An overview of the psych package

William Revelle
Department of Psychology
Northwestern University

January 7, 2017

Contents

0.1 Jump starting the psych package—a guide for the impatient ............ 4

1 Overview of this and related documents 6

2 Getting started 7

3 Basic data analysis 8
  3.1 Data input from the clipboard ................................................. 8
  3.2 Basic descriptive statistics .................................................... 9
    3.2.1 Outlier detection using outlier ....................................... 10
    3.2.2 Basic data cleaning using scrub ..................................... 12
    3.2.3 Recoding categorical variables into dummy coded variables ...... 12
  3.3 Simple descriptive graphics .................................................. 12
    3.3.1 Scatter Plot Matrices .................................................... 13
    3.3.2 Density or violin plots ................................................ 13
    3.3.3 Means and error bars ................................................... 17
    3.3.4 Error bars for tabular data ........................................... 17
    3.3.5 Two dimensional displays of means and errors ..................... 21
    3.3.6 Back to back histograms .............................................. 23
    3.3.7 Correlational structure ................................................. 24
    3.3.8 Heatmap displays of correlational structure ......................... 25
  3.4 Testing correlations .......................................................... 25
  3.5 Polychoric, tetrachoric, polyserial, and biserial correlations .......... 31
  3.6 Multiple regression from data or correlation matrices .................. 31
  3.7 Mediation and Moderation analysis ........................................ 35
4 Item and scale analysis
  4.1 Dimension reduction through factor analysis and cluster analysis ....... 39
    4.1.1 Minimum Residual Factor Analysis ................................. 41
    4.1.2 Principal Axis Factor Analysis .................................. 42
    4.1.3 Weighted Least Squares Factor Analysis .......................... 42
    4.1.4 Principal Components analysis (PCA) .............................. 48
    4.1.5 Hierarchical and bi-factor solutions .............................. 48
    4.1.6 Item Cluster Analysis: iclust ................................... 52
  4.2 Confidence intervals using bootstrapping techniques .................... 55
  4.3 Comparing factor/component/cluster solutions ........................... 55
  4.4 Determining the number of dimensions to extract ........................ 61
    4.4.1 Very Simple Structure ........................................... 63
    4.4.2 Parallel Analysis .................................................. 64
  4.5 Factor extension ....................................................... 66
  4.6 Exploratory Structural Equation Modeling (ESEM) .......................... 66

5 Classical Test Theory and Reliability ..................................... 70
  5.1 Reliability of a single scale .......................................... 71
  5.2 Using $\omega$ to find the reliability of a single scale ................ 75
  5.3 Estimating $\omega_h$ using Confirmatory Factor Analysis .................. 79
    5.3.1 Other estimates of reliability .................................... 80
  5.4 Reliability and correlations of multiple scales within an inventory ...... 81
    5.4.1 Scoring from raw data .............................................. 81
    5.4.2 Forming scales from a correlation matrix ......................... 83
  5.5 Scoring Multiple Choice Items ......................................... 85
  5.6 Item analysis ........................................................... 87
    5.6.1 Exploring the item structure of scales ............................. 87
    5.6.2 Empirical scale construction ...................................... 89

6 Item Response Theory analysis .............................................. 90
  6.1 Factor analysis and Item Response Theory ................................ 91
  6.2 Speeding up analyses .................................................... 95
  6.3 IRT based scoring ....................................................... 96
    6.3.1 1 versus 2 parameter IRT scoring .................................. 100

7 Multilevel modeling ....................................................... 102
  7.1 Decomposing data into within and between level correlations using statsBy 103
  7.2 Generating and displaying multilevel data ................................ 103
  7.3 Factor analysis by groups .............................................. 104

8 Set Correlation and Multiple Regression from the correlation matrix ...... 104
9 Simulation functions 107
10 Graphical Displays 109
11 Converting output to APA style tables using \LaTeX 109
12 Miscellaneous functions 111
13 Data sets 112
14 Development version and a users guide 114
15 Psychometric Theory 114
16 SessionInfo 114
0.1 Jump starting the *psych* package—a guide for the impatient

You have installed *psych* (section 2) and you want to use it without reading much more. What should you do?

1. Activate the *psych* package:
   ```r
   library(psych)
   ```

2. Input your data (section 3.1). 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:
   ```r
   myData <- read.clipboard.tab()
   ```

3. Make sure that what you just read is right. Describe it (section 3.2) and perhaps look at the first and last few lines:
   ```r
   describe(myData)
   headTail(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 3.3.1) Even better, use the `outlier` to detect outliers.
   ```r
   pairs.panels(myData)
   outlier(myData)
   ```

5. Note that you have some weird subjects, probably due to data entry errors. Either edit the data by hand (use the `edit` command) or just `scrub` the data (section 3.2.2).
   ```r
   cleaned <- scrub(myData, max=9)  #e.g., change anything greater than 9 to NA
   ```

6. Graph the data with error bars for each variable (section 3.3.3).
   ```r
   error.bars(myData)
   ```

7. Find the correlations of all of your data.
   - Descriptively (just the values) (section 3.3.7)
     ```r
     r <- lowerCor(myData)
     ```
   - Graphically (section 3.3.8)
     ```r
     corPlot(r)
     ```
   - Inferentially (the values, the ns, and the p values) (section 3.4)
     ```r
     corr.test(myData)
     ```

8. Test for the number of factors in your data using parallel analysis (`fa.parallel`, section 4.4.2) or Very Simple Structure (`vss`, 4.4.1).
   ```r
   fa.parallel(myData)
   vss(myData)
   ```
9. Factor analyze (see section 4.1) 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 4.1.1-4.1.3). Compare the solution to a hierarchical cluster analysis using the ICLUST algorithm (Revelle, 1979) (see section 4.1.6). Also consider a hierarchical factor solution to find coefficient $\omega$ (see 4.1.5).

\[
\text{fa}(\text{myData})
\]
\[
\text{iclust}(\text{myData})
\]
\[
\text{omega}(\text{myData})
\]

If you prefer to do a principal components analysis you may use the \texttt{principal} function. The default is one component.

\[
\text{principal}(\text{myData})
\]

10. Some people like to find coefficient $\alpha$ as an estimate of reliability. This may be done for a single scale using the \texttt{alpha} function (see 5.1). Perhaps more useful is the ability to create several scales as unweighted averages of specified items using the \texttt{scoreItems} function (see 5.4) and to find various estimates of internal consistency for these scales, find their intercorrelations, and find scores for all the subjects.

\[
\text{alpha}(\text{myData}) \quad \texttt{#score all of the items as part of one scale.}
\]
\[
\text{myKeys <- make.keys(nvar=20, list(first = c(1,-3,5,-7,8:10), second=c(2,4,-6,11:15,-16)))}
\]
\[
\text{my.scores <- scoreItems(myKeys,myData) \quad \texttt{#form several scales}}
\]
\[
\text{my.scores} \quad \texttt{#show the highlights of the results}
\]

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

The \textit{psych} package (Revelle, 2015) 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 \url{http://personality-project.org/r/book/}.

Some of the functions (e.g., \texttt{read.clipboard}, \texttt{describe}, \texttt{pairs.panels}, \texttt{scatter.hist}, \texttt{error.bars}, \texttt{multi.hist}, \texttt{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 \texttt{fa} function includes five methods of \textit{factor analysis} (\textit{minimum residual}, \textit{principal axis}, \textit{weighted least squares}, \textit{generalized least squares} and \textit{maximum likelihood} factor analysis). Principal Components Analysis (PCA) is also available through the use of the \texttt{principal} function. Determining the number of factors or components to extract may be done by using the Very Simple Structure (Revelle and Rocklin, 1979) (\texttt{vss}), Minimum Average Partial correlation (Velicer, 1976) (\texttt{MAP}) or parallel analysis (\texttt{fa.parallel}) criteria. These and several other criteria are included in the \texttt{nfactors} function. Two parameter Item Response Theory (IRT) models for dichotomous or polytomous items may be found by factoring \textit{tetrachoric} or \textit{polychoric} correlation matrices and expressing the resulting parameters in terms of location and discrimination using \texttt{irt.fa}.

Bifactor and hierarchical factor structures may be estimated by using Schmid Leiman transformations (Schmid and Leiman, 1957) (\texttt{schmid}) to transform a hierarchical factor structure into a \textit{bifactor} solution (Holzinger and Swineford, 1937).

Scale construction can be done using the Item Cluster Analysis (Revelle, 1979) (\texttt{iclust}) function to determine the structure and to calculate reliability coefficients $\alpha$ (Cronbach, 1951) (\texttt{alpha}, \texttt{scoreItems}, \texttt{score.multiple.choice}), $\beta$ (Revelle, 1979; Revelle and Zinbarg, 2009) (\texttt{iclust}) and McDonald's $\omega_h$ and $\omega_t$ (McDonald, 1999) (\texttt{omega}). Guttman's six estimates of internal consistency reliability (Guttman (1945), as well as additional estimates (Revelle and Zinbarg, 2009) are in the \texttt{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 \texttt{pairs.panels}, correlation “heat maps” (\texttt{corPlot}) factor, cluster, and structural diagrams using \texttt{fa.diagram}, \texttt{iclust.diagram}, \texttt{structure.diagram} and \texttt{het.diagram}, as well as item response characteristics and item and test information characteristic curves \texttt{plot.irt} and \texttt{plot.poly}.

This vignette is meant to give an overview of the \textit{psych} package. That is, it is meant to give a summary of the main functions in the \textit{psych} package with examples of how
they are used for data description, dimension reduction, and scale construction. The extended user manual at psych_manual.pdf includes examples of graphic output and more extensive demonstrations than are found in the help menus. (Also available at http://personality-project.org/r/psych_manual.pdf). The vignette, psych for sem, at psych_for_sem.pdf, discusses how to use psych as a front end to the sem package of John Fox (Fox et al., 2012). (The vignette is also available at http://personality-project.org/r/book/psych_for_sem.pdf).

For a step by step tutorial in the use of the psych package and the base functions in R for basic personality research, see the guide for using R for personality research at http://personalitytheory.org/r/r.short.html. For an introduction to psychometric theory with applications in R, see the draft chapters at http://personality-project.org/r/book).

2 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, factor.minres, factor.pa, factor.wls, 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.

install.packages("ctv")
library(ctv)
task.views("Psychometrics")

The “Psychometrics” task view will install a large number of useful packages. To install the bare minimum for the examples in this vignette, it is necessary to install just 3 packages:

install.packages(list(c("GPArotation","mnormt"))

Because of the difficulty of installing the package Rgraphviz, alternative graphics have been developed and are available as diagram functions. If Rgraphviz is available, some functions will take advantage of it. An alternative is to use “dot” output of commands for any external graphics package that uses the dot language.
3 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)}

3.1 Data input from the clipboard

There are of course many ways to enter data into R. Reading from a local file using read.table 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

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

### 3.2 Basic descriptive statistics

Once the data are read in, then `describe` or `describeBy` will provide basic descriptive 
statistics arranged in a data frame format. Consider the data set `sat.act` which in-
cludes 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 *.

`describeBy` reports descriptive statistics broken down by some categorizing variable (e.g., 
gender, age, etc.)

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

<table>
<thead>
<tr>
<th>vars</th>
<th>n</th>
<th>mean</th>
<th>sd</th>
<th>median</th>
<th>trimmed</th>
<th>mad</th>
<th>min</th>
<th>max</th>
<th>range</th>
<th>skew</th>
</tr>
</thead>
<tbody>
<tr>
<td>gender</td>
<td>700</td>
<td>1.65</td>
<td>0.48</td>
<td>2.00</td>
<td>1.68</td>
<td>0.00</td>
<td>1.00</td>
<td>2.00</td>
<td>1.00</td>
<td>-0.61</td>
</tr>
<tr>
<td>education</td>
<td>2700</td>
<td>3.16</td>
<td>1.43</td>
<td>3.00</td>
<td>3.31</td>
<td>1.48</td>
<td>0.00</td>
<td>5.00</td>
<td>5.00</td>
<td>-0.68</td>
</tr>
<tr>
<td>age</td>
<td>3700</td>
<td>26.59</td>
<td>9.50</td>
<td>22.00</td>
<td>23.86</td>
<td>5.93</td>
<td>13.00</td>
<td>65.00</td>
<td>52.00</td>
<td>1.64</td>
</tr>
<tr>
<td>ACT</td>
<td>4700</td>
<td>28.55</td>
<td>4.82</td>
<td>29.00</td>
<td>28.84</td>
<td>4.45</td>
<td>3.00</td>
<td>36.00</td>
<td>33.00</td>
<td>-0.66</td>
</tr>
<tr>
<td>SATV</td>
<td>5700</td>
<td>612.23</td>
<td>112.90</td>
<td>620.00</td>
<td>619.45</td>
<td>118.61</td>
<td>200.00</td>
<td>800.00</td>
<td>600.00</td>
<td>-0.64</td>
</tr>
<tr>
<td>SATQ</td>
<td>6687</td>
<td>610.22</td>
<td>115.64</td>
<td>620.00</td>
<td>617.25</td>
<td>118.61</td>
<td>200.00</td>
<td>800.00</td>
<td>600.00</td>
<td>-0.59</td>
</tr>
</tbody>
</table>

These data may then be analyzed by groups defined in a logical statement or by some other 
variable. E.g., break down the descriptive data for males or females. These descriptive 
data can also be seen graphically using the `error.bars.by` function (Figure 5). By setting 
`skew=FALSE` and `ranges=FALSE`, the output is limited to the most basic statistics.
> # basic descriptive statistics by a grouping variable.
> describeBy(sat.act,sat.act$gender,skew=FALSE,ranges=FALSE)

$'1'$

<table>
<thead>
<tr>
<th>vars</th>
<th>n</th>
<th>mean</th>
<th>sd</th>
<th>se</th>
</tr>
</thead>
<tbody>
<tr>
<td>gender</td>
<td>1</td>
<td>247</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>education</td>
<td>2</td>
<td>247</td>
<td>3.00</td>
<td>1.54</td>
</tr>
<tr>
<td>age</td>
<td>3</td>
<td>247</td>
<td>25.86</td>
<td>9.74</td>
</tr>
<tr>
<td>ACT</td>
<td>4</td>
<td>247</td>
<td>28.79</td>
<td>5.06</td>
</tr>
<tr>
<td>SATV</td>
<td>5</td>
<td>247</td>
<td>615.11</td>
<td>114.16</td>
</tr>
<tr>
<td>SATQ</td>
<td>6</td>
<td>245</td>
<td>635.87</td>
<td>116.02</td>
</tr>
</tbody>
</table>

$'2'$

<table>
<thead>
<tr>
<th>vars</th>
<th>n</th>
<th>mean</th>
<th>sd</th>
<th>se</th>
</tr>
</thead>
<tbody>
<tr>
<td>gender</td>
<td>1</td>
<td>453</td>
<td>2.00</td>
<td>0.00</td>
</tr>
<tr>
<td>education</td>
<td>2</td>
<td>453</td>
<td>3.26</td>
<td>1.35</td>
</tr>
<tr>
<td>age</td>
<td>3</td>
<td>453</td>
<td>25.45</td>
<td>9.37</td>
</tr>
<tr>
<td>ACT</td>
<td>4</td>
<td>453</td>
<td>28.42</td>
<td>4.69</td>
</tr>
<tr>
<td>SATV</td>
<td>5</td>
<td>453</td>
<td>610.66</td>
<td>112.31</td>
</tr>
<tr>
<td>SATQ</td>
<td>6</td>
<td>442</td>
<td>596.00</td>
<td>113.07</td>
</tr>
</tbody>
</table>

attr("call")
by.data.frame(data = x, INDICES = group, FUN = describe, type = type,
                   skew = FALSE, ranges = FALSE)

The output from the `describeBy` function can be forced into a matrix form for easy analysis by other programs. In addition, describeBy can group by several grouping variables at the same time.

> sa.mat <- describeBy(sat.act,list(sat.act$gender,sat.act$education),
                      + skew=FALSE,ranges=FALSE,mat=TRUE)
> headTail(sa.mat)

<table>
<thead>
<tr>
<th>item</th>
<th>group1</th>
<th>group2</th>
<th>vars</th>
<th>n</th>
<th>mean</th>
<th>sd</th>
<th>se</th>
</tr>
</thead>
<tbody>
<tr>
<td>gender1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>27</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>gender2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>30</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>gender3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>20</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>gender4</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>25</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SATV9</td>
<td>69</td>
<td>1</td>
<td>4</td>
<td>6</td>
<td>51</td>
<td>635.9</td>
<td>104.12</td>
</tr>
<tr>
<td>SATV10</td>
<td>70</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>86</td>
<td>597.59</td>
<td>106.24</td>
</tr>
<tr>
<td>SATV11</td>
<td>71</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>46</td>
<td>657.83</td>
<td>89.61</td>
</tr>
<tr>
<td>SATV12</td>
<td>72</td>
<td>2</td>
<td>5</td>
<td>6</td>
<td>93</td>
<td>606.72</td>
<td>105.58</td>
</tr>
</tbody>
</table>

### 3.2.1 Outlier detection using `outlier`

One way to detect unusual data is to consider how far each data point is from the multivariate centroid of the data. That is, find the squared Mahalanobis distance for each data point and then compare these to the expected values of $\chi^2$. This produces a Q-Q (quantile-quantile) plot with the n most extreme data points labeled (Figure 1). The outlier values are in the vector d2.
> png( 'outlier.png' )
> d2 <- outlier(sat.act,cex=.8)
> dev.off()
null device
1

Figure 1: Using the `outlier` function to graphically show outliers. The y axis is the Mahalanobis $D^2$, the X axis is the distribution of $\chi^2$ for the same number of degrees of freedom. The outliers detected here may be shown graphically using `pairs.panels` (see 2, and may be found by sorting d2.
3.2.2 Basic data cleaning using scrub

If, after describing the data it is apparent that there were data entry errors that need to be globally replaced with NA, or only certain ranges of data will be analyzed, the data can be “cleaned” using the scrub function.

Consider a data set of 10 rows of 12 columns with values from 1 - 120. All values of columns 3 - 5 that are less than 30, 40, or 50 respectively, or greater than 70 in any of the three columns will be replaced with NA. In addition, any value exactly equal to 45 will be set to NA. (max and isvalue are set to one value here, but they could be a different value for every column).

