

How To: Use the psych package for Factor Analysis and data reduction

William Revelle
Department of Psychology
Northwestern University

July 3, 2018

Contents

| | |
|--|-----------|
| 1 Overview of this and related documents | 3 |
| 1.1 Jump starting the <i>psych</i> package—a guide for the impatient | 3 |
| 2 Overview of this and related documents | 6 |
| 3 Getting started | 6 |
| 4 Basic data analysis | 7 |
| 4.1 Getting the data by using read.file | 7 |
| 4.2 Data input from the clipboard | 8 |
| 4.3 Basic descriptive statistics | 9 |
| 4.4 Simple descriptive graphics | 9 |
| 4.4.1 Scatter Plot Matrices | 10 |
| 4.4.2 Correlational structure | 10 |
| 4.4.3 Heatmap displays of correlational structure | 12 |
| 4.5 Polychoric, tetrachoric, polyserial, and biserial correlations | 15 |
| 5 Item and scale analysis | 15 |
| 5.1 Dimension reduction through factor analysis and cluster analysis | 15 |
| 5.1.1 Minimum Residual Factor Analysis | 17 |
| 5.1.2 Principal Axis Factor Analysis | 18 |
| 5.1.3 Weighted Least Squares Factor Analysis | 18 |
| 5.1.4 Principal Components analysis (PCA) | 24 |
| 5.1.5 Hierarchical and bi-factor solutions | 24 |
| 5.1.6 Item Cluster Analysis: iclust | 28 |

| | | |
|-----------|--|-----------|
| 5.2 | Confidence intervals using bootstrapping techniques | 31 |
| 5.3 | Comparing factor/component/cluster solutions | 31 |
| 5.4 | Determining the number of dimensions to extract. | 37 |
| 5.4.1 | Very Simple Structure | 38 |
| 5.4.2 | Parallel Analysis | 40 |
| 5.5 | Factor extension | 40 |
| 6 | Classical Test Theory and Reliability | 42 |
| 6.1 | Reliability of a single scale | 45 |
| 6.2 | Using ω to find the reliability of a single scale | 48 |
| 6.3 | Estimating ω_h using Confirmatory Factor Analysis | 52 |
| 6.3.1 | Other estimates of reliability | 54 |
| 6.4 | Reliability and correlations of multiple scales within an inventory | 54 |
| 6.4.1 | Scoring from raw data | 54 |
| 6.4.2 | Forming scales from a correlation matrix | 57 |
| 6.5 | Scoring Multiple Choice Items | 59 |
| 6.6 | Item analysis | 60 |
| 7 | Item Response Theory analysis | 60 |
| 7.1 | Factor analysis and Item Response Theory | 62 |
| 7.2 | Speeding up analyses | 66 |
| 7.3 | IRT based scoring | 68 |
| 8 | Multilevel modeling | 70 |
| 8.1 | Decomposing data into within and between level correlations using <code>statsBy</code> | 70 |
| 8.2 | Generating and displaying multilevel data | 72 |
| 9 | Set Correlation and Multiple Regression from the correlation matrix | 72 |
| 10 | Simulation functions | 75 |
| 11 | Graphical Displays | 77 |
| 12 | Miscellaneous functions | 79 |
| 13 | Data sets | 80 |
| 14 | Development version and a users guide | 81 |
| 15 | Psychometric Theory | 82 |
| 16 | SessionInfo | 82 |

1 Overview of this and related documents

To do basic and advanced personality and psychological research using R is not as complicated as some think. This is one of a set of “How To” to do various things using R (R Core Team, 2018), particularly using the *psych* (Revelle, 2018) package.

The current list of How To’s includes:

1. An [introduction](#) (vignette) of the *psych* package
2. An [overview](#) (vignette) of the *psych* package
3. [Installing](#) R and some useful packages
4. Using R and the *psych* package to find ω_h and ω_t .
5. Using R and the *psych* for [factor analysis](#) and principal components analysis (this document).
6. Using the `scoreItems` function to find [scale scores and scale statistics](#).
7. Using `mediate` and `setCor` to do [mediation, moderation and regression analysis](#)

1.1 Jump starting the *psych* package—a guide for the impatient

You have installed *psych* (section 3) and you want to use it without reading much more. What should you do?

1. Activate the *psych* package:

```
library(psych)
```

R code

2. Input your data (section 4.2). If your file name ends in .sav, .text, .txt, .csv, .xpt, .rds, .Rds, .rda, or .RDATA, then just read it in directly using `read.file`. Or you can go to your friendly text editor or data manipulation program (e.g., Excel) and copy the data to the clipboard. Include a first line that has the variable labels. Paste it into *psych* using the `read.clipboard.tab` command:

```
myData <- read.file()      #this will open a search window on your machine
#                          and read or load the file.
#or
#first copy your file to your clipboard and then
myData <- read.clipboard.tab() #if you have an excel file
```

R code

3. Make sure that what you just read is right. Describe it (section 4.3) and perhaps look at the first and last few lines. If you want to “view” the first and last few lines using a spreadsheet like viewer, use `quickView`.

R code

```
describe(myData)
headTail(myData)
#or
quickView(myData)
```

4. Look at the patterns in the data. If you have fewer than about 10 variables, look at the SPLOM (Scatter Plot Matrix) of the data using `pairs.panels` (section 4.4.1).

R code

```
pairs.panels(myData)
```

5. Find the correlations of all of your data.

- Descriptively (just the values) (section 4.4.2)

R code

```
lowerCor(myData)
```

- Graphically (section 4.4.3)

R code

```
corPlot(r)
```

6. Test for the number of factors in your data using parallel analysis (`fa.parallel`, section 5.4.2) or Very Simple Structure (`vss`, 5.4.1) .

R code

```
fa.parallel(myData)
vss(myData)
```

7. Factor analyze (see section 5.1) the data with a specified number of factors (the default is 1), the default method is minimum residual, the default rotation for more than one factor is oblimin. There are many more possibilities (see sections 5.1.1-5.1.3). Compare the solution to a hierarchical cluster analysis using the ICLUST algorithm (Revelle, 1979) (see section 5.1.6). Also consider a hierarchical factor solution to find coefficient ω (see 5.1.5).

R code

```
fa(myData)
iclust(myData)
omega(myData)
```

8. Some people like to find coefficient α as an estimate of reliability. This may be done for a single scale using the `alpha` function (see 6.1). Perhaps more useful is the ability to create several scales as unweighted averages of specified items using the `scoreItems` function (see 6.4) and to find various estimates of internal consistency for these scales, find their intercorrelations, and find scores for all the subjects.

R code

```
alpha(myData) #score all of the items as part of one scale.
myKeys <- make.keys(nvar=20,list(first = c(1,-3,5,-7,8:10),second=c(2,4,-6,11:15,-16)))
```

```
my.scores <- scoreItems(myKeys,myData) #form several scales
my.scores #show the highlights of the results
```

At this point you have had a chance to see the highlights of the *psych* package and to do some basic (and advanced) data analysis. You might find reading the entire [overview](#) vignette helpful to get a broader understanding of what can be done in R using the *psych*. Remember that the help command (?) is available for every function. Try running the examples for each help page.

2 Overview of this and related documents

The *psych* package (Revelle, 2018) has been developed at Northwestern University since 2005 to include functions most useful for personality, psychometric, and psychological research. The package is also meant to supplement a text on psychometric theory (Revelle, prep), a draft of which is available at <http://personality-project.org/r/book/>.

Some of the functions (e.g., `read.file`, `read.clipboard`, `describe`, `pairs.panels`, `scatter.hist`, `error.bars`, `multi.hist`, `bi.bars`) are useful for basic data entry and descriptive analyses.

Psychometric applications emphasize techniques for dimension reduction including factor analysis, cluster analysis, and principal components analysis. The `fa` function includes five methods of *factor analysis* (*minimum residual*, *principal axis*, *weighted least squares*, *generalized least squares* and *maximum likelihood* factor analysis). Determining the number of factors or components to extract may be done by using the Very Simple Structure (Revelle and Rocklin, 1979) (`vss`), Minimum Average Partial correlation (Velicer, 1976) (MAP) or parallel analysis (`fa.parallel`) criteria. Item Response Theory (IRT) models for dichotomous or polytomous items may be found by factoring `tetrachoric` or `polychoric` correlation matrices and expressing the resulting parameters in terms of location and discrimination using `irt.fa`. Bifactor and hierarchical factor structures may be estimated by using Schmid Leiman transformations (Schmid and Leiman, 1957) (`schmid`) to transform a hierarchical factor structure into a *bifactor* solution (Holzinger and Swineford, 1937). Scale construction can be done using the Item Cluster Analysis (Revelle, 1979) (`iclust`) function to determine the structure and to calculate reliability coefficients α (Cronbach, 1951) (`alpha`, `scoreItems`, `score.multiple.choice`), β (Revelle, 1979; Revelle and Zinbarg, 2009) (`iclust`) and McDonald's ω_h and ω_t (McDonald, 1999) (`omega`). Guttman's six estimates of internal consistency reliability (Guttman (1945)), as well as additional estimates (Revelle and Zinbarg, 2009) are in the `guttman` function. The six measures of Intraclass correlation coefficients (ICC) discussed by Shrout and Fleiss (1979) are also available.

Graphical displays include Scatter Plot Matrix (SPLOM) plots using `pairs.panels`, correlation "heat maps" (`cor.plot`) factor, cluster, and structural diagrams using `fa.diagram`, `iclust.diagram`, `structure.diagram`, as well as item response characteristics and item and test information characteristic curves `plot.irt` and `plot.poly`.

3 Getting started

Some of the functions described in this overview require other packages. Particularly useful for rotating the results of factor analyses (from e.g., `fa` or `principal`) or hierarchical factor models using `omega` or `schmid`, is the *GPArotation* package. These and other useful

packages may be installed by first installing and then using the task views (*ctv*) package to install the “Psychometrics” task view, but doing it this way is not necessary.

4 Basic data analysis

A number of *psych* functions facilitate the entry of data and finding basic descriptive statistics.

Remember, to run any of the *psych* functions, it is necessary to make the package active by using the `library` command:

```
library(psych)
```

The other packages, once installed, will be called automatically by *psych*.

It is possible to automatically load *psych* and other functions by creating and then saving a “.First” function: e.g.,

```
.First <- function(x) {library(psych)}
```

4.1 Getting the data by using `read.file`

Although many find copying the data to the clipboard and then using the `read.clipboard` functions (see below), a helpful alternative is to read the data in directly. This can be done using the `read.file` function which calls `file.choose` to find the file and then based upon the suffix of the file, chooses the appropriate way to read it. For files with suffixes of `.txt`, `.text`, `.r`, `.rds`, `.rda`, `.csv`, `.xpt`, or `.sav`, the file will be read correctly.

```
my.data <- read.file()
```

If the file contains Fixed Width Format (fww) data, the column information can be specified with the `widths` command.

```
my.data <- read.file(widths = c(4,rep(1,35))) #will read in a file without a header row  
# and 36 fields, the first of which is 4 columns, the rest of which are 1 column each.
```

If the file is a `.RData` file (with suffix of `.RData`, `.Rda`, `.rda`, `.Rdata`, or `.rdata`) the object will be loaded. Depending what was stored, this might be several objects. If the file is a `.sav` file from SPSS, it will be read with the most useful default options (converting the file to a `data.frame` and converting character fields to numeric). Alternative options may be specified. If it is an export file from SAS (`.xpt` or `.XPT`) it will be read. `.csv` files (comma separated files), normal `.txt` or `.text` files, `.data`, or `.dat` files will be read as well. These are

assumed to have a header row of variable labels (`header=TRUE`). If the data do not have a header row, you must specify `read.file(header=FALSE)`.

To read SPSS files and to keep the value labels, specify `use.value.labels=TRUE`. Unfortunately, this will make it difficult to do any analyses, and you might want to read it again converting the value labels to numeric values.

```
R code
my.spss <- read.file(use.value.labels=TRUE) #this will keep the value labels for .sav files
my.data <- read.file() #just read the file and convert the value labels to numeric.
```

You can `quickView` the `my.spss` object to see what it looks like as well as the `my.data` to see the numeric encoding.

```
R code
quickView(my.spss) #to see the original encoding
quickkView(mydata) #to see the numeric encoding
```

4.2 Data input from the clipboard

There are of course many ways to enter data into R. Reading from a local file using `read.file` is perhaps the most preferred. However, many users will enter their data in a text editor or spreadsheet program and then want to copy and paste into R. This may be done by using `read.table` and specifying the input file as “clipboard” (PCs) or “pipe(pbpaste)” (Macs). Alternatively, the `read.clipboard` set of functions are perhaps more user friendly:

`read.clipboard` is the base function for reading data from the clipboard.

`read.clipboard.csv` for reading text that is comma delimited.

`read.clipboard.tab` for reading text that is tab delimited (e.g., copied directly from an Excel file).

`read.clipboard.lower` for reading input of a lower triangular matrix with or without a diagonal. The resulting object is a square matrix.

`read.clipboard.upper` for reading input of an upper triangular matrix.

`read.clipboard.fwf` for reading in fixed width fields (some very old data sets)

For example, given a data set copied to the clipboard from a spreadsheet, just enter the command

```
R code
my.data <- read.clipboard()
```

This will work if every data field has a value and even missing data are given some values (e.g., NA or -999). If the data were entered in a spreadsheet and the missing values

were just empty cells, then the data should be read in as a tab delimited or by using the `read.clipboard.tab` function.

```
R code
my.data <- read.clipboard(sep="\t") #define the tab option, or
my.tab.data <- read.clipboard.tab() #just use the alternative function
```

For the case of data in fixed width fields (some old data sets tend to have this format), copy to the clipboard and then specify the width of each field (in the example below, the first variable is 5 columns, the second is 2 columns, the next 5 are 1 column the last 4 are 3 columns).

```
R code
my.data <- read.clipboard.fwf(widths=c(5,2,rep(1,5),rep(3,4)))
```

4.3 Basic descriptive statistics

Once the data are read in, then `describe` will provide basic descriptive statistics arranged in a data frame format. Consider the data set `sat.act` which includes data from 700 web based participants on 3 demographic variables and 3 ability measures.

`describe` reports means, standard deviations, medians, min, max, range, skew, kurtosis and standard errors for integer or real data. Non-numeric data, although the statistics are meaningless, will be treated as if numeric (based upon the categorical coding of the data), and will be flagged with an `*`.

It is very important to describe your data before you continue on doing more complicated multivariate statistics. The problem of outliers and bad data can not be overemphasized.

```
> library(psych)
> data(sat.act)
> describe(sat.act) #basic descriptive statistics
```

| | vars | n | mean | sd | median | trimmed | mad | min | max | range | skew | kurtosis | se |
|-----------|------|-----|--------|--------|--------|---------|--------|-----|-----|-------|-------|----------|------|
| gender | 1 | 700 | 1.65 | 0.48 | 2 | 1.68 | 0.00 | 1 | 2 | 1 | -0.61 | -1.62 | 0.02 |
| education | 2 | 700 | 3.16 | 1.43 | 3 | 3.31 | 1.48 | 0 | 5 | 5 | -0.68 | -0.07 | 0.05 |
| age | 3 | 700 | 25.59 | 9.50 | 22 | 23.86 | 5.93 | 13 | 65 | 52 | 1.64 | 2.42 | 0.36 |
| ACT | 4 | 700 | 28.55 | 4.82 | 29 | 28.84 | 4.45 | 3 | 36 | 33 | -0.66 | 0.53 | 0.18 |
| SATV | 5 | 700 | 612.23 | 112.90 | 620 | 619.45 | 118.61 | 200 | 800 | 600 | -0.64 | 0.33 | 4.27 |
| SATQ | 6 | 687 | 610.22 | 115.64 | 620 | 617.25 | 118.61 | 200 | 800 | 600 | -0.59 | -0.02 | 4.41 |

4.4 Simple descriptive graphics

Graphic descriptions of data are very helpful both for understanding the data as well as communicating important results. Scatter Plot Matrices (SPLOMS) using the `pairs.panels`

function are useful ways to look for strange effects involving outliers and non-linearities. `error.bars.by` will show group means with 95% confidence boundaries.

4.4.1 Scatter Plot Matrices

Scatter Plot Matrices (SPLOMS) are very useful for describing the data. The `pairs.panels` function, adapted from the help menu for the `pairs` function produces xy scatter plots of each pair of variables below the diagonal, shows the histogram of each variable on the diagonal, and shows the *lowess* locally fit regression line as well. An ellipse around the mean with the axis length reflecting one standard deviation of the x and y variables is also drawn. The x axis in each scatter plot represents the column variable, the y axis the row variable (Figure 1). When plotting many subjects, it is both faster and cleaner to set the plot character (`pch`) to be `'.'`. (See Figure 1 for an example.)

`pairs.panels` will show the pairwise scatter plots of all the variables as well as histograms, locally smoothed regressions, and the Pearson correlation. When plotting many data points (as in the case of the `sat.act` data, it is possible to specify that the plot character is a period to get a somewhat cleaner graphic.

4.4.2 Correlational structure

There are many ways to display correlations. Tabular displays are probably the most common. The output from the `cor` function in core R is a rectangular matrix. `lowerMat` will round this to (2) digits and then display as a lower off diagonal matrix. `lowerCor` calls `cor` with `use='pairwise'`, `method='pearson'` as default values and returns (invisibly) the full correlation matrix and displays the lower off diagonal matrix.

```
> lowerCor(sat.act)
      gender  edctn age  ACT  SATV  SATQ
gender      1.00
education  0.09  1.00
age        -0.02  0.55  1.00
ACT        -0.04  0.15  0.11  1.00
SATV       -0.02  0.05 -0.04  0.56  1.00
SATQ       -0.17  0.03 -0.03  0.59  0.64  1.00
```

When comparing results from two different groups, it is convenient to display them as one matrix, with the results from one group below the diagonal, and the other group above the diagonal. Use `lowerUpper` to do this:

```
> female <- subset(sat.act, sat.act$gender==2)
> male <- subset(sat.act, sat.act$gender==1)
> lower <- lowerCor(male[-1])
```

```

> png( 'pairspanels.png' )
> pairs.panels(sat.act,pch='.')
> dev.off()
null device
1

```

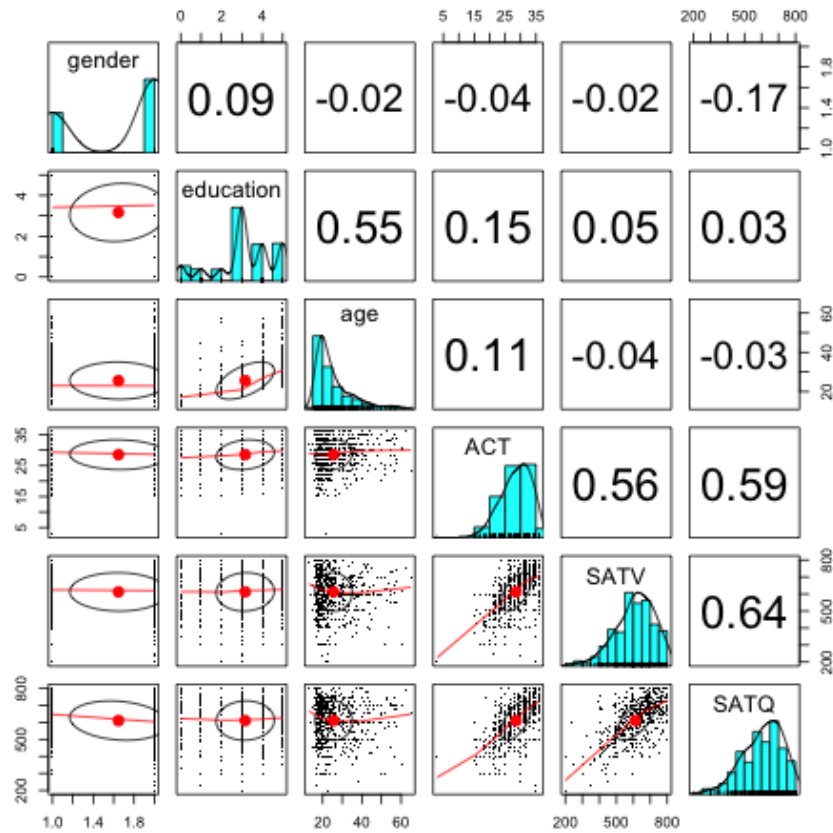


Figure 1: Using the `pairs.panels` function to graphically show relationships. The x axis in each scatter plot represents the column variable, the y axis the row variable. Note the extreme outlier for the ACT. The plot character was set to a period (`pch='.'`) in order to make a cleaner graph.

```

      edctn age  ACT  SATV  SATQ
education 1.00
age       0.61 1.00
ACT       0.16 0.15 1.00
SATV     0.02 -0.06 0.61 1.00
SATQ     0.08 0.04 0.60 0.68 1.00

> upper <- lowerCor(female[-1])

      edctn age  ACT  SATV  SATQ
education 1.00
age       0.52 1.00
ACT       0.16 0.08 1.00
SATV     0.07 -0.03 0.53 1.00
SATQ     0.03 -0.09 0.58 0.63 1.00

> both <- lowerUpper(lower, upper)
> round(both, 2)

      education  age  ACT  SATV  SATQ
education      NA 0.52 0.16 0.07 0.03
age            0.61  NA 0.08 -0.03 -0.09
ACT            0.16 0.15  NA 0.53 0.58
SATV           0.02 -0.06 0.61  NA 0.63
SATQ           0.08 0.04 0.60 0.68  NA

```

It is also possible to compare two matrices by taking their differences and displaying one (below the diagonal) and the difference of the second from the first above the diagonal:

```

> diffs <- lowerUpper(lower, upper, diff=TRUE)
> round(diffs, 2)

      education  age  ACT  SATV  SATQ
education      NA 0.09 0.00 -0.05 0.05
age            0.61  NA 0.07 -0.03 0.13
ACT            0.16 0.15  NA 0.08 0.02
SATV           0.02 -0.06 0.61  NA 0.05
SATQ           0.08 0.04 0.60 0.68  NA

```

4.4.3 Heatmap displays of correlational structure

Perhaps a better way to see the structure in a correlation matrix is to display a *heat map* of the correlations. This is just a matrix color coded to represent the magnitude of the correlation. This is useful when considering the number of factors in a data set. Consider the **Thurstone** data set which has a clear 3 factor solution (Figure 2) or a simulated data set of 24 variables with a circumplex structure (Figure 3). The color coding represents a “heat map” of the correlations, with darker shades of red representing stronger negative and darker shades of blue stronger positive correlations. As an option, the value of the correlation can be shown.

```

> png('corplot.png')
> cor.plot(Thurstone,numbers=TRUE,main="9 cognitive variables from Thurstone")
> dev.off()
null device
      1

```

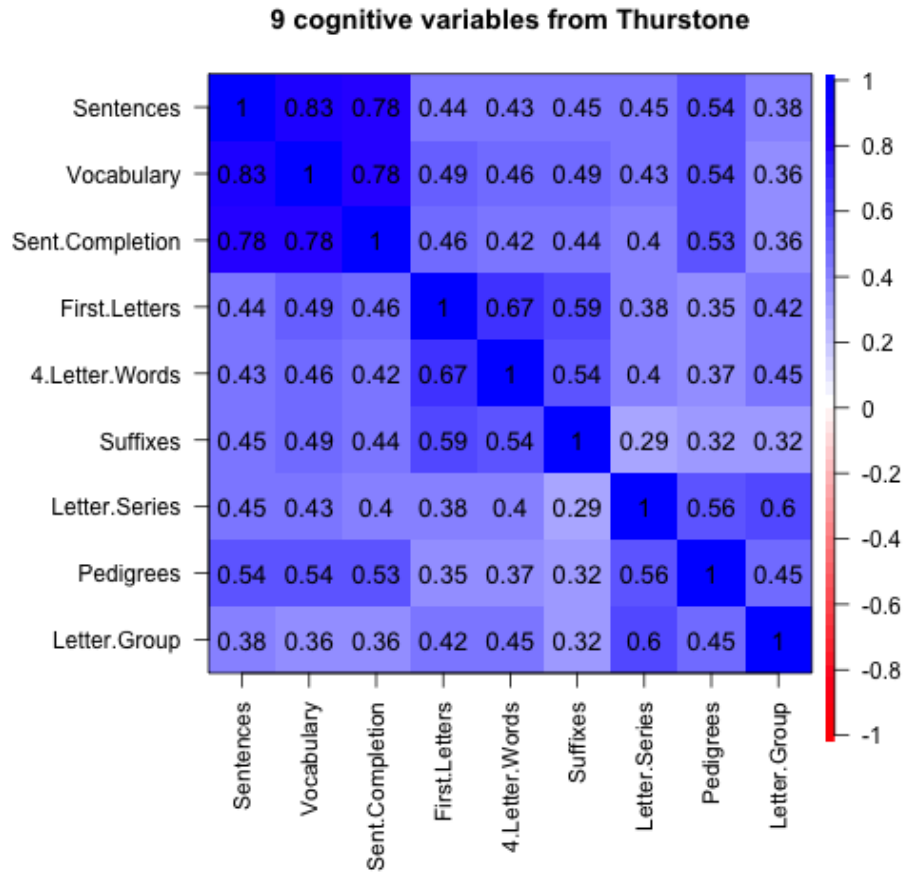


Figure 2: The structure of correlation matrix can be seen more clearly if the variables are grouped by factor and then the correlations are shown by color. By using the 'numbers' option, the values are displayed as well.

```
> png('circplot.png')
> circ <- sim.circ(24)
> r.circ <- cor(circ)
> cor.plot(r.circ,main='24 variables in a circumplex')
> dev.off()
null device
      1
```

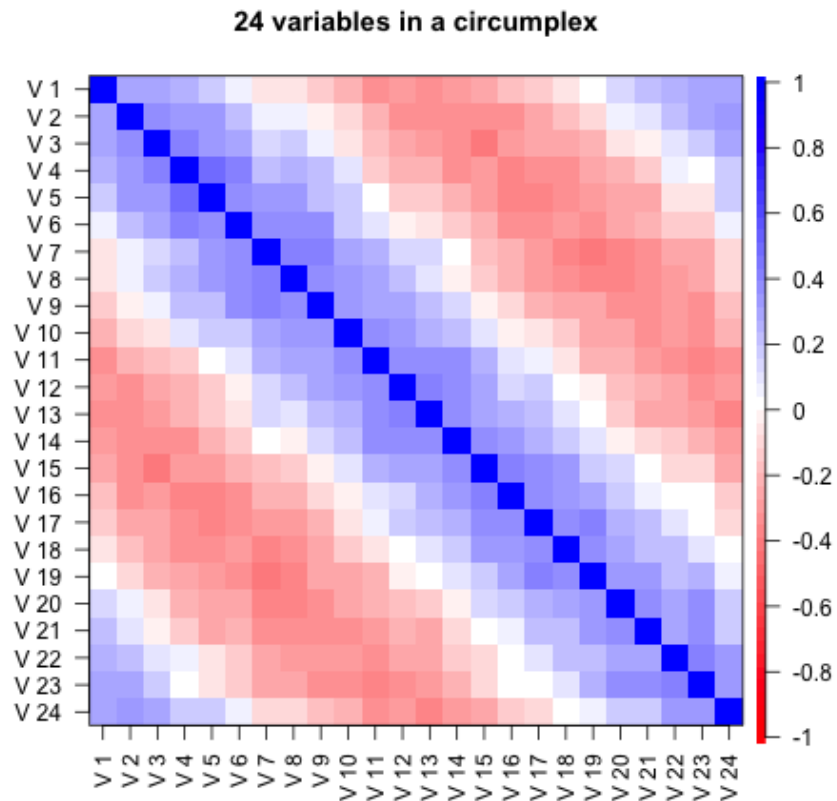


Figure 3: Using the `cor.plot` function to show the correlations in a circumplex. Correlations are highest near the diagonal, diminish to zero further from the diagonal, and the increase again towards the corners of the matrix. Circumplex structures are common in the study of affect.

