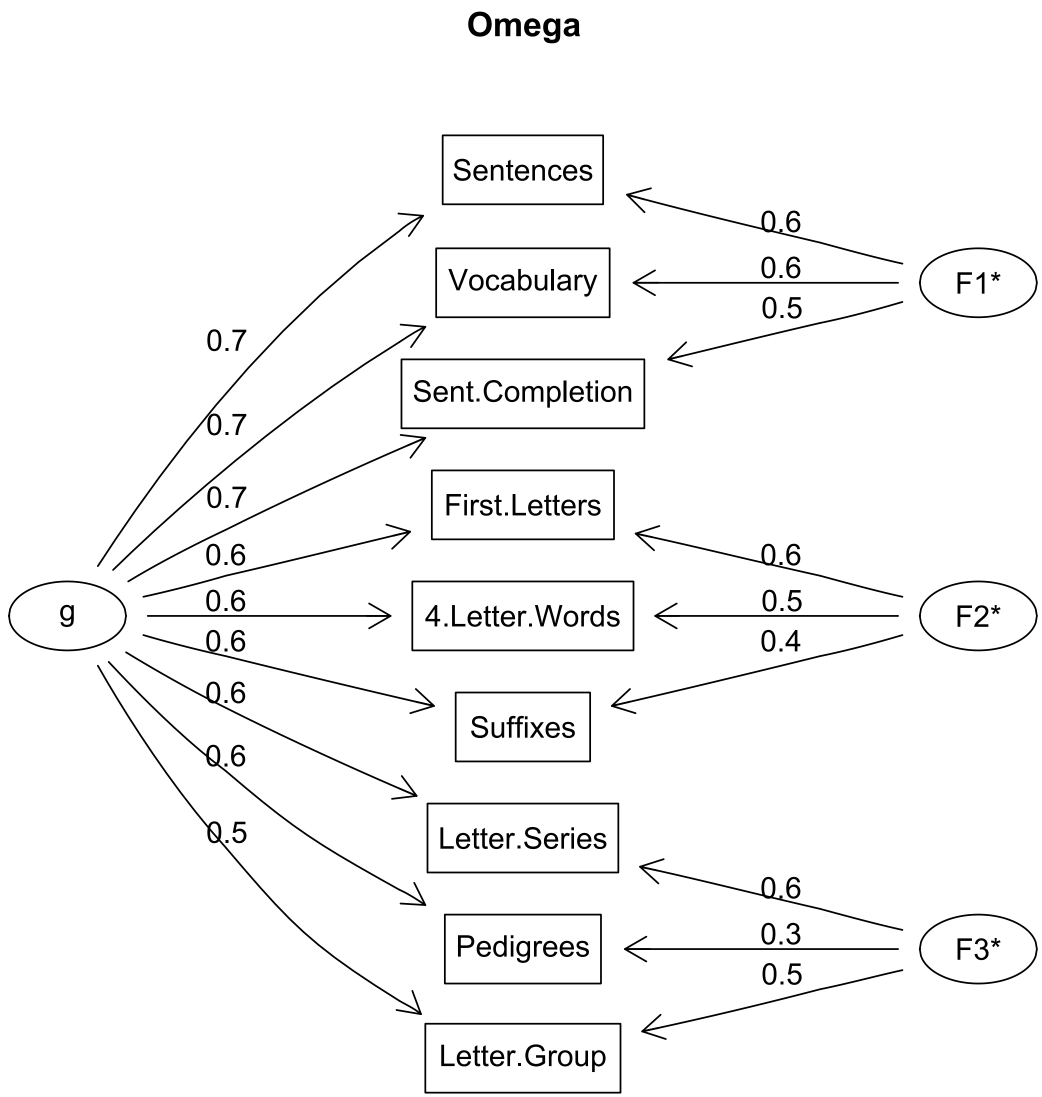
Structural Equation modeling in R

I. sem by John Fox
II. does not do multiple group analyses
III. Mx in R is a coming attraction
IV. Using psych as a front end to sem to generate the model commands