```r
> x <- matrix(1:120,ncol=10,byrow=TRUE)
> colnames(x) <- paste('V',1:10,sep='')
> new.x <- scrub(x,3:5,min=c(30,40,50),max=70,isvalue=45,newvalue=NA)
> new.x
```

Note that the number of subjects for those columns has decreased, and the minimums have gone up but the maximums down. Data cleaning and examination for outliers should be a routine part of any data analysis.

3.2.3 Recoding categorical variables into dummy coded variables

Sometimes categorical variables (e.g., college major, occupation, ethnicity) are to be analyzed using correlation or regression. To do this, one can form “dummy codes” which are merely binary variables for each category. This may be done using dummy.code. Subsequent analyses using these dummy coded variables may be using biserial or point biserial (regular Pearson r) to show effect sizes and may be plotted in e.g., spider plots.

3.3 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. By default, `error.bars.by` and `error.bars` will show “cats eyes” to graphically show the confidence limits (Figure 5) This may be turned off by specifying `eyes=FALSE`. `densityBy` or `violinBy` may be used to show the distribution of the data in “violin” plots (Figure 4). (These are sometimes called “lava-lamp” plots.)

### 3.3.1 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 2). When plotting many subjects, it is both faster and cleaner to set the plot character (pch) to be ‘.’. (See Figure 2 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. However, in this figure, to show the outliers, we use colors and a larger plot character.

Another example of `pairs.panels` is to show differences between experimental groups. Consider the data in the `affect` data set. The scores reflect post test scores on positive and negative affect and energetic and tense arousal. The colors show the results for four movie conditions: depressing, frightening movie, neutral, and a comedy.

### 3.3.2 Density or violin plots

Graphical presentation of data may be shown using box plots to show the median and 25th and 75th percentiles. A powerful alternative is to show the density distribution using the `violinBy` function (Figure 4).
> png( 'pairspanels.png' )
> sat.d2 <- data.frame(sat.act,d2) #combine the d2 statistics from before with the sat.act data.frame
> pairs.panels(sat.d2,bg=c("yellow","blue")[d2 > 25]+1,pch=21)
> dev.off()
null device

Figure 2: 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. If the plot character were set to a period (pch='.'), it would make a cleaner graphic, but in to show the outliers in color we use the plot characters 21 and 22.
Figure 3: 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. The coloring represent four different movie conditions.
> data(sat.act)
> violinBy(sat.act[,5:6], sat.act$gender, grp.name=c(“M”, “F”), main=“Density Plot by gender for SAT V and Q”)

Figure 4: Using the `violinBy` function to show the distribution of SAT V and Q for males and females. The plot shows the medians, and 25th and 75th percentiles, as well as the entire range and the density distribution.
3.3.3 Means and error bars

Additional descriptive graphics include the ability to draw error bars on sets of data, as well as to draw error bars in both the x and y directions for paired data. These are the functions `error.bars`, `error.bars.by`, `error.bars.tab`, and `error.crosses`.

`error.bars` show the 95% confidence intervals for each variable in a data frame or matrix. These errors are based upon normal theory and the standard errors of the mean. Alternative options include +/- one standard deviation or 1 standard error. If the data are repeated measures, the error bars will be reflect the between variable correlations. By default, the confidence intervals are displayed using a “cats eyes” plot which emphasizes the distribution of confidence within the confidence interval.

`error.bars.by` does the same, but grouping the data by some condition.

`error.bars.tab` draws bar graphs from tabular data with error bars based upon the standard error of proportion ($\sigma_p = \sqrt{pq/N}$)

`error.crosses` draw the confidence intervals for an x set and a y set of the same size.

The use of the `error.bars.by` function allows for graphic comparisons of different groups (see Figure 5). Five personality measures are shown as a function of high versus low scores on a “lie” scale. People with higher lie scores tend to report being more agreeable, conscientious and less neurotic than people with lower lie scores. The error bars are based upon normal theory and thus are symmetric rather than reflect any skewing in the data.

Although not recommended, it is possible to use the `error.bars` function to draw bar graphs with associated error bars. (This kind of dynamite plot (Figure 7) can be very misleading in that the scale is arbitrary. Go to a discussion of the problems in presenting data this way at http://emdbolker.wikidot.com/blog:dynamite. In the example shown, note that the graph starts at 0, although is out of the range. This is a function of using bars, which always are assumed to start at zero. Consider other ways of showing your data.

3.3.4 Error bars for tabular data

However, it is sometimes useful to show error bars for tabular data, either found by the `table` function or just directly input. These may be found using the `error.bars.tab` function.
> data(epi.bfi)
> error.bars.by(epi.bfi[,6:10],epi.bfi$epilie<4)

Figure 5: Using the `error.bars.by` function shows that self reported personality scales on the Big Five Inventory vary as a function of the Lie scale on the EPI. The “cats eyes” show the distribution of the confidence.
> error.bars.by(sat.act[5:6], sat.act$gender, bars=TRUE,  
+ labels=c("Male", "Female"), ylab="SAT score", xlab="")

Figure 6: A “Dynamite plot” of SAT scores as a function of gender is one way of misleading the reader. By using a bar graph, the range of scores is ignored. Bar graphs start from 0.
```r
> T <- with(sat.act.table, table(gender, education))
> rownames(T) <- c("M", "F")
> error.bars.tab(T, way="both", ylab="Proportion of Education Level", xlab="Level of Education",
+ main="Proportion of sample by education level")
```

Figure 7: The proportion of each education level that is Male or Female. By using the way="both" option, the percentages and errors are based upon the grand total. Alternatively, way="columns" finds column wise percentages, way="rows" finds rowwise percentages. The data can be converted to percentages (as shown) or by total count (raw=TRUE). The function invisibly returns the probabilities and standard errors. See the help menu for an example of entering the data as a data.frame.
3.3.5 Two dimensional displays of means and errors

Yet another way to display data for different conditions is to use the `errorCrosses` function. For instance, the effect of various movies on both “Energetic Arousal” and “Tense Arousal” can be seen in one graph and compared to the same movie manipulations on “Positive Affect” and “Negative Affect”. Note how Energetic Arousal is increased by three of the movie manipulations, but that Positive Affect increases following the Happy movie only.
Figure 8: The use of the `errorCircles` function allows for two dimensional displays of means and error bars. The first call to `errorCircles` finds descriptive statistics for the `affect` data.frame based upon the grouping variable of Film. These data are returned and then used by the second call which examines the effect of the same grouping variable upon different measures. The size of the circles represent the relative sample sizes for each group. The data are from the PMC lab and reported in Smillie et al. (2012).
### 3.3.6 Back to back histograms

The `bi.bars` function summarize the characteristics of two groups (e.g., males and females) on a second variable (e.g., age) by drawing back to back histograms (see Figure 9).

```r
> data(bfi)
> with(bfi, {bi.bars(age, gender, ylab="Age", main="Age by males and females")})
```

![Age by males and females](attachment:image.png)

Figure 9: A bar plot of the age distribution for males and females shows the use of `bi.bars`. The data are males and females from 2800 cases collected using the `SAPA` procedure and are available as part of the `bfi` data set.
3.3.7 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.

```r
> 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:

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

> upper <- lowerCor(female[-1])
> both <- lowerUpper(lower,upper)
> round(both,2)

edctn 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:


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

<table>
<thead>
<tr>
<th></th>
<th>education</th>
<th>age</th>
<th>ACT</th>
<th>SATV</th>
<th>SATQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>education</td>
<td>NA</td>
<td>0.09</td>
<td>0.00</td>
<td>-0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>age</td>
<td>0.61</td>
<td>NA</td>
<td>0.07</td>
<td>-0.03</td>
<td>0.13</td>
</tr>
<tr>
<td>ACT</td>
<td>0.16</td>
<td>0.15</td>
<td>NA</td>
<td>0.08</td>
<td>0.02</td>
</tr>
<tr>
<td>SATV</td>
<td>0.02</td>
<td>-0.06</td>
<td>0.61</td>
<td>NA</td>
<td>0.05</td>
</tr>
<tr>
<td>SATQ</td>
<td>0.08</td>
<td>0.04</td>
<td>0.60</td>
<td>0.68</td>
<td>NA</td>
</tr>
</tbody>
</table>
```

### 3.3.8 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 10) or a simulated data set of 24 variables with a circumplex structure (Figure 11). 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.

Yet another way to show structure is to use “spider” plots. Particularly if variables are ordered in some meaningful way (e.g., in a circumplex), a spider plot will show this structure easily. This is just a plot of the magnitude of the correlation as a radial line, with length ranging from 0 (for a correlation of -1) to 1 (for a correlation of 1). (See Figure 12).

### 3.4 Testing correlations

Correlations are wonderful descriptive statistics of the data but some people like to test whether these correlations differ from zero, or differ from each other. The `cor.test` function (in the *stats* package) will test the significance of a single correlation, and the `rcorr` function in the *Hmisc* package will do this for many correlations. In the *psych* package, the `corr.test` function reports the correlation (Pearson, Spearman, or Kendall) between all variables in either one or two data frames or matrices, as well as the number of observations for each case, and the (two-tailed) probability for each correlation. Unfortunately, these probability values have not been corrected for multiple comparisons and so should be taken with a great deal of salt. Thus, in `corr.test` and `corr.p` the raw probabilities are reported below the diagonal and the probabilities adjusted for multiple comparisons using (by default) the Holm correction are reported above the diagonal (Table 1). (See the `p.adjust` function for a discussion of Holm (1979) and other corrections.)

Testing the difference between any two correlations can be done using the `r.test` function. The function actually does four different tests (based upon an article by Steiger (1980),
> png('corplot.png')
> corPlot(Thurstone,numbers=TRUE,upper=FALSE,diag=FALSE,main="9 cognitive variables from Thurstone")
> dev.off()

null device

Figure 10: 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. By default, the complete matrix is shown. Setting upper=FALSE and diag=FALSE shows a cleaner figure.
Figure 11: Using the corPlot 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. For circumplex structures, it is perhaps useful to show the complete matrix.
Figure 12: A spider plot can show circumplex structure very clearly. Circumplex structures are common in the study of affect.
Table 1: The \texttt{corr.test} function reports correlations, cell sizes, and raw and adjusted probability values. \texttt{corr.p} reports the probability values for a correlation matrix. By default, the adjustment used is that of Holm (1979).