4.5 Polychoric, tetrachoric, polyserial, and biserial correlations

The Pearson correlation of dichotomous data is also known as the ϕ coefficient. If the data, e.g., ability items, are thought to represent an underlying continuous although latent variable, the ϕ will underestimate the value of the Pearson applied to these latent variables. One solution to this problem is to use the `tetrachoric` correlation which is based upon the assumption of a bivariate normal distribution that has been cut at certain points. The `draw.tetra` function demonstrates the process, A simple generalization of this to the case of the multiple cuts is the `polychoric` correlation.

Other estimated correlations based upon the assumption of bivariate normality with cut points include the `biserial` and `polyserial` correlation.

If the data are a mix of continuous, polytomous and dichotomous variables, the `mixed.cor` function will calculate the appropriate mixture of Pearson, polychoric, tetrachoric, biserial, and polyserial correlations.

The correlation matrix resulting from a number of tetrachoric or polychoric correlation matrix sometimes will not be positive semi-definite. This will also happen if the correlation matrix is formed by using pair-wise deletion of cases. The `cor.smooth` function will adjust the smallest eigen values of the correlation matrix to make them positive, rescale all of them to sum to the number of variables, and produce a “smoothed” correlation matrix. An example of this problem is a data set of `burt` which probably had a typo in the original correlation matrix. Smoothing the matrix corrects this problem.

5 Item and scale analysis

The main functions in the *psych* package are for analyzing the structure of items and of scales and for finding various estimates of scale reliability. These may be considered as problems of dimension reduction (e.g., factor analysis, cluster analysis, principal components analysis) and of forming and estimating the reliability of the resulting composite scales.

5.1 Dimension reduction through factor analysis and cluster analysis

Parsimony of description has been a goal of science since at least the famous dictum commonly attributed to William of Ockham to not multiply entities beyond necessity¹. The goal for parsimony is seen in psychometrics as an attempt either to describe (components)

¹Although probably neither original with Ockham nor directly stated by him ([Thorburn, 1918](#)), Ockham’s razor remains a fundamental principal of science.

or to explain (factors) the relationships between many observed variables in terms of a more limited set of components or latent factors.

The typical data matrix represents multiple items or scales usually thought to reflect fewer underlying constructs². At the most simple, a set of items can be thought to represent a random sample from one underlying domain or perhaps a small set of domains. The question for the psychometrician is how many domains are represented and how well does each item represent the domains. Solutions to this problem are examples of *factor analysis* (FA), *principal components analysis* (PCA), and *cluster analysis* (CA). All of these procedures aim to reduce the complexity of the observed data. In the case of FA, the goal is to identify fewer underlying constructs to explain the observed data. In the case of PCA, the goal can be mere data reduction, but the interpretation of components is frequently done in terms similar to those used when describing the latent variables estimated by FA. Cluster analytic techniques, although usually used to partition the subject space rather than the variable space, can also be used to group variables to reduce the complexity of the data by forming fewer and more homogeneous sets of tests or items.

At the data level the data reduction problem may be solved as a *Singular Value Decomposition* of the original matrix, although the more typical solution is to find either the *principal components* or *factors* of the covariance or correlation matrices. Given the pattern of regression weights from the variables to the components or from the factors to the variables, it is then possible to find (for components) individual *component* or *cluster scores* or estimate (for factors) *factor scores*.

Several of the functions in *psych* address the problem of data reduction.

fa incorporates five alternative algorithms: *minres factor analysis*, *principal axis factor analysis*, *weighted least squares factor analysis*, *generalized least squares factor analysis* and *maximum likelihood factor analysis*. That is, it includes the functionality of three other functions that will be eventually phased out.

principal Principal Components Analysis reports the largest *n* eigen vectors rescaled by the square root of their eigen values.

factor.congruence The congruence between two factors is the cosine of the angle between them. This is just the cross products of the loadings divided by the sum of the squared loadings. This differs from the correlation coefficient in that the mean loading is not subtracted before taking the products. **factor.congruence** will find the cosines between two (or more) sets of factor loadings.

vss Very Simple Structure [Revelle and Rocklin \(1979\)](#) applies a goodness of fit test to determine the optimal number of factors to extract. It can be thought of as a quasi-

²[Cattell \(1978\)](#) as well as [MacCallum et al. \(2007\)](#) argue that the data are the result of many more factors than observed variables, but are willing to estimate the major underlying factors.

confirmatory model, in that it fits the very simple structure (all except the biggest c loadings per item are set to zero where c is the level of complexity of the item) of a factor pattern matrix to the original correlation matrix. For items where the model is usually of complexity one, this is equivalent to making all except the largest loading for each item 0. This is typically the solution that the user wants to interpret. The analysis includes the MAP criterion of [Velicer \(1976\)](#) and a χ^2 estimate.

- `fa.parallel` The parallel factors technique compares the observed eigen values of a correlation matrix with those from random data.
- `fa.plot` will plot the loadings from a factor, principal components, or cluster analysis (just a call to plot will suffice). If there are more than two factors, then a SPLOM of the loadings is generated.
- `nfactors` A number of different tests for the number of factors problem are run.
- `fa.diagram` replaces `fa.graph` and will draw a path diagram representing the factor structure. It does not require `Rgraphviz` and thus is probably preferred.
- `fa.graph` requires `Rgraphviz` and will draw a graphic representation of the factor structure. If factors are correlated, this will be represented as well.
- `iclust` is meant to do item cluster analysis using a hierarchical clustering algorithm specifically asking questions about the reliability of the clusters ([Revelle, 1979](#)). Clusters are formed until either coefficient α [Cronbach \(1951\)](#) or β [Revelle \(1979\)](#) fail to increase.

5.1.1 Minimum Residual Factor Analysis

The factor model is an approximation of a correlation matrix by a matrix of lower rank. That is, can the correlation matrix, ${}_nR_n$ be approximated by the product of a factor matrix, ${}_nF_k$ and its transpose plus a diagonal matrix of uniqueness.

$$R = FF' + U^2 \tag{1}$$

The maximum likelihood solution to this equation is found by `factanal` in the *stats* package. Five alternatives are provided in *psych*, all of them are included in the `fa` function and are called by specifying the factor method (e.g., `fm="minres"`, `fm="pa"`, `fm=""` "wls", `fm="gls"` and `fm="ml"`). In the discussion of the other algorithms, the calls shown are to the `fa` function specifying the appropriate method.

`factor.minres` attempts to minimize the off diagonal residual correlation matrix by adjusting the eigen values of the original correlation matrix. This is similar to what is done

in `factanal`, but uses an ordinary least squares instead of a maximum likelihood fit function. The solutions tend to be more similar to the MLE solutions than are the `factor.pa` solutions. `min.res` is the default for the `fa` function.

A classic data set, collected by [Thurstone and Thurstone \(1941\)](#) and then reanalyzed by [Bechtoldt \(1961\)](#) and discussed by [McDonald \(1999\)](#), is a set of 9 cognitive variables with a clear bi-factor structure [Holzinger and Swineford \(1937\)](#). The minimum residual solution was transformed into an oblique solution using the default option on `rotate` which uses an oblimin transformation (Table 1). Alternative rotations and transformations include “none”, “varimax”, “quartimax”, “bentlerT”, and “geominT” (all of which are orthogonal rotations). as well as “promax”, “oblimin”, “simplimax”, “bentlerQ”, and “geominQ” and “cluster” which are possible oblique transformations of the solution. The default is to do a oblimin transformation, although prior versions defaulted to varimax. The measures of factor adequacy reflect the multiple correlations of the factors with the best fitting linear regression estimates of the factor scores ([Grice, 2001](#)).

5.1.2 Principal Axis Factor Analysis

An alternative, least squares algorithm, `factor.pa`, (incorporated into `fa` as an option (`fm = “pa”`) does a Principal Axis factor analysis by iteratively doing an eigen value decomposition of the correlation matrix with the diagonal replaced by the values estimated by the factors of the previous iteration. This OLS solution is not as sensitive to improper matrices as is the maximum likelihood method, and will sometimes produce more interpretable results. It seems as if the SAS example for PA uses only one iteration. Setting the `max.iter` parameter to 1 produces the SAS solution.

The solutions from the `fa`, the `factor.minres` and `factor.pa` as well as the `principal` functions can be rotated or transformed with a number of options. Some of these call the *GPArotation* package. Orthogonal rotations are `varimax` and `quartimax`. Oblique transformations include `oblimin`, `quartimin` and then two targeted rotation functions `Promax` and `target.rot`. The latter of these will transform a loadings matrix towards an arbitrary target matrix. The default is to transform towards an independent cluster solution.

Using the Thurstone data set, three factors were requested and then transformed into an independent clusters solution using `target.rot` (Table 2).

5.1.3 Weighted Least Squares Factor Analysis

Similar to the minres approach of minimizing the squared residuals, factor method “wls” weights the squared residuals by their uniquenesses. This tends to produce slightly smaller

Table 1: Three correlated factors from the Thurstone 9 variable problem. By default, the solution is transformed obliquely using oblimin. The extraction method is (by default) minimum residual.

```

> f3t <- fa(Thurstone, 3, n.obs=213)
> f3t
Factor Analysis using method = minres
Call: fa(r = Thurstone, nfactors = 3, n.obs = 213)
Standardized loadings (pattern matrix) based upon correlation matrix
      MR1  MR2  MR3  h2  u2 com
Sentences      0.91 -0.04  0.04  0.82  0.18  1.0
Vocabulary     0.89  0.06 -0.03  0.84  0.16  1.0
Sent.Completion 0.83  0.04  0.00  0.73  0.27  1.0
First.Letters  0.00  0.86  0.00  0.73  0.27  1.0
4.Letter.Words -0.01  0.74  0.10  0.63  0.37  1.0
Suffixes       0.18  0.63 -0.08  0.50  0.50  1.2
Letter.Series  0.03 -0.01  0.84  0.72  0.28  1.0
Pedigrees     0.37 -0.05  0.47  0.50  0.50  1.9
Letter.Group  -0.06  0.21  0.64  0.53  0.47  1.2

      MR1  MR2  MR3
SS loadings      2.64  1.86  1.50
Proportion Var   0.29  0.21  0.17
Cumulative Var   0.29  0.50  0.67
Proportion Explained 0.44  0.31  0.25
Cumulative Proportion 0.44  0.75  1.00

With factor correlations of
      MR1  MR2  MR3
MR1  1.00  0.59  0.54
MR2  0.59  1.00  0.52
MR3  0.54  0.52  1.00

Mean item complexity = 1.2
Test of the hypothesis that 3 factors are sufficient.

The degrees of freedom for the null model are 36 and the objective function was 5.2 with Chi Square of 10
The degrees of freedom for the model are 12 and the objective function was 0.01

The root mean square of the residuals (RMSR) is 0.01
The df corrected root mean square of the residuals is 0.01

The harmonic number of observations is 213 with the empirical chi square 0.58 with prob < 1
The total number of observations was 213 with MLE Chi Square = 2.82 with prob < 1

Tucker Lewis Index of factoring reliability = 1.027
RMSEA index = 0 and the 90 % confidence intervals are NA NA
BIC = -61.51
Fit based upon off diagonal values = 1
Measures of factor score adequacy
      MR1  MR2  MR3
Correlation of scores with factors      0.96  0.92  0.90
Multiple R square of scores with factors  0.93  0.85  0.81
Minimum correlation of possible factor scores 0.86  0.71  0.63

```

Table 2: The 9 variable problem from Thurstone is a classic example of factoring where there is a higher order factor, g, that accounts for the correlation between the factors. The extraction method was principal axis. The transformation was a targeted transformation to a simple cluster solution.

```

> f3 <- fa(Thurstone, 3, n.obs = 213, fm="pa")
> f3o <- target.rot(f3)
> f3o
Call: NULL
Standardized loadings (pattern matrix) based upon correlation matrix
      PA1  PA2  PA3  h2  u2
Sentences      0.89 -0.03  0.07  0.81  0.19
Vocabulary     0.89  0.07  0.00  0.80  0.20
Sent.Completion 0.83  0.04  0.03  0.70  0.30
First.Letters  -0.02  0.85 -0.01  0.73  0.27
4.Letter.Words -0.05  0.74  0.09  0.57  0.43
Suffixes       0.17  0.63 -0.09  0.43  0.57
Letter.Series  -0.06 -0.08  0.84  0.69  0.31
Pedigrees      0.33 -0.09  0.48  0.37  0.63
Letter.Group   -0.14  0.16  0.64  0.45  0.55

      PA1  PA2  PA3
SS loadings      2.45  1.72  1.37
Proportion Var   0.27  0.19  0.15
Cumulative Var   0.27  0.46  0.62
Proportion Explained 0.44  0.31  0.25
Cumulative Proportion 0.44  0.75  1.00
      PA1  PA2  PA3
PA1 1.00  0.02  0.08
PA2 0.02  1.00  0.09
PA3 0.08  0.09  1.00

```

overall residuals. In the example of weighted least squares, the output is shown by using the `print` function with the `cut` option set to 0. That is, all loadings are shown (Table 3).

The unweighted least squares solution may be shown graphically using the `fa.plot` function which is called by the generic `plot` function (Figure 4). Factors were transformed obliquely using a `oblimin`. These solutions may be shown as item by factor plots (Figure 4) or by a structure diagram (Figure 5).

```
> plot(f3t)
```

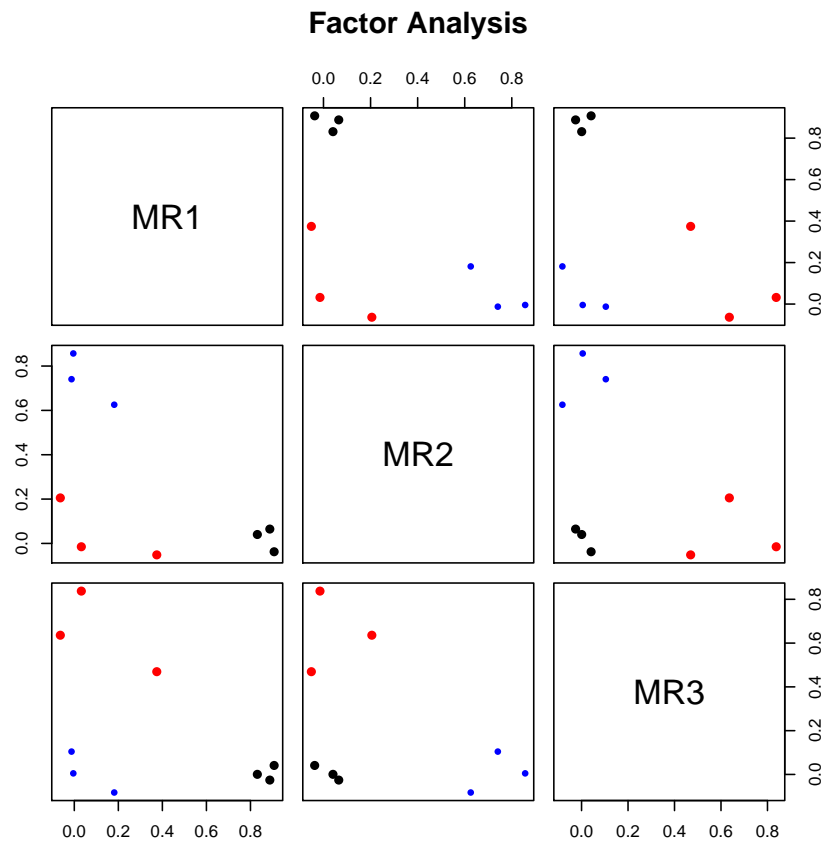


Figure 4: A graphic representation of the 3 oblique factors from the Thurstone data using `plot`. Factors were transformed to an oblique solution using the `oblimin` function from the `GPArotation` package.

A comparison of these three approaches suggests that the minres solution is more similar to a maximum likelihood solution and fits slightly better than the `pa` or `wls` solutions. Comparisons with SPSS suggest that the `pa` solution matches the SPSS OLS solution, but

Table 3: The 9 variable problem from Thurstone is a classic example of factoring where there is a higher order factor, g , that accounts for the correlation between the factors. The factors were extracted using a weighted least squares algorithm. All loadings are shown by using the `cut=0` option in the `print.psych` function.

```
> f3w <- fa(Thurstone, 3, n.obs = 213, fm="wls")
> print(f3w, cut=0, digits=3)
Factor Analysis using method = wls
Call: fa(r = Thurstone, nfactors = 3, n.obs = 213, fm = "wls")
Standardized loadings (pattern matrix) based upon correlation matrix
```

| | WLS1 | WLS2 | WLS3 | h2 | u2 | com |
|-----------------|--------|--------|--------|-------|-------|------|
| Sentences | 0.905 | -0.034 | 0.040 | 0.822 | 0.178 | 1.01 |
| Vocabulary | 0.890 | 0.066 | -0.031 | 0.835 | 0.165 | 1.01 |
| Sent.Completion | 0.833 | 0.034 | 0.007 | 0.735 | 0.265 | 1.00 |
| First.Letters | -0.002 | 0.855 | 0.003 | 0.731 | 0.269 | 1.00 |
| 4.Letter.Words | -0.016 | 0.743 | 0.106 | 0.629 | 0.371 | 1.04 |
| Suffixes | 0.180 | 0.626 | -0.082 | 0.496 | 0.504 | 1.20 |
| Letter.Series | 0.033 | -0.015 | 0.838 | 0.719 | 0.281 | 1.00 |
| Pedigrees | 0.381 | -0.051 | 0.464 | 0.505 | 0.495 | 1.95 |
| Letter.Group | -0.062 | 0.209 | 0.632 | 0.527 | 0.473 | 1.24 |

```

                WLS1  WLS2  WLS3
SS loadings    2.647  1.864  1.488
Proportion Var 0.294  0.207  0.165
Cumulative Var 0.294  0.501  0.667
Proportion Explained 0.441  0.311  0.248
Cumulative Proportion 0.441  0.752  1.000

With factor correlations of
      WLS1  WLS2  WLS3
WLS1 1.000  0.591  0.535
WLS2 0.591  1.000  0.516
WLS3 0.535  0.516  1.000

Mean item complexity = 1.2
Test of the hypothesis that 3 factors are sufficient.

The degrees of freedom for the null model are 36 and the objective function was 5.198 with Chi Square of
The degrees of freedom for the model are 12 and the objective function was 0.014

The root mean square of the residuals (RMSR) is 0.006
The df corrected root mean square of the residuals is 0.01

The harmonic number of observations is 213 with the empirical chi square 0.531 with prob < 1
The total number of observations was 213 with MLE Chi Square = 2.886 with prob < 0.996

Tucker Lewis Index of factoring reliability = 1.0264
RMSEA index = 0 and the 90 % confidence intervals are NA NA
BIC = -61.45
Fit based upon off diagonal values = 1
Measures of factor score adequacy
```

| | WLS1 | WLS2 | WLS3 |
|---|-------|-------|-------|
| Correlation of scores with factors | 0.964 | 0.923 | 0.902 |
| Multiple R square of scores with factors | 0.929 | 0.853 | 0.814 |
| Minimum correlation of possible factor scores | 0.858 | 0.706 | 0.627 |

```
> fa.diagram(f3t)
```

Factor Analysis

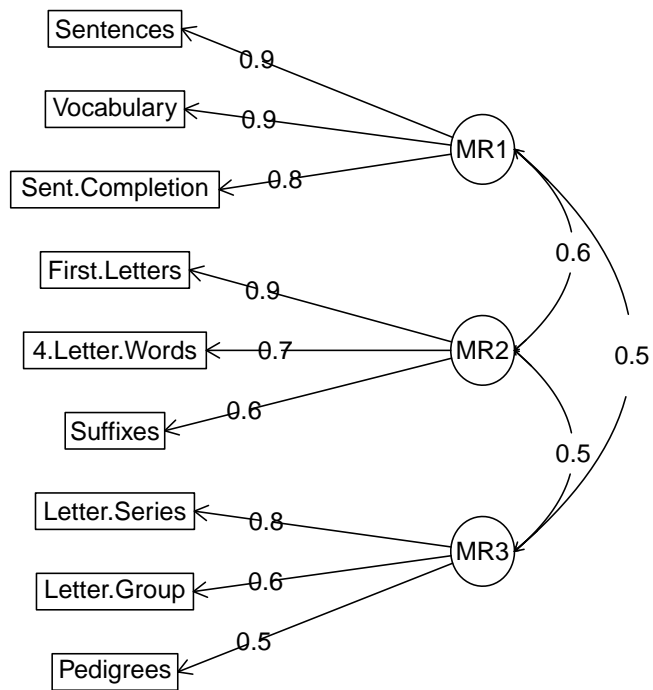


Figure 5: A graphic representation of the 3 oblique factors from the Thurstone data using `fa.diagram`. Factors were transformed to an oblique solution using `oblmin`.

that the minres solution is slightly better. At least in one test data set, the weighted least squares solutions, although fitting equally well, had slightly different structure loadings. Note that the rotations used by SPSS will sometimes use the “Kaiser Normalization”. By default, the rotations used in psych do not normalize, but this can be specified as an option in `fa`.

5.1.4 Principal Components analysis (PCA)

An alternative to factor analysis, which is unfortunately frequently confused with *factor analysis*, is *principal components analysis*. Although the goals of *PCA* and *FA* are similar, *PCA* is a descriptive model of the data, while *FA* is a structural model. Psychologists typically use *PCA* in a manner similar to factor analysis and thus the `principal` function produces output that is perhaps more understandable than that produced by `princomp` in the *stats* package. Table 4 shows a *PCA* of the Thurstone 9 variable problem rotated using the `Promax` function. Note how the loadings from the factor model are similar but smaller than the principal component loadings. This is because the *PCA* model attempts to account for the entire variance of the correlation matrix, while *FA* accounts for just the *common variance*. This distinction becomes most important for small correlation matrices. (Factor loadings and *PCA* loadings converge as the number of items increases. Factor loadings may be thought of as the asymptotic principal component loadings as the number of items on the component grows towards infinity.) Also note how the goodness of fit statistics, based upon the residual off diagonal elements, is much worse than the `fa` solution.

5.1.5 Hierarchical and bi-factor solutions

For a long time structural analysis of the ability domain have considered the problem of factors that are themselves correlated. These correlations may themselves be factored to produce a higher order, general factor. An alternative (Holzinger and Swineford, 1937; Jensen and Weng, 1994) is to consider the general factor affecting each item, and then to have group factors account for the residual variance. Exploratory factor solutions to produce a hierarchical or a bifactor solution are found using the `omega` function. This technique has more recently been applied to the personality domain to consider such things as the structure of neuroticism (treated as a general factor, with lower order factors of anxiety, depression, and aggression).