The model

```
> om <- omega(Thurstone) #creates the path model and the model commands
> om$model
      Path                                  Parameter             Initial
Value
 [1,] "g->Sentences"                        "Sentences"           NA
 [2,] "g->Vocabulary"                       "Vocabulary"          NA
 [3,] "g->Sent.Completion"                  "Sent.Completion"     NA
 [4,] "g->First.Letters"                    "First.Letters"       NA
 [5,] "g->4.Letter.Words"                   "4.Letter.Words"      NA
 [6,] "g->Suffixes"                         "Suffixes"            NA
 [7,] "g->Letter.Series"                    "Letter.Series"       NA
 [8,] "g->Pedigrees"                        "Pedigrees"           NA
 [9,] "g->Letter.Group"                     "Letter.Group"        NA
[10,] "F1*->Sentences"                      "F1*Sentences"        NA
[11,] "F1*->Vocabulary"                     "F1*Vocabulary"       NA
[12,] "F1*->Sent.Completion"                "F1*Sent.Completion" NA
[13,] "F2*->First.Letters"                  "F2*First.Letters"    NA
[14,] "F2*->4.Letter.Words"                 "F2*4.Letter.Words"   NA
[15,] "F2*->Suffixes"                       "F2*Suffixes"         NA
[16,] "F3*->Letter.Series"                  "F3*Letter.Series"    NA
[17,] "F3*->Pedigrees"                      "F3*Pedigrees"        NA
[18,] "F3*->Letter.Group"                   "F3*Letter.Group"     NA
[19,] "Sentences<->Sentences"               "e1"                  NA
[20,] "Vocabulary<->Vocabulary"             "e2"                  NA
[21,] "Sent.Completion<->Sent.Completion"   "e3"                  NA
[22,] "First.Letters<->First.Letters"       "e4"                  NA
[23,] "4.Letter.Words<->4.Letter.Words"     "e5"                  NA
[24,] "Suffixes<->Suffixes"                 "e6"                  NA
[25,] "Letter.Series<->Letter.Series"       "e7"                  NA
[26,] "Pedigrees<->Pedigrees"               "e8"                  NA
[27,] "Letter.Group<->Letter.Group"         "e9"                  NA
```

The model


Omega
Sentences
Vocabulary
Sent.Completion
First.Letters
4.Letter.Words
Suffixes
Letter.Series
Pedigrees
Letter.Group
g
F1*
F2*
F3*
0.7
0.7
0.7
0.6
0.6
0.6
0.6
0.6
0.5
0.6
0.6
0.5
0.6
0.5
0.4
0.6
0.3
0.5

Do the sem

```
> library(sem)
> sem.bf <- sem(om$model,Thurstone,213)
> summary(sem.bf,digits=2)

 Model Chisquare =  24   Df =  18 Pr(>Chisq) = 0.15
 Chisquare (null model) =  1102   Df =  36
 Goodness-of-fit index =  0.98
 Adjusted goodness-of-fit index =  0.94
 RMSEA index =  0.04   90% CI: (NA, 0.078)
 Bentler-Bonnett NFI =  0.98
 Tucker-Lewis NNFI =  0.99
 Bentler CFI =  1
 SRMR =  0.035
 BIC =  -72

 Normalized Residuals
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-8.2e-01 -3.3e-01 -8.9e-07  2.8e-02  1.6e-01  1.8e+00
```

Parameter values

```
 Parameter Estimates
                   Estimate Std Error z value Pr(>|z|)
Sentences           0.77     0.073     10.57   0.0e+00  Sentences <--- g
Vocabulary          0.79     0.072     10.92   0.0e+00  Vocabulary <--- g
Sent.Completion     0.75     0.073     10.27   0.0e+00  Sent.Completion <--- g
First.Letters       0.61     0.072      8.43   0.0e+00  First.Letters <--- g
4.Letter.Words      0.60     0.074      8.09   6.7e-16  4.Letter.Words <--- g
Suffixes            0.57     0.071      8.00   1.3e-15  Suffixes <--- g
Letter.Series       0.57     0.074      7.63   2.3e-14  Letter.Series <--- g
Pedigrees           0.66     0.069      9.55   0.0e+00  Pedigrees <--- g
Letter.Group        0.53     0.079      6.71   2.0e-11  Letter.Group <--- g
F1*Sentences        0.49     0.085      5.71   1.1e-08  Sentences <--- F1*
F1*Vocabulary       0.45     0.090      5.00   5.7e-07  Vocabulary <--- F1*
F1*Sent.Completion  0.40     0.093      4.33   1.5e-05  Sent.Completion <--- F1*
F2*First.Letters    0.61     0.086      7.16   8.2e-13  First.Letters <--- F2*
F2*4.Letter.Words   0.51     0.085      5.96   2.5e-09  4.Letter.Words <--- F2*
F2*Suffixes         0.39     0.078      5.04   4.7e-07  Suffixes <--- F2*
F3*Letter.Series    0.73     0.159      4.56   5.1e-06  Letter.Series <--- F3*
F3*Pedigrees        0.25     0.089      2.77   5.6e-03  Pedigrees <--- F3*
F3*Letter.Group     0.41     0.122      3.35   8.1e-04  Letter.Group <--- F3*
e1                  0.17     0.034      5.05   4.4e-07  Sentences <--> Sentences
e2                  0.17     0.030      5.65   1.6e-08  Vocabulary <--> Vocabulary
e3                  0.27     0.033      8.09   6.7e-16  Sent.Completion <-->Sent.Comp
e4                  0.25     0.079      3.18   1.5e-03  First.Letters <-First.Letters
e5                  0.39     0.063      6.13   8.8e-10  4.Letter.Words <-->4.Letter.W
e6                  0.52     0.060      8.68   0.0e+00  Suffixes <--> Suffixes
e7                  0.15     0.223      0.67   5.0e-01  Letter.Series <--> Letter.Ser
e8                  0.50     0.060      8.39   0.0e+00  Pedigrees <--> Pedigrees
```