```r
> corr.test(sat.act)
Call:corr.test(x = sat.act)
Correlation matrix
          gender education age ACT SATV SATQ
gender  1.00  0.09 -0.02 -0.04 -0.02 -0.17
education 0.09  1.00  0.55  0.15  0.05  0.03
age -0.02  0.55  1.00  0.11 -0.04 -0.03
ACT -0.04  0.15  0.11  1.00  0.56  0.59
SATV -0.02  0.05 -0.04  0.56  1.00  0.64
SATQ -0.17  0.03 -0.03  0.59  0.64  1.00
Sample Size
          gender education age ACT SATV SATQ
gender  700  700  700  700  700  687
education 700  700  700  700  700  687
age  700  700  700  700  700  687
ACT  700  700  700  700  700  687
SATV  700  700  700  700  700  687
SATQ  687  687  687  687  687  687
Probability values (Entries above the diagonal are adjusted for multiple tests.)
          gender education age ACT SATV SATQ
gender  0.00  0.17  1.00  1.00  1  0
education 0.02  0.00  0.00  0.00  1  1
age  0.58  0.00  0.00  0.03  1  1
ACT  0.33  0.00  0.00  0.00  0  0
SATV  0.62  0.22  0.26  0.00  0  0
SATQ  0.00  0.36  0.37  0.00  0  0
```

To see confidence intervals of the correlations, print with the short=FALSE option.
depending upon the input:

1) For a sample size n, find the t and p value for a single correlation as well as the confidence interval.

   > r.test(50,.3)

   Correlation tests
   Call:r.test(n = 50, r12 = 0.3)
   Test of significance of a correlation
   t value 2.18 with probability < 0.034
   and confidence interval 0.02 0.53

2) For sample sizes of n and n2 (n2 = n if not specified) find the z of the difference between the z transformed correlations divided by the standard error of the difference of two z scores.

   > r.test(30,.4,.6)

   Correlation tests
   Call:r.test(n = 30, r12 = 0.4, r34 = 0.6)
   Test of difference between two independent correlations
   z value 0.99 with probability 0.32

3) For sample size n, and correlations r_a= r12, r_b= r23 and r13 specified, test for the difference of two dependent correlations (Steiger case A).

   > r.test(103,.4,.5,.1)

   Correlation tests
   Call:[1] "r.test(n = 103 , r12 = 0.4 , r23 = 0.1 , r13 = 0.5 )"
   Test of difference between two correlated correlations
   t value -0.89 with probability < 0.37

4) For sample size n, test for the difference between two dependent correlations involving different variables. (Steiger case B).

   > r.test(103,.5,.6,.7,.5,.5,.8) #steiger Case B

   Correlation tests
   Call:r.test(n = 103, r12 = 0.5, r34 = 0.6, r23 = 0.7, r13 = 0.5, r14 = 0.5,
   r24 = 0.8)
   Test of difference between two dependent correlations
   z value -1.2 with probability 0.23

To test whether a matrix of correlations differs from what would be expected if the population correlations were all zero, the function cortest follows Steiger (1980) who pointed out that the sum of the squared elements of a correlation matrix, or the Fisher z score equivalents, is distributed as chi square under the null hypothesis that the values are zero (i.e., elements of the identity matrix). This is particularly useful for examining whether correlations in a single matrix differ from zero or for comparing two matrices. Although obvious, cortest can be used to test whether the sat.act data matrix produces non-zero correlations (it does). This is a much more appropriate test when testing whether a residual matrix differs from zero.

   > cortest(sat.act)
3.5 Polychoric, tetrachoric, polyserial, and biserial correlations

The Pearson correlation of dichotomous data is also known as the $\phi$ coefficient. If the data, e.g., ability items, are thought to represent an underlying continuous although latent variable, the $\phi$ 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 (Figure 13). This is also shown in terms of dichotomizing the bivariate normal density function using the draw.cor function (Figure 14). 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 sometimes 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.

3.6 Multiple regression from data or correlation matrices

The typical application of the lm function is to do a linear model of one Y variable as a function of multiple X variables. Because lm is designed to analyze complex interactions, it requires raw data as input. It is, however, sometimes convenient to do multiple regression from a correlation or covariance matrix. The setCor function will do this, taking a set of y variables predicted from a set of x variables, perhaps with a set of z covariates removed from both x and y. Consider the Thurstone correlation matrix and find the multiple correlation of the last five variables as a function of the first 4.

```r
> setCor(y = 5:9,x=1:4,data=Thurstone)
```
Figure 13: The tetrachoric correlation estimates what a Pearson correlation would be given a two by two table of observed values assumed to be sampled from a bivariate normal distribution. The \( \phi \) correlation is just a Pearson r performed on the observed values.
Figure 14: The tetrachoric correlation estimates what a Pearson correlation would be given a two by two table of observed values assumed to be sampled from a bivariate normal distribution. The $\phi$ correlation is just a Pearson $r$ performed on the observed values. It is found (laboriously) by optimizing the fit of the bivariate normal for various values of the correlation to the observed cell frequencies.
Call: setCor(y = 5:9, x = 1:4, data = Thurstone)

Multiple Regression from matrix input

Beta weights

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentences</td>
<td>0.09</td>
<td>0.07</td>
<td>0.25</td>
<td>0.21</td>
<td>0.20</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.09</td>
<td>0.17</td>
<td>0.09</td>
<td>0.16</td>
<td>-0.02</td>
</tr>
<tr>
<td>Sent.Completion</td>
<td>0.02</td>
<td>0.05</td>
<td>0.04</td>
<td>0.21</td>
<td>0.08</td>
</tr>
<tr>
<td>First.Letters</td>
<td>0.58</td>
<td>0.45</td>
<td>0.21</td>
<td>0.08</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Multiple R

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentences</td>
<td>0.69</td>
<td>0.63</td>
<td>0.50</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

multiple R2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentences</td>
<td>0.48</td>
<td>0.40</td>
<td>0.25</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Unweighted multiple R

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentences</td>
<td>0.59</td>
<td>0.58</td>
<td>0.49</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Unweighted multiple R2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentences</td>
<td>0.34</td>
<td>0.34</td>
<td>0.24</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Various estimates of between set correlations

Squared Canonical Correlations

[1] 0.6280 0.1478 0.0076 0.0049

Average squared canonical correlation = 0.2
Cohen's Set Correlation R2 = 0.69
Unweighted correlation between the two sets = 0.73

By specifying the number of subjects in correlation matrix, appropriate estimates of standard errors, t-values, and probabilities are also found. The next example finds the regressions with variables 1 and 2 used as covariates. The $\hat{\beta}$ weights for variables 3 and 4 do not change, but the multiple correlation is much less. It also shows how to find the residual correlations between variables 5-9 with variables 1-4 removed.

```r
> sc <- setCor(y = 5:9, x = 3:4, data = Thurstone, z = 1:2)

Call: setCor(y = 5:9, x = 3:4, data = Thurstone, z = 1:2)

Multiple Regression from matrix input

Beta weights

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sent.Completion</td>
<td>0.02</td>
<td>0.05</td>
<td>0.04</td>
<td>0.21</td>
<td>0.08</td>
</tr>
<tr>
<td>First.Letters</td>
<td>0.58</td>
<td>0.45</td>
<td>0.21</td>
<td>0.08</td>
<td>0.31</td>
</tr>
</tbody>
</table>
```

34
Multiple R

<table>
<thead>
<tr>
<th>Four.Letter.Words</th>
<th>Suffixes</th>
<th>Letter.Series</th>
<th>Pedigrees</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.58</td>
<td>0.46</td>
<td>0.21</td>
<td>0.18</td>
</tr>
<tr>
<td>Letter.Group</td>
<td>0.30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Multiple R2

<table>
<thead>
<tr>
<th>Four.Letter.Words</th>
<th>Suffixes</th>
<th>Letter.Series</th>
<th>Pedigrees</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.331</td>
<td>0.210</td>
<td>0.043</td>
<td>0.032</td>
</tr>
<tr>
<td>Letter.Group</td>
<td>0.092</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Unweighted multiple R

<table>
<thead>
<tr>
<th>Four.Letter.Words</th>
<th>Suffixes</th>
<th>Letter.Series</th>
<th>Pedigrees</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.44</td>
<td>0.35</td>
<td>0.17</td>
<td>0.14</td>
</tr>
<tr>
<td>Letter.Group</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Unweighted multiple R2

<table>
<thead>
<tr>
<th>Four.Letter.Words</th>
<th>Suffixes</th>
<th>Letter.Series</th>
<th>Pedigrees</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.19</td>
<td>0.12</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>Letter.Group</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Various estimates of between set correlations

Squared Canonical Correlations

1 0.405 0.023

Average squared canonical correlation = 0.21
Cohen’s Set Correlation R2 = 0.42
Unweighted correlation between the two sets = 0.48

> round(sc$residual,2)

<table>
<thead>
<tr>
<th>Four.Letter.Words</th>
<th>Suffixes</th>
<th>Letter.Series</th>
<th>Pedigrees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Four.Letter.Words</td>
<td>0.52</td>
<td>0.11</td>
<td>0.09</td>
</tr>
<tr>
<td>Suffixes</td>
<td>0.11</td>
<td>0.60</td>
<td>-0.01</td>
</tr>
<tr>
<td>Letter.Series</td>
<td>0.09</td>
<td>-0.01</td>
<td>0.75</td>
</tr>
<tr>
<td>Pedigrees</td>
<td>0.06</td>
<td>0.01</td>
<td>0.28</td>
</tr>
<tr>
<td>Letter.Group</td>
<td>0.13</td>
<td>0.03</td>
<td>0.37</td>
</tr>
</tbody>
</table>

3.7 Mediation and Moderation analysis

Although multiple regression is a straightforward method for determining the effect of multiple predictors \((x_{1,2,...})\) on a criterion variable, \(y\), some prefer to think of the effect of one predictor, \(x\), as mediated by another variable, \(m\) (Preacher and Hayes, 2004). Thus, we may find the indirect path from \(x\) to \(m\), and then from \(m\) to \(y\) as well as the direct path from \(x\) to \(y\). Call these paths a, b, and c, respectively. Then the indirect effect of \(x\)
on y through m is just ab and the direct effect is c. Statistical tests of the ab effect are best done by bootstrapping.

Consider the example from Preacher and Hayes (2004) as analyzed using the `mediate` function and the subsequent graphic from `mediate.diagram`. The data are found in the example for `mediate`.

Call: `mediate(y = 1, x = 2, m = 3, data = sobel)`

The DV (Y) was SATIS. The IV (X) was THERAPY. The mediating variable(s) = ATTRIB.

Total Direct effect (c) of THERAPY on SATIS = 0.76  S.E. = 0.31  t direct = 2.5  with probability = 0.019
Direct effect (c') of THERAPY on SATIS removing ATTRIB = 0.43  S.E. = 0.32  t direct = 1.35  with probability =
Indirect effect (ab) of THERAPY on SATIS through ATTRIB = 0.33
Mean bootstrapped indirect effect = 0.33 with standard error = 0.17  Lower CI = 0.04  Upper CI = 0.71
R2 of model = 0.31
To see the longer output, specify `short = FALSE` in the print statement

Full output

<table>
<thead>
<tr>
<th>Total effect estimates (c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SATIS 0.76 0.31 2.5 0.0186</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Direct effect estimates (c')</th>
</tr>
</thead>
<tbody>
<tr>
<td>SATIS 0.43 0.32 1.35 0.190</td>
</tr>
<tr>
<td>THERAPY 0.40 0.18 2.23 0.034</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>'a' effect estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>THERAPY 0.82 0.3 2.74 0.0106</td>
</tr>
<tr>
<td>ATTRIB 0.4 0.18 2.23 0.034</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>'b' effect estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>SATIS boot 0.33 0.33 0.17 0.04 0.71</td>
</tr>
<tr>
<td>THERAPY 0.33 0.33 0.17 0.04 0.71</td>
</tr>
</tbody>
</table>

### 4 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.
Figure 15: A mediated model taken from Preacher and Hayes, 2004 and solved using the `mediate` function. The direct path from Therapy to Satisfaction has an effect of .76, while the indirect path through Attribution has an effect of .33. Compare this to the normal regression graphic created by `setCor.diagram`. 
Figure 16: The conventional regression model for the Preacher and Hayes, 2004 data set solved using the `sector` function. Compare this to the previous figure.
4.1 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 necessity. The goal for parsimony is seen in psychometrics as an attempt either to describe (components) 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 constructs. 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.

fa.poly is useful when finding the factor structure of categorical items. fa.poly first finds the tetrachoric or polychoric correlations between the categorical variables and then proceeds to do a normal factor analysis. By setting the n.iter option to be greater

---

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

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.
than 1, it will also find confidence intervals for the factor solution. Warning. Finding polychoric correlations is very slow, so think carefully before doing so.

**factor.minres (deprecated)** Minimum residual factor analysis is a least squares, iterative solution to the factor problem. minres attempts to minimize the residual (off-diagonal) correlation matrix. It produces solutions similar to maximum likelihood solutions, but will work even if the matrix is singular.

**factor.pa (deprecated)** Principal Axis factor analysis is a least squares, iterative solution to the factor problem. PA will work for cases where maximum likelihood techniques (factanal) will not work. The original communality estimates are either the squared multiple correlations (smc) for each item or 1.

**factor.wls (deprecated)** Weighted least squares factor analysis is a least squares, iterative solution to the factor problem. It minimizes the (weighted) squared residual matrix. The weights are based upon the independent contribution of each variable.

**principal** Principal Components Analysis reports the largest n eigen vectors rescaled by the square root of their eigen values. Note that PCA is not the same as factor analysis and the two should not be confused.

**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-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 $\chi^2$ estimate.

**nfactors** combines VSS, MAP, and a number of other fit statistics. The depressing reality is that frequently these conventional fit estimates of the number of factors do not agree.

**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.

\texttt{fa.diagram} replaces \texttt{fa.graph} and will draw a path diagram representing the factor structure. It does not require \texttt{Rgraphviz} and thus is probably preferred.

\texttt{fa.graph} requires \texttt{Rgraphviz} and will draw a graphic representation of the factor structure. If factors are correlated, this will be represented as well.

\texttt{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 \( \alpha \) Cronbach (1951) or \( \beta \) Revelle (1979) fail to increase.

\subsection*{4.1.1 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, \( \hat{\mathbf{R}}_n \), be approximated by the product of a factor matrix, \( \hat{\mathbf{F}}_k \), and its transpose plus a diagonal matrix of uniqueness.

\[ R = FF' + U^2 \]  \hfill (1)

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

\texttt{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 \texttt{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 \texttt{factor.pa} solutions. \texttt{min.res} is the default for the \texttt{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 oblivim transformation (Table 2). Alternative rotations and transformations include “none”, “varimax”, “quartimax”, “bentlerT”, “varimin” 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 oblivim transformation. 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).

Note that if extracting more than one factor, and doing any oblique rotation, it is necessary to have the GPArotation installed. This is checked for in the appropriate functions.

### 4.1.2 Principal Axis Factor Analysis

An alternative, least squares algorithm (included in fa with the fm=pa option or as a standalone function (factor.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 include varimax, quartimax, varimin, bifactor. Oblique transformations include oblimin, quartimin, biquartimin 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 3).

### 4.1.3 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 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 4).

The unweighted least squares solution may be shown graphically using the fa.plot function which is called by the generic plot function (Figure 17). Factors were transformed obliquely using an oblimin. These solutions may be shown as item by factor plots (Figure 17) or by a structure diagram (Figure 18).

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 2: 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.