Consider the 9 Thurstone variables analyzed in the prior factor analyses. The correlations between the factors (as shown in Figure 5) can themselves be factored. This results in a higher order factor model (Figure 6). An alternative solution is to take this higher order model and then solve for the general factor loadings as well as the loadings on the residualized lower order factors using the *Schmid-Leiman* procedure. (Figure 7). Yet

Table 4: The Thurstone problem can also be analyzed using Principal Components Analysis. Compare this to Table 2. The loadings are higher for the PCA because the model accounts for the unique as well as the common variance. The fit of the off diagonal elements, however, is much worse than the `fa` results.

```
> p3p <-principal(Thurstone,3,n.obs = 213,rotate="Promax")
> p3p
Principal Components Analysis
Call: principal(r = Thurstone, nfactors = 3, rotate = "Promax", n.obs = 213)
Standardized loadings (pattern matrix) based upon correlation matrix
```

| | RC1 | RC2 | RC3 | h2 | u2 | com |
|-----------------|-------|-------|-------|------|------|-----|
| Sentences | 0.92 | 0.01 | 0.01 | 0.86 | 0.14 | 1.0 |
| Vocabulary | 0.90 | 0.10 | -0.05 | 0.86 | 0.14 | 1.0 |
| Sent.Completion | 0.91 | 0.04 | -0.04 | 0.83 | 0.17 | 1.0 |
| First.Letters | 0.01 | 0.84 | 0.07 | 0.78 | 0.22 | 1.0 |
| 4.Letter.Words | -0.05 | 0.81 | 0.17 | 0.75 | 0.25 | 1.1 |
| Suffixes | 0.18 | 0.79 | -0.15 | 0.70 | 0.30 | 1.2 |
| Letter.Series | 0.03 | -0.03 | 0.88 | 0.78 | 0.22 | 1.0 |
| Pedigrees | 0.45 | -0.16 | 0.57 | 0.67 | 0.33 | 2.1 |
| Letter.Group | -0.19 | 0.19 | 0.86 | 0.75 | 0.25 | 1.2 |

| | RC1 | RC2 | RC3 |
|-----------------------|------|------|------|
| SS loadings | 2.83 | 2.19 | 1.96 |
| Proportion Var | 0.31 | 0.24 | 0.22 |
| Cumulative Var | 0.31 | 0.56 | 0.78 |
| Proportion Explained | 0.41 | 0.31 | 0.28 |
| Cumulative Proportion | 0.41 | 0.72 | 1.00 |


```
With component correlations of
  RC1  RC2  RC3
RC1 1.00 0.51 0.53
RC2 0.51 1.00 0.44
RC3 0.53 0.44 1.00

Mean item complexity = 1.2
Test of the hypothesis that 3 components are sufficient.

The root mean square of the residuals (RMSR) is 0.06
with the empirical chi square 56.17 with prob < 1.1e-07

Fit based upon off diagonal values = 0.98
```

another solution is to use structural equation modeling to directly solve for the general and group factors.

```
> om.h <- omega(Thurstone, n.obs=213, sl=FALSE)
> op <- par(mfrow=c(1,1))
```

Omega

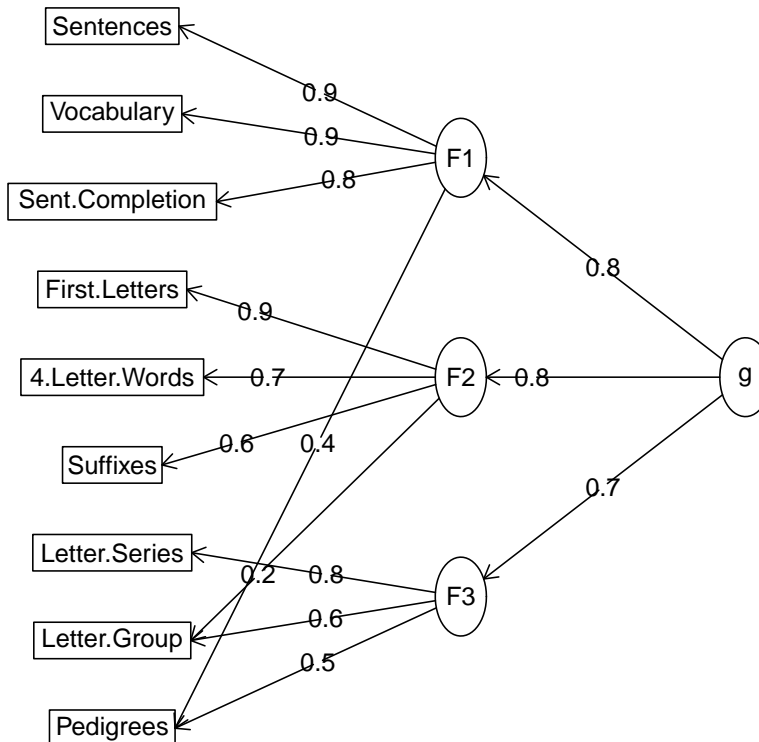


Figure 6: A higher order factor solution to the Thurstone 9 variable problem

Yet another approach to the bifactor structure is to use the `bifactor` rotation function in either `psych` or in `GPArotation`. This does the rotation discussed in [Jennrich and Bentler \(2011\)](#). Unfortunately, this rotation will frequently achieve a local rather than global optimum and it is recommended to try multiple original configurations.

```
> om <- omega(Thurstone, n.obs=213)
```

Omega

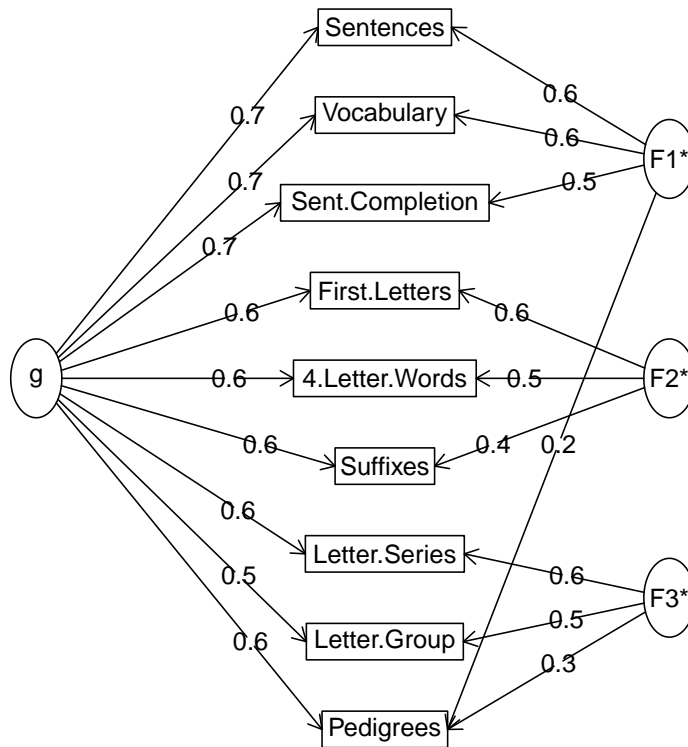


Figure 7: A bifactor factor solution to the Thurstone 9 variable problem

5.1.6 Item Cluster Analysis: `iclust`

An alternative to factor or components analysis is *cluster analysis*. The goal of cluster analysis is the same as factor or components analysis (reduce the complexity of the data and attempt to identify homogeneous subgroupings). Mainly used for clustering people or objects (e.g., projectile points if an anthropologist, DNA if a biologist, galaxies if an astronomer), clustering may be used for clustering items or tests as well. Introduced to psychologists by Tryon (1939) in the 1930's, the cluster analytic literature exploded in the 1970s and 1980s (Blashfield, 1980; Blashfield and Aldenderfer, 1988; Everitt, 1974; Hartigan, 1975). Much of the research is in taxonomic applications in biology (Sneath and Sokal, 1973; Sokal and Sneath, 1963) and marketing (Cooksey and Soutar, 2006) where clustering remains very popular. It is also used for taxonomic work in forming clusters of people in family (Henry et al., 2005) and clinical psychology (Martinet and Ferrand, 2007; Mun et al., 2008). Interestingly enough it has had limited applications to psychometrics. This is unfortunate, for as has been pointed out by e.g. (Tryon, 1935; Loevinger et al., 1953), the theory of factors, while mathematically compelling, offers little that the geneticist or behaviorist or perhaps even non-specialist finds compelling. Cooksey and Soutar (2006) reviews why the `iclust` algorithm is particularly appropriate for scale construction in marketing.

Hierarchical cluster analysis forms clusters that are nested within clusters. The resulting *tree diagram* (also known somewhat pretentiously as a *rooted dendritic structure*) shows the nesting structure. Although there are many hierarchical clustering algorithms in R (e.g., `agnes`, `hclust`, and `iclust`), the one most applicable to the problems of scale construction is `iclust` (Revelle, 1979).

1. Find the proximity (e.g. correlation) matrix,
2. Identify the most similar pair of items
3. Combine this most similar pair of items to form a new variable (cluster),
4. Find the similarity of this cluster to all other items and clusters,
5. Repeat steps 2 and 3 until some criterion is reached (e.g., typically, if only one cluster remains or in `iclust` if there is a failure to increase reliability coefficients α or β).
6. Purify the solution by reassigning items to the most similar cluster center.

`iclust` forms clusters of items using a hierarchical clustering algorithm until one of two measures of internal consistency fails to increase (Revelle, 1979). The number of clusters may be specified a priori, or found empirically. The resulting statistics include the average split half reliability, α (Cronbach, 1951), as well as the worst split half reliability, β (Revelle, 1979), which is an estimate of the general factor saturation of the resulting scale (Figure 8).

Cluster loadings (corresponding to the structure matrix of factor analysis) are reported when printing (Table 7). The pattern matrix is available as an object in the results.

```
> data(bfi)
> ic <- iclust(bfi[1:25])
```

ICLUST

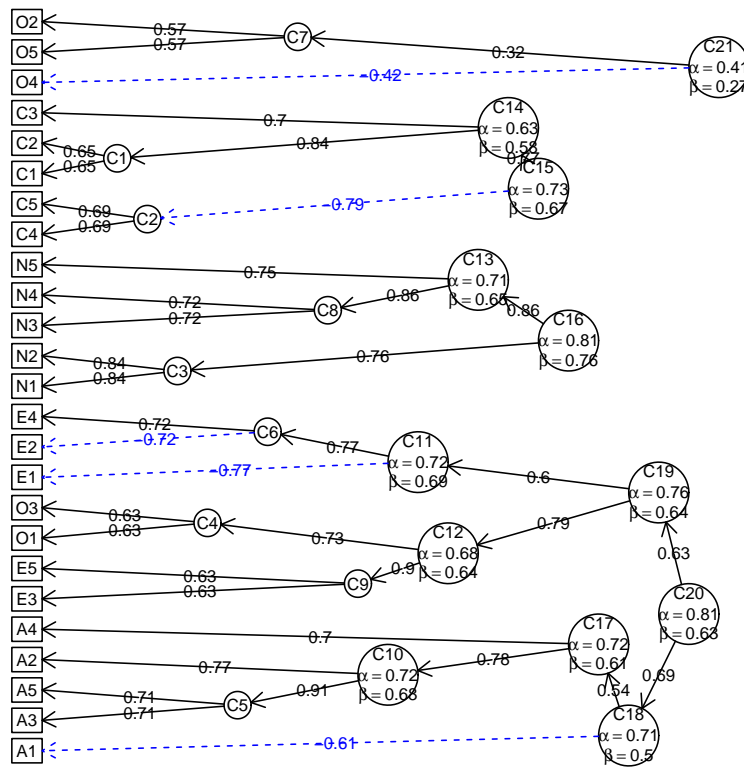


Figure 8: Using the `iclust` function to find the cluster structure of 25 personality items (the three demographic variables were excluded from this analysis). When analyzing many variables, the tree structure may be seen more clearly if the graphic output is saved as a pdf and then enlarged using a pdf viewer.

The previous analysis (Figure 8) was done using the Pearson correlation. A somewhat cleaner structure is obtained when using the `polychoric` function to find polychoric correlations (Figure 9). Note that the first time finding the polychoric correlations some time, but the next three analyses were done using that correlation matrix (`r.poly$rho`). When using the console for input, `polychoric` will report on its progress while working using

Table 5: The summary statistics from an iclust analysis shows three large clusters and smaller cluster.

```
> summary(ic) #show the results
ICLUST (Item Cluster Analysis)Call: iclust(r.mat = bfi[1:25])
ICLUST
```

Purified Alpha:

| C20 | C16 | C15 | C21 |
|------|------|------|------|
| 0.80 | 0.81 | 0.73 | 0.61 |

Guttman Lambda6*

| C20 | C16 | C15 | C21 |
|------|------|------|------|
| 0.82 | 0.81 | 0.72 | 0.61 |

Original Beta:

| C20 | C16 | C15 | C21 |
|------|------|------|------|
| 0.63 | 0.76 | 0.67 | 0.27 |

Cluster size:

| C20 | C16 | C15 | C21 |
|-----|-----|-----|-----|
| 10 | 5 | 5 | 5 |

Purified scale intercorrelations

reliabilities on diagonal

correlations corrected for attenuation above diagonal:

| | C20 | C16 | C15 | C21 |
|-----|-------|--------|-------|-------|
| C20 | 0.80 | -0.291 | 0.40 | -0.33 |
| C16 | -0.24 | 0.815 | -0.29 | 0.11 |
| C15 | 0.30 | -0.221 | 0.73 | -0.30 |
| C21 | -0.23 | 0.074 | -0.20 | 0.61 |

`progressBar`. `polychoric` and `tetrachoric`, when running on OSX will take advantage of the multiple cores on the machine and run somewhat faster.

Table 6: The `polychoric` and the `tetrachoric` functions can take a long time to finish and report their progress by a series of dots as they work. The dots are suppressed when creating a Sweave document.

```
> data(bfi)
> r.poly <- polychoric(bfi[1:25]) #the ... indicate the progress of the function
```

A comparison of these four cluster solutions suggests both a problem and an advantage of clustering techniques. The problem is that the solutions differ. The advantage is that the structure of the items may be seen more clearly when examining the clusters rather than a simple factor solution.

5.2 Confidence intervals using bootstrapping techniques

Exploratory factoring techniques are sometimes criticized because of the lack of statistical information on the solutions. Overall estimates of goodness of fit including χ^2 and RMSEA are found in the `fa` and `omega` functions. Confidence intervals for the factor loadings may be found by doing multiple bootstrapped iterations of the original analysis. This is done by setting the `n.iter` parameter to the desired number of iterations. This can be done for factoring of Pearson correlation matrices as well as polychoric/tetrachoric matrices (See Table 8). Although the example value for the number of iterations is set to 20, more conventional analyses might use 1000 bootstraps. This will take much longer.

5.3 Comparing factor/component/cluster solutions

Cluster analysis, factor analysis, and principal components analysis all produce structure matrices (matrices of correlations between the dimensions and the variables) that can in turn be compared in terms of the *congruence coefficient* which is just the cosine of the angle between the dimensions

$$c_{f_i f_j} = \frac{\sum_{k=1}^n f_{ik} f_{jk}}{\sum f_{ik}^2 \sum f_{jk}^2}.$$

Consider the case of a four factor solution and four cluster solution to the Big Five problem.

```
> f4 <- fa(bfi[1:25], 4, fm="pa")
> factor.congruence(f4, ic)
```

| | C20 | C16 | C15 | C21 |
|-----|-------|-------|-------|-------|
| PA1 | 0.92 | -0.32 | 0.44 | -0.40 |
| PA2 | -0.26 | 0.95 | -0.33 | 0.12 |

```

> ic.poly <- iclust(r.poly$rho,title="ICLUST using polychoric correlations")
> iclust.diagram(ic.poly)

```

ICLUST using polychoric correlations

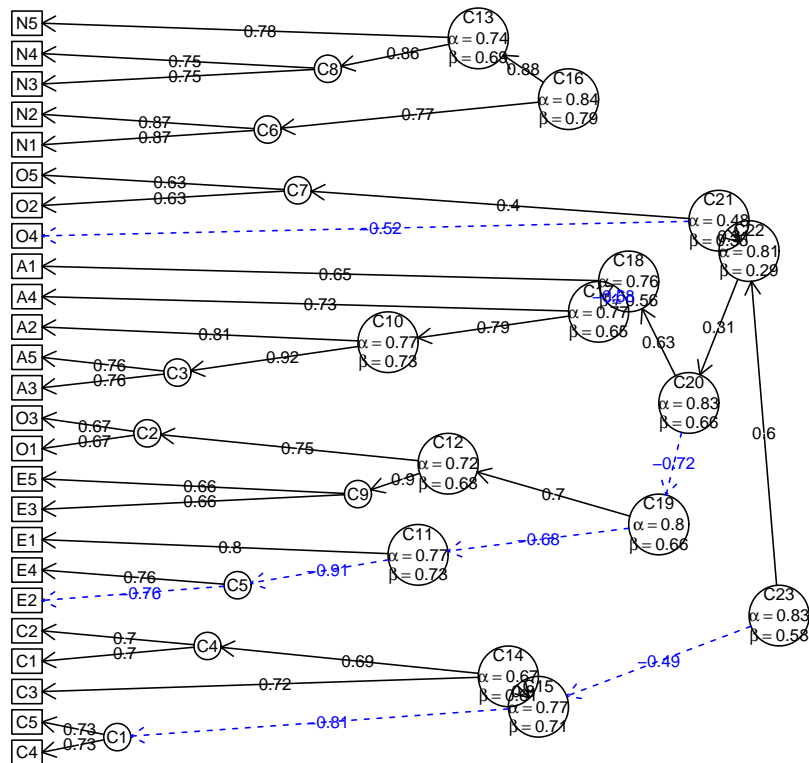


Figure 9: ICLUST of the BFI data set using polychoric correlations. Compare this solution to the previous one (Figure 8) which was done using Pearson correlations.


```
> ic.poly <- iclust(r.poly$rho,5,title="ICLUST using polychoric correlations for nclusters=5")
> iclust.diagram(ic.poly)
```

ICLUST using polychoric correlations for nclusters=5

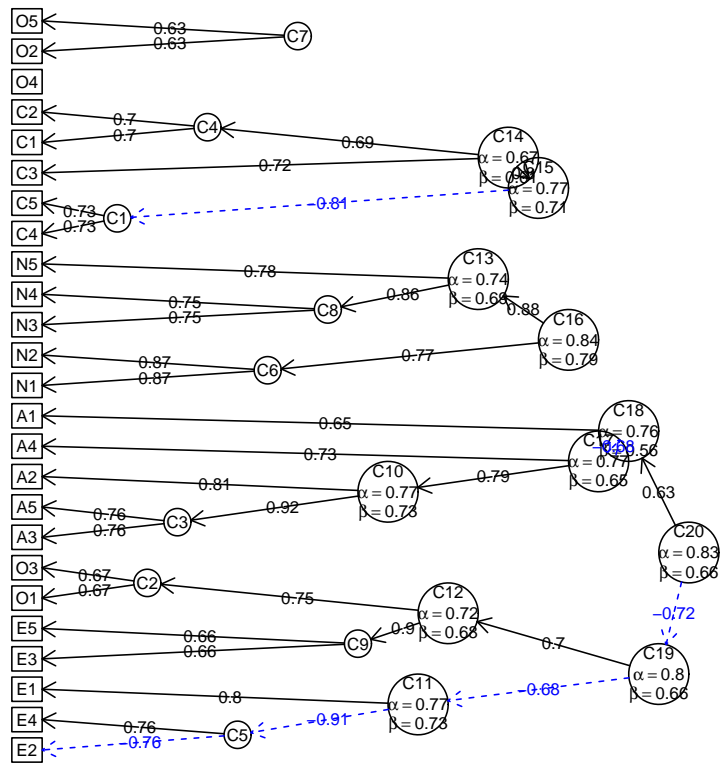


Figure 10: ICLUST of the BFI data set using polychoric correlations with the solution set to 5 clusters. Compare this solution to the previous one (Figure 9) which was done without specifying the number of clusters and to the next one (Figure 11) which was done by changing the beta criterion.

```
> ic.poly <- iclust(r.poly$rho,beta.size=3,title="ICLUST beta.size=3")
```

ICLUST beta.size=3

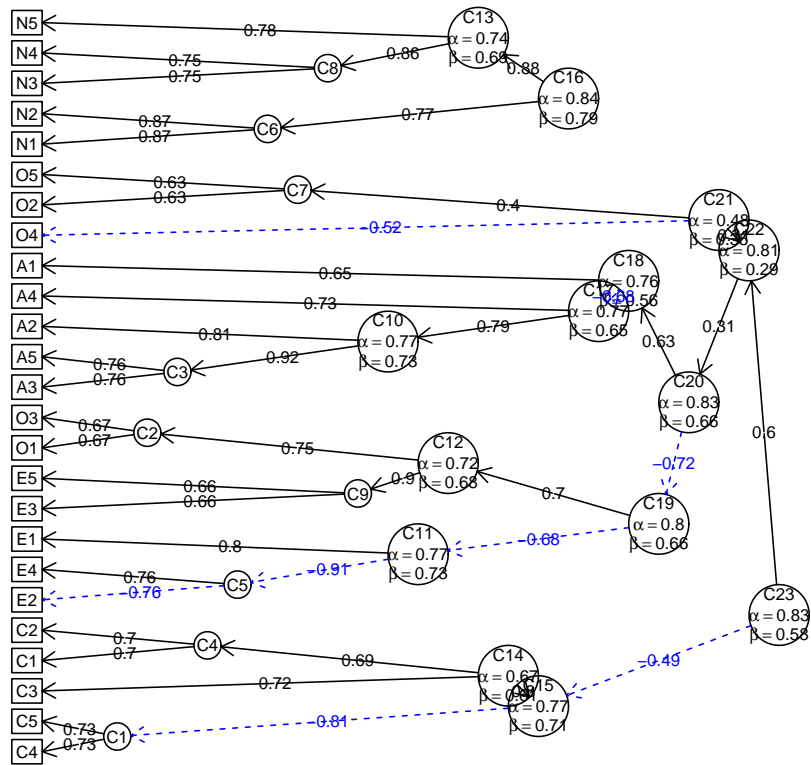


Figure 11: ICLUST of the BFI data set using polychoric correlations with the beta criterion set to 3. Compare this solution to the previous three (Figure 8, 9, 10).