```r
> if(!require('GPArotation')) {stop('GPArotation must be installed to do rotations')}
+ if3t <- 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

<table>
<thead>
<tr>
<th></th>
<th>MR1</th>
<th>MR2</th>
<th>MR3</th>
<th>h2</th>
<th>u2</th>
<th>com</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentences</td>
<td>0.91</td>
<td>-0.04</td>
<td>0.04</td>
<td>0.82</td>
<td>0.18</td>
<td>1.0</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.89</td>
<td>-0.06</td>
<td>-0.03</td>
<td>0.84</td>
<td>0.16</td>
<td>1.0</td>
</tr>
<tr>
<td>Sent.Completion</td>
<td>0.83</td>
<td>0.04</td>
<td>0.00</td>
<td>0.73</td>
<td>0.27</td>
<td>1.0</td>
</tr>
<tr>
<td>First.Letters</td>
<td>0.00</td>
<td>0.86</td>
<td>0.00</td>
<td>0.73</td>
<td>0.27</td>
<td>1.0</td>
</tr>
<tr>
<td>Four.Letter.Words</td>
<td>-0.01</td>
<td>0.74</td>
<td>0.10</td>
<td>0.63</td>
<td>0.37</td>
<td>1.0</td>
</tr>
<tr>
<td>Suffixes</td>
<td>0.18</td>
<td>0.63</td>
<td>-0.08</td>
<td>0.50</td>
<td>0.50</td>
<td>1.2</td>
</tr>
<tr>
<td>Letter.Series</td>
<td>0.03</td>
<td>-0.01</td>
<td>0.84</td>
<td>0.72</td>
<td>0.28</td>
<td>1.0</td>
</tr>
<tr>
<td>Pedigrees</td>
<td>0.37</td>
<td>-0.05</td>
<td>0.47</td>
<td>0.50</td>
<td>0.50</td>
<td>1.9</td>
</tr>
<tr>
<td>Letter.Group</td>
<td>-0.06</td>
<td>0.21</td>
<td>0.64</td>
<td>0.53</td>
<td>0.47</td>
<td>1.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>MR1</th>
<th>MR2</th>
<th>MR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS loadings</td>
<td>2.64</td>
<td>1.86</td>
<td>1.50</td>
</tr>
<tr>
<td>Proportion Var</td>
<td>0.29</td>
<td>0.21</td>
<td>0.17</td>
</tr>
<tr>
<td>Cumulative Var</td>
<td>0.29</td>
<td>0.50</td>
<td>0.67</td>
</tr>
<tr>
<td>Proportion Explained</td>
<td>0.44</td>
<td>0.31</td>
<td>0.25</td>
</tr>
<tr>
<td>Cumulative Proportion</td>
<td>0.44</td>
<td>0.75</td>
<td>1.00</td>
</tr>
</tbody>
</table>

With factor correlations of

<table>
<thead>
<tr>
<th></th>
<th>MR1</th>
<th>MR2</th>
<th>MR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR1</td>
<td>1.00</td>
<td>0.59</td>
<td>0.54</td>
</tr>
<tr>
<td>MR2</td>
<td>0.59</td>
<td>1.00</td>
<td>0.52</td>
</tr>
<tr>
<td>MR3</td>
<td>0.54</td>
<td>0.52</td>
<td>1.00</td>
</tr>
</tbody>
</table>

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 1081.97
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 Likelihood 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

<table>
<thead>
<tr>
<th></th>
<th>MR1</th>
<th>MR2</th>
<th>MR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation of scores with factors</td>
<td>0.96</td>
<td>0.92</td>
<td>0.90</td>
</tr>
<tr>
<td>Multiple R square of scores with factors</td>
<td>0.93</td>
<td>0.85</td>
<td>0.81</td>
</tr>
<tr>
<td>Minimum correlation of possible factor scores</td>
<td>0.86</td>
<td>0.71</td>
<td>0.63</td>
</tr>
</tbody>
</table>
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 extraction method was principal axis. The transformation was a targeted transformation to a simple cluster solution.

```r
> if(!require('GPArotation')) {stop('GPArotation must be installed to do rotations')} else {
+  f3 <- fa(Thurstone,3,n.obs = 213,fm="pa")
+  f3o <- target.rot(f3)
+}
Call: NULL

Standardized loadings (pattern matrix) based upon correlation matrix

<table>
<thead>
<tr>
<th></th>
<th>PA1</th>
<th>PA2</th>
<th>PA3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentences</td>
<td>0.89</td>
<td>-0.03</td>
<td>0.07</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.89</td>
<td>0.07</td>
<td>0.00</td>
</tr>
<tr>
<td>Sent.Completion</td>
<td>0.83</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>First.Letters</td>
<td>-0.02</td>
<td>0.85</td>
<td>-0.01</td>
</tr>
<tr>
<td>Four.Letter.Words</td>
<td>-0.05</td>
<td>0.74</td>
<td>0.09</td>
</tr>
<tr>
<td>Suffixes</td>
<td>0.17</td>
<td>0.63</td>
<td>-0.09</td>
</tr>
<tr>
<td>Letter.Series</td>
<td>-0.06</td>
<td>-0.08</td>
<td>0.84</td>
</tr>
<tr>
<td>Pedigrees</td>
<td>0.33</td>
<td>-0.09</td>
<td>0.48</td>
</tr>
<tr>
<td>Letter.Group</td>
<td>-0.14</td>
<td>0.16</td>
<td>0.64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>PA1</th>
<th>PA2</th>
<th>PA3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS loadings</td>
<td>2.45</td>
<td>1.72</td>
<td>1.37</td>
</tr>
<tr>
<td>Proportion Var</td>
<td>0.27</td>
<td>0.19</td>
<td>0.15</td>
</tr>
<tr>
<td>Cumulative Var</td>
<td>0.27</td>
<td>0.46</td>
<td>0.62</td>
</tr>
<tr>
<td>Proportion Explained</td>
<td>0.44</td>
<td>0.31</td>
<td>0.25</td>
</tr>
<tr>
<td>Cumulative Proportion</td>
<td>0.44</td>
<td>0.75</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>PA1</th>
<th>PA2</th>
<th>PA3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA1</td>
<td>1.00</td>
<td>0.02</td>
<td>0.08</td>
</tr>
<tr>
<td>PA2</td>
<td>0.02</td>
<td>1.00</td>
<td>0.09</td>
</tr>
<tr>
<td>PA3</td>
<td>0.08</td>
<td>0.09</td>
<td>1.00</td>
</tr>
</tbody>
</table>

44
Table 4: 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.

```r
> 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

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

<table>
<thead>
<tr>
<th></th>
<th>WLS1</th>
<th>WLS2</th>
<th>WLS3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS loadings</td>
<td>2.647</td>
<td>1.864</td>
<td>1.488</td>
</tr>
<tr>
<td>Proportion Var</td>
<td>0.294</td>
<td>0.207</td>
<td>0.165</td>
</tr>
<tr>
<td>Cumulative Var</td>
<td>0.294</td>
<td>0.501</td>
<td>0.667</td>
</tr>
<tr>
<td>Proportion Explained</td>
<td>0.441</td>
<td>0.311</td>
<td>0.248</td>
</tr>
<tr>
<td>Cumulative Proportion</td>
<td>0.441</td>
<td>0.752</td>
<td>1.000</td>
</tr>
</tbody>
</table>

With factor correlations of

<table>
<thead>
<tr>
<th></th>
<th>WLS1</th>
<th>WLS2</th>
<th>WLS3</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLS1</td>
<td>1.000</td>
<td>0.591</td>
<td>0.535</td>
</tr>
<tr>
<td>WLS2</td>
<td>0.591</td>
<td>1.000</td>
<td>0.516</td>
</tr>
<tr>
<td>WLS3</td>
<td>0.535</td>
<td>0.516</td>
<td>1.000</td>
</tr>
</tbody>
</table>

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 1081.968
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 Likelihood 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

<table>
<thead>
<tr>
<th></th>
<th>WLS1</th>
<th>WLS2</th>
<th>WLS3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation of scores with factors</td>
<td>0.964</td>
<td>0.923</td>
<td>0.902</td>
</tr>
<tr>
<td>Multiple R square of scores with factors</td>
<td>0.929</td>
<td>0.853</td>
<td>0.814</td>
</tr>
<tr>
<td>Minimum correlation of possible factor scores</td>
<td>0.858</td>
<td>0.706</td>
<td>0.627</td>
</tr>
</tbody>
</table>
Figure 17: 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.
Factor Analysis

Figure 18: A graphic representation of the 3 oblique factors from the Thurstone data using \texttt{fa.diagram}. Factors were transformed to an oblique solution using \texttt{oblimin}.
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 \texttt{fa}.

4.1.4 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. Some psychologists use PCA in a manner similar to factor analysis and thus the \texttt{principal} function produces output that is perhaps more understandable than that produced by \texttt{princomp} in the \texttt{stats} package. Table 5 shows a PCA of the Thurstone 9 variable problem rotated using the \texttt{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. Also note how the goodness of fit statistics, based upon the residual off diagonal elements, is much worse than the \texttt{fa} solution.

4.1.5 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 \texttt{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 18 can themselves be factored. This results in a higher order factor model (Figure 19). 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 \texttt{Schmid-Leiman} procedure. (Figure 20). Yet another solution is to use structural equation modeling to directly solve for the general and group factors.
Table 5: The Thurstone problem can also be analyzed using Principal Components Analysis. Compare this to Table 3. 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")

Principal Components Analysis

Call: principal(r = Thurstone, nfactors = 3, rotate = "Promax", n.obs = 213)

Standardized loadings (pattern matrix) based upon correlation matrix

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

<table>
<thead>
<tr>
<th></th>
<th>RC1</th>
<th>RC2</th>
<th>RC3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS loadings</td>
<td>2.83</td>
<td>2.19</td>
<td>1.96</td>
</tr>
<tr>
<td>Proportion Var</td>
<td>0.31</td>
<td>0.24</td>
<td>0.22</td>
</tr>
<tr>
<td>Cumulative Var</td>
<td>0.31</td>
<td>0.56</td>
<td>0.78</td>
</tr>
<tr>
<td>Proportion Explained</td>
<td>0.41</td>
<td>0.31</td>
<td>0.28</td>
</tr>
<tr>
<td>Cumulative Proportion</td>
<td>0.41</td>
<td>0.72</td>
<td>1.00</td>
</tr>
</tbody>
</table>

With component correlations of

<table>
<thead>
<tr>
<th>RC1</th>
<th>RC2</th>
<th>RC3</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC1</td>
<td>1.00</td>
<td>0.51</td>
</tr>
<tr>
<td>RC2</td>
<td>0.51</td>
<td>1.00</td>
</tr>
<tr>
<td>RC3</td>
<td>0.53</td>
<td>0.44</td>
</tr>
</tbody>
</table>

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

Omega

Figure 19: A higher order factor solution to the Thurstone 9 variable problem
> om <- omega(Thurstone, n.obs=213)

Omega

Figure 20: A bifactor factor solution to the Thurstone 9 variable problem
Yet another approach to the bifactor structure is due use the **bifactor** rotation function in either *psych* or in *GPArotation*. This does the rotation discussed in Jennrich and Bentler (2011).

### 4.1.6 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 taxonmetric 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 (Martinent and Ferrand, 2007; Mun et al., 2008). Interestingly enough it has 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 $\alpha$ or $\beta$).
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, $\alpha$ (Cronbach, 1951), as well as the worst split half reliability, $\beta$ (Revelle, 1979), which is an estimate of the general factor saturation of the resulting scale (Figure 21). Cluster loadings (corresponding to the structure matrix of factor analysis) are reported when printing (Table 8). The pattern matrix is available as an object in the results.

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

![ICLUST](image)

Figure 21: 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.
Table 6: 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
```
The previous analysis (Figure 21) was done using the Pearson correlation. A somewhat cleaner structure is obtained when using the polychoric function to find polychoric correlations (Figure 22). 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 progressBar.

Table 7: 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.

```r
> 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.

### 4.2 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 $\chi^2$ 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 9). 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.

Bootstrapped confidence intervals can also be found for the loadings of a factoring of a polychoric matrix. fa.poly will find the polychoric correlation matrix and if the n.iter option is greater than 1, will then randomly resample the data (case wise) to give bootstrapped samples. This will take a long time for large number of items or iterations.

### 4.3 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 Burt’s congruence coefficient (also known as Tucker’s
ICLUST using polychoric correlations

Figure 22: ICLUST of the BFI data set using polychoric correlations. Compare this solution to the previous one (Figure 21) which was done using Pearson correlations.
Figure 23: ICLUST of the BFI data set using polychoric correlations with the solution set to 5 clusters. Compare this solution to the previous one (Figure 22) which was done without specifying the number of clusters and to the next one (Figure 24) 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 24: ICLUST of the BFI data set using polychoric correlations with the beta criterion set to 3. Compare this solution to the previous three (Figure 21, 22, 23).
Table 8: 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=0.5

```r
> 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:

<table>
<thead>
<tr>
<th></th>
<th>O</th>
<th>P</th>
<th>C20</th>
<th>C16</th>
<th>C15</th>
<th>C21</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td></td>
<td></td>
<td>C20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td></td>
<td></td>
<td>C20</td>
<td>0.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td></td>
<td></td>
<td>C20</td>
<td>0.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td></td>
<td></td>
<td>C20</td>
<td>0.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td></td>
<td></td>
<td>C20</td>
<td>0.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td></td>
<td></td>
<td>C15</td>
<td>0.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td></td>
<td></td>
<td>C15</td>
<td>0.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td></td>
<td>C15</td>
<td>C15</td>
<td>0.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td></td>
<td>C15</td>
<td>C15</td>
<td>0.31</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td></td>
<td>C15</td>
<td>C15</td>
<td>0.30</td>
<td>0.36</td>
<td>0.59</td>
</tr>
<tr>
<td>E1</td>
<td></td>
<td></td>
<td>C20</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2</td>
<td></td>
<td></td>
<td>C20</td>
<td>0.61</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>E3</td>
<td></td>
<td></td>
<td>C20</td>
<td>0.59</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>E4</td>
<td></td>
<td></td>
<td>C20</td>
<td>0.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E5</td>
<td></td>
<td></td>
<td>C20</td>
<td>0.50</td>
<td>0.40</td>
<td>0.32</td>
</tr>
<tr>
<td>N1</td>
<td></td>
<td>C16</td>
<td>C16</td>
<td>0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2</td>
<td></td>
<td>C16</td>
<td>C16</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N3</td>
<td></td>
<td>C16</td>
<td>C16</td>
<td>0.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N4</td>
<td></td>
<td>C16</td>
<td>C16</td>
<td>0.34</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>N5</td>
<td></td>
<td>C16</td>
<td>C16</td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O1</td>
<td></td>
<td>C20</td>
<td>C21</td>
<td>0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2</td>
<td></td>
<td>C21</td>
<td>C21</td>
<td>0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O3</td>
<td></td>
<td>C20</td>
<td>C21</td>
<td>0.39</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>O4</td>
<td></td>
<td>C21</td>
<td>C21</td>
<td>0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O5</td>
<td></td>
<td>C21</td>
<td>C21</td>
<td>0.53</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th></th>
<th>C20</th>
<th>C16</th>
<th>C15</th>
<th>C21</th>
</tr>
</thead>
<tbody>
<tr>
<td>C20</td>
<td>0.80</td>
<td>0.29</td>
<td>0.40</td>
<td>0.33</td>
</tr>
<tr>
<td>C16</td>
<td>-0.24</td>
<td>0.81</td>
<td>0.29</td>
<td>0.11</td>
</tr>
<tr>
<td>C15</td>
<td>0.30</td>
<td>-0.22</td>
<td>0.73</td>
<td>0.30</td>
</tr>
<tr>
<td>C21</td>
<td>0.23</td>
<td>0.07</td>
<td>-0.20</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Cluster fit = 0.68 Pattern fit = 0.96 RMSR = 0.05
NULL
Table 9: 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.

```r
> 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

<table>
<thead>
<tr>
<th></th>
<th>MR2</th>
<th>MR1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>0.07</td>
<td>-0.40</td>
</tr>
<tr>
<td>A2</td>
<td>0.02</td>
<td>0.65</td>
</tr>
<tr>
<td>A3</td>
<td>-0.03</td>
<td>0.77</td>
</tr>
<tr>
<td>A4</td>
<td>0.15</td>
<td>0.44</td>
</tr>
<tr>
<td>A5</td>
<td>0.02</td>
<td>0.62</td>
</tr>
<tr>
<td>C1</td>
<td>0.57</td>
<td>-0.05</td>
</tr>
<tr>
<td>C2</td>
<td>0.62</td>
<td>-0.01</td>
</tr>
<tr>
<td>C3</td>
<td>0.54</td>
<td>0.03</td>
</tr>
<tr>
<td>C4</td>
<td>-0.66</td>
<td>0.01</td>
</tr>
<tr>
<td>C5</td>
<td>-0.57</td>
<td>-0.05</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>MR2</th>
<th>MR1</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR2</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>MR1</td>
<td>0.52</td>
<td></td>
</tr>
</tbody>
</table>

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 Likelihood Chi Square = 464.04 with prob < 9.2e-82

Tucker Lewis Index of factoring reliability = 0.865
RMSEA index = 0.006 and the 90% confidence intervals are 0.006 0.084
BIC = 257.67
Fit based upon off diagonal values = 0.98

Measures of factor score adequacy

<table>
<thead>
<tr>
<th></th>
<th>MR2</th>
<th>MR1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation of scores with factors</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>Multiple R square of scores with factors</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Minimum correlation of possible factor scores</td>
<td>0.49</td>
<td></td>
</tr>
</tbody>
</table>

Coefficients and bootstrapped confidence intervals

<table>
<thead>
<tr>
<th></th>
<th>low</th>
<th>MR2 upper</th>
<th>low</th>
<th>MR1 upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>0.03</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>A2</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>A3</td>
<td>-0.07</td>
<td>-0.03</td>
<td>0.00</td>
<td>0.03</td>
</tr>
<tr>
<td>A4</td>
<td>0.10</td>
<td>0.15</td>
<td>0.20</td>
<td>0.40</td>
</tr>
<tr>
<td>A5</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.06</td>
<td>0.57</td>
</tr>
<tr>
<td>C1</td>
<td>0.54</td>
<td>0.57</td>
<td>0.60</td>
<td>0.69</td>
</tr>
<tr>
<td>C2</td>
<td>0.58</td>
<td>0.62</td>
<td>0.67</td>
<td>0.76</td>
</tr>
<tr>
<td>C3</td>
<td>0.48</td>
<td>0.54</td>
<td>0.58</td>
<td>0.66</td>
</tr>
<tr>
<td>C4</td>
<td>-0.69</td>
<td>-0.66</td>
<td>-0.60</td>
<td>-0.63</td>
</tr>
<tr>
<td>C5</td>
<td>-0.62</td>
<td>-0.57</td>
<td>-0.52</td>
<td>-0.48</td>
</tr>
</tbody>
</table>

Interfactor correlations and bootstrapped confidence intervals

<table>
<thead>
<tr>
<th></th>
<th>lower estimate</th>
<th>upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR2-MR1</td>
<td>0.27</td>
<td>0.32</td>
</tr>
</tbody>
</table>

>

60
coefficient) which is just the cosine of the angle between the dimensions

\[ c_{ij} = \frac{\sum_{k=1}^{K} f_{ik} f_{jk}}{\sqrt{\sum f_{ik}^2 \sum f_{jk}^2}}. \]

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

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

<table>
<thead>
<tr>
<th></th>
<th>C20</th>
<th>C16</th>
<th>C15</th>
<th>C21</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA1</td>
<td>0.92</td>
<td>-0.32</td>
<td>0.44</td>
<td>-0.40</td>
</tr>
<tr>
<td>PA2</td>
<td>-0.26</td>
<td>0.95</td>
<td>-0.33</td>
<td>0.12</td>
</tr>
<tr>
<td>PA3</td>
<td>0.35</td>
<td>-0.24</td>
<td>0.88</td>
<td>-0.37</td>
</tr>
<tr>
<td>PA4</td>
<td>0.29</td>
<td>-0.12</td>
<td>0.27</td>
<td>-0.90</td>
</tr>
</tbody>
</table>

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 10).

Table 10: Congruence coefficients for oblique factor, bifactor and component solutions for the Thurstone problem.

```r
> factor.congruence(list(f3t,f3o,om,p3p))
```

<table>
<thead>
<tr>
<th></th>
<th>MR1</th>
<th>MR2</th>
<th>MR3</th>
<th>PA1</th>
<th>PA2</th>
<th>PA3</th>
<th>g</th>
<th>F1*</th>
<th>F2*</th>
<th>F3*</th>
<th>h2</th>
<th>RC1</th>
<th>RC2</th>
<th>RC3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR1</td>
<td>1.00</td>
<td>0.06</td>
<td>0.09</td>
<td>1.00</td>
<td>0.06</td>
<td>0.13</td>
<td>0.72</td>
<td>1.00</td>
<td>0.06</td>
<td>0.09</td>
<td>0.74</td>
<td>1.00</td>
<td>0.08</td>
<td>0.04</td>
</tr>
<tr>
<td>MR2</td>
<td>0.06</td>
<td>1.00</td>
<td>0.08</td>
<td>0.03</td>
<td>1.00</td>
<td>0.06</td>
<td>0.60</td>
<td>1.00</td>
<td>0.08</td>
<td>0.57</td>
<td>0.04</td>
<td>0.99</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>MR3</td>
<td>0.09</td>
<td>0.08</td>
<td>1.00</td>
<td>0.01</td>
<td>0.01</td>
<td>1.00</td>
<td>0.52</td>
<td>0.09</td>
<td>0.08</td>
<td>1.00</td>
<td>0.51</td>
<td>0.06</td>
<td>0.02</td>
<td>0.99</td>
</tr>
<tr>
<td>PA1</td>
<td>1.00</td>
<td>0.03</td>
<td>0.01</td>
<td>1.00</td>
<td>0.04</td>
<td>0.05</td>
<td>0.67</td>
<td>1.00</td>
<td>0.03</td>
<td>0.01</td>
<td>0.69</td>
<td>1.00</td>
<td>0.06</td>
<td>-0.04</td>
</tr>
<tr>
<td>PA2</td>
<td>0.06</td>
<td>1.00</td>
<td>0.01</td>
<td>0.04</td>
<td>1.00</td>
<td>0.00</td>
<td>0.57</td>
<td>0.06</td>
<td>1.00</td>
<td>0.01</td>
<td>0.54</td>
<td>0.04</td>
<td>0.99</td>
<td>0.05</td>
</tr>
<tr>
<td>PA3</td>
<td>0.13</td>
<td>0.06</td>
<td>1.00</td>
<td>0.05</td>
<td>0.00</td>
<td>1.00</td>
<td>0.54</td>
<td>0.13</td>
<td>0.06</td>
<td>1.00</td>
<td>0.53</td>
<td>0.10</td>
<td>0.01</td>
<td>0.99</td>
</tr>
<tr>
<td>g</td>
<td>0.72</td>
<td>0.60</td>
<td>0.52</td>
<td>0.67</td>
<td>0.57</td>
<td>0.54</td>
<td>1.00</td>
<td>0.72</td>
<td>0.60</td>
<td>0.52</td>
<td>0.99</td>
<td>0.69</td>
<td>0.58</td>
<td>0.50</td>
</tr>
<tr>
<td>F1*</td>
<td>1.00</td>
<td>0.06</td>
<td>0.09</td>
<td>1.00</td>
<td>0.06</td>
<td>0.13</td>
<td>0.72</td>
<td>1.00</td>
<td>0.06</td>
<td>0.09</td>
<td>0.74</td>
<td>1.00</td>
<td>0.08</td>
<td>0.04</td>
</tr>
<tr>
<td>F2*</td>
<td>0.06</td>
<td>1.00</td>
<td>0.08</td>
<td>0.03</td>
<td>1.00</td>
<td>0.06</td>
<td>0.60</td>
<td>0.06</td>
<td>1.00</td>
<td>0.08</td>
<td>0.57</td>
<td>0.04</td>
<td>0.99</td>
<td>0.12</td>
</tr>
<tr>
<td>F3*</td>
<td>0.09</td>
<td>0.08</td>
<td>1.00</td>
<td>0.01</td>
<td>0.01</td>
<td>1.00</td>
<td>0.52</td>
<td>0.09</td>
<td>0.08</td>
<td>1.00</td>
<td>0.51</td>
<td>0.06</td>
<td>0.02</td>
<td>0.99</td>
</tr>
<tr>
<td>h2</td>
<td>0.74</td>
<td>0.57</td>
<td>0.51</td>
<td>0.69</td>
<td>0.54</td>
<td>0.53</td>
<td>0.99</td>
<td>0.74</td>
<td>0.57</td>
<td>0.51</td>
<td>1.00</td>
<td>0.71</td>
<td>0.56</td>
<td>0.49</td>
</tr>
<tr>
<td>RC1</td>
<td>1.00</td>
<td>0.04</td>
<td>0.06</td>
<td>1.00</td>
<td>0.04</td>
<td>0.10</td>
<td>0.69</td>
<td>1.00</td>
<td>0.04</td>
<td>0.06</td>
<td>0.71</td>
<td>1.00</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>RC2</td>
<td>0.08</td>
<td>0.99</td>
<td>0.02</td>
<td>0.06</td>
<td>0.99</td>
<td>0.01</td>
<td>0.58</td>
<td>0.08</td>
<td>0.99</td>
<td>0.02</td>
<td>0.56</td>
<td>0.06</td>
<td>1.00</td>
<td>0.05</td>
</tr>
<tr>
<td>RC3</td>
<td>0.04</td>
<td>0.12</td>
<td>0.99</td>
<td>-0.04</td>
<td>0.05</td>
<td>0.99</td>
<td>0.50</td>
<td>0.04</td>
<td>0.12</td>
<td>0.99</td>
<td>0.49</td>
<td>0.00</td>
<td>0.05</td>
<td>1.00</td>
</tr>
</tbody>
</table>

4.4 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) \texttt{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.


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 the eigen values of random factors will all tend towards 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. \texttt{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 \texttt{sim.structure} and of small extraneous factors using the \texttt{sim.minor} function.
4.4.1 Very Simple Structure

The \texttt{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 \texttt{vss} is the MAP criterion (Minimum Absolute Partial correlation) of Velicer (1976).

Using the Very Simple Structure criterion for the \texttt{bfi} data suggests that 4 factors are optimal (Figure 25). However, the MAP criterion suggests that 5 is optimal.

\begin{verbatim}
> vss <- vss(bfi[1:25],title="Very Simple Structure of a Big 5 inventory")
\end{verbatim}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figures/very_simple_structure.png}
\caption{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.}
\end{figure}
> vss

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

<table>
<thead>
<tr>
<th></th>
<th>vss1</th>
<th>vss2</th>
<th>map</th>
<th>dof</th>
<th>chisq</th>
<th>prob</th>
<th>sqresid</th>
<th>fit</th>
<th>RMSEA</th>
<th>BIC</th>
<th>SABIC</th>
<th>complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.49</td>
<td>0.00</td>
<td>0.024</td>
<td>275</td>
<td>11831</td>
<td>0.0e+00</td>
<td>26.0</td>
<td>0.49</td>
<td>0.0150</td>
<td>9648</td>
<td>10522.1</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>0.54</td>
<td>0.63</td>
<td>0.018</td>
<td>251</td>
<td>7279</td>
<td>0.0e+00</td>
<td>18.9</td>
<td>0.63</td>
<td>0.0100</td>
<td>5287</td>
<td>6084.5</td>
<td>1.2</td>
</tr>
<tr>
<td>3</td>
<td>0.57</td>
<td>0.69</td>
<td>0.017</td>
<td>228</td>
<td>5010</td>
<td>0.0e+00</td>
<td>14.8</td>
<td>0.71</td>
<td>0.0075</td>
<td>3200</td>
<td>3924.3</td>
<td>1.3</td>
</tr>
<tr>
<td>4</td>
<td>0.58</td>
<td>0.74</td>
<td>0.015</td>
<td>206</td>
<td>3366</td>
<td>0.0e+00</td>
<td>11.7</td>
<td>0.77</td>
<td>0.0055</td>
<td>1731</td>
<td>2385.1</td>
<td>1.4</td>
</tr>
<tr>
<td>5</td>
<td>0.53</td>
<td>0.73</td>
<td>0.015</td>
<td>185</td>
<td>1750</td>
<td>1.4e-252</td>
<td>9.5</td>
<td>0.81</td>
<td>0.0030</td>
<td>281</td>
<td>869.3</td>
<td>1.6</td>
</tr>
<tr>
<td>6</td>
<td>0.54</td>
<td>0.72</td>
<td>0.016</td>
<td>165</td>
<td>1014</td>
<td>4.4e-122</td>
<td>8.4</td>
<td>0.84</td>
<td>0.0018</td>
<td>-296</td>
<td>228.5</td>
<td>1.7</td>
</tr>
<tr>
<td>7</td>
<td>0.52</td>
<td>0.70</td>
<td>0.019</td>
<td>146</td>
<td>696</td>
<td>1.4e-72</td>
<td>7.9</td>
<td>0.84</td>
<td>0.0013</td>
<td>-463</td>
<td>1.2</td>
<td>1.9</td>
</tr>
<tr>
<td>8</td>
<td>0.52</td>
<td>0.69</td>
<td>0.022</td>
<td>128</td>
<td>492</td>
<td>4.7e-44</td>
<td>7.4</td>
<td>0.85</td>
<td>0.0010</td>
<td>-524</td>
<td>-117.6</td>
<td>1.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>eChisq</th>
<th>SRMR</th>
<th>eCRMS</th>
<th>eBIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23881</td>
<td>0.119</td>
<td>0.125</td>
<td>21698</td>
</tr>
<tr>
<td>2</td>
<td>12432</td>
<td>0.086</td>
<td>0.094</td>
<td>10440</td>
</tr>
<tr>
<td>3</td>
<td>7232</td>
<td>0.066</td>
<td>0.075</td>
<td>5422</td>
</tr>
<tr>
<td>4</td>
<td>3750</td>
<td>0.047</td>
<td>0.057</td>
<td>2115</td>
</tr>
<tr>
<td>5</td>
<td>1495</td>
<td>0.030</td>
<td>0.038</td>
<td>27</td>
</tr>
<tr>
<td>6</td>
<td>670</td>
<td>0.020</td>
<td>0.027</td>
<td>-639</td>
</tr>
<tr>
<td>7</td>
<td>448</td>
<td>0.016</td>
<td>0.023</td>
<td>-711</td>
</tr>
<tr>
<td>8</td>
<td>289</td>
<td>0.013</td>
<td>0.020</td>
<td>-727</td>
</tr>
</tbody>
</table>

### 4.4.2 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 26). 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`. By default the number of replications is 20. This is appropriate when choosing the number of factors from dichotomous or polytomous data matrices.
Parallel analysis suggests that the number of factors = 6 and the number of components = 6

Figure 26: 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.
4.5 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 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 27).

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

4.6 Exploratory Structural Equation Modeling (ESEM)

Generalizing the procedures of factor extension, we can do Exploratory Structural Equation Modeling (ESEM). Traditional Exploratory Factor Analysis (EFA) examines how latent variables can account for the correlations within a data set. All loadings and cross loadings are found and rotation is done to some approximation of simple structure. Traditional Confirmatory Factor Analysis (CFA) tests such models by fitting just a limited number of loadings and typically does not allow any (or many) cross loadings. Structural Equation Modeling then applies two such measurement models, one to a set of X variables, another to a set of Y variables, and then tries to estimate the correlation between these two sets of latent variables. (Some SEM procedures estimate all the parameters from the same model, thus making the loadings in set Y affect those in set X.) It is possible to do a similar, exploratory modeling (ESEM) by conducting two Exploratory Factor Analyses, one in set X, one in set Y, and then finding the correlations of the X factors with the Y factors, as well as the correlations of the Y variables with the X factors and the X variables with the Y factors.