Table 7: The output from `iclust` includes the loadings of each item on each cluster. These are equivalent to factor structure loadings. By specifying the value of `cut`, small loadings are suppressed. The default is for `cut=0.su`

```

> print(ic, cut=.3)
ICLUST (Item Cluster Analysis)
Call: iclust(r.mat = bfi[1:25])

Purified Alpha:
  C20 C16 C15 C21
0.80 0.81 0.73 0.61

G6* reliability:
  C20 C16 C15 C21
0.83 1.00 0.67 0.38

Original Beta:
  C20 C16 C15 C21
0.63 0.76 0.67 0.27

Cluster size:
C20 C16 C15 C21
 10   5   5   5

Item by Cluster Structure matrix:
   O  P  C20  C16  C15  C21
A1 C20 C20
A2 C20 C20  0.59
A3 C20 C20  0.65
A4 C20 C20  0.43
A5 C20 C20  0.65
C1 C15 C15           0.54
C2 C15 C15           0.62
C3 C15 C15           0.54
C4 C15 C15  0.31 -0.66
C5 C15 C15 -0.30  0.36 -0.59
E1 C20 C20 -0.50
E2 C20 C20 -0.61  0.34
E3 C20 C20  0.59           -0.39
E4 C20 C20  0.66
E5 C20 C20  0.50           0.40 -0.32
N1 C16 C16           0.76
N2 C16 C16           0.75
N3 C16 C16           0.74
N4 C16 C16 -0.34  0.62
N5 C16 C16           0.55
O1 C20 C21           -0.53
O2 C21 C21           0.44
O3 C20 C21  0.39           -0.62
O4 C21 C21           -0.33
O5 C21 C21           0.53

With eigenvalues of:
C20 C16 C15 C21
3.2 2.6 1.9 1.5

Purified scale intercorrelations
reliabilities on diagonal
correlations corrected for attenuation above diagonal:
  C20  C16  C15  C21
C20  0.80 -0.29  0.40 -0.33
C16 -0.24  0.81 -0.29  0.11
C15  0.30 -0.22  0.73 -0.30
C21 -0.23  0.07 -0.20  0.61

Cluster fit = 0.68  Pattern fit = 0.96  RMSR = 0.05
NULL

```

Table 8: An example of bootstrapped confidence intervals on 10 items from the Big 5 inventory. The number of bootstrapped samples was set to 20. More conventional bootstrapping would use 100 or 1000 replications.

```

> fa(bfi[1:10],2,n.iter=20)
Factor Analysis with confidence intervals using method = fa(r = bfi[1:10], nfactors = 2, n.iter = 20)
Factor Analysis using method = minres
Call: fa(r = bfi[1:10], nfactors = 2, n.iter = 20)
Standardized loadings (pattern matrix) based upon correlation matrix
  MR2  MR1  h2  u2 com
A1  0.07 -0.40 0.15 0.85 1.1
A2  0.02  0.65 0.44 0.56 1.0
A3 -0.03  0.77 0.57 0.43 1.0
A4  0.15  0.44 0.26 0.74 1.2
A5  0.02  0.62 0.39 0.61 1.0
C1  0.57 -0.05 0.30 0.70 1.0
C2  0.62 -0.01 0.39 0.61 1.0
C3  0.54  0.03 0.30 0.70 1.0
C4 -0.66  0.01 0.43 0.57 1.0
C5 -0.57 -0.05 0.35 0.65 1.0

  MR2  MR1
SS loadings      1.80 1.77
Proportion Var   0.18 0.18
Cumulative Var   0.18 0.36
Proportion Explained 0.50 0.50
Cumulative Proportion 0.50 1.00

With factor correlations of
  MR2  MR1
MR2 1.00 0.32
MR1 0.32 1.00

Mean item complexity = 1
Test of the hypothesis that 2 factors are sufficient.

The degrees of freedom for the null model are 45 and the objective function was 2.03 with Chi Square of 5664.89
The degrees of freedom for the model are 26 and the objective function was 0.17

The root mean square of the residuals (RMSR) is 0.04
The df corrected root mean square of the residuals is 0.05

The harmonic number of observations is 2762 with the empirical chi square 403.38 with prob < 2.6e-69
The total number of observations was 2800 with MLE Chi Square = 464.04 with prob < 9.2e-82

Tucker Lewis Index of factoring reliability = 0.865
RMSEA index = 0.078 and the 90 % confidence intervals are 0.071 0.084
BIC = 257.67
Fit based upon off diagonal values = 0.98
Measures of factor score adequacy
  MR2  MR1
Correlation of scores with factors      0.86 0.88
Multiple R square of scores with factors 0.74 0.77
Minimum correlation of possible factor scores 0.49 0.54

Coefficients and bootstrapped confidence intervals
  low  MR2 upper  low  MR1 upper
A1  0.03 0.07 0.11 -0.46 -0.40 -0.36
A2 -0.02 0.02 0.05 0.61 0.65 0.70
A3 -0.06 -0.03 0.00 0.74 0.77 0.80
A4 0.10 0.15 0.18 0.41 0.44 0.48
A5 -0.02 0.02 0.07 0.58 0.62 0.66
C1 0.52 0.57 0.60 -0.08 -0.05 -0.02
C2 0.59 0.62 0.66 -0.04 -0.01 0.02
C3 0.50 0.54 0.57 -0.01 0.03 0.08
C4 -0.70 -0.66 -0.62 -0.02 0.01 0.03
C5 -0.62 -0.57 -0.53 -0.08 -0.05 -0.01

Interfactor correlations and bootstrapped confidence intervals
  lower estimate upper
MR2-MR1 0.28 0.32 0.37
>

```

```
PA3  0.35 -0.24  0.88 -0.37
PA4  0.29 -0.12  0.27 -0.90
```

A more complete comparison of oblique factor solutions (both minres and principal axis), bifactor and component solutions to the Thurstone data set is done using the `factor.congruence` function. (See table 9).

Table 9: Congruence coefficients for oblique factor, bifactor and component solutions for the Thurstone problem.

```
> factor.congruence(list(f3t, f3o, om, p3p))
      MR1 MR2 MR3 PA1 PA2 PA3 g F1* F2* F3* h2 RC1 RC2 RC3
MR1  1.00 0.06 0.09  1.00 0.06 0.13 0.72 1.00 0.06 0.09 0.74 1.00 0.08  0.04
MR2  0.06 1.00 0.08  0.03 1.00 0.06 0.60 0.06 1.00 0.08 0.57 0.04 0.99  0.12
MR3  0.09 0.08 1.00  0.01 0.01 1.00 0.52 0.09 0.08 1.00 0.51 0.06 0.02  0.99
PA1  1.00 0.03 0.01  1.00 0.04 0.05 0.67 1.00 0.03 0.01 0.69 1.00 0.06 -0.04
PA2  0.06 1.00 0.01  0.04 1.00 0.00 0.57 0.06 1.00 0.01 0.54 0.04 0.99  0.05
PA3  0.13 0.06 1.00  0.05 0.00 1.00 0.54 0.13 0.06 1.00 0.53 0.10 0.01  0.99
g    0.72 0.60 0.52  0.67 0.57 0.54 1.00 0.72 0.60 0.52 0.99 0.69 0.58  0.50
F1*  1.00 0.06 0.09  1.00 0.06 0.13 0.72 1.00 0.06 0.09 0.74 1.00 0.08  0.04
F2*  0.06 1.00 0.08  0.03 1.00 0.06 0.60 0.06 1.00 0.08 0.57 0.04 0.99  0.12
F3*  0.09 0.08 1.00  0.01 0.01 1.00 0.52 0.09 0.08 1.00 0.51 0.06 0.02  0.99
h2   0.74 0.57 0.51  0.69 0.54 0.53 0.99 0.74 0.57 0.51 1.00 0.71 0.56  0.49
RC1  1.00 0.04 0.06  1.00 0.04 0.10 0.69 1.00 0.04 0.06 0.71 1.00 0.06  0.00
RC2  0.08 0.99 0.02  0.06 0.99 0.01 0.58 0.08 0.99 0.02 0.56 0.06 1.00  0.05
RC3  0.04 0.12 0.99 -0.04 0.05 0.99 0.50 0.04 0.12 0.99 0.49 0.00 0.05  1.00
```

5.4 Determining the number of dimensions to extract.

How many dimensions to use to represent a correlation matrix is an unsolved problem in psychometrics. There are many solutions to this problem, none of which is uniformly the best. Henry Kaiser once said that “a solution to the number-of factors problem in factor analysis is easy, that he used to make up one every morning before breakfast. But the problem, of course is to find *the* solution, or at least a solution that others will regard quite highly not as the best” [Horn and Engstrom \(1979\)](#).

Techniques most commonly used include

- 1) Extracting factors until the chi square of the residual matrix is not significant.
- 2) Extracting factors until the change in chi square from factor n to factor n+1 is not significant.
- 3) Extracting factors until the eigen values of the real data are less than the corresponding eigen values of a random data set of the same size (parallel analysis) `fa.parallel` ([Horn, 1965](#)).
- 4) Plotting the magnitude of the successive eigen values and applying the scree test (a sudden drop in eigen values analogous to the change in slope seen when scrambling up the

talus slope of a mountain and approaching the rock face (Cattell, 1966).

- 5) Extracting factors as long as they are interpretable.
- 6) Using the Very Structure Criterion (`vss`) (Revelle and Rocklin, 1979).
- 7) Using Wayne Velicer's Minimum Average Partial (`MAP`) criterion (Velicer, 1976).
- 8) Extracting principal components until the eigen value < 1 .

Each of the procedures has its advantages and disadvantages. Using either the chi square test or the change in square test is, of course, sensitive to the number of subjects and leads to the nonsensical condition that if one wants to find many factors, one simply runs more subjects. Parallel analysis is partially sensitive to sample size in that for large samples ($N > 1,000$) all of the eigen values of random factors will be very close to 1. The scree test is quite appealing but can lead to differences of interpretation as to when the scree "breaks". Extracting interpretable factors means that the number of factors reflects the investigators creativity more than the data. `vss`, while very simple to understand, will not work very well if the data are very factorially complex. (Simulations suggests it will work fine if the complexities of some of the items are no more than 2). The eigen value of 1 rule, although the default for many programs, seems to be a rough way of dividing the number of variables by 3 and is probably the worst of all criteria.

An additional problem in determining the number of factors is what is considered a factor. Many treatments of factor analysis assume that the residual correlation matrix after the factors of interest are extracted is composed of just random error. An alternative concept is that the matrix is formed from major factors of interest but that there are also numerous minor factors of no substantive interest but that account for some of the shared covariance between variables. The presence of such minor factors can lead one to extract too many factors and to reject solutions on statistical grounds of misfit that are actually very good fits to the data. This problem is partially addressed later in the discussion of simulating complex structures using `sim.structure` and of small extraneous factors using the `sim.minor` function.

5.4.1 Very Simple Structure

The `vss` function compares the fit of a number of factor analyses with the loading matrix "simplified" by deleting all except the c greatest loadings per item, where c is a measure of factor complexity (Revelle and Rocklin (1979)). Included in `vss` is the `MAP` criterion (Minimum Absolute Partial correlation) of (Velicer (1976)).

Using the Very Simple Structure criterion for the `bfi` data suggests that 4 factors are optimal (Figure 12). However, the `MAP` criterion suggests that 5 is optimal.

```
> vss <- vss(bfi[1:25],title="Very Simple Structure of a Big 5 inventory")
```

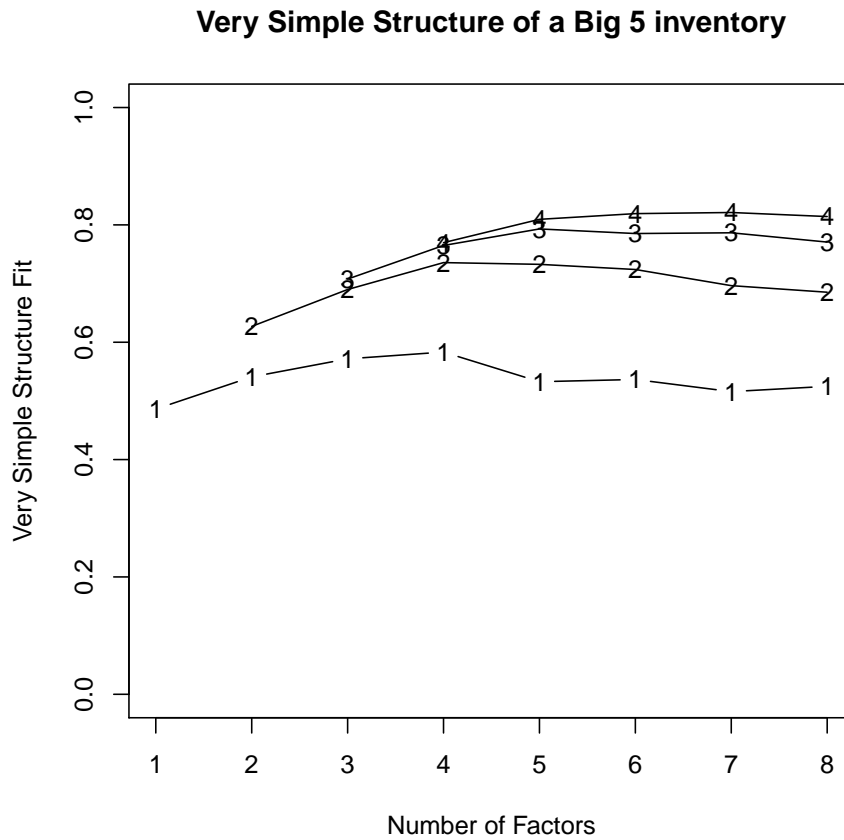


Figure 12: The Very Simple Structure criterion for the number of factors compares solutions for various levels of item complexity and various numbers of factors. For the Big 5 Inventory, the complexity 1 and 2 solutions both achieve their maxima at four factors. This is in contrast to parallel analysis which suggests 6 and the MAP criterion which suggests 5.

```

> vss
Very Simple Structure of Very Simple Structure of a Big 5 inventory
Call: vss(x = bfi[1:25], title = "Very Simple Structure of a Big 5 inventory")
VSS complexity 1 achieves a maximum of 0.58 with 4 factors
VSS complexity 2 achieves a maximum of 0.74 with 4 factors

The Velicer MAP achieves a minimum of 0.01 with 5 factors
BIC achieves a minimum of -524.26 with 8 factors
Sample Size adjusted BIC achieves a minimum of -117.56 with 8 factors

Statistics by number of factors
  vss1 vss2  map dof chisq  prob sqresid  fit RMSEA  BIC  SABIC complex eChisq  SRMR eCRMS  eBIC
1 0.49 0.00 0.024 275 11831 0.0e+00 26.0 0.49 0.123 9648 10522.1 1.0 23881 0.119 0.125 21698
2 0.54 0.63 0.018 251 7279 0.0e+00 18.9 0.63 0.100 5287 6084.5 1.2 12432 0.086 0.094 10440
3 0.57 0.69 0.017 228 5010 0.0e+00 14.8 0.71 0.087 3200 3924.3 1.3 7232 0.066 0.075 5422
4 0.58 0.74 0.015 206 3366 0.0e+00 11.7 0.77 0.074 1731 2385.1 1.4 3750 0.047 0.057 2115
5 0.53 0.73 0.015 185 1750 1.4e-252 9.5 0.81 0.055 281 869.3 1.6 1495 0.030 0.038 27
6 0.54 0.72 0.016 165 1014 4.4e-122 8.4 0.84 0.043 -296 228.5 1.7 670 0.020 0.027 -639
7 0.52 0.70 0.019 146 696 1.4e-72 7.9 0.84 0.037 -463 1.2 1.9 448 0.016 0.023 -711
8 0.52 0.69 0.022 128 492 4.7e-44 7.4 0.85 0.032 -524 -117.6 1.9 289 0.013 0.020 -727

```

5.4.2 Parallel Analysis

An alternative way to determine the number of factors is to compare the solution to random data with the same properties as the real data set. If the input is a data matrix, the comparison includes random samples from the real data, as well as normally distributed random data with the same number of subjects and variables. For the BFI data, parallel analysis suggests that 6 factors might be most appropriate (Figure 13). It is interesting to compare `fa.parallel` with the `paran` from the *paran* package. This latter uses `smcs` to estimate communalities. Simulations of known structures with a particular number of major factors but with the presence of trivial, minor (but not zero) factors, show that using `smcs` will tend to lead to too many factors.

A more tedious problem in terms of computation is to do parallel analysis of *polychoric* correlation matrices. This is done by `fa.parallel.poly` or `fa.parallel` with the `option="poly"`. By default the number of replications is 20. This is appropriate when choosing the number of factors from dichotomous or polytomous data matrices.

5.5 Factor extension

Sometimes we are interested in the relationship of the factors in one space with the variables in a different space. One solution is to find factors in both spaces separately and then find the structural relationships between them. This is the technique of structural equation modeling in packages such as *sem* or *lavaan*. An alternative is to use the concept of *factor extension* developed by (Dwyer, 1937). Consider the case of 16 variables created


```
> fa.parallel(bfi[1:25],main="Parallel Analysis of a Big 5 inventory")
Parallel analysis suggests that the number of factors = 6 and the number of components = 6
```

Parallel Analysis of a Big 5 inventory

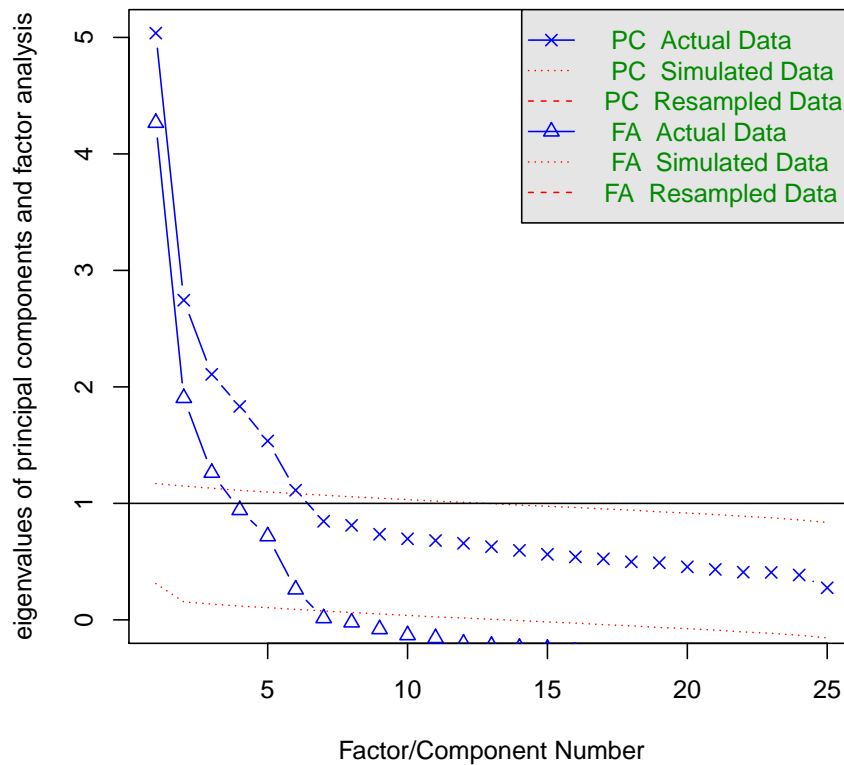


Figure 13: Parallel analysis compares factor and principal components solutions to the real data as well as resampled data. Although vss suggests 4 factors, MAP 5, parallel analysis suggests 6. One more demonstration of Kaiser’s dictum.

to represent one two dimensional space. If factors are found from eight of these variables, they may then be extended to the additional eight variables (See Figure 14).

Another way to examine the overlap between two sets is the use of *set correlation* found by `set.cor` (discussed later).

6 Classical Test Theory and Reliability

Surprisingly, 112 years after Spearman (1904) introduced the concept of reliability to psychologists, there are still multiple approaches for measuring it. Although very popular, Cronbach’s α (Cronbach, 1951) underestimates the reliability of a test and over estimates the first factor saturation (Revelle and Zinbarg, 2009).

α (Cronbach, 1951) is the same as Guttman’s λ_3 (Guttman, 1945) and may be found by

$$\lambda_3 = \frac{n}{n-1} \left(1 - \frac{\text{tr}(V)_x}{V_x} \right) = \frac{n}{n-1} \frac{V_x - \text{tr}(V_x)}{V_x} = \alpha$$

Perhaps because it is so easy to calculate and is available in most commercial programs, alpha is without doubt the most frequently reported measure of internal consistency reliability. For τ equivalent items, alpha is the mean of all possible split half reliabilities (corrected for test length). For a unifactorial test, it is a reasonable estimate of the first factor saturation, although if the test has any microstructure (i.e., if it is “lumpy”) coefficients β (Revelle, 1979) (see `iclust`) and ω_h (see `omega`) are more appropriate estimates of the general factor saturation. ω is a better estimate of the reliability of the total test.

Guttman’s λ_6 (G6) considers the amount of variance in each item that can be accounted for the linear regression of all of the other items (the squared multiple correlation or `smc`), or more precisely, the variance of the errors, e_j^2 , and is

$$\lambda_6 = 1 - \frac{\sum e_j^2}{V_x} = 1 - \frac{\sum (1 - r_{smc}^2)}{V_x}.$$

The squared multiple correlation is a lower bound for the item communality and as the number of items increases, becomes a better estimate.

G6 is also sensitive to lumpiness in the test and should not be taken as a measure of unifactorial structure. For lumpy tests, it will be greater than alpha. For tests with equal item loadings, $\alpha > G6$, but if the loadings are unequal or if there is a general factor, $G6 > \alpha$. G6 estimates item reliability by the squared multiple correlation of the other items in a scale. A modification of G6, $G6^*$, takes as an estimate of an item reliability

```

> v16 <- sim.item(16)
> s <- c(1,3,5,7,9,11,13,15)
> f2 <- fa(v16[s],2)
> fe <- fa.extension(cor(v16)[s,-s],f2)
> fa.diagram(f2,fe=fe)

```

Factor analysis and extension

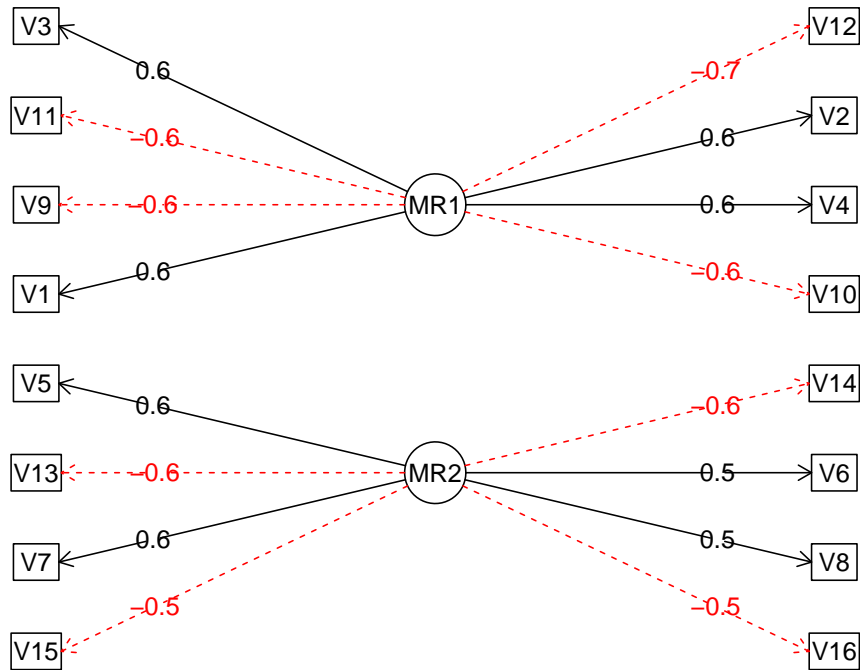


Figure 14: Factor extension applies factors from one set (those on the left) to another set of variables (those on the right). `fa.extension` is particularly useful when one wants to define the factors with one set of variables and then apply those factors to another set. `fa.diagram` is used to show the structure.

the smc with all the items in an inventory, including those not keyed for a particular scale. This will lead to a better estimate of the reliable variance of a particular item.

Alpha, G6 and G6* are positive functions of the number of items in a test as well as the average intercorrelation of the items in the test. When calculated from the item variances and total test variance, as is done here, raw alpha is sensitive to differences in the item variances. Standardized alpha is based upon the correlations rather than the covariances.

More complete reliability analyses of a single scale can be done using the **omega** function which finds ω_h and ω_t based upon a hierarchical factor analysis.

Alternative functions **scoreItems** and **cluster.cor** will also score multiple scales and report more useful statistics. “Standardized” alpha is calculated from the inter-item correlations and will differ from raw alpha.

Functions for examining the reliability of a single scale or a set of scales include:

alpha Internal consistency measures of reliability range from ω_h to α to ω_t . The **alpha** function reports two estimates: Cronbach’s coefficient α and Guttman’s λ_6 . Also reported are item - whole correlations, α if an item is omitted, and item means and standard deviations.

guttman Eight alternative estimates of test reliability include the six discussed by [Guttman \(1945\)](#), four discussed by ten Berge and Zegers (1978) ($\mu_0 \dots \mu_3$) as well as β (the worst split half, [Revelle, 1979](#)), the glb (greatest lowest bound) discussed by Bentler and Woodward (1980), and ω_h and ω_t ([McDonald, 1999](#); [Zinbarg et al., 2005](#)).

omega Calculate McDonald’s omega estimates of general and total factor saturation. ([Revelle and Zinbarg \(2009\)](#) compare these coefficients with real and artificial data sets.)

cluster.cor Given a $n \times c$ cluster definition matrix of -1s, 0s, and 1s (the keys) , and a $n \times n$ correlation matrix, find the correlations of the composite clusters.

scoreItems Given a matrix or data.frame of k keys for m items (-1, 0, 1), and a matrix or data.frame of items scores for m items and n people, find the sum scores or average scores for each person and each scale. If the input is a square matrix, then it is assumed that correlations or covariances were used, and the raw scores are not available. In addition, report Cronbach’s alpha, coefficient G6*, the average r, the scale intercorrelations, and the item by scale correlations (both raw and corrected for item overlap and scale reliability). Replace missing values with the item median or mean if desired. Will adjust scores for reverse scored items.

score.multiple.choice Ability tests are typically multiple choice with one right answer. **score.multiple.choice** takes a scoring key and a data matrix (or data.frame) and finds

total or average number right for each participant. Basic test statistics (alpha, average r, item means, item-whole correlations) are also reported.