Consider the simulated data set of three ability variables, two motivational variables, and three outcome variables:

```
Call: sim.structural(fx = fx, Phi = Phi, fy = fy)

$model (Population correlation matrix)

       V   Q   A nach Anx  gpa  Pre  MA
V 1.00 0.72 0.54 0.00 0.00 0.38 0.32 0.25
Q 0.72 1.00 0.48 0.00 0.00 0.34 0.28 0.22
A 0.54 0.48 1.00 0.48 -0.42 0.50 0.42 0.34
nach 0.00 0.00 0.48 1.00 -0.56 0.34 0.28 0.22
```

66
> 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 27: 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.
Anx  0.00  0.00  -0.42  -0.56  1.00  -0.29  -0.24  -0.20  
gpa  0.38  0.34  0.50  0.34  -0.29  1.00  0.30  0.24  
Pre  0.32  0.28  0.42  0.28  -0.24  0.30  1.00  0.20  
MA   0.25  0.22  0.34  0.22  -0.20  0.24  0.20  1.00  

$reliability (population reliability)  
  V   Q  A nach Anx gpa Pre MA  
0.81  0.64  0.72  0.64  0.49  0.36  0.25  0.16  

We can fit this by using the esem function and then draw the solution (see Figure 28) using the esem.diagram function (which is normally called automatically by esem).

Exploratory Structural Equation Modeling Analysis using method = minres  
Call: esem(r = gre.gpa$gpa$model, varsX = 1:5, varsY = 6:8, nfX = 2, nfY = 1,  
n.obs = 1000, plot = FALSE)  

For the 'X' set:  
  MR1  MR2  
V   0.91  -0.06  
Q   0.81  -0.05  
A   0.53   0.57  
nach -0.10   0.81  
Anx  0.08  -0.71  

For the 'Y' set:  
  MR1  
gpa 0.6  
Pre 0.5  
MA 0.4  

Correlations between the X and Y sets.  
  X1  X2  Y1  
X1 1.00 0.19 0.68  
X2 0.19 1.00 0.67  
Y1 0.68 0.67 1.00  

The degrees of freedom for the null model are 56 and the empirical chi square function was 6930.29  
The degrees of freedom for the model are 7 and the empirical chi square function was 21.83  
  with prob < 0.0027  

The root mean square of the residuals (RMSR) is 0.02  
The df corrected root mean square of the residuals is 0.04
with the empirical chi square  21.83  with prob <  0.0027  
The total number of observations was  1000  with fitted Chi Square =  2175.06  with prob <

Empirical BIC = -26.53  
ESABIC = -4.29  
Fit based upon off diagonal values = 1  
To see the item loadings for the X and Y sets combined, and the associated fa output, print

**Exploratory Structural Model**

Figure 28: An example of a Exploratory Structure Equation Model.
5 Classical Test Theory and Reliability

Surprisingly, 110 years after Spearman (1904) introduced the concept of reliability to psychologists, there are still multiple approaches for measuring it. Although very popular, Cronbach’s $\alpha$ (Cronbach, 1951) underestimates the reliability of a test and overestimates the first factor saturation (Revelle and Zinbarg, 2009).

$\alpha$ (Cronbach, 1951) is the same as Guttman’s $\lambda_3$ (Guttman, 1945) and may be found by

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

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. 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 $\beta$ (Revelle, 1979) (see iclust) and $\omega_h$ (see omega) are more appropriate estimates of the general factor saturation. $\omega$ is a better estimate of the reliability of the total test.

Guttman’s $\lambda_6$ (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, $e^2_j$, and is

$$\lambda_6 = 1 - \frac{\sum e^2_j}{V_x} = 1 - \frac{\sum (1 - r^2_{smc})}{V_x}.$$  

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 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 $\omega_h$ and $\omega_t$ 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 $\omega_h$ to $\alpha$ to $\omega_t$. The `alpha` function reports two estimates: Cronbach’s coefficient $\alpha$ and Guttman’s $\lambda_6$. Also reported are item - whole correlations, $\alpha$ 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) ($\mu_0$...$\mu_3$) as well as $\beta$ (the worst split half, Revelle, 1979), the glb (greatest lowest bound) discussed by Bentler and Woodward (1980), and $\omega_h$ and $\omega_t$ ((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 x c cluster definition matrix of -1s, 0s, and 1s (the keys) , and a n x 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.

### 5.1 Reliability of a single scale

A conventional (but non-optimal) estimate of the internal consistency reliability of a test is coefficient $\alpha$ (Cronbach, 1951). Alternative estimates are Guttman’s $\lambda_6$, Revelle’s $\beta$, 

71
McDonald's $\omega_h$ and $\omega_t$. Consider a simulated data set, representing 9 items with a hierarchical structure and the following correlation matrix. Then using the \texttt{alpha} function, the $\alpha$ and $\lambda_6$ estimates of reliability may be found for all 9 items, as well as the if one item is dropped at a time.

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

<table>
<thead>
<tr>
<th></th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
<th>V5</th>
<th>V6</th>
<th>V7</th>
<th>V8</th>
<th>V9</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>1.00</td>
<td>0.58</td>
<td>0.59</td>
<td>0.41</td>
<td>0.44</td>
<td>0.30</td>
<td>0.40</td>
<td>0.31</td>
<td>0.25</td>
</tr>
<tr>
<td>V2</td>
<td>0.58</td>
<td>1.00</td>
<td>0.48</td>
<td>0.35</td>
<td>0.36</td>
<td>0.22</td>
<td>0.24</td>
<td>0.29</td>
<td>0.19</td>
</tr>
<tr>
<td>V3</td>
<td>0.59</td>
<td>0.48</td>
<td>1.00</td>
<td>0.32</td>
<td>0.35</td>
<td>0.23</td>
<td>0.29</td>
<td>0.20</td>
<td>0.17</td>
</tr>
<tr>
<td>V4</td>
<td>0.41</td>
<td>0.35</td>
<td>0.32</td>
<td>1.00</td>
<td>0.44</td>
<td>0.36</td>
<td>0.26</td>
<td>0.27</td>
<td>0.22</td>
</tr>
<tr>
<td>V5</td>
<td>0.44</td>
<td>0.36</td>
<td>0.35</td>
<td>0.44</td>
<td>1.00</td>
<td>0.32</td>
<td>0.24</td>
<td>0.23</td>
<td>0.20</td>
</tr>
<tr>
<td>V6</td>
<td>0.30</td>
<td>0.22</td>
<td>0.36</td>
<td>0.32</td>
<td>1.00</td>
<td>0.26</td>
<td>0.26</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>V7</td>
<td>0.40</td>
<td>0.24</td>
<td>0.29</td>
<td>0.26</td>
<td>0.24</td>
<td>1.00</td>
<td>0.38</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>V8</td>
<td>0.31</td>
<td>0.29</td>
<td>0.20</td>
<td>0.27</td>
<td>0.23</td>
<td>0.26</td>
<td>1.00</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>V9</td>
<td>0.25</td>
<td>0.19</td>
<td>0.17</td>
<td>0.22</td>
<td>0.20</td>
<td>0.12</td>
<td>0.25</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

> alpha(r9)

Reliability analysis
Call: alpha(x = r9)

<table>
<thead>
<tr>
<th>raw_alpha</th>
<th>std.alpha</th>
<th>G6(smc)</th>
<th>average_r</th>
<th>S/N</th>
<th>ase</th>
<th>mean</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.31</td>
<td>4.0</td>
<td>0.013</td>
<td>0.017</td>
<td>0.63</td>
</tr>
</tbody>
</table>

lower alpha upper 95% confidence boundaries
0.77 0.8 0.83

Reliability if an item is dropped:

<table>
<thead>
<tr>
<th>raw_alpha</th>
<th>std.alpha</th>
<th>G6(smc)</th>
<th>average_r</th>
<th>S/N</th>
<th>ase</th>
<th>mean</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>0.75</td>
<td>0.75</td>
<td>0.75</td>
<td>0.28</td>
<td>3.1</td>
<td>0.017</td>
<td></td>
</tr>
<tr>
<td>V2</td>
<td>0.77</td>
<td>0.77</td>
<td>0.77</td>
<td>0.30</td>
<td>3.4</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>V3</td>
<td>0.77</td>
<td>0.77</td>
<td>0.77</td>
<td>0.30</td>
<td>3.4</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>V4</td>
<td>0.77</td>
<td>0.77</td>
<td>0.77</td>
<td>0.30</td>
<td>3.4</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>V5</td>
<td>0.78</td>
<td>0.78</td>
<td>0.77</td>
<td>0.30</td>
<td>3.5</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>V6</td>
<td>0.79</td>
<td>0.79</td>
<td>0.79</td>
<td>0.32</td>
<td>3.8</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>V7</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
<td>0.31</td>
<td>3.6</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>V8</td>
<td>0.79</td>
<td>0.79</td>
<td>0.78</td>
<td>0.32</td>
<td>3.7</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>V9</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
<td>0.34</td>
<td>4.0</td>
<td>0.013</td>
<td></td>
</tr>
</tbody>
</table>

Item statistics

<table>
<thead>
<tr>
<th>n</th>
<th>raw.r</th>
<th>r.cor</th>
<th>r.drop</th>
<th>mean</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>500</td>
<td>0.77</td>
<td>0.76</td>
<td>0.67</td>
<td>0.0327</td>
</tr>
<tr>
<td>V2</td>
<td>500</td>
<td>0.67</td>
<td>0.66</td>
<td>0.62</td>
<td>0.1071</td>
</tr>
<tr>
<td>V3</td>
<td>500</td>
<td>0.65</td>
<td>0.65</td>
<td>0.61</td>
<td>0.53</td>
</tr>
<tr>
<td>V4</td>
<td>500</td>
<td>0.65</td>
<td>0.65</td>
<td>0.59</td>
<td>0.53</td>
</tr>
<tr>
<td>V5</td>
<td>500</td>
<td>0.64</td>
<td>0.64</td>
<td>0.58</td>
<td>0.52</td>
</tr>
<tr>
<td>V6</td>
<td>500</td>
<td>0.55</td>
<td>0.55</td>
<td>0.46</td>
<td>0.41</td>
</tr>
<tr>
<td>V7</td>
<td>500</td>
<td>0.60</td>
<td>0.60</td>
<td>0.52</td>
<td>0.46</td>
</tr>
<tr>
<td>V8</td>
<td>500</td>
<td>0.58</td>
<td>0.57</td>
<td>0.49</td>
<td>0.43</td>
</tr>
<tr>
<td>V9</td>
<td>500</td>
<td>0.47</td>
<td>0.48</td>
<td>0.36</td>
<td>0.32</td>
</tr>
</tbody>
</table>

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

72
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 items 2 and 6 (complaints and critical) are to be scored (incorrectly) negatively.

```r
> alpha(attitude, keys=c("complaints", "critical"))
```

Reliability analysis
Call: alpha(x = attitude, keys = c("complaints", "critical"))

```
<table>
<thead>
<tr>
<th>raw_alpha</th>
<th>std.alpha</th>
<th>G6(smc)</th>
<th>average_r</th>
<th>S/N</th>
<th>ase</th>
<th>mean</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.18</td>
<td>0.27</td>
<td>0.66</td>
<td>0.05</td>
<td>0.37</td>
<td>0.19</td>
<td>53</td>
<td>4.7</td>
</tr>
</tbody>
</table>
```

lower alpha upper 95% confidence boundaries
-0.18 0.18 0.55

Reliability if an item is dropped:
```
<table>
<thead>
<tr>
<th>raw_alpha std.alpha G6(smc) average_r S/N ase mean sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>rating -0.023 0.090 0.52 0.0162 0.099 0.265</td>
</tr>
<tr>
<td>complaints- 0.689 0.666 0.76 0.2496 1.995 0.078</td>
</tr>
<tr>
<td>privileges -0.133 0.021 0.58 0.0036 0.022 0.282</td>
</tr>
<tr>
<td>learning -0.459 -0.251 0.41 -0.0346 -0.201 0.363</td>
</tr>
<tr>
<td>raises -0.178 -0.062 0.47 -0.0098 -0.058 0.295</td>
</tr>
<tr>
<td>critical- 0.364 0.473 0.76 0.1299 0.596 0.139</td>
</tr>
<tr>
<td>advance -0.137 -0.016 0.52 -0.0026 -0.016 0.258</td>
</tr>
</tbody>
</table>
```

Item statistics
```
<table>
<thead>
<tr>
<th>n</th>
<th>raw.r</th>
<th>std.r</th>
<th>r.cor</th>
<th>r.drop</th>
<th>mean</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>rating</td>
<td>30</td>
<td>0.600</td>
<td>0.601</td>
<td>0.28</td>
<td>65</td>
<td>12.2</td>
</tr>
<tr>
<td>complaints</td>
<td>30</td>
<td>-0.542</td>
<td>-0.559</td>
<td>-0.78</td>
<td>-0.76</td>
<td>50</td>
</tr>
<tr>
<td>privileges</td>
<td>30</td>
<td>0.679</td>
<td>0.663</td>
<td>0.57</td>
<td>53</td>
<td>12.2</td>
</tr>
<tr>
<td>learning</td>
<td>30</td>
<td>0.857</td>
<td>0.853</td>
<td>0.90</td>
<td>56</td>
<td>11.7</td>
</tr>
<tr>
<td>raises</td>
<td>30</td>
<td>0.719</td>
<td>0.730</td>
<td>0.77</td>
<td>50</td>
<td>10.4</td>
</tr>
<tr>
<td>critical</td>
<td>30</td>
<td>0.015</td>
<td>0.036</td>
<td>-0.29</td>
<td>-0.27</td>
<td>42</td>
</tr>
<tr>
<td>advance</td>
<td>30</td>
<td>0.688</td>
<td>0.694</td>
<td>0.66</td>
<td>46</td>
<td>10.3</td>
</tr>
</tbody>
</table>
```

Note how the reliability of the 7 item scales with an incorrectly reversed item is very poor, but if items 2 and 6 is dropped then the reliability is improved substantially. This suggests that items 2 and 6 were incorrectly scored. Doing the analysis again with the items positively scored produces much more favorable results.

```r
> alpha(attitude)
```