6.1 Reliability of a single scale

A conventional (but non-optimal) estimate of the internal consistency reliability of a test is coefficient α (Cronbach, 1951). Alternative estimates are Guttman's λ_6 , Revelle's β , McDonald's ω_h and ω_t . Consider a simulated data set, representing 9 items with a hierarchical structure and the following correlation matrix. Then using the `alpha` function, the α and λ_6 estimates of reliability may be found for all 9 items, as well as the if one item is dropped at a time.

```
> set.seed(17)
> r9 <- sim.hierarchical(n=500,raw=TRUE)$observed
> round(cor(r9),2)

      V1  V2  V3  V4  V5  V6  V7  V8  V9
V1 1.00 0.58 0.59 0.41 0.44 0.30 0.40 0.31 0.25
V2 0.58 1.00 0.48 0.35 0.36 0.22 0.24 0.29 0.19
V3 0.59 0.48 1.00 0.32 0.35 0.23 0.29 0.20 0.17
V4 0.41 0.35 0.32 1.00 0.44 0.36 0.26 0.27 0.22
V5 0.44 0.36 0.35 0.44 1.00 0.32 0.24 0.23 0.20
V6 0.30 0.22 0.23 0.36 0.32 1.00 0.26 0.26 0.12
V7 0.40 0.24 0.29 0.26 0.24 0.26 1.00 0.38 0.25
V8 0.31 0.29 0.20 0.27 0.23 0.26 0.38 1.00 0.25
V9 0.25 0.19 0.17 0.22 0.20 0.12 0.25 0.25 1.00

> alpha(r9)

Reliability analysis
Call: alpha(x = r9)

      raw_alpha std.alpha G6(smc) average_r S/N  ase  mean  sd
      0.8      0.8      0.8      0.31  4 0.013 0.017 0.63

lower alpha upper      95% confidence boundaries
0.77 0.8 0.83

Reliability if an item is dropped:
      raw_alpha std.alpha G6(smc) average_r S/N alpha se
V1      0.75      0.75  0.75      0.28 3.1  0.017
V2      0.77      0.77  0.77      0.30 3.4  0.015
V3      0.77      0.77  0.77      0.30 3.4  0.015
V4      0.77      0.77  0.77      0.30 3.4  0.015
V5      0.78      0.78  0.77      0.30 3.5  0.015
V6      0.79      0.79  0.79      0.32 3.8  0.014
V7      0.78      0.78  0.78      0.31 3.6  0.015
V8      0.79      0.79  0.78      0.32 3.7  0.014
V9      0.80      0.80  0.80      0.34 4.0  0.013

Item statistics
      n raw.r std.r r.cor r.drop  mean  sd
V1 500 0.77 0.76 0.76 0.67 0.0327 1.02
V2 500 0.67 0.66 0.62 0.55 0.1071 1.00
```

```

V3 500 0.65 0.65 0.61 0.53 -0.0462 1.01
V4 500 0.65 0.65 0.59 0.53 -0.0091 0.98
V5 500 0.64 0.64 0.58 0.52 -0.0012 1.03
V6 500 0.55 0.55 0.46 0.41 -0.0499 0.99
V7 500 0.60 0.60 0.52 0.46 0.0481 1.01
V8 500 0.58 0.57 0.49 0.43 0.0526 1.07
V9 500 0.47 0.48 0.36 0.32 0.0164 0.97

```

Some scales have items that need to be reversed before being scored. Rather than reversing the items in the raw data, it is more convenient to just specify which items need to be reversed scored. This may be done in `alpha` by specifying a `keys` vector of 1s and -1s. (This concept of keys vector is more useful when scoring multiple scale inventories, see below.) As an example, consider scoring the 7 attitude items in the attitude data set. Assume a conceptual mistake in that item 2 is to be scored (incorrectly) negatively.

```

> keys <- c(1,-1,1,1,1,1,1)
> alpha(attitude,keys)

```

Reliability analysis

Call: `alpha(x = attitude, keys = keys)`

| raw_alpha | std.alpha | G6(smc) | average_r | S/N | ase | mean | sd |
|-----------|-----------|---------|-----------|-----|------|------|-----|
| 0.43 | 0.52 | 0.75 | 0.14 | 1.1 | 0.13 | 58 | 5.5 |

| lower | alpha | upper | 95% confidence boundaries |
|-------|-------|-------|---------------------------|
| 0.18 | 0.43 | 0.68 | |

Reliability if an item is dropped:

| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | alpha | se |
|-------------|-----------|-----------|---------|-----------|------|-------|----|
| rating | 0.32 | 0.44 | 0.67 | 0.114 | 0.77 | 0.170 | |
| complaints- | 0.80 | 0.80 | 0.82 | 0.394 | 3.89 | 0.057 | |
| privileges | 0.27 | 0.41 | 0.72 | 0.103 | 0.69 | 0.176 | |
| learning | 0.14 | 0.31 | 0.64 | 0.069 | 0.44 | 0.207 | |
| raises | 0.14 | 0.27 | 0.61 | 0.059 | 0.38 | 0.205 | |
| critical | 0.36 | 0.47 | 0.76 | 0.130 | 0.90 | 0.139 | |
| advance | 0.21 | 0.34 | 0.66 | 0.079 | 0.51 | 0.171 | |

Item statistics

| | n | raw.r | std.r | r.cor | r.drop | mean | sd |
|-------------|----|-------|-------|-------|--------|------|------|
| rating | 30 | 0.60 | 0.60 | 0.60 | 0.33 | 65 | 12.2 |
| complaints- | 30 | -0.57 | -0.58 | -0.74 | -0.74 | 50 | 13.3 |
| privileges | 30 | 0.66 | 0.65 | 0.54 | 0.42 | 53 | 12.2 |
| learning | 30 | 0.80 | 0.79 | 0.78 | 0.64 | 56 | 11.7 |
| raises | 30 | 0.82 | 0.83 | 0.85 | 0.69 | 65 | 10.4 |
| critical | 30 | 0.50 | 0.53 | 0.35 | 0.27 | 75 | 9.9 |
| advance | 30 | 0.74 | 0.75 | 0.71 | 0.58 | 43 | 10.3 |

Note how the reliability of the 7 item scales with an incorrectly reversed item is very poor, but if the item 2 is dropped then the reliability is improved substantially. This suggests that item 2 was incorrectly scored. Doing the analysis again with item 2 positively scored produces much more favorable results.

```

> keys <- c(1,1,1,1,1,1,1)
> alpha(attitude,keys)

```

Reliability analysis

Call: alpha(x = attitude, keys = keys)

```
raw_alpha std.alpha G6(smc) average_r S/N ase mean sd
0.84      0.84      0.88      0.43 5.2 0.042 60 8.2
```

```
lower alpha upper      95% confidence boundaries
0.76 0.84 0.93
```

Reliability if an item is dropped:

```
raw_alpha std.alpha G6(smc) average_r S/N alpha se
rating      0.81      0.81      0.83      0.41 4.2 0.052
complaints  0.80      0.80      0.82      0.39 3.9 0.057
privileges  0.83      0.82      0.87      0.44 4.7 0.048
learning    0.80      0.80      0.84      0.40 4.0 0.054
raises      0.80      0.78      0.83      0.38 3.6 0.056
critical    0.86      0.86      0.89      0.51 6.3 0.038
advance     0.84      0.83      0.86      0.46 5.0 0.043
```

Item statistics

```
n raw.r std.r r.cor r.drop mean sd
rating 30 0.78 0.76 0.75 0.67 65 12.2
complaints 30 0.84 0.81 0.82 0.74 67 13.3
privileges 30 0.70 0.68 0.60 0.56 53 12.2
learning 30 0.81 0.80 0.78 0.71 56 11.7
raises 30 0.85 0.86 0.85 0.79 65 10.4
critical 30 0.42 0.45 0.31 0.27 75 9.9
advance 30 0.60 0.62 0.56 0.46 43 10.3
```

It is useful when considering items for a potential scale to examine the item distribution.

This is done in `scoreItems` as well as in `alpha`.

```
> items <- sim.congeneric(N=500,short=FALSE,low=-2,high=2,categorical=TRUE) #500 responses to 4 discrete items
> alpha(items$observed) #item response analysis of congeneric measures
```

Reliability analysis

Call: alpha(x = items\$observed)

```
raw_alpha std.alpha G6(smc) average_r S/N ase mean sd
0.72      0.72      0.67      0.39 2.6 0.021 0.056 0.73
```

```
lower alpha upper      95% confidence boundaries
0.68 0.72 0.76
```

Reliability if an item is dropped:

```
raw_alpha std.alpha G6(smc) average_r S/N alpha se
V1      0.59      0.59      0.49      0.32 1.4 0.032
V2      0.65      0.65      0.56      0.38 1.9 0.027
V3      0.67      0.67      0.59      0.41 2.0 0.026
V4      0.72      0.72      0.64      0.46 2.5 0.022
```

Item statistics

```
n raw.r std.r r.cor r.drop mean sd
V1 500 0.81 0.81 0.74 0.62 0.058 0.97
V2 500 0.74 0.75 0.63 0.52 0.012 0.98
V3 500 0.72 0.72 0.57 0.48 0.056 0.99
V4 500 0.68 0.67 0.48 0.41 0.098 1.02
```

```

Non missing response frequency for each item
      -2  -1   0   1   2 miss
V1 0.04 0.24 0.40 0.25 0.07   0
V2 0.06 0.23 0.40 0.24 0.06   0
V3 0.05 0.25 0.37 0.26 0.07   0
V4 0.06 0.21 0.37 0.29 0.07   0

```

6.2 Using omega to find the reliability of a single scale

Two alternative estimates of reliability that take into account the hierarchical structure of the inventory are McDonald’s ω_h and ω_r . These may be found using the `omega` function for an exploratory analysis (See Figure 15) or `omegaSem` for a confirmatory analysis using the `sem` package (Fox et al., 2013) based upon the exploratory solution from `omega`.

McDonald has proposed coefficient omega (hierarchical) (ω_h) as an estimate of the general factor saturation of a test. Zinbarg et al. (2005) <http://personality-project.org/revelle/publications/zinbarg.revelle.pmet.05.pdf> compare McDonald’s ω_h to Cronbach’s α and Revelle’s β . They conclude that ω_h is the best estimate. (See also Zinbarg et al. (2006) and Revelle and Zinbarg (2009) <http://personality-project.org/revelle/publications/revelle.zinbarg.08.pdf>).

One way to find ω_h is to do a factor analysis of the original data set, rotate the factors obliquely, factor that correlation matrix, do a Schmid-Leiman (`schmid`) transformation to find general factor loadings, and then find ω_h .

ω_h differs slightly as a function of how the factors are estimated. Four options are available, the default will do a minimum residual factor analysis, `fm="pa"` does a principal axes factor analysis (`factor.pa`), `fm="mle"` uses the factanal function, and `fm="pc"` does a principal components analysis (`principal`). 84 For ability items, it is typically the case that all items will have positive loadings on the general factor. However, for non-cognitive items it is frequently the case that some items are to be scored positively, and some negatively. Although probably better to specify which directions the items are to be scored by specifying a key vector, if `flip = TRUE` (the default), items will be reversed so that they have positive loadings on the general factor. The keys are reported so that scores can be found using the `scoreItems` function. Arbitrarily reversing items this way can overestimate the general factor. (See the example with a simulated circumplex).

β , an alternative to ω , is defined as the worst split half reliability. It can be estimated by using `iclust` (Item Cluster analysis: a hierarchical clustering algorithm). For a very complimentary review of why the `iclust` algorithm is useful in scale construction, see Cooksey and Soutar (2006).

The `omega` function uses exploratory factor analysis to estimate the ω_h coefficient. It is important to remember that “A recommendation that should be heeded, regardless of the

method chosen to estimate ω_h , is to always examine the pattern of the estimated general factor loadings prior to estimating ω_h . Such an examination constitutes an informal test of the assumption that there is a latent variable common to all of the scale's indicators that can be conducted even in the context of EFA. If the loadings were salient for only a relatively small subset of the indicators, this would suggest that there is no true general factor underlying the covariance matrix. Just such an informal assumption test would have afforded a great deal of protection against the possibility of misinterpreting the misleading ω_h estimates occasionally produced in the simulations reported here." (Zinbarg et al., 2006, p 137).

Although ω_h is uniquely defined only for cases where 3 or more subfactors are extracted, it is sometimes desired to have a two factor solution. By default this is done by forcing the `schmid` extraction to treat the two subfactors as having equal loadings.

There are three possible options for this condition: setting the general factor loadings between the two lower order factors to be "equal" which will be the $\sqrt{r_{ab}}$ where r_{ab} is the oblique correlation between the factors) or to "first" or "second" in which case the general factor is equated with either the first or second group factor. A message is issued suggesting that the model is not really well defined. This solution discussed in Zinbarg et al., 2007. To do this in `omega`, add the `option="first"` or `option="second"` to the call.

Although obviously not meaningful for a 1 factor solution, it is of course possible to find the sum of the loadings on the first (and only) factor, square them, and compare them to the overall matrix variance. This is done, with appropriate complaints.

In addition to ω_h , another of McDonald's coefficients is ω_t . This is an estimate of the total reliability of a test.

McDonald's ω_t , which is similar to Guttman's λ_6 , (see `guttman`) uses the estimates of uniqueness u^2 from factor analysis to find e_j^2 . This is based on a decomposition of the variance of a test score, V_x into four parts: that due to a general factor, g , that due to a set of group factors, f , (factors common to some but not all of the items), specific factors, s unique to each item, and e , random error. (Because specific variance can not be distinguished from random error unless the test is given at least twice, some combine these both into error).

Letting $x = cg + Af + Ds + e$ then the communality of item $_j$, based upon general as well as group factors, $h_j^2 = c_j^2 + \sum f_{ij}^2$ and the unique variance for the item $u_j^2 = \sigma_j^2(1 - h_j^2)$ may be used to estimate the test reliability. That is, if h_j^2 is the communality of item $_j$, based upon general as well as group factors, then for standardized items, $e_j^2 = 1 - h_j^2$ and

$$\omega_t = \frac{1cc'1 + 1AA'1'}{V_x} = 1 - \frac{\sum(1 - h_j^2)}{V_x} = 1 - \frac{\sum u^2}{V_x}$$

Because $h_j^2 \geq r_{smc}^2$, $\omega_t \geq \lambda_6$.

It is important to distinguish here between the two ω coefficients of McDonald, 1978 and Equation 6.20a of McDonald, 1999, ω_t and ω_h . While the former is based upon the sum of squared loadings on all the factors, the latter is based upon the sum of the squared loadings on the general factor.

$$\omega_h = \frac{1cc'1}{V_x}$$

Another estimate reported is the omega for an infinite length test with a structure similar to the observed test. This is found by

$$\omega_{inf} = \frac{1cc'1}{1cc'1 + 1AA'1'}$$

In the case of these simulated 9 variables, the amount of variance attributable to a general factor (ω_h) is quite large, and the reliability of the set of 9 items is somewhat greater than that estimated by α or λ_6 .

```
> om.9

9 simulated variables
Call: omega(m = r9, title = "9 simulated variables")
Alpha:          0.8
G.6:            0.8
Omega Hierarchical: 0.69
Omega H asymptotic: 0.82
Omega Total     0.83

Schmid Leiman Factor loadings greater than 0.2
   g  F1*  F2*  F3*  h2  u2  p2
V1 0.72  0.45          0.72 0.28 0.71
V2 0.58  0.37          0.47 0.53 0.71
V3 0.57  0.41          0.49 0.51 0.66
V4 0.57          0.41 0.50 0.50 0.66
V5 0.55          0.32 0.41 0.59 0.73
V6 0.43          0.27 0.28 0.72 0.66
V7 0.47          0.47          0.44 0.56 0.50
V8 0.43          0.42          0.36 0.64 0.51
V9 0.32          0.23          0.16 0.84 0.63

With eigenvalues of:
   g  F1*  F2*  F3*
2.48 0.52 0.47 0.36

general/max 4.76  max/min = 1.46
mean percent general = 0.64  with sd = 0.08 and cv of 0.13
Explained Common Variance of the general factor = 0.65

The degrees of freedom are 12 and the fit is 0.03
The number of observations was 500 with Chi Square = 15.86 with prob < 0.2
The root mean square of the residuals is 0.02
The df corrected root mean square of the residuals is 0.03
```

```
> om.9 <- omega(r9,title="9 simulated variables")
```

9 simulated variables

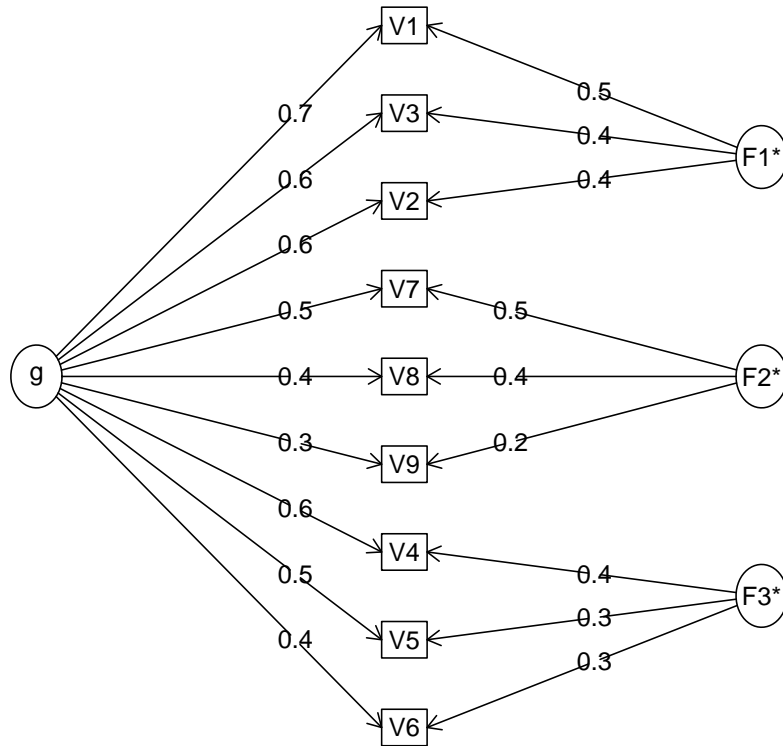


Figure 15: A bifactor solution for 9 simulated variables with a hierarchical structure.

RMSEA index = 0.026 and the 90 % confidence intervals are NA 0.055
 BIC = -58.71

Compare this with the adequacy of just a general factor and no group factors
 The degrees of freedom for just the general factor are 27 and the fit is 0.32
 The number of observations was 500 with Chi Square = 159.82 with prob < 8.3e-21
 The root mean square of the residuals is 0.08
 The df corrected root mean square of the residuals is 0.09

RMSEA index = 0.1 and the 90 % confidence intervals are 0.085 0.114
 BIC = -7.97

Measures of factor score adequacy

| | g | F1* | F2* | F3* |
|---|------|-------|-------|-------|
| Correlation of scores with factors | 0.84 | 0.59 | 0.61 | 0.54 |
| Multiple R square of scores with factors | 0.71 | 0.35 | 0.37 | 0.29 |
| Minimum correlation of factor score estimates | 0.43 | -0.30 | -0.25 | -0.41 |

Total, General and Subset omega for each subset

| | g | F1* | F2* | F3* |
|--|------|------|------|------|
| Omega total for total scores and subscales | 0.83 | 0.79 | 0.56 | 0.65 |
| Omega general for total scores and subscales | 0.69 | 0.55 | 0.30 | 0.46 |
| Omega group for total scores and subscales | 0.12 | 0.24 | 0.26 | 0.19 |

6.3 Estimating ω_h using Confirmatory Factor Analysis

The `omegaSem` function will do an exploratory analysis and then take the highest loading items on each factor and do a confirmatory factor analysis using the `sem` package. These results can produce slightly different estimates of ω_h , primarily because cross loadings are modeled as part of the general factor.

```
> omegaSem(r9, n.obs=500)
```

```
Call: omegaSem(m = r9, n.obs = 500)
```

```
Omega
```

```
Call: omega(m = m, nfactors = nfactors, fm = fm, key = key, flip = flip,
  digits = digits, title = title, sl = sl, labels = labels,
  plot = plot, n.obs = n.obs, rotate = rotate, Phi = Phi, option = option)
```

```
Alpha:          0.8
G.6:            0.8
Omega Hierarchical: 0.69
Omega H asymptotic: 0.82
Omega Total     0.83
```

```
Schmid Leiman Factor loadings greater than 0.2
```

| | g | F1* | F2* | F3* | h2 | u2 | p2 |
|----|------|------|------|------|------|------|------|
| V1 | 0.72 | 0.45 | | | 0.72 | 0.28 | 0.71 |
| V2 | 0.58 | 0.37 | | | 0.47 | 0.53 | 0.71 |
| V3 | 0.57 | 0.41 | | | 0.49 | 0.51 | 0.66 |
| V4 | 0.57 | | | 0.41 | 0.50 | 0.50 | 0.66 |
| V5 | 0.55 | | | 0.32 | 0.41 | 0.59 | 0.73 |
| V6 | 0.43 | | | 0.27 | 0.28 | 0.72 | 0.66 |
| V7 | 0.47 | | 0.47 | | 0.44 | 0.56 | 0.50 |
| V8 | 0.43 | | 0.42 | | 0.36 | 0.64 | 0.51 |
| V9 | 0.32 | | 0.23 | | 0.16 | 0.84 | 0.63 |

With eigenvalues of:

| | g | F1* | F2* | F3* |
|--|------|------|------|------|
| | 2.48 | 0.52 | 0.47 | 0.36 |

general/max 4.76 max/min = 1.46
mean percent general = 0.64 with sd = 0.08 and cv of 0.13
Explained Common Variance of the general factor = 0.65

The degrees of freedom are 12 and the fit is 0.03
The number of observations was 500 with Chi Square = 15.86 with prob < 0.2
The root mean square of the residuals is 0.02
The df corrected root mean square of the residuals is 0.03
RMSEA index = 0.026 and the 90 % confidence intervals are NA 0.055
BIC = -58.71

Compare this with the adequacy of just a general factor and no group factors
The degrees of freedom for just the general factor are 27 and the fit is 0.32
The number of observations was 500 with Chi Square = 159.82 with prob < 8.3e-21
The root mean square of the residuals is 0.08
The df corrected root mean square of the residuals is 0.09

RMSEA index = 0.1 and the 90 % confidence intervals are 0.085 0.114
BIC = -7.97

Measures of factor score adequacy

| | g | F1* | F2* | F3* |
|---|------|-------|-------|-------|
| Correlation of scores with factors | 0.84 | 0.59 | 0.61 | 0.54 |
| Multiple R square of scores with factors | 0.71 | 0.35 | 0.37 | 0.29 |
| Minimum correlation of factor score estimates | 0.43 | -0.30 | -0.25 | -0.41 |

Total, General and Subset omega for each subset

| | g | F1* | F2* | F3* |
|--|------|------|------|------|
| Omega total for total scores and subscales | 0.83 | 0.79 | 0.56 | 0.65 |
| Omega general for total scores and subscales | 0.69 | 0.55 | 0.30 | 0.46 |
| Omega group for total scores and subscales | 0.12 | 0.24 | 0.26 | 0.19 |

Omega Hierarchical from a confirmatory model using sem = 0.72

Omega Total from a confirmatory model using sem = 0.83

With loadings of

| | g | F1* | F2* | F3* | h2 | u2 |
|----|------|------|------|------|------|------|
| V1 | 0.74 | 0.40 | | | 0.71 | 0.29 |
| V2 | 0.58 | 0.36 | | | 0.47 | 0.53 |
| V3 | 0.56 | 0.42 | | | 0.50 | 0.50 |
| V4 | 0.57 | | 0.45 | | 0.53 | 0.47 |
| V5 | 0.58 | | 0.25 | 0.40 | 0.60 | |
| V6 | 0.43 | | 0.26 | 0.26 | 0.74 | |
| V7 | 0.49 | | 0.38 | | 0.38 | 0.62 |
| V8 | 0.44 | | 0.45 | | 0.39 | 0.61 |
| V9 | 0.34 | | 0.24 | | 0.17 | 0.83 |

With eigenvalues of:

| | g | F1* | F2* | F3* |
|--|------|------|------|------|
| | 2.60 | 0.47 | 0.40 | 0.33 |

6.3.1 Other estimates of reliability

Other estimates of reliability are found by the `splitHalf` function. These are described in more detail in [Revelle and Zinbarg \(2009\)](#). They include the 6 estimates from Guttman, four from TenBerge, and an estimate of the greatest lower bound. See also [Revelle and Condon \(2018a,b\)](#) for detailed discussions of the many ways of estimating reliability.

```
> splitHalf(r9)

Split half reliabilities
Call: splitHalf(r = r9)

Maximum split half reliability (lambda 4) = 0.84
Guttman lambda 6                        = 0.8
Average split half reliability          = 0.79
Guttman lambda 3 (alpha)                = 0.8
Minimum split half reliability (beta)   = 0.72
```

6.4 Reliability and correlations of multiple scales within an inventory

A typical research question in personality involves an inventory of multiple items purporting to measure multiple constructs. For example, the data set `bfi` includes 25 items thought to measure five dimensions of personality (Extraversion, Emotional Stability, Conscientiousness, Agreeableness, and Openness). The data may either be the raw data or a correlation matrix (`scoreItems`) or just a correlation matrix of the items (`cluster.cor` and `cluster.loadings`). When finding reliabilities for multiple scales, item reliabilities can be estimated using the squared multiple correlation of an item with all other items, not just those that are keyed for a particular scale. This leads to an estimate of $G6^*$.

6.4.1 Scoring from raw data

To score these five scales from the 25 items, use the `scoreItems` function with the helper function `make.keys`. Logically, scales are merely the weighted composites of a set of items. The weights used are -1, 0, and 1. 0 implies do not use that item in the scale, 1 implies a positive weight (add the item to the total score), -1 a negative weight (subtract the item from the total score, i.e., reverse score the item). Reverse scoring an item is equivalent to subtracting the item from the maximum + minimum possible value for that item. The minima and maxima can be estimated from all the items, or can be specified by the user.

There are two different ways that scale scores tend to be reported. Social psychologists and educational psychologists tend to report the scale score as the *average item score* while many personality psychologists tend to report the *total item score*. The default option for `scoreItems` is to report item averages (which thus allows interpretation in the same metric

as the items) but totals can be found as well. Personality researchers should be encouraged to report scores based upon item means and avoid using the total score although some reviewers are adamant about the following the tradition of total scores.

The printed output includes coefficients α and $G6^*$, the average correlation of the items within the scale (corrected for item overlap and scale reliability), as well as the correlations between the scales (below the diagonal, the correlations above the diagonal are corrected for attenuation. As is the case for most of the *psych* functions, additional information is returned as part of the object.

First, create keys matrix using the `make.keys` function. (The keys matrix could also be prepared externally using a spreadsheet and then copying it into R). Although not normally necessary, show the keys to understand what is happening.

Note that the number of items to specify in the `make.keys` function is the total number of items in the inventory. That is, if scoring just 5 items from a 25 item inventory, `make.keys` should be told that there are 25 items. `make.keys` just changes a list of items on each scale to make up a scoring matrix. Because the `bfi` data set has 25 items as well as 3 demographic items, the number of variables is specified as 28.

```
> keys <- make.keys(nvars=28, list(Agree=c(-1,2:5),Conscientious=c(6:8,-9,-10),
+ Extraversion=c(-11,-12,13:15),Neuroticism=c(16:20),
+ Openness = c(21,-22,23,24,-25)),
+ item.labels=colnames(bfi))
> keys
```

| | Agree | Conscientious | Extraversion | Neuroticism | Openness |
|--------|-------|---------------|--------------|-------------|----------|
| A1 | -1 | 0 | 0 | 0 | 0 |
| A2 | 1 | 0 | 0 | 0 | 0 |
| A3 | 1 | 0 | 0 | 0 | 0 |
| A4 | 1 | 0 | 0 | 0 | 0 |
| A5 | 1 | 0 | 0 | 0 | 0 |
| C1 | 0 | 1 | 0 | 0 | 0 |
| C2 | 0 | 1 | 0 | 0 | 0 |
| C3 | 0 | 1 | 0 | 0 | 0 |
| C4 | 0 | -1 | 0 | 0 | 0 |
| C5 | 0 | -1 | 0 | 0 | 0 |
| E1 | 0 | 0 | -1 | 0 | 0 |
| E2 | 0 | 0 | -1 | 0 | 0 |
| E3 | 0 | 0 | 1 | 0 | 0 |
| E4 | 0 | 0 | 1 | 0 | 0 |
| E5 | 0 | 0 | 1 | 0 | 0 |
| N1 | 0 | 0 | 0 | 1 | 0 |
| N2 | 0 | 0 | 0 | 1 | 0 |
| N3 | 0 | 0 | 0 | 1 | 0 |
| N4 | 0 | 0 | 0 | 1 | 0 |
| N5 | 0 | 0 | 0 | 1 | 0 |
| O1 | 0 | 0 | 0 | 0 | 1 |
| O2 | 0 | 0 | 0 | 0 | -1 |
| O3 | 0 | 0 | 0 | 0 | 1 |
| O4 | 0 | 0 | 0 | 0 | 1 |
| O5 | 0 | 0 | 0 | 0 | -1 |
| gender | 0 | 0 | 0 | 0 | 0 |

```
education    0          0          0          0          0
age          0          0          0          0          0
```

The use of multiple key matrices for different inventories is facilitated by using the `superMatrix` function to combine two or more matrices. This allows convenient scoring of large data sets combining multiple inventories with keys based upon each individual inventory. Pretend for the moment that the big 5 items were made up of two inventories, one consisting of the first 10 items, the second the last 18 items. (15 personality items + 3 demographic items.) Then the following code would work:

```
> keys.1<- make.keys(10,list(Agree=c(-1,2:5),Conscientious=c(6:8,-9,-10)))
> keys.2 <- make.keys(15,list(Extraversion=c(-1,-2,3:5),Neuroticism=c(6:10),
+ Openness = c(11,-12,13,14,-15)))
> keys.25 <- superMatrix(list(keys.1,keys.2))
```

The resulting keys matrix is identical to that found above except that it does not include the extra 3 demographic items. This is useful when scoring the raw items because the response frequencies for each category are reported, and for the demographic data,

This use of making multiple key matrices and then combining them into one super matrix of keys is particularly useful when combining demographic information with items to be scores. A set of demographic keys can be made and then these can be combined with the keys for the particular scales.

Now use these keys in combination with the raw data to score the items, calculate basic reliability and intercorrelations, and find the item-by scale correlations for each item and each scale. By default, missing data are replaced by the median for that variable.

```
> scores <- scoreItems(keys,bfi)
> scores

Call: scoreItems(keys = keys, items = bfi)

(Unstandardized) Alpha:
  Agree Conscientious Extraversion Neuroticism Openness
alpha  0.7          0.72          0.76          0.81          0.6

Standard errors of unstandardized Alpha:
  Agree Conscientious Extraversion Neuroticism Openness
ASE  0.014          0.014          0.013          0.011          0.017

Average item correlation:
  Agree Conscientious Extraversion Neuroticism Openness
average.r 0.32          0.34          0.39          0.46          0.23

Guttman 6* reliability:
  Agree Conscientious Extraversion Neuroticism Openness
Lambda.6 0.7          0.72          0.76          0.81          0.6

Signal/Noise based upon av.r :
  Agree Conscientious Extraversion Neuroticism Openness
Signal/Noise 2.3          2.6          3.2          4.3          1.5
```



```

Scale intercorrelations corrected for attenuation
raw correlations below the diagonal, alpha on the diagonal
corrected correlations above the diagonal:
      Agree Conscientious Extraversion Neuroticism Openness
Agree      0.70          0.36          0.63         -0.245      0.23
Conscientious 0.26          0.72          0.35         -0.305      0.30
Extraversion 0.46          0.26          0.76         -0.284      0.32
Neuroticism  -0.18         -0.23         -0.22          0.812     -0.12
Openness     0.15          0.19          0.22         -0.086      0.60

```

In order to see the item by scale loadings and frequency counts of the data print with the short option = FALSE

To see the additional information (the raw correlations, the individual scores, etc.), they may be specified by name. Then, to visualize the correlations between the raw scores, use the `pairs.panels` function on the scores values of scores. (See figure 16

6.4.2 Forming scales from a correlation matrix

There are some situations when the raw data are not available, but the correlation matrix between the items is available. In this case, it is not possible to find individual scores, but it is possible to find the reliability and intercorrelations of the scales. This may be done using the `cluster.cor` function or the `scoreItems` function. The use of a keys matrix is the same as in the raw data case.

Consider the same `bfi` data set, but first find the correlations, and then use `cluster.cor`.

```

> r.bfi <- cor(bfi,use="pairwise")
> scales <- cluster.cor(keys,r.bfi)
> summary(scales)

```

Call: cluster.cor(keys = keys, r.mat = r.bfi)

```

Scale intercorrelations corrected for attenuation
raw correlations below the diagonal, (standardized) alpha on the diagonal
corrected correlations above the diagonal:
      Agree Conscientious Extraversion Neuroticism Openness
Agree      0.71          0.35          0.64         -0.242      0.25
Conscientious 0.25          0.73          0.36         -0.287      0.30
Extraversion 0.47          0.27          0.76         -0.278      0.35
Neuroticism  -0.18         -0.22         -0.22          0.815     -0.11
Openness     0.16          0.20          0.24         -0.074      0.61

```

To find the correlations of the items with each of the scales (the “structure” matrix) or the correlations of the items controlling for the other scales (the “pattern” matrix), use the `cluster.loadings` function. To do both at once (e.g., the correlations of the scales as well as the item by scale correlations), it is also possible to just use `scoreItems`.

```

> png('scores.png')
> pairs.panels(scores$scores,pch='.',jiggle=TRUE)
> dev.off()
quartz
  2

```

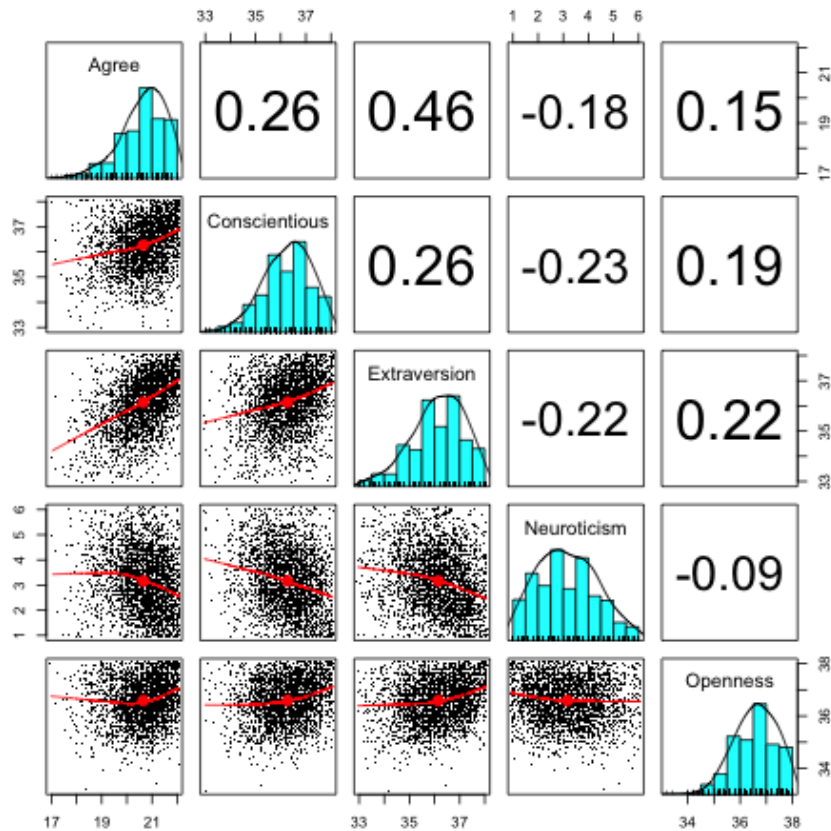


Figure 16: A graphic analysis of the Big Five scales found by using the scoreItems function. The pair.wise plot allows us to see that some participants have reached the ceiling of the scale for these 5 items scales. Using the pch='.' option in pairs.panels is recommended when plotting many cases. The data points were “jittered” by setting jiggle=TRUE. Jiggling this way shows the density more clearly. To save space, the figure was done as a png. For a clearer figure, save as a pdf.

6.5 Scoring Multiple Choice Items

Some items (typically associated with ability tests) are not themselves mini-scales ranging from low to high levels of expression of the item of interest, but are rather multiple choice where one response is the correct response. Two analyses are useful for this kind of item: examining the response patterns to all the alternatives (looking for good or bad distractors) and scoring the items as correct or incorrect. Both of these operations may be done using the `score.multiple.choice` function. Consider the 16 example items taken from an online ability test at the Personality Project: <http://test.personality-project.org>. This is part of the Synthetic Aperture Personality Assessment (*SAPA*) study discussed in [Revelle et al. \(2011, 2010\)](#).

```
> data(iqitems)
> iq.keys <- c(4,4,4, 6,6,3,4,4, 5,2,2,4, 3,2,6,7)
> score.multiple.choice(iq.keys,iqitems)

Call: score.multiple.choice(key = iq.keys, data = iqitems)

(Unstandardized) Alpha:
[1] 0.84

Average item correlation:
[1] 0.25

item statistics
      key  0  1  2  3  4  5  6  7  8 miss  r  n mean  sd
reason.4  4 0.05 0.05 0.11 0.10 0.64 0.03 0.02 0.00 0.00 0 0.59 1523 0.64 0.48
reason.16  4 0.04 0.06 0.08 0.10 0.70 0.01 0.00 0.00 0.00 0 0.53 1524 0.70 0.46
reason.17  4 0.05 0.03 0.05 0.03 0.70 0.03 0.11 0.00 0.00 0 0.59 1523 0.70 0.46
reason.19  6 0.04 0.02 0.13 0.03 0.06 0.10 0.62 0.00 0.00 0 0.56 1523 0.62 0.49
letter.7    6 0.05 0.01 0.05 0.03 0.11 0.14 0.60 0.00 0.00 0 0.58 1524 0.60 0.49
letter.33   3 0.06 0.10 0.13 0.57 0.04 0.09 0.02 0.00 0.00 0 0.56 1523 0.57 0.50
letter.34   4 0.04 0.09 0.07 0.11 0.61 0.05 0.02 0.00 0.00 0 0.59 1523 0.61 0.49
letter.58   4 0.06 0.14 0.09 0.09 0.44 0.16 0.01 0.00 0.00 0 0.58 1525 0.44 0.50
matrix.45   5 0.04 0.01 0.06 0.14 0.18 0.53 0.04 0.00 0.00 0 0.51 1523 0.53 0.50
matrix.46   2 0.04 0.12 0.55 0.07 0.11 0.06 0.05 0.00 0.00 0 0.52 1524 0.55 0.50
matrix.47   2 0.04 0.05 0.61 0.07 0.11 0.06 0.06 0.00 0.00 0 0.55 1523 0.61 0.49
matrix.55   4 0.04 0.02 0.18 0.14 0.37 0.07 0.18 0.00 0.00 0 0.45 1524 0.37 0.48
rotate.3    3 0.04 0.03 0.04 0.19 0.22 0.15 0.05 0.12 0.15 0 0.51 1523 0.19 0.40
rotate.4    2 0.04 0.03 0.21 0.05 0.18 0.04 0.04 0.25 0.15 0 0.56 1523 0.21 0.41
rotate.6    6 0.04 0.22 0.02 0.05 0.14 0.05 0.30 0.04 0.14 0 0.55 1523 0.30 0.46
rotate.8    7 0.04 0.03 0.21 0.07 0.16 0.05 0.13 0.19 0.13 0 0.48 1524 0.19 0.39

> #just convert the items to true or false
> iq.tf <- score.multiple.choice(iq.keys,iqitems,score=FALSE)
> describe(iq.tf) #compare to previous results

      vars  n mean  sd median trimmed mad min max range skew kurtosis  se
reason.4  1 1523 0.64 0.48      1  0.68  0  0  1      1 -0.58  -1.66 0.01
reason.16  2 1524 0.70 0.46      1  0.75  0  0  1      1 -0.86  -1.26 0.01
reason.17  3 1523 0.70 0.46      1  0.75  0  0  1      1 -0.86  -1.26 0.01
reason.19  4 1523 0.62 0.49      1  0.64  0  0  1      1 -0.47  -1.78 0.01
letter.7   5 1524 0.60 0.49      1  0.62  0  0  1      1 -0.41  -1.84 0.01
letter.33  6 1523 0.57 0.50      1  0.59  0  0  1      1 -0.29  -1.92 0.01
letter.34  7 1523 0.61 0.49      1  0.64  0  0  1      1 -0.46  -1.79 0.01
```

| | | | | | | | | | | | | | |
|-----------|----|------|------|------|---|------|---|---|---|---|-------|-------|------|
| letter.58 | 8 | 1525 | 0.44 | 0.50 | 0 | 0.43 | 0 | 0 | 1 | 1 | 0.23 | -1.95 | 0.01 |
| matrix.45 | 9 | 1523 | 0.53 | 0.50 | 1 | 0.53 | 0 | 0 | 1 | 1 | -0.10 | -1.99 | 0.01 |
| matrix.46 | 10 | 1524 | 0.55 | 0.50 | 1 | 0.56 | 0 | 0 | 1 | 1 | -0.20 | -1.96 | 0.01 |
| matrix.47 | 11 | 1523 | 0.61 | 0.49 | 1 | 0.64 | 0 | 0 | 1 | 1 | -0.47 | -1.78 | 0.01 |
| matrix.55 | 12 | 1524 | 0.37 | 0.48 | 0 | 0.34 | 0 | 0 | 1 | 1 | 0.52 | -1.73 | 0.01 |
| rotate.3 | 13 | 1523 | 0.19 | 0.40 | 0 | 0.12 | 0 | 0 | 1 | 1 | 1.55 | 0.40 | 0.01 |
| rotate.4 | 14 | 1523 | 0.21 | 0.41 | 0 | 0.14 | 0 | 0 | 1 | 1 | 1.40 | -0.03 | 0.01 |
| rotate.6 | 15 | 1523 | 0.30 | 0.46 | 0 | 0.25 | 0 | 0 | 1 | 1 | 0.88 | -1.24 | 0.01 |
| rotate.8 | 16 | 1524 | 0.19 | 0.39 | 0 | 0.11 | 0 | 0 | 1 | 1 | 1.62 | 0.63 | 0.01 |

Once the items have been scored as true or false (assigned scores of 1 or 0), they made then be scored into multiple scales using the normal `scoreItems` function.

6.6 Item analysis

Basic item analysis starts with describing the data (`describe`, finding the number of dimensions using factor analysis (`fa`) and cluster analysis `iclust` perhaps using the Very Simple Structure criterion (`vss`), or perhaps parallel analysis `fa.parallel`. Item whole correlations may then be found for scales scored on one dimension (`alpha` or many scales simultaneously (`scoreItems`)). Scales can be modified by changing the keys matrix (i.e., dropping particular items, changing the scale on which an item is to be scored). This analysis can be done on the normal Pearson correlation matrix or by using polychoric correlations. Validities of the scales can be found using multiple correlation of the raw data or based upon correlation matrices using the `set.cor` function. However, more powerful item analysis tools are now available by using Item Response Theory approaches.

Although the `response.frequencies` output from `score.multiple.choice` is useful to examine in terms of the probability of various alternatives being endorsed, it is even better to examine the pattern of these responses as a function of the underlying latent trait or just the total score. This may be done by using `irt.responses` (Figure 17).

7 Item Response Theory analysis

The use of Item Response Theory has become is said to be the “new psychometrics”. The emphasis is upon item properties, particularly those of item difficulty or location and item discrimination. These two parameters are easily found from classic techniques when using factor analyses of correlation matrices formed by `polychoric` or `tetrachoric` correlations. The `irt.fa` function does this and then graphically displays item discrimination and item location as well as item and test information (see Figure 18).

```

> data(iqitems)
> iq.keys <- c(4,4,4, 6,6,3,4,4, 5,2,2,4, 3,2,6,7)
> scores <- score.multiple.choice(iq.keys,iqitems,score=TRUE,short=FALSE)
> #note that for speed we can just do this on simple item counts rather than IRT based scores.
> op <- par(mfrow=c(2,2)) #set this to see the output for multiple items
> irt.responses(scores$scores,iqitems[1:4],breaks=11)

```

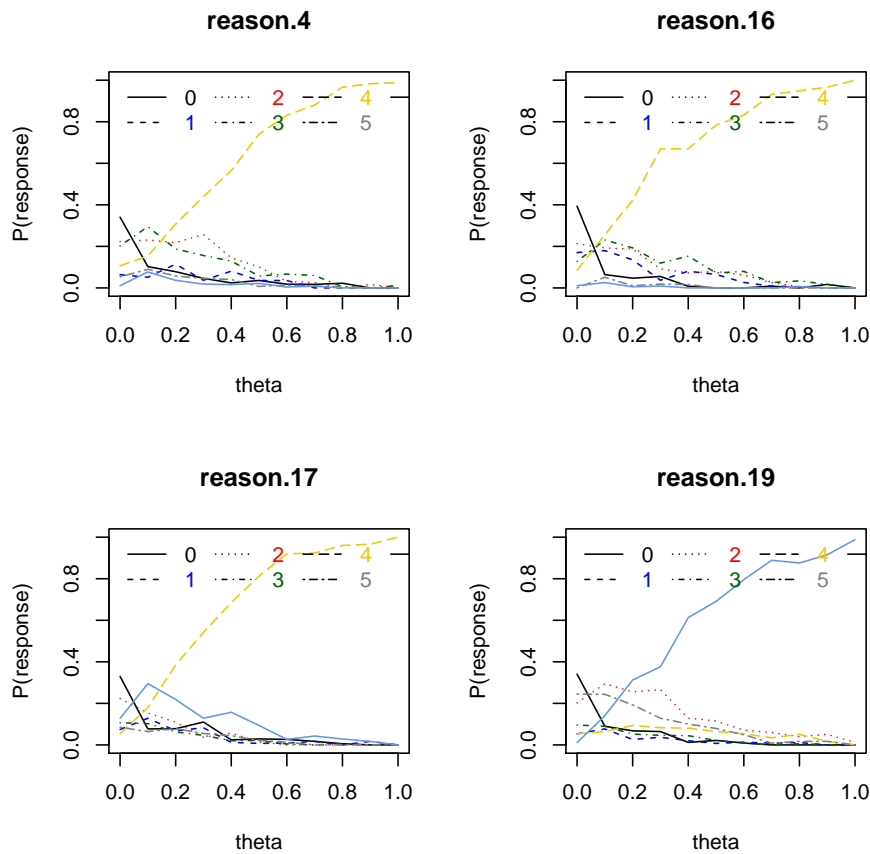


Figure 17: The pattern of responses to multiple choice ability items can show that some items have poor distractors. This may be done by using the the `irt.responses` function. A good distractor is one that is negatively related to ability.

7.1 Factor analysis and Item Response Theory

If the correlations of all of the items reflect one underlying latent variable, then factor analysis of the matrix of tetrachoric correlations should allow for the identification of the regression slopes (α) of the items on the latent variable. These regressions are, of course just the factor loadings. Item difficulty, δ_j and item discrimination, α_j may be found from factor analysis of the tetrachoric correlations where λ_j is just the factor loading on the first factor and τ_j is the normal threshold reported by the `tetrachoric` function.

$$\delta_j = \frac{D\tau}{\sqrt{1-\lambda_j^2}}, \quad \alpha_j = \frac{\lambda_j}{\sqrt{1-\lambda_j^2}} \quad (2)$$

where D is a scaling factor used when converting to the parameterization of *logistic* model and is 1.702 in that case and 1 in the case of the normal ogive model. Thus, in the case of the normal model, factor loadings (λ_j) and item thresholds (τ) are just

$$\lambda_j = \frac{\alpha_j}{\sqrt{1+\alpha_j^2}}, \quad \tau_j = \frac{\delta_j}{\sqrt{1+\alpha_j^2}}.$$

Consider 9 dichotomous items representing one factor but differing in their levels of difficulty

```
> set.seed(17)
> d9 <- sim.irt(9,1000,-2.5,2.5,mod="normal") #dichotomous items
> test <- irt.fa(d9$items)
```

```
> test
```

```
Item Response Analysis using Factor Analysis
```

```
Call: irt.fa(x = d9$items)
```

```
Item Response Analysis using Factor Analysis
```

```
Summary information by factor and item
```

```
Factor = 1
      -3  -2  -1   0   1   2   3
V1    0.48 0.59 0.24 0.06 0.01 0.00 0.00
V2    0.30 0.68 0.45 0.12 0.02 0.00 0.00
V3    0.12 0.50 0.74 0.29 0.06 0.01 0.00
V4    0.05 0.26 0.74 0.55 0.14 0.03 0.00
V5    0.01 0.07 0.44 1.05 0.41 0.06 0.01
V6    0.00 0.03 0.15 0.60 0.74 0.24 0.04
V7    0.00 0.01 0.04 0.22 0.73 0.63 0.16
V8    0.00 0.00 0.02 0.12 0.45 0.69 0.31
V9    0.00 0.01 0.02 0.08 0.25 0.47 0.36
Test Info  0.98 2.14 2.85 3.09 2.81 2.14 0.89
SEM       1.01 0.68 0.59 0.57 0.60 0.68 1.06
Reliability -0.02 0.53 0.65 0.68 0.64 0.53 -0.12
```

```
Factor analysis with Call: fa(r = r, nfactors = nfactors, n.obs = n.obs, rotate = rotate,
fm = fm)
```

```
Test of the hypothesis that 1 factor is sufficient.
The degrees of freedom for the model is 27 and the objective function was 1.2
The number of observations was 1000 with Chi Square = 1195.58 with prob < 7.3e-235
```

```
The root mean square of the residuals (RMSA) is 0.09
The df corrected root mean square of the residuals is 0.1
```

```
Tucker Lewis Index of factoring reliability = 0.699
RMSEA index = 0.209 and the 90 % confidence intervals are 0.198 0.218
BIC = 1009.07
```

Similar analyses can be done for polytomous items such as those of the bfi extraversion scale:

```
> data(bfi)
> e.irt <- irt.fa(bfi[11:15])
> e.irt
```

Item Response Analysis using Factor Analysis

```
Call: irt.fa(x = bfi[11:15])
Item Response Analysis using Factor Analysis
```

```
Summary information by factor and item
Factor = 1
```

| | -3 | -2 | -1 | 0 | 1 | 2 | 3 |
|-------------|------|------|------|------|------|------|------|
| E1 | 0.37 | 0.58 | 0.75 | 0.76 | 0.61 | 0.39 | 0.21 |
| E2 | 0.48 | 0.89 | 1.18 | 1.24 | 0.99 | 0.59 | 0.25 |
| E3 | 0.34 | 0.49 | 0.58 | 0.57 | 0.49 | 0.36 | 0.23 |
| E4 | 0.63 | 0.98 | 1.11 | 0.96 | 0.67 | 0.37 | 0.16 |
| E5 | 0.33 | 0.42 | 0.47 | 0.45 | 0.37 | 0.27 | 0.17 |
| Test Info | 2.15 | 3.36 | 4.09 | 3.98 | 3.13 | 1.97 | 1.02 |
| SEM | 0.68 | 0.55 | 0.49 | 0.50 | 0.56 | 0.71 | 0.99 |
| Reliability | 0.53 | 0.70 | 0.76 | 0.75 | 0.68 | 0.49 | 0.02 |

```
Factor analysis with Call: fa(r = r, nfactors = nfactors, n.obs = n.obs, rotate = rotate,
fm = fm)
```

```
Test of the hypothesis that 1 factor is sufficient.
The degrees of freedom for the model is 5 and the objective function was 0.05
The number of observations was 2800 with Chi Square = 135.92 with prob < 1.3e-27
```

```
The root mean square of the residuals (RMSA) is 0.04
The df corrected root mean square of the residuals is 0.06
```

```
Tucker Lewis Index of factoring reliability = 0.932
RMSEA index = 0.097 and the 90 % confidence intervals are 0.083 0.111
BIC = 96.23
```