Reliability analysis
Call: alpha(x = attitude)

```
<table>
<thead>
<tr>
<th>raw_alpha</th>
<th>std.alpha</th>
<th>G6(smc)</th>
<th>average_r</th>
<th>S/N</th>
<th>ase</th>
<th>mean</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.84</td>
<td>0.84</td>
<td>0.88</td>
<td>0.43</td>
<td>5.2</td>
<td>0.042</td>
<td>60</td>
<td>8.2</td>
</tr>
</tbody>
</table>
```

lower alpha upper 95% confidence boundaries
0.76 0.84 0.93

Reliability if an item is dropped:
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`.

```r
> 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 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

<table>
<thead>
<tr>
<th>n</th>
<th>raw.r</th>
<th>std.r</th>
<th>r.cor</th>
<th>r.drop</th>
<th>mean</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>rating</td>
<td>30</td>
<td>0.78</td>
<td>0.76</td>
<td>0.75</td>
<td>0.67</td>
<td>65</td>
</tr>
<tr>
<td>complaints</td>
<td>30</td>
<td>0.84</td>
<td>0.81</td>
<td>0.82</td>
<td>0.74</td>
<td>67</td>
</tr>
<tr>
<td>privileges</td>
<td>30</td>
<td>0.80</td>
<td>0.81</td>
<td>0.82</td>
<td>0.74</td>
<td>67</td>
</tr>
<tr>
<td>learning</td>
<td>30</td>
<td>0.81</td>
<td>0.80</td>
<td>0.78</td>
<td>0.71</td>
<td>66</td>
</tr>
<tr>
<td>raises</td>
<td>30</td>
<td>0.85</td>
<td>0.86</td>
<td>0.85</td>
<td>0.79</td>
<td>65</td>
</tr>
<tr>
<td>critical</td>
<td>30</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
<td>0.74</td>
<td>67</td>
</tr>
<tr>
<td>advance</td>
<td>30</td>
<td>0.80</td>
<td>0.78</td>
<td>0.83</td>
<td>0.38</td>
<td>3.6</td>
</tr>
</tbody>
</table>

Non missing response frequency for each item

<table>
<thead>
<tr>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>miss</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>0.04</td>
<td>0.24</td>
<td>0.40</td>
<td>0.25</td>
<td>0.07</td>
</tr>
<tr>
<td>V2</td>
<td>0.06</td>
<td>0.23</td>
<td>0.40</td>
<td>0.24</td>
<td>0.06</td>
</tr>
<tr>
<td>V3</td>
<td>0.05</td>
<td>0.25</td>
<td>0.37</td>
<td>0.26</td>
<td>0.07</td>
</tr>
<tr>
<td>V4</td>
<td>0.06</td>
<td>0.21</td>
<td>0.37</td>
<td>0.29</td>
<td>0.07</td>
</tr>
</tbody>
</table>
```
5.2 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 $\omega_h$ and $\omega_t$. These may be found using the omega function for an exploratory analysis (See Figure 29) or omegaSem for a confirmatory analysis using the sem based upon the exploratory solution from omega.

McDonald has proposed coefficient omega (hierarchical) ($\omega_h$) 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 $\omega_h$ to Cronbach’s $\alpha$ and Revelle’s $\beta$. They conclude that $\omega_h$ 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 $\omega_h$ 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 $\omega_h$.

$\omega_h$ 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).

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).

$\beta$, an alternative to $\omega$, 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 $\omega_h$ coefficient. It is important to remember that “A recommendation that should be heeded, regardless of the method chosen to estimate $\omega_h$, is to always examine the pattern of the estimated general factor loadings prior to estimating $\omega_h$. 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 $\omega_h$ estimates occasionally produced in the simulations reported here.” (Zinbarg et al., 2006, p 137).

Although $\omega_h$ 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 $\omega_h$, another of McDonald’s coefficients is $\omega_t$. This is an estimate of the total reliability of a test.

McDonald’s $\omega_t$, which is similar to Guttman’s $\lambda_6$, (see guttmann) uses the estimates of uniqueness $u^2$ from factor analysis to find $e^2_j$. This is based on a decomposition of the variance of a test score, $V_x$ into four parts: that due to a general factor, $\tilde{g}$, that due to a set of group factors, $\tilde{f}$, (factors common to some but not all of the items), specific factors, $\tilde{s}$ unique to each item, and $\tilde{e}$, 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 $\tilde{x} = \tilde{c}\tilde{g} + \tilde{A}\tilde{f} + \tilde{D}s + \tilde{e}$ then the communality of item $j$, based upon general as well as group factors, $h^2_j = c^2_j + \sum f^2_{ij}$ and the unique variance for the item $u^2_j = \sigma^2_j(1 - h^2_j)$ may be used to estimate the test reliability. That is, if $h^2_j$ is the communality of item $j$, based upon general as well as group factors, then for standardized items, $e^2_j = 1 - h^2_j$ and

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

Because $h^2_j \geq r^2_{smc}$, $\omega_t \geq \lambda_6$.

It is important to distinguish here between the two $\omega$ coefficients of McDonald, 1978 and Equation 6.20a of McDonald, 1999, $\omega_h$ and $\omega_t$. 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

76
on the general factor.

\[ \omega_h = \frac{1c^cI}{V_x} \]

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

\[ \omega_{inf} = \frac{1c^cI}{1c^cI + IA'A'I'} \]

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

9 simulated variables
```

![Diagram of a bifactor solution](image.png)

Figure 29: A bifactor solution for 9 simulated variables with a hierarchical structure.

In the case of these simulated 9 variables, the amount of variance attributable to a general factor (\(\omega_h\)) is quite large, and the reliability of the set of 9 items is somewhat greater than that estimated by \(\alpha\) or \(\lambda_6\).
> 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

<table>
<thead>
<tr>
<th></th>
<th>F1*</th>
<th>F2*</th>
<th>F3*</th>
<th>h2</th>
<th>u2</th>
<th>p2</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>0.72</td>
<td>0.45</td>
<td>0.72</td>
<td>0.28</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>V2</td>
<td>0.58</td>
<td>0.37</td>
<td>0.47</td>
<td>0.53</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>V3</td>
<td>0.57</td>
<td>0.41</td>
<td>0.49</td>
<td>0.51</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>V4</td>
<td>0.57</td>
<td>0.41</td>
<td>0.50</td>
<td>0.50</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>V5</td>
<td>0.55</td>
<td>0.32</td>
<td>0.41</td>
<td>0.59</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>V6</td>
<td>0.43</td>
<td>0.27</td>
<td>0.28</td>
<td>0.72</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>V7</td>
<td>0.47</td>
<td>0.47</td>
<td>0.44</td>
<td>0.56</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>V8</td>
<td>0.43</td>
<td>0.42</td>
<td>0.36</td>
<td>0.64</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>V9</td>
<td>0.32</td>
<td>0.23</td>
<td>0.16</td>
<td>0.84</td>
<td>0.63</td>
<td></td>
</tr>
</tbody>
</table>

With eigenvalues of:

<table>
<thead>
<tr>
<th></th>
<th>F1*</th>
<th>F2*</th>
<th>F3*</th>
</tr>
</thead>
<tbody>
<tr>
<td>g</td>
<td>2.48</td>
<td>0.52</td>
<td>0.47</td>
</tr>
<tr>
<td>max/min</td>
<td>1.46</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Explain 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 = 0.001 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 = 0.01 and the 90% confidence intervals are 0.01 0.114
BIC = -7.97

Measures of factor score adequacy

<table>
<thead>
<tr>
<th></th>
<th>g F1* F2* F3*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation of scores with factors</td>
<td>0.84 0.59 0.61 0.54</td>
</tr>
<tr>
<td>Multiple R square of scores with factors</td>
<td>0.71 0.35 0.37 0.29</td>
</tr>
<tr>
<td>Minimum correlation of factor score estimates</td>
<td>0.43 -0.30 -0.25 -0.41</td>
</tr>
</tbody>
</table>

Total, General and Subset omega for each subset

<table>
<thead>
<tr>
<th></th>
<th>g F1* F2* F3*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omega total for total scores and subscales</td>
<td>0.83 0.79 0.56 0.65</td>
</tr>
<tr>
<td>Omega general for total scores and subscales</td>
<td>0.69 0.55 0.30 0.46</td>
</tr>
<tr>
<td>Omega group for total scores and subscales</td>
<td>0.12 0.24 0.26 0.19</td>
</tr>
</tbody>
</table>
5.3 Estimating $\omega_h$ 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 $\omega_h$, primarily because cross loadings are modeled as part of the general factor.

```r
> omegaSem(r9,n.obs=500,lavaan=FALSE)
Call: omegaSem(m = r9, n.obs = 500, lavaan = FALSE)
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 = 0.001 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 = 0.01 and the 90 % confidence intervals are 0.01 0.114
BIC = -7.97
Measures of factor score adequacy

<table>
<thead>
<tr>
<th></th>
<th>g</th>
<th>F1*</th>
<th>F2*</th>
<th>F3*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation of scores with factors</td>
<td>0.84</td>
<td>0.59</td>
<td>0.61</td>
<td>0.54</td>
</tr>
<tr>
<td>Multiple R square of scores with factors</td>
<td>0.71</td>
<td>0.35</td>
<td>0.37</td>
<td>0.29</td>
</tr>
<tr>
<td>Minimum correlation of factor score estimates</td>
<td>0.43</td>
<td>-0.30</td>
<td>-0.25</td>
<td>-0.41</td>
</tr>
</tbody>
</table>

Total, General and Subset omega for each subset

<table>
<thead>
<tr>
<th></th>
<th>g</th>
<th>F1*</th>
<th>F2*</th>
<th>F3*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omega total for total scores and subscales</td>
<td>0.83</td>
<td>0.79</td>
<td>0.56</td>
<td>0.65</td>
</tr>
<tr>
<td>Omega general for total scores and subscales</td>
<td>0.69</td>
<td>0.55</td>
<td>0.30</td>
<td>0.46</td>
</tr>
<tr>
<td>Omega group for total scores and subscales</td>
<td>0.12</td>
<td>0.24</td>
<td>0.26</td>
<td>0.19</td>
</tr>
</tbody>
</table>

The following analyses were done using the sem package

Omega Hierarchical from a confirmatory model using sem = 0.72

Omega Total from a confirmatory model using sem = 0.83

With loadings of:

<table>
<thead>
<tr>
<th></th>
<th>g</th>
<th>F1*</th>
<th>F2*</th>
<th>F3*</th>
<th>h2</th>
<th>u2</th>
<th>p2</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>0.74</td>
<td>0.40</td>
<td></td>
<td></td>
<td>0.71</td>
<td>0.29</td>
<td>0.77</td>
</tr>
<tr>
<td>V2</td>
<td>0.58</td>
<td>0.36</td>
<td>0.47</td>
<td>0.53</td>
<td>0.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V3</td>
<td>0.56</td>
<td>0.42</td>
<td>0.50</td>
<td>0.50</td>
<td>0.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V4</td>
<td>0.57</td>
<td></td>
<td>0.45</td>
<td>0.53</td>
<td>0.47</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>V5</td>
<td>0.58</td>
<td>0.25</td>
<td>0.40</td>
<td>0.60</td>
<td>0.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V6</td>
<td>0.43</td>
<td>0.26</td>
<td>0.26</td>
<td>0.74</td>
<td>0.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V7</td>
<td>0.49</td>
<td>0.38</td>
<td>0.38</td>
<td>0.62</td>
<td>0.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V8</td>
<td>0.44</td>
<td>0.45</td>
<td>0.39</td>
<td>0.61</td>
<td>0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V9</td>
<td>0.34</td>
<td>0.24</td>
<td>0.17</td>
<td>0.83</td>
<td>0.68</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

With eigenvalues of:

<table>
<thead>
<tr>
<th></th>
<th>g</th>
<th>F1*</th>
<th>F2*</th>
<th>F3*</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.60</td>
<td>0.47</td>
<td>0.40</td>
<td>0.33</td>
<td></td>
</tr>
</tbody>
</table>

general/max 5.51  max/min = 1.41
mean percent general = 0.68  with sd = 0.1 and cv of 0.15
Explained Common Variance of the general factor = 0.68

5.3.1 Other estimates of reliability

Other estimates of reliability are found by the splitHalf and guttman functions. These are described in more detail in Revelle and Zinbarg (2009) and in Revelle and Condon
(2014). They include the 6 estimates from Guttman, four from TenBerge, and an estimate of the greatest lower bound.

```r
> 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
```

### 5.4 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 $G^6$.

#### 5.4.1 Scoring from raw data

To score these five scales from the 25 items, use the `scoreItems` function and a list of items to be scored on each scale (a `keys.list`). Items may be listed by location (convenient but dangerous), or name (probably safer).

Make a `keys.list` by by specifying the items for each scale, preceding items to be negatively keyed with a `-` sign:

```r
> keys.list <- list(agree=c("-A1","A2","A3","A4","A5"),
+ conscientious=c("C1","C2","C2","-C4","-C5"),
+ extraversion=c("-E1","-E2","E3","E4","E5"),
+ neuroticism=c("N1","N2","N3","N4","N5"),
+ openness = c("O1","-O2","O3","O4","-O5"))
> keys.list <- 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))
> keys.list <- list(agree=c("-A1","A2","A3","A4","-A5"),
+ conscientious=c("C1","C2","C2","-C4","-C5"),
+ extraversion=c("-E1","-E2","E3","E4","E5"),
+ neuroticism=c("N1","N2","N3","N4","N5"),
+ openness = c("O1","-O2","O3","O4","-O5"))
```

81
In the past (prior to version 1.6.9, the keys.list was then converted a keys matrix using the helper function `make.keys`. This is no longer necessary. 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. There are two ways to make up the keys. You can specify the items by location (the old way) or by name (the newer and probably preferred way). To use the newer way you must specify the file on which you will use the keys. The example below shows how to construct keys either way.

Note that the number of items to specify in the `make.keys` function is the total number of items in the inventory. This is done automatically in the new way of forming keys, but if using the older way, the number must be specified. 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.

Then, use this keys list to score the items.

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

Call: `scoreItems(keys = keys.list, items = bfi)`

(Unstandardized) Alpha:

<table>
<thead>
<tr>
<th></th>
<th>agree</th>
<th>conscientious</th>
<th>extraversion</th>
<th>neuroticism</th>
<th>openness</th>
</tr>
</thead>
<tbody>
<tr>
<td>alpha</td>
<td>0.7</td>
<td>0.69</td>
<td>0.76</td>
<td>0.81</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Standard errors of unstandardized Alpha:

<table>
<thead>
<tr>
<th></th>
<th>agree</th>
<th>conscientious</th>
<th>extraversion</th>
<th>neuroticism</th>
<th>openness</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASE</td>
<td>0.014</td>
<td>0.016</td>
<td>0.013</td>
<td>0.011</td>
<td>0.017</td>
</tr>
</tbody>
</table>

Average item correlation:

<table>
<thead>
<tr>
<th></th>
<th>agree</th>
<th>conscientious</th>
<th>extraversion</th>
<th>neuroticism</th>
<th>openness</th>
</tr>
</thead>
<tbody>
<tr>
<td>average.r</td>
<td>0.32</td>
<td>0.35</td>
<td>0.39</td>
<td>0.46</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Guttman 6* reliability:

<table>
<thead>
<tr>
<th></th>
<th>agree</th>
<th>conscientious</th>
<th>extraversion</th>
<th>neuroticism</th>
<th>openness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>agree</td>
<td>conscientious</td>
<td>extraversion</td>
<td>neuroticism</td>
<td>openness</td>
</tr>
</tbody>
</table>

82
Lambda.6 0.7 0.69 0.76 0.81 0.6

Signal/Noise based upon av.r:
agree conscientious extraversion neuroticism openness
Signal/Noise 2.3 2.2 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
conscienitious 0.25 0.69 0.37 -0.334 0.33
extraversion 0.46 0.27 0.76 -0.284 0.32
neuroticism -0.18 -0.25 -0.22 0.812 -0.12
openness 0.15 0.21 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 30

5.4.2 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 scoreItems.