The item information functions show that not all of items are equally good (Figure 19):

These procedures can be generalized to more than one factor by specifying the number of

```

> op <- par(mfrow=c(3,1))
> plot(test,type="ICC")
> plot(test,type="IIC")
> plot(test,type="test")
> op <- par(mfrow=c(1,1))

```

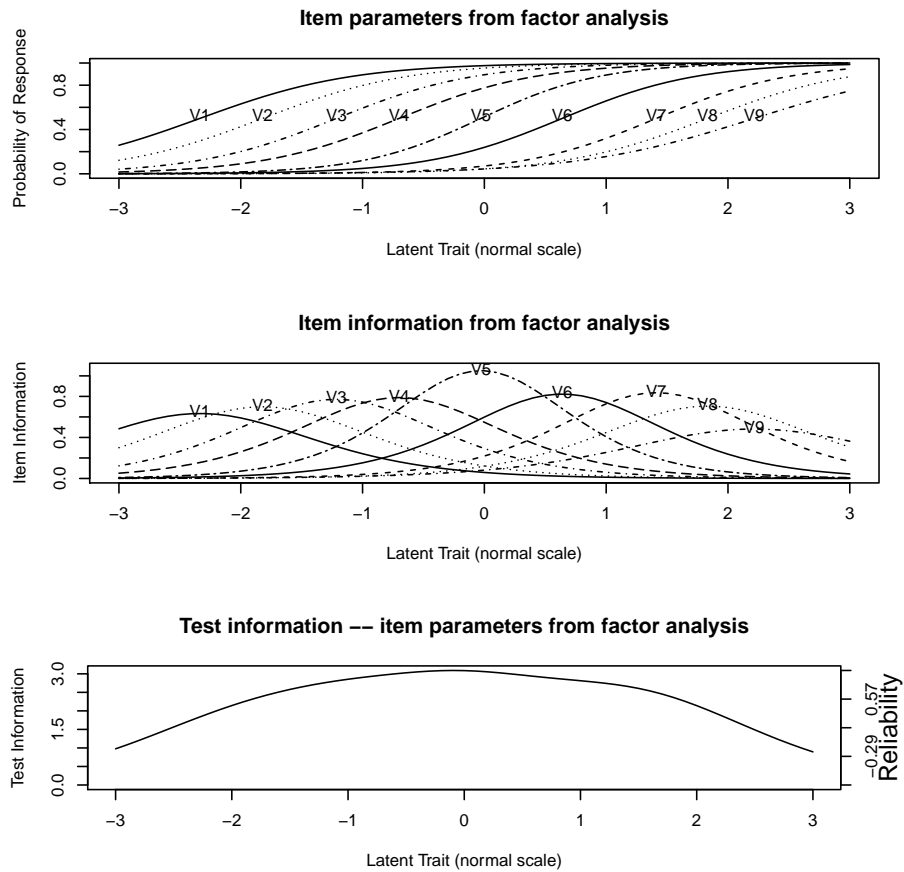


Figure 18: A graphic analysis of 9 dichotomous (simulated) items. The top panel shows the probability of item endorsement as the value of the latent trait increases. Items differ in their location (difficulty) and discrimination (slope). The middle panel shows the information in each item as a function of latent trait level. An item is most informative when the probability of endorsement is 50%. The lower panel shows the total test information. These items form a test that is most informative (most accurate) at the middle range of the latent trait.


```
> e.info <- plot(e.irt, type="IIC")
```

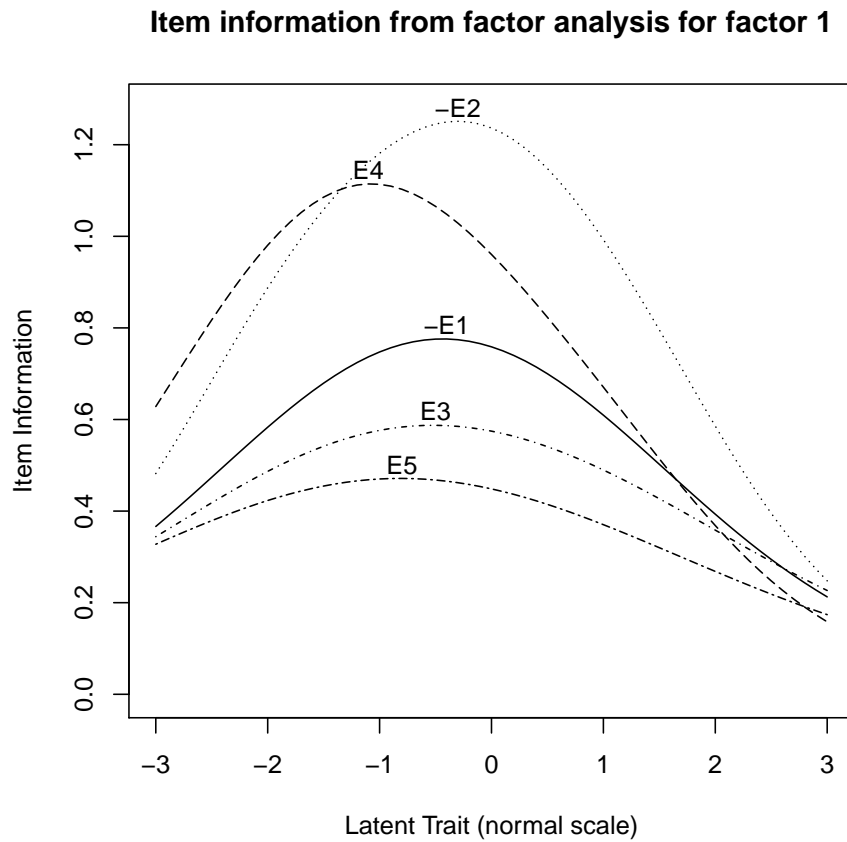


Figure 19: A graphic analysis of 5 extraversion items from the bfi. The curves represent the amount of information in the item as a function of the latent score for an individual. That is, each item is maximally discriminating at a different part of the latent continuum. Print `e.info` to see the average information for each item.

factors in `irt.fa`. The plots can be limited to those items with discriminations greater than some value of `cut`. An invisible object is returned when plotting the output from `irt.fa` that includes the average information for each item that has loadings greater than `cut`.

```
> print(e.info, sort=TRUE)
```

```
Item Response Analysis using Factor Analysis

Summary information by factor and item
Factor = 1
      -3  -2  -1   0   1   2   3
E2      0.48 0.89 1.18 1.24 0.99 0.59 0.25
E4      0.63 0.98 1.11 0.96 0.67 0.37 0.16
E1      0.37 0.58 0.75 0.76 0.61 0.39 0.21
E3      0.34 0.49 0.58 0.57 0.49 0.36 0.23
E5      0.33 0.42 0.47 0.45 0.37 0.27 0.17
Test Info 2.15 3.36 4.09 3.98 3.13 1.97 1.02
SEM      0.68 0.55 0.49 0.50 0.56 0.71 0.99
Reliability 0.53 0.70 0.76 0.75 0.68 0.49 0.02
```

More extensive IRT packages include the *ltm* and *eRm* and should be used for serious Item Response Theory analysis.

7.2 Speeding up analyses

Finding tetrachoric or polychoric correlations is very time consuming. Thus, to speed up the process of analysis, the original correlation matrix is saved as part of the output of both `irt.fa` and `omega`. Subsequent analyses may be done by using this correlation matrix. This is done by doing the analysis not on the original data, but rather on the output of the previous analysis.

For example, taking the output from the 16 ability items from the *SAPA* project when scored for True/False using `score.multiple.choice` we can first do a simple IRT analysis of one factor (Figure 21) and then use that correlation matrix to do an `omega` analysis to show the sub-structure of the ability items .

```
> iq.irt
```

```
Item Response Analysis using Factor Analysis

Call: irt.fa(x = iq.tf)
Item Response Analysis using Factor Analysis

Summary information by factor and item
Factor = 1
      -3  -2  -1   0   1   2   3
reason.4  0.04 0.20 0.59 0.59 0.19 0.04 0.01
reason.16 0.07 0.23 0.46 0.39 0.16 0.05 0.01
reason.17 0.06 0.27 0.69 0.50 0.14 0.03 0.01
reason.19 0.05 0.17 0.41 0.46 0.22 0.07 0.02
```

```
> iq.irt <- irt.fa(iq.tf)
```

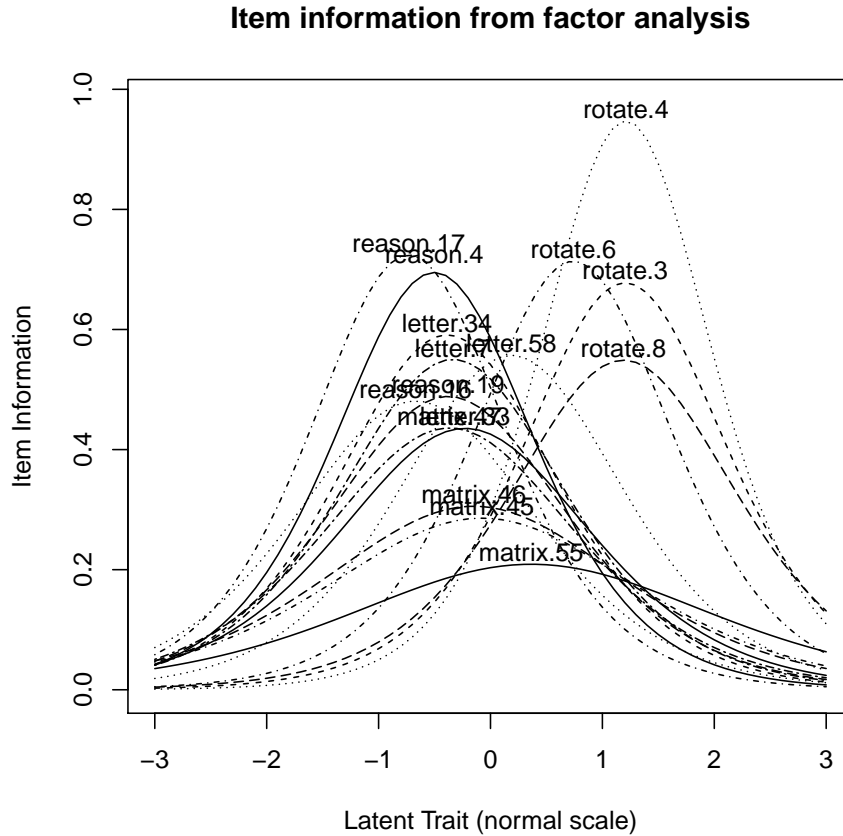


Figure 20: A graphic analysis of 16 ability items sampled from the *SAPA* project. The curves represent the amount of information in the item as a function of the latent score for an individual. That is, each item is maximally discriminating at a different part of the latent continuum. Print `iq.irt` to see the average information for each item. Partly because this is a power test (it is given on the web) and partly because the items have not been carefully chosen, the items are not very discriminating at the high end of the ability dimension.

```

letter.7      0.04 0.16 0.44 0.52 0.23 0.06 0.02
letter.33     0.04 0.14 0.34 0.43 0.24 0.08 0.02
letter.34     0.04 0.17 0.48 0.54 0.22 0.06 0.01
letter.58     0.02 0.08 0.28 0.55 0.39 0.13 0.03
matrix.45     0.05 0.11 0.23 0.29 0.21 0.10 0.04
matrix.46     0.05 0.12 0.25 0.30 0.21 0.10 0.04
matrix.47     0.05 0.16 0.37 0.41 0.21 0.07 0.02
matrix.55     0.04 0.08 0.14 0.20 0.19 0.13 0.06
rotate.3      0.00 0.01 0.07 0.29 0.66 0.45 0.13
rotate.4      0.00 0.01 0.05 0.30 0.91 0.55 0.11
rotate.6      0.01 0.03 0.14 0.49 0.68 0.27 0.06
rotate.8      0.00 0.02 0.08 0.27 0.54 0.39 0.13
Test Info    0.55 1.95 4.99 6.53 5.41 2.57 0.71
SEM          1.34 0.72 0.45 0.39 0.43 0.62 1.18
Reliability -0.80 0.49 0.80 0.85 0.82 0.61 -0.40

```

```

Factor analysis with Call: fa(r = r, nfactors = nfactors, n.obs = n.obs, rotate = rotate,
fm = fm)

```

```

Test of the hypothesis that 1 factor is sufficient.
The degrees of freedom for the model is 104 and the objective function was 1.85
The number of observations was 1525 with Chi Square = 2801.05 with prob < 0

```

```

The root mean square of the residuals (RMSA) is 0.08
The df corrected root mean square of the residuals is 0.09

```

```

Tucker Lewis Index of factoring reliability = 0.742
RMSEA index = 0.131 and the 90 % confidence intervals are 0.126 0.135
BIC = 2038.76

```

7.3 IRT based scoring

The primary advantage of IRT analyses is examining the item properties (both difficulty and discrimination). With complete data, the scores based upon simple total scores and based upon IRT are practically identical (this may be seen in the examples for `score.irt`). However, when working with data such as those found in the Synthetic Aperture Personality Assessment (*SAPA*) project, it is advantageous to use IRT based scoring. *SAPA* data might have 2-3 items/person sampled from scales with 10-20 items. Simply finding the average of the three (classical test theory) fails to consider that the items might differ in either discrimination or in difficulty. The `score.irt` function applies basic IRT to this problem.

Consider 1000 randomly generated subjects with scores on 9 true/false items differing in difficulty. Selectively drop the hardest items for the 1/3 lowest subjects, and the 4 easiest items for the 1/3 top subjects (this is a crude example of what tailored testing would do). Then score these subjects:

```

> v9 <- sim.irt(9,1000,-2.,2.,mod="normal") #dichotomous items
> items <- v9$items
> test <- irt.fa(items)

```

```
> om <- omega(iq.irt$rho, 4)
```

Omega

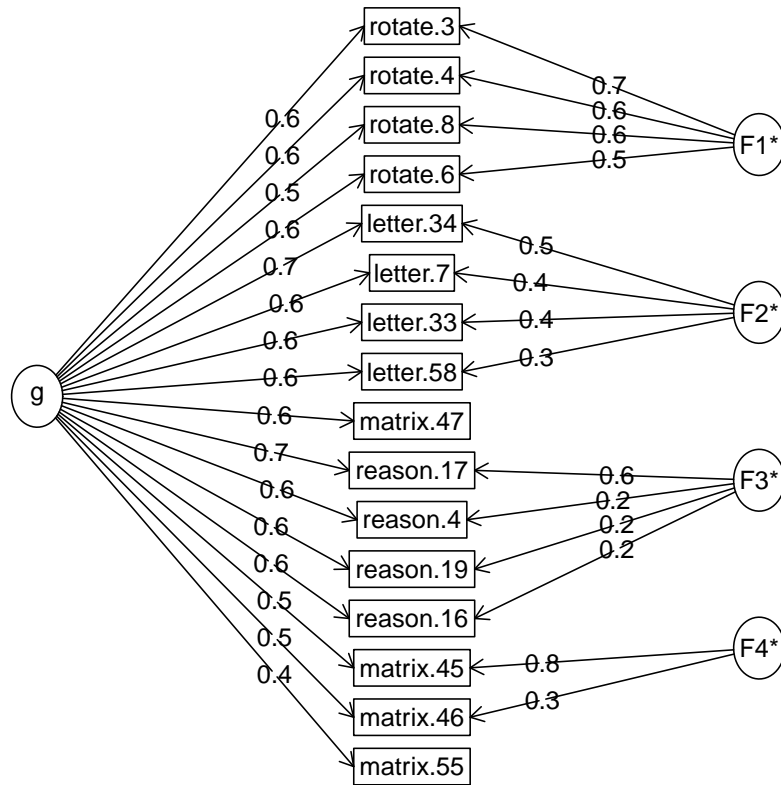


Figure 21: An Omega analysis of 16 ability items sampled from the SAPA project. The items represent a general factor as well as four lower level factors. The analysis is done using the tetrachoric correlations found in the previous `irt.fa` analysis. The four matrix items have some serious problems, which may be seen later when examine the item response functions.

```

> total <- rowSums(items)
> ord <- order(total)
> items <- items[ord,]
> #now delete some of the data - note that they are ordered by score
> items[1:333,5:9] <- NA
> items[334:666,3:7] <- NA
> items[667:1000,1:4] <- NA
> scores <- score.irt(test,items)
> unitweighted <- score.irt(items=items,keys=rep(1,9))
> scores.df <- data.frame(true=v9$theta[ord],scores,unitweighted)
> colnames(scores.df) <- c("True theta","irt theta","total","fit","rasch","total","fit")

```

These results are seen in Figure 22.

8 Multilevel modeling

Correlations between individuals who belong to different natural groups (based upon e.g., ethnicity, age, gender, college major, or country) reflect an unknown mixture of the pooled correlation within each group as well as the correlation of the means of these groups. These two correlations are independent and do not allow inferences from one level (the group) to the other level (the individual). When examining data at two levels (e.g., the individual and by some grouping variable), it is useful to find basic descriptive statistics (means, sds, ns per group, within group correlations) as well as between group statistics (over all descriptive statistics, and overall between group correlations). Of particular use is the ability to decompose a matrix of correlations at the individual level into correlations within group and correlations between groups.

8.1 Decomposing data into within and between level correlations using statsBy

There are at least two very powerful packages (*nlme* and *multilevel*) which allow for complex analysis of hierarchical (multilevel) data structures. `statsBy` is a much simpler function to give some of the basic descriptive statistics for two level models.

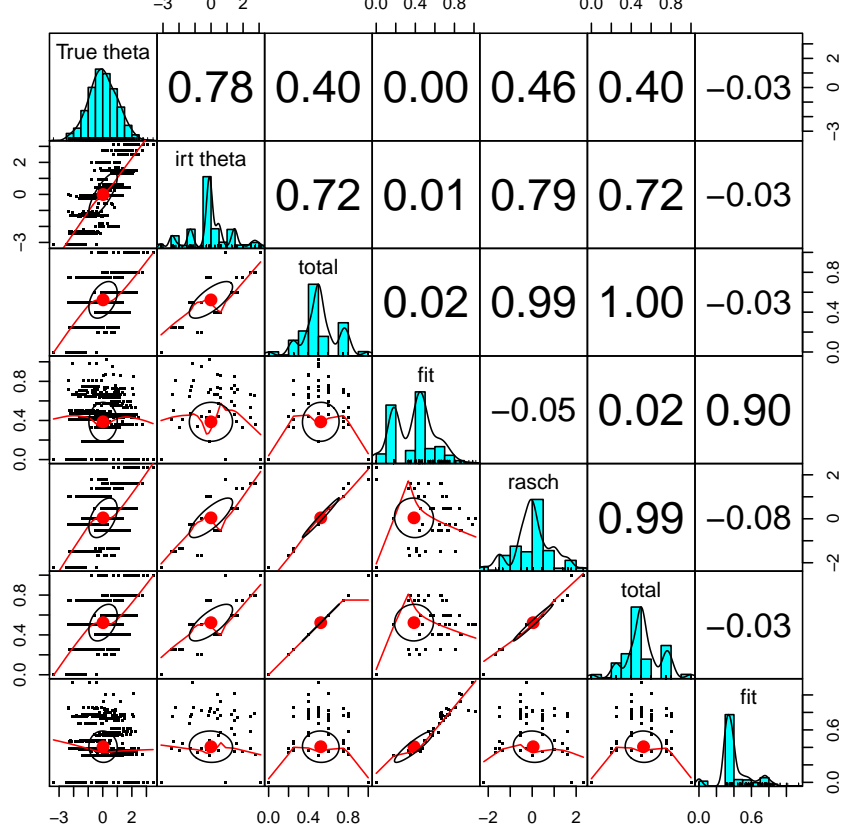
This follows the decomposition of an observed correlation into the pooled correlation within groups (rwg) and the weighted correlation of the means between groups which is discussed by Pedhazur (1997) and by Bliese (2009) in the multilevel package.

$$r_{xy} = \eta_{x_{wg}} * \eta_{y_{wg}} * r_{xy_{wg}} + \eta_{x_{bg}} * \eta_{y_{bg}} * r_{xy_{bg}} \quad (3)$$

Figure 22: IRT based scoring and total test scores for 1000 simulated subjects. True theta values are reported and then the IRT and total scoring systems.

```
> pairs.panels(scores.df,pch='.',gap=0)
> title('Comparing true theta for IRT, Rasch and classically based scoring',1)
```

Comparing true theta for IRT, Rasch and classically based scoring



where r_{xy} is the normal correlation which may be decomposed into a within group and between group correlations $r_{xy_{wg}}$ and $r_{xy_{bg}}$ and η (eta) is the correlation of the data with the within group values, or the group means.

8.2 Generating and displaying multilevel data

`withinBetween` is an example data set of the mixture of within and between group correlations. The within group correlations between 9 variables are set to be 1, 0, and -1 while those between groups are also set to be 1, 0, -1. These two sets of correlations are crossed such that V1, V4, and V7 have within group correlations of 1, as do V2, V5 and V8, and V3, V6 and V9. V1 has a within group correlation of 0 with V2, V5, and V8, and a -1 within group correlation with V3, V6 and V9. V1, V2, and V3 share a between group correlation of 1, as do V4, V5 and V6, and V7, V8 and V9. The first group has a 0 between group correlation with the second and a -1 with the third group. See the help file for `withinBetween` to display these data.

`sim.multilevel` will generate simulated data with a multilevel structure.

The `statsBy.boot` function will randomize the grouping variable `ntrials` times and find the `statsBy` output. This can take a long time and will produce a great deal of output. This output can then be summarized for relevant variables using the `statsBy.boot.summary` function specifying the variable of interest.

Consider the case of the relationship between various tests of ability when the data are grouped by level of education (`statsBy(sat.act)`) or when affect data are analyzed within and between an affect manipulation (`statsBy(affect)`).

9 Set Correlation and Multiple Regression from the correlation matrix

An important generalization of multiple regression and multiple correlation is *set correlation* developed by Cohen (1982) and discussed by Cohen et al. (2003). Set correlation is a multivariate generalization of multiple regression and estimates the amount of variance shared between two sets of variables. Set correlation also allows for examining the relationship between two sets when controlling for a third set. This is implemented in the `set.cor` function. Set correlation is

$$R^2 = 1 - \prod_{i=1}^n (1 - \lambda_i)$$

where λ_i is the i th eigen value of the eigen value decomposition of the matrix

$$R = R_{xx}^{-1} R_{xy} R_{xx}^{-1} R_{xy}^{-1}.$$

Unfortunately, there are several cases where set correlation will give results that are much too high. This will happen if some variables from the first set are highly related to those in the second set, even though most are not. In this case, although the set correlation can be very high, the degree of relationship between the sets is not as high. In this case, an alternative statistic, based upon the average canonical correlation might be more appropriate.

`set.cor` has the additional feature that it will calculate multiple and partial correlations from the correlation or covariance matrix rather than the original data.

Consider the correlations of the 6 variables in the `sat.act` data set. First do the normal multiple regression, and then compare it with the results using `set.cor`. Two things to notice. `set.cor` works on the *correlation* or *covariance* or *raw data* matrix, and thus if using the correlation matrix, will report standardized $\hat{\beta}$ weights. Secondly, it is possible to do several multiple regressions simultaneously. If the number of observations is specified, or if the analysis is done on raw data, statistical tests of significance are applied.

For this example, the analysis is done on the correlation matrix rather than the raw data.

```
> C <- cov(sat.act,use="pairwise")
> modell <- lm(ACT~ gender + education + age, data=sat.act)
> summary(modell)

Call:
lm(formula = ACT ~ gender + education + age, data = sat.act)

Residuals:
    Min       1Q   Median       3Q      Max
-25.2458  -3.2133   0.7769   3.5921   9.2630

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  27.41706    0.82140  33.378 < 2e-16 ***
gender       -0.48606    0.37984  -1.280  0.20110
education     0.47890    0.15235   3.143  0.00174 **
age           0.01623    0.02278   0.712  0.47650
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.768 on 696 degrees of freedom
Multiple R-squared:  0.0272,    Adjusted R-squared:  0.02301
F-statistic: 6.487 on 3 and 696 DF,  p-value: 0.0002476

Compare this with the output from set.cor.

> #compare with mat.regress
> set.cor(c(4:6),c(1:3),C, n.obs=700)
```

```
Call: setCor(y = y, x = x, data = data, z = z, n.obs = n.obs, use = use,
  std = std, square = square, main = main, plot = plot)
```

Multiple Regression from matrix input

Beta weights

| | ACT | SATV | SATQ |
|-----------|-------|-------|-------|
| gender | -0.05 | -0.03 | -0.18 |
| education | 0.14 | 0.10 | 0.10 |
| age | 0.03 | -0.10 | -0.09 |

Multiple R

| | ACT | SATV | SATQ |
|--|------|------|------|
| | 0.16 | 0.10 | 0.19 |

multiple R2

| | ACT | SATV | SATQ |
|--|--------|--------|--------|
| | 0.0272 | 0.0096 | 0.0359 |

Unweighted multiple R

| | ACT | SATV | SATQ |
|--|------|------|------|
| | 0.15 | 0.05 | 0.11 |

Unweighted multiple R2

| | ACT | SATV | SATQ |
|--|------|------|------|
| | 0.02 | 0.00 | 0.01 |

SE of Beta weights

| | ACT | SATV | SATQ |
|-----------|------|------|------|
| gender | 0.18 | 4.29 | 4.34 |
| education | 0.22 | 5.13 | 5.18 |
| age | 0.22 | 5.11 | 5.16 |

t of Beta Weights

| | ACT | SATV | SATQ |
|-----------|-------|-------|-------|
| gender | -0.27 | -0.01 | -0.04 |
| education | 0.65 | 0.02 | 0.02 |
| age | 0.15 | -0.02 | -0.02 |

Probability of t <

| | ACT | SATV | SATQ |
|-----------|------|------|------|
| gender | 0.79 | 0.99 | 0.97 |
| education | 0.51 | 0.98 | 0.98 |
| age | 0.88 | 0.98 | 0.99 |

Shrunken R2

| | ACT | SATV | SATQ |
|--|--------|--------|--------|
| | 0.0230 | 0.0054 | 0.0317 |

Standard Error of R2

| | ACT | SATV | SATQ |
|--|--------|--------|--------|
| | 0.0120 | 0.0073 | 0.0137 |

F

| | ACT | SATV | SATQ |
|--|------|------|------|
| | 6.49 | 2.26 | 8.63 |

Probability of F <

| | ACT | SATV | SATQ |
|--|----------|----------|----------|
| | 2.48e-04 | 8.08e-02 | 1.24e-05 |

```

degrees of freedom of regression
[1] 3 696

Various estimates of between set correlations
Squared Canonical Correlations
[1] 0.050 0.033 0.008
Chisq of canonical correlations
[1] 35.8 23.1 5.6

Average squared canonical correlation = 0.03
Cohen's Set Correlation R2 = 0.09
Shrunken Set Correlation R2 = 0.08
F and df of Cohen's Set Correlation 7.26 9 1681.86
Unweighted correlation between the two sets = 0.01

```

Note that the `set.cor` analysis also reports the amount of shared variance between the predictor set and the criterion (dependent) set. This set correlation is symmetric. That is, the R^2 is the same independent of the direction of the relationship.

10 Simulation functions

It is particularly helpful, when trying to understand psychometric concepts, to be able to generate sample data sets that meet certain specifications. By knowing “truth” it is possible to see how well various algorithms can capture it. Several of the `sim` functions create artificial data sets with known structures.

A number of functions in the `psych` package will generate simulated data. These functions include `sim` for a factor simplex, and `sim.simplex` for a data simplex, `sim.circ` for a circumplex structure, `sim.congeneric` for a one factor factor congeneric model, `sim.dichot` to simulate dichotomous items, `sim.hierarchical` to create a hierarchical factor model, `sim.item` is a more general item simulation, `sim.minor` to simulate major and minor factors, `sim.omega` to test various examples of omega, `sim.parallel` to compare the efficiency of various ways of determining the number of factors, `sim.rasch` to create simulated rasch data, `sim.irt` to create general 1 to 4 parameter IRT data by calling `sim.npl` 1 to 4 parameter logistic IRT or `sim.npn` 1 to 4 parameter normal IRT, `sim.structural` a general simulation of structural models, and `sim.anova` for ANOVA and lm simulations, and `sim.vss`. Some of these functions are separately documented and are listed here for ease of the help function. See each function for more detailed help.

`sim` The default version is to generate a four factor simplex structure over three occasions, although more general models are possible.

`sim.simple` Create major and minor factors. The default is for 12 variables with 3 major factors and 6 minor factors.

`sim.structure` To combine a measurement and structural model into one data matrix. Useful for understanding structural equation models.

`sim.hierarchical` To create data with a hierarchical (bifactor) structure.

`sim.congeneric` To create congeneric items/tests for demonstrating classical test theory. This is just a special case of `sim.structure`.

`sim.circ` To create data with a circumplex structure.

`sim.item` To create items that either have a simple structure or a circumplex structure.

`sim.dichot` Create dichotomous item data with a simple or circumplex structure.

`sim.rasch` Simulate a 1 parameter logistic (Rasch) model.

`sim.irt` Simulate a 2 parameter logistic (2PL) or 2 parameter Normal model. Will also do 3 and 4 PL and PN models.

`sim.multilevel` Simulate data with different within group and between group correlational structures.

Some of these functions are described in more detail in the companion vignette: [psych for sem](#).

The default values for `sim.structure` is to generate a 4 factor, 12 variable data set with a simplex structure between the factors.

Two data structures that are particular challenges to exploratory factor analysis are the simplex structure and the presence of minor factors. Simplex structures `sim.simplex` will typically occur in developmental or learning contexts and have a correlation structure of r between adjacent variables and r^n for variables n apart. Although just one latent variable (r) needs to be estimated, the structure will have $nvar-1$ factors.

Many simulations of factor structures assume that except for the major factors, all residuals are normally distributed around 0. An alternative, and perhaps more realistic situation, is that there are a few major (big) factors and many minor (small) factors. The challenge is thus to identify the major factors. `sim.minor` generates such structures. The structures generated can be thought of as having a major factor structure with some small correlated residuals.

Although coefficient ω_h is a very useful indicator of the general factor saturation of a unifactorial test (one with perhaps several sub factors), it has problems with the case of multiple, independent factors. In this situation, one of the factors is labelled as “general” and the omega estimate is too large. This situation may be explored using the `sim.omega` function.

The four irt simulations, `sim.rasch`, `sim.irt`, `sim.npl` and `sim.npn`, simulate dichotomous items following the Item Response model. `sim.irt` just calls either `sim.npl` (for logistic models) or `sim.npn` (for normal models) depending upon the specification of the model.

The logistic model is

$$P(x|\theta_i, \delta_j, \gamma_j, \zeta_j) = \gamma_j + \frac{\zeta_j - \gamma_j}{1 + e^{\alpha_j(\delta_j - \theta_i)}}. \quad (4)$$

where γ is the lower asymptote or guessing parameter, ζ is the upper asymptote (normally 1), α_j is item discrimination and δ_j is item difficulty. For the 1 Parameter Logistic (Rasch) model, $\gamma=0$, $\zeta=1$, $\alpha=1$ and item difficulty is the only free parameter to specify.

(Graphics of these may be seen in the demonstrations for the logistic function.)

The normal model (`irt.npn` calculates the probability using `pnorm` instead of the logistic function used in `irt.npl`, but the meaning of the parameters are otherwise the same. With the $\alpha = 1.702$ in the logistic model the two models are practically identical.

11 Graphical Displays

Many of the functions in the *psych* package include graphic output and examples have been shown in the previous figures. After running `fa`, `iclust`, `omega`, `irt.fa`, plotting the resulting object is done by the `plot.psych` function as well as specific diagram functions. e.g., (but not shown)

```
f3 <- fa(Thurstone, 3)
plot(f3)
fa.diagram(f3)
c <- iclust(Thurstone)
plot(c) #a pretty boring plot
iclust.diagram(c) #a better diagram
c3 <- iclust(Thurstone, 3)
plot(c3) #a more interesting plot
data(bfi)
e.irt <- irt.fa(bfi[11:15])
plot(e.irt)
ot <- omega(Thurstone)
plot(ot)
omega.diagram(ot)
```

The ability to show path diagrams to represent factor analytic and structural models is discussed in somewhat more detail in the accompanying vignette, [psych for sem](#). Basic routines to draw path diagrams are included in the `dia.rect` and accompanying functions. These are used by the `fa.diagram`, `structure.diagram` and `iclust.diagram` functions.

```

> xlim=c(0,10)
> ylim=c(0,10)
> plot(NA,xlim=xlim,ylim=ylim,main="Demonstration of dia functions",axes=FALSE,xlab="",ylab="")
> ul <- dia.rect(1,9,labels="upper left",xlim=xlim,ylim=ylim)
> ll <- dia.rect(1,3,labels="lower left",xlim=xlim,ylim=ylim)
> lr <- dia.ellipse(9,3,"lower right",xlim=xlim,ylim=ylim)
> ur <- dia.ellipse(9,9,"upper right",xlim=xlim,ylim=ylim)
> ml <- dia.ellipse(3,6,"middle left",xlim=xlim,ylim=ylim)
> mr <- dia.ellipse(7,6,"middle right",xlim=xlim,ylim=ylim)
> bl <- dia.ellipse(1,1,"bottom left",xlim=xlim,ylim=ylim)
> br <- dia.ellipse(9,1,"bottom right",xlim=xlim,ylim=ylim)
> dia.arrow(from=lr,to=ul,labels="right to left")
> dia.arrow(from=ul,to=ur,labels="left to right")
> dia.curved.arrow(from=lr,to=ll$right,labels="right to left")
> dia.curved.arrow(to=ur,from=ul$right,labels="left to right")
> dia.curve(ll$top,ul$bottom,"double") #for rectangles, specify where to point
> dia.curved.arrow(mr,ur,"up") #but for ellipses, just point to it.
> dia.curve(ml,mr,"across")
> dia.arrow(ur,lr,"top down")
> dia.curved.arrow(br$top,lr$bottom,"up")
> dia.curved.arrow(bl,br,"left to right")
> dia.arrow(bl,ll$bottom)
> dia.curved.arrow(ml,ll$right)
> dia.curved.arrow(mr,lr$top)

```

Demonstration of dia functions

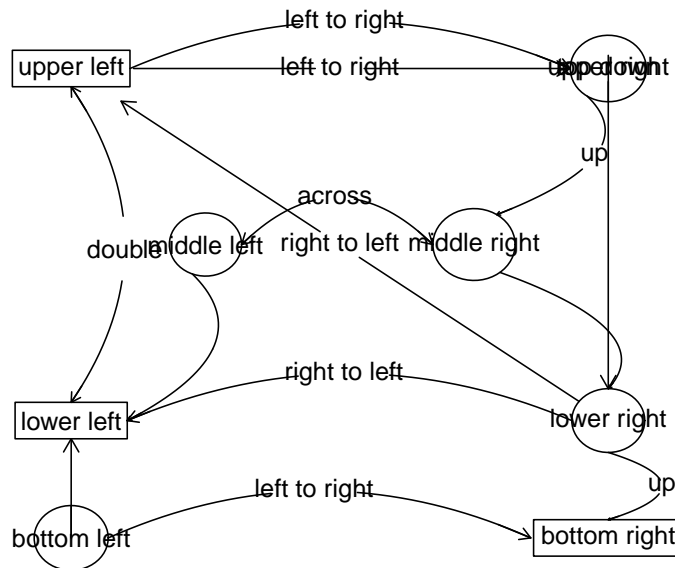


Figure 23: The basic graphic capabilities of the dia functions are shown in this figure.

12 Miscellaneous functions

A number of functions have been developed for some very specific problems that don't fit into any other category. The following is an incomplete list. Look at the *Index* for *psych* for a list of all of the functions.

`block.random` Creates a block randomized structure for `n` independent variables. Useful for teaching block randomization for experimental design.

`df2latex` is useful for taking tabular output (such as a correlation matrix or that of `describe` and converting it to a L^AT_EX table. May be used when Sweave is not convenient.

`cor2latex` Will format a correlation matrix in APA style in a L^AT_EX table. See also `fa2latex` and `irt2latex`.

`cosinor` One of several functions for doing *circular statistics*. This is important when studying mood effects over the day which show a diurnal pattern. See also `circadian.mean`, `circadian.cor` and `circadian.linear.cor` for finding circular means, circular correlations, and correlations of circular with linear data.

`fisherz` Convert a correlation to the corresponding Fisher *z* score.

`geometric.mean` also `harmonic.mean` find the appropriate mean for working with different kinds of data.

`ICC` and `cohen.kappa` are typically used to find the reliability for raters.

`headtail` combines the `head` and `tail` functions to show the first and last lines of a data set or output.

`topBottom` Same as `headtail`. Combines the `head` and `tail` functions to show the first and last lines of a data set or output, but does not add ellipsis between.

`mardia` calculates univariate or multivariate (Mardia's test) skew and kurtosis for a vector, matrix, or `data.frame`

`p.rep` finds the probability of replication for an *F*, *t*, or *r* and estimate effect size.

`partial.r` partials a *y* set of variables out of an *x* set and finds the resulting partial correlations. (See also `set.cor`.)

`rangeCorrection` will correct correlations for restriction of range.

`reverse.code` will reverse code specified items. Done more conveniently in most *psych* functions, but supplied here as a helper function when using other packages.

superMatrix Takes two or more matrices, e.g., A and B, and combines them into a “Super matrix” with A on the top left, B on the lower right, and 0s for the other two quadrants. A useful trick when forming complex keys, or when forming example problems.

13 Data sets

A number of data sets for demonstrating psychometric techniques are included in the *psych* package. These include six data sets showing a hierarchical factor structure (five cognitive examples, **Thurstone**, **Thurstone.33**, **Holzinger**, **Bechtoldt.1**, **Bechtoldt.2**, and one from health psychology **Reise**). One of these (**Thurstone**) is used as an example in the *sem* package as well as **McDonald (1999)**. The original data are from **Thurstone and Thurstone (1941)** and reanalyzed by **Bechtoldt (1961)**. Personality item data representing five personality factors on 25 items (**bfi**) or 13 personality inventory scores (**epi.bfi**), and 14 multiple choice iq items (**iqitems**). The **vegetables** example has paired comparison preferences for 9 vegetables. This is an example of Thurstonian scaling used by **Guilford (1954)** and **Nunnally (1967)**. Other data sets include **cubits**, **peas**, and **heights** from Galton.

Thurstone Holzinger-Swineford (1937) introduced the bifactor model of a general factor and uncorrelated group factors. The Holzinger correlation matrix is a 14 * 14 matrix from their paper. The Thurstone correlation matrix is a 9 * 9 matrix of correlations of ability items. The Reise data set is 16 * 16 correlation matrix of mental health items. The Bechtoldt data sets are both 17 x 17 correlation matrices of ability tests.

bfi 25 personality self report items taken from the International Personality Item Pool (ipip.ori.org) were included as part of the Synthetic Aperture Personality Assessment (*SAPA*) web based personality assessment project. The data from 2800 subjects are included here as a demonstration set for scale construction, factor analysis and Item Response Theory analyses.

sat.act Self reported scores on the SAT Verbal, SAT Quantitative and ACT were collected as part of the Synthetic Aperture Personality Assessment (*SAPA*) web based personality assessment project. Age, gender, and education are also reported. The data from 700 subjects are included here as a demonstration set for correlation and analysis.

epi.bfi A small data set of 5 scales from the Eysenck Personality Inventory, 5 from a Big 5 inventory, a Beck Depression Inventory, and State and Trait Anxiety measures. Used for demonstrations of correlations, regressions, graphic displays.

iq 14 multiple choice ability items were included as part of the Synthetic Aperture Personality Assessment (*SAPA*) web based personality assessment project. The data from 1000 subjects are included here as a demonstration set for scoring multiple choice inventories and doing basic item statistics.

galton Two of the earliest examples of the correlation coefficient were Francis Galton's data sets on the relationship between mid parent and child height and the similarity of parent generation peas with child peas. **galton** is the data set for the Galton height. **peas** is the data set Francis Galton used to introduce the correlation coefficient with an analysis of the similarities of the parent and child generation of 700 sweet peas.

Dwyer [Dwyer \(1937\)](#) introduced a method for *factor extension* (see **fa.extension** that finds loadings on factors from an original data set for additional (extended) variables. This data set includes his example.

miscellaneous **cities** is a matrix of airline distances between 11 US cities and may be used for demonstrating multiple dimensional scaling. **vegetables** is a classic data set for demonstrating Thurstonian scaling and is the preference matrix of 9 vegetables from [Guilford \(1954\)](#). Used by [Guilford \(1954\)](#); [Nunnally \(1967\)](#); [Nunnally and Bernstein \(1994\)](#), this data set allows for examples of basic scaling techniques.

14 Development version and a users guide

The most recent development version is available as a source file at the repository maintained at <http://personality-project.org/r>. That version will have removed the most recently discovered bugs (but perhaps introduced other, yet to be discovered ones). To install this development version, either for PCs or Macs,

R code

```
install.packages("psych", repos = "http://personality-project.org/r", type = "source")
```

Although the individual help pages for the *psych* package are available as part of R and may be accessed directly (e.g. `?psych`), the full manual for the **psych** package is also available as a pdf at http://personality-project.org/r/psych_manual.pdf

News and a history of changes are available in the NEWS and CHANGES files in the source files. To view the most recent news,

```
> news(Version >= "1.8.4", package="psych")
```

15 Psychometric Theory

The *psych* package has been developed to help psychologists do basic research. Many of the functions were developed to supplement a book (<http://personality-project.org/r/book> An introduction to Psychometric Theory with Applications in R (Revelle, prep) More information about the use of some of the functions may be found in the book .

For more extensive discussion of the use of *psych* in particular and R in general, consult <http://personality-project.org/r/r.guide.html> A short guide to R.

16 SessionInfo

This document was prepared using the following settings.

```
> sessionInfo()

R Under development (unstable) (2018-04-05 r74542)
Platform: x86_64-apple-darwin15.6.0 (64-bit)
Running under: macOS High Sierra 10.13.5

Matrix products: default
BLAS: /Library/Frameworks/R.framework/Versions/3.6/Resources/lib/libRblas.0.dylib
LAPACK: /Library/Frameworks/R.framework/Versions/3.6/Resources/lib/libRlapack.dylib

locale:
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

attached base packages:
[1] stats    graphics  grDevices  utils      datasets  methods   base

other attached packages:
[1] psych_1.8.6.30

loaded via a namespace (and not attached):
[1] compiler_3.6.0  tools_3.6.0    parallel_3.6.0  foreign_0.8-70  nlme_3.1-136   mnormt_1.5-5    grid_3.6.0      lattice_0.20-35
```

References

- Bechtoldt, H. (1961). An empirical study of the factor analysis stability hypothesis. *Psychometrika*, 26(4):405–432.
- Blashfield, R. K. (1980). The growth of cluster analysis: Tryon, Ward, and Johnson. *Multivariate Behavioral Research*, 15(4):439 – 458.
- Blashfield, R. K. and Aldenderfer, M. S. (1988). The methods and problems of cluster analysis. In Nesselroade, J. R. and Cattell, R. B., editors, *Handbook of multivariate experimental psychology (2nd ed.)*, pages 447–473. Plenum Press, New York, NY.
- Bliese, P. D. (2009). *Multilevel Modeling in R (2.3) A Brief Introduction to R, the multilevel package and the nlme package*.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2):245–276.
- Cattell, R. B. (1978). *The scientific use of factor analysis*. Plenum Press, New York.
- Cohen, J. (1982). Set correlation as a general multivariate data-analytic method. *Multivariate Behavioral Research*, 17(3).
- Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. L. Erlbaum Associates, Mahwah, N.J., 3rd ed edition.
- Cooksey, R. and Soutar, G. (2006). Coefficient beta and hierarchical item clustering - an analytical procedure for establishing and displaying the dimensionality and homogeneity of summated scales. *Organizational Research Methods*, 9:78–98.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16:297–334.
- Dwyer, P. S. (1937). The determination of the factor loadings of a given test from the known factor loadings of other tests. *Psychometrika*, 2(3):173–178.
- Everitt, B. (1974). *Cluster analysis*. John Wiley & Sons, Cluster analysis. 122 pp. Oxford, England.
- Fox, J., Nie, Z., and Byrnes, J. (2013). *sem: Structural Equation Models*. R package version 3.1-3.
- Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods*, 6(4):430–450.
- Guilford, J. P. (1954). *Psychometric Methods*. McGraw-Hill, New York, 2nd edition.

- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10(4):255–282.
- Hartigan, J. A. (1975). *Clustering Algorithms*. John Wiley & Sons, Inc., New York, NY, USA.
- Henry, D. B., Tolan, P. H., and Gorman-Smith, D. (2005). Cluster analysis in family psychology research. *Journal of Family Psychology*, 19(1):121–132.
- Holzinger, K. and Swineford, F. (1937). The bi-factor method. *Psychometrika*, 2(1):41–54.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2):179–185.
- Horn, J. L. and Engstrom, R. (1979). Cattell’s scree test in relation to Bartlett’s chi-square test and other observations on the number of factors problem. *Multivariate Behavioral Research*, 14(3):283–300.
- Jennrich, R. and Bentler, P. (2011). Exploratory bi-factor analysis. *Psychometrika*, 76(4):537–549.
- Jensen, A. R. and Weng, L.-J. (1994). What is a good g? *Intelligence*, 18(3):231–258.
- Loevinger, J., Gleser, G., and DuBois, P. (1953). Maximizing the discriminating power of a multiple-score test. *Psychometrika*, 18(4):309–317.
- MacCallum, R. C., Browne, M. W., and Cai, L. (2007). Factor analysis models as approximations. In Cudeck, R. and MacCallum, R. C., editors, *Factor analysis at 100: Historical developments and future directions*, pages 153–175. Lawrence Erlbaum Associates Publishers, Mahwah, NJ.
- Martinent, G. and Ferrand, C. (2007). A cluster analysis of precompetitive anxiety: Relationship with perfectionism and trait anxiety. *Personality and Individual Differences*, 43(7):1676–1686.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. L. Erlbaum Associates, Mahwah, N.J.
- Mun, E. Y., von Eye, A., Bates, M. E., and Vaschillo, E. G. (2008). Finding groups using model-based cluster analysis: Heterogeneous emotional self-regulatory processes and heavy alcohol use risk. *Developmental Psychology*, 44(2):481–495.
- Nunnally, J. C. (1967). *Psychometric theory*. McGraw-Hill, New York,.
- Nunnally, J. C. and Bernstein, I. H. (1994). *Psychometric theory*. McGraw-Hill, New York,, 3rd edition.

- Pedhazur, E. (1997). *Multiple regression in behavioral research: explanation and prediction*. Harcourt Brace College Publishers.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Revelle, W. (1979). Hierarchical cluster-analysis and the internal structure of tests. *Multivariate Behavioral Research*, 14(1):57–74.
- Revelle, W. (2018). *psych: Procedures for Personality and Psychological Research*. Northwestern University, Evanston, <https://cran.r-project.org/web/packages=psych>. R package version 1.8.4.
- Revelle, W. (in prep). *An introduction to psychometric theory with applications in R*. Springer.
- Revelle, W., Condon, D., and Wilt, J. (2011). Methodological advances in differential psychology. In Chamorro-Premuzic, T., Furnham, A., and von Stumm, S., editors, *Handbook of Individual Differences*, chapter 2, pages 39–73. Wiley-Blackwell.
- Revelle, W. and Condon, D. M. (2018a). Reliability. In Irwing, P., Booth, T., and Hughes, D. J., editors, *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development*. John Wiley & Sons, London.
- Revelle, W. and Condon, D. M. (2018b). Reliability from α to ω : A tutorial. (*under review*: <https://osf.io/e685p/>).
- Revelle, W. and Rocklin, T. (1979). Very Simple Structure - alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research*, 14(4):403–414.
- Revelle, W., Wilt, J., and Rosenthal, A. (2010). Individual differences in cognition: New methods for examining the personality-cognition link. In Gruszka, A., Matthews, G., and Szymura, B., editors, *Handbook of Individual Differences in Cognition: Attention, Memory and Executive Control*, chapter 2, pages 27–49. Springer, New York, N.Y.
- Revelle, W. and Zinbarg, R. E. (2009). Coefficients alpha, beta, omega and the glb: comments on Sijsma. *Psychometrika*, 74(1):145–154.
- Schmid, J. J. and Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, 22(1):83–90.
- Shrout, P. E. and Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2):420–428.
- Sneath, P. H. A. and Sokal, R. R. (1973). *Numerical taxonomy: the principles and practice of numerical classification*. A Series of books in biology. W. H. Freeman, San Francisco.

- Sokal, R. R. and Sneath, P. H. A. (1963). *Principles of numerical taxonomy*. A Series of books in biology. W. H. Freeman, San Francisco.
- Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101.
- Thorburn, W. M. (1918). The myth of Occam’s razor. *Mind*, 27:345–353.
- Thurstone, L. L. and Thurstone, T. G. (1941). *Factorial studies of intelligence*. The University of Chicago press, Chicago, Ill.
- Tryon, R. C. (1935). A theory of psychological components—an alternative to ”mathematical factors.”. *Psychological Review*, 42(5):425–454.
- Tryon, R. C. (1939). *Cluster analysis*. Edwards Brothers, Ann Arbor, Michigan.
- Velicer, W. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41(3):321–327.
- Zinbarg, R. E., Revelle, W., Yovel, I., and Li, W. (2005). Cronbach’s α , Revelle’s β , and McDonald’s ω_H : Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, 70(1):123–133.
- Zinbarg, R. E., Yovel, I., Revelle, W., and McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale’s indicators: A comparison of estimators for ω_h . *Applied Psychological Measurement*, 30(2):121–144.