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

Call: scoreItems(keys = keys.list, items = 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 0.24 0.69 0.38 -0.314 0.33
extraversion 0.47 0.27 0.76 -0.278 0.35
neuroticism -0.18 -0.25 -0.22 0.815 -0.11
openness 0.16 0.22 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
> png('scores.png')
> pairs.panels(scores$scores,pch='.',jiggle=TRUE)
> dev.off()

Figure 30: 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.
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.

### 5.5 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).

```r
> data(iqitems)
> iq.keys <- c(4,4,4, 6,6,3,4, 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: 0.84

Average item correlation: 0.25

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

sd

| reason.4        | 0.48 |
| reason.16       | 0.46 |
| reason.17       | 0.46 |
| reason.19       | 0.49 |
| letter.7        | 0.49 |
| letter.33       | 0.50 |
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.
5.6 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 31).

5.6.1 Exploring the item structure of scales

The Big Five scales found above can be understood in terms of the item - whole correlations, but it is also useful to think of the endorsement frequency of the items. The item.lookup function will sort items by their factor loading/item-whole correlation, and then resort those above a certain threshold in terms of the item means. Item content is shown by using the dictionary developed for those items. This allows one to see the structure of each scale in terms of its endorsement range. This is a simple way of thinking of items that is also possible to do using the various IRT approaches discussed later.

```r
> m <- colMeans(bfi, na.rm=TRUE)
> item.lookup(scales$item.corrected[,1:3], m, dictionary=bfi.dictionary[1:2])

<table>
<thead>
<tr>
<th>Item</th>
<th>Agree</th>
<th>Conscientious</th>
<th>Extraversion</th>
<th>Mean</th>
<th>ItemLabel</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>-0.40</td>
<td>-0.06</td>
<td>-0.10</td>
<td>2.41</td>
<td>q_146</td>
</tr>
<tr>
<td>E1</td>
<td>-0.31</td>
<td>-0.07</td>
<td>-0.59</td>
<td>2.97</td>
<td>q_712</td>
</tr>
<tr>
<td>E2</td>
<td>-0.39</td>
<td>-0.26</td>
<td>-0.70</td>
<td>3.14</td>
<td>q_501</td>
</tr>
<tr>
<td>E3</td>
<td>0.46</td>
<td>0.21</td>
<td>0.61</td>
<td>4.00</td>
<td>q_1205</td>
</tr>
<tr>
<td>E4</td>
<td>0.51</td>
<td>0.24</td>
<td>0.68</td>
<td>4.42</td>
<td>q_1410</td>
</tr>
<tr>
<td>E5</td>
<td>0.35</td>
<td>0.40</td>
<td>0.56</td>
<td>4.42</td>
<td>q_1768</td>
</tr>
<tr>
<td>A5</td>
<td>0.62</td>
<td>0.22</td>
<td>0.55</td>
<td>4.56</td>
<td>q_1419</td>
</tr>
<tr>
<td>A3</td>
<td>0.70</td>
<td>0.22</td>
<td>0.48</td>
<td>4.60</td>
<td>q_1206</td>
</tr>
<tr>
<td>A4</td>
<td>0.49</td>
<td>0.30</td>
<td>0.30</td>
<td>4.70</td>
<td>q_1364</td>
</tr>
<tr>
<td>A2</td>
<td>0.67</td>
<td>0.21</td>
<td>0.40</td>
<td>4.80</td>
<td>q_1162</td>
</tr>
<tr>
<td>C4</td>
<td>-0.23</td>
<td>-0.67</td>
<td>-0.23</td>
<td>2.66</td>
<td>q_626</td>
</tr>
<tr>
<td>B4</td>
<td>-0.22</td>
<td>-0.31</td>
<td>-0.39</td>
<td>3.19</td>
<td>q_1479</td>
</tr>
<tr>
<td>C5</td>
<td>-0.26</td>
<td>-0.58</td>
<td>-0.29</td>
<td>3.30</td>
<td>q_1949</td>
</tr>
<tr>
<td>C3</td>
<td>0.21</td>
<td>0.56</td>
<td>0.15</td>
<td>4.30</td>
<td>q_619</td>
</tr>
<tr>
<td>C2</td>
<td>0.21</td>
<td>0.61</td>
<td>0.18</td>
<td>4.37</td>
<td>q_530</td>
</tr>
<tr>
<td>ES1</td>
<td>0.35</td>
<td>0.40</td>
<td>0.56</td>
<td>4.42</td>
<td>q_1768</td>
</tr>
<tr>
<td>C1</td>
<td>0.13</td>
<td>0.54</td>
<td>0.20</td>
<td>4.50</td>
<td>q_124</td>
</tr>
<tr>
<td>E11</td>
<td>-0.31</td>
<td>-0.07</td>
<td>-0.59</td>
<td>2.97</td>
<td>q_712</td>
</tr>
<tr>
<td>S41</td>
<td>-0.22</td>
<td>-0.31</td>
<td>-0.39</td>
<td>3.19</td>
<td>q_1479</td>
</tr>
<tr>
<td>E31</td>
<td>0.45</td>
<td>0.21</td>
<td>0.61</td>
<td>4.00</td>
<td>q_1206</td>
</tr>
<tr>
<td>E41</td>
<td>0.51</td>
<td>0.24</td>
<td>0.68</td>
<td>4.42</td>
<td>q_1410</td>
</tr>
</tbody>
</table>
```
Figure 31: 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.
5.6.2 Empirical scale construction

There are some situations where one wants to identify those items that most relate to a particular criterion. Although this will capitalize on chance and the results should interpreted cautiously, it does give a feel for what is being measured. Consider the following example from the bfi data set. The items that best predicted gender, education, and age may be found using the bestScales function. This also shows the use of a dictionary that has the item content.

> data(bfi)
> bestScales(bfi, criteria=c("gender", "education", "age"), cut=.1, dictionary=bfi.dictionary[,1:3])

The items most correlated with the criteria yield \( r' \)s of

<table>
<thead>
<tr>
<th>correlation</th>
<th>n.items</th>
</tr>
</thead>
<tbody>
<tr>
<td>gender</td>
<td>0.32</td>
</tr>
<tr>
<td>education</td>
<td>0.14</td>
</tr>
<tr>
<td>age</td>
<td>0.24</td>
</tr>
</tbody>
</table>

The best items, their correlations and content are

<table>
<thead>
<tr>
<th>Row.names</th>
<th>gender</th>
<th>ItemLabel</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>0.21</td>
<td>q_1505</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.19</td>
<td>q_1162</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>-0.16</td>
<td>q_146</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0.14</td>
<td>q_1206</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>0.13</td>
<td>q_1364</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>-0.13</td>
<td>q_712</td>
</tr>
</tbody>
</table>
7 N3 0.1207718 q_1099 Have frequent mood swings.
8 O1 -0.1031059 q_128 Am full of ideas.
9 A5 0.1007091 q_1419 Make people feel at ease.

$education
Row.names education ItemLabel Item
1 A1 -0.1415734 q_146 Am indifferent to the feelings of others.

$age
Row.names age ItemLabel Item
1 A1 -0.1609637 q_146 Am indifferent to the feelings of others.
2 C4 -0.1482659 q_626 Do things in a half-way manner.
3 A4 0.1442473 q_1364 Love children.
4 A5 0.1290935 q_1419 Make people feel at ease.
5 E5 0.1146922 q_1768 Take charge.
6 A2 0.1142767 q_1364 Inquire about others’ well-being.
7 N3 -0.1108648 q_1099 Have frequent mood swings.
8 E2 -0.1051612 q_901 Find it difficult to approach others.
9 N5 -0.1043322 q_1505 Panic easily.

6 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 32).
6.1 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 \( (\alpha) \) of the items on the latent variable. These regressions are, of course just the factor loadings. Item difficulty, \( \delta_j \) and item discrimination, \( \alpha_j \) may be found from factor analysis of the tetrachoric correlations where \( \lambda_j \) is just the factor loading on the first factor and \( \tau_j \) is the normal threshold reported by the tetrachoric function.

\[
\delta_j = \frac{D\tau}{\sqrt{1 - \lambda_j^2}}, \quad \alpha_j = \frac{\lambda_j}{\sqrt{1 - \lambda_j^2}} \tag{2}
\]

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 \( (\lambda_j) \) and item thresholds \( (\tau) \) 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

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

```r
> test
Item Response Analysis using Factor Analysis

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

Summary information by factor and item
Factor = 1

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

91
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 = 0.043 and the 90 % confidence intervals are 0.043 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])
>
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 0.59 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 = 0.009 and the 90 % confidence intervals are 0.009 0.111
BIC = 96.23

The item information functions show that not all of items are equally good (Figure 33):

These procedures can be generalized to more than one factor by specifying the number of
Figure 32: 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 33: 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 \texttt{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 \texttt{irt.fa} that includes the average information for each item that has loadings greater than cut.

\begin{verbatim}
> print(e.info, sort=TRUE)

Item Response Analysis using Factor Analysis

Summary information by factor and item
Factor = 1

<table>
<thead>
<tr>
<th></th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2</td>
<td>0.48</td>
<td>0.89</td>
<td>1.18</td>
<td>1.24</td>
<td>0.99</td>
<td>0.59</td>
<td>0.25</td>
</tr>
<tr>
<td>E4</td>
<td>0.63</td>
<td>0.98</td>
<td>1.11</td>
<td>0.96</td>
<td>0.67</td>
<td>0.37</td>
<td>0.16</td>
</tr>
<tr>
<td>E1</td>
<td>0.37</td>
<td>0.58</td>
<td>0.75</td>
<td>0.76</td>
<td>0.61</td>
<td>0.39</td>
<td>0.21</td>
</tr>
<tr>
<td>E3</td>
<td>0.34</td>
<td>0.49</td>
<td>0.58</td>
<td>0.57</td>
<td>0.49</td>
<td>0.36</td>
<td>0.23</td>
</tr>
<tr>
<td>E5</td>
<td>0.33</td>
<td>0.42</td>
<td>0.47</td>
<td>0.45</td>
<td>0.37</td>
<td>0.27</td>
<td>0.17</td>
</tr>
</tbody>
</table>

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 \textit{ltm} and \textit{eRm} and should be used for serious Item Response Theory analysis.

6.2 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 \texttt{irt.fa} and \texttt{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.

In addition, recent releases of the \textit{psych} take advantage of the \textit{parallels} package and use multi-cores. The default for Macs and Unix machines is to use two cores, but this can be increased using the \texttt{options} command. The biggest step up in improvement is from 1 to 2 cores, but for large problems using polychoric correlations, the more cores available, the better.

For example of taking the output from the 16 ability items from the \textit{SAPA} project when scored for True/False using \texttt{score.multiple.choice} we can first do a simple IRT analysis of one factor (Figure 36) and then use that correlation matrix to do an \texttt{omega} analysis to show the sub-structure of the ability items. We can also show the total test information (merely the sum of the item information. This shows that even with just 16 items, the test is very reliable for most of the range of ability. The \texttt{fa.irt} function saves the correlation matrix and item statistics so that they can be redrawn with other options.

\begin{verbatim}
detectCores() # how many are available
options("mc.cores") # how many have been set to be used
\end{verbatim}
options(mc.cores=4) # set to use 4 cores

> iq.irt

Item Response Analysis using Factor Analysis

Call: irt.fa(x = ability)

Item Response Analysis using Factor Analysis

Summary information by factor and item
Factor = 1

<table>
<thead>
<tr>
<th></th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>reason.4</td>
<td>0.05</td>
<td>0.24</td>
<td>0.64</td>
<td>0.53</td>
<td>0.16</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>reason.16</td>
<td>0.08</td>
<td>0.22</td>
<td>0.38</td>
<td>0.31</td>
<td>0.14</td>
<td>0.05</td>
<td>0.01</td>
</tr>
<tr>
<td>reason.17</td>
<td>0.08</td>
<td>0.33</td>
<td>0.69</td>
<td>0.42</td>
<td>0.11</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>reason.19</td>
<td>0.06</td>
<td>0.17</td>
<td>0.35</td>
<td>0.36</td>
<td>0.19</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>letter.7</td>
<td>0.05</td>
<td>0.18</td>
<td>0.41</td>
<td>0.44</td>
<td>0.20</td>
<td>0.06</td>
<td>0.02</td>
</tr>
<tr>
<td>letter.33</td>
<td>0.05</td>
<td>0.15</td>
<td>0.31</td>
<td>0.36</td>
<td>0.20</td>
<td>0.08</td>
<td>0.02</td>
</tr>
<tr>
<td>letter.34</td>
<td>0.05</td>
<td>0.19</td>
<td>0.45</td>
<td>0.46</td>
<td>0.20</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>letter.58</td>
<td>0.02</td>
<td>0.09</td>
<td>0.30</td>
<td>0.53</td>
<td>0.35</td>
<td>0.12</td>
<td>0.03</td>
</tr>
<tr>
<td>matrix.45</td>
<td>0.05</td>
<td>0.11</td>
<td>0.19</td>
<td>0.23</td>
<td>0.17</td>
<td>0.09</td>
<td>0.04</td>
</tr>
<tr>
<td>matrix.46</td>
<td>0.05</td>
<td>0.12</td>
<td>0.22</td>
<td>0.26</td>
<td>0.18</td>
<td>0.09</td>
<td>0.04</td>
</tr>
<tr>
<td>matrix.47</td>
<td>0.06</td>
<td>0.17</td>
<td>0.33</td>
<td>0.34</td>
<td>0.19</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>matrix.55</td>
<td>0.04</td>
<td>0.08</td>
<td>0.13</td>
<td>0.17</td>
<td>0.16</td>
<td>0.11</td>
<td>0.06</td>
</tr>
<tr>
<td>rotate.3</td>
<td>0.00</td>
<td>0.01</td>
<td>0.06</td>
<td>0.30</td>
<td>0.75</td>
<td>0.49</td>
<td>0.12</td>
</tr>
<tr>
<td>rotate.4</td>
<td>0.00</td>
<td>0.01</td>
<td>0.05</td>
<td>0.31</td>
<td>1.00</td>
<td>0.56</td>
<td>0.10</td>
</tr>
<tr>
<td>rotate.6</td>
<td>0.01</td>
<td>0.03</td>
<td>0.15</td>
<td>0.53</td>
<td>0.69</td>
<td>0.25</td>
<td>0.05</td>
</tr>
<tr>
<td>rotate.8</td>
<td>0.00</td>
<td>0.02</td>
<td>0.08</td>
<td>0.29</td>
<td>0.59</td>
<td>0.41</td>
<td>0.13</td>
</tr>
<tr>
<td>Test Info</td>
<td>0.67</td>
<td>2.11</td>
<td>4.73</td>
<td>5.83</td>
<td>5.28</td>
<td>2.55</td>
<td>0.69</td>
</tr>
<tr>
<td>SEM</td>
<td>1.22</td>
<td>0.69</td>
<td>0.46</td>
<td>0.41</td>
<td>0.44</td>
<td>0.63</td>
<td>1.20</td>
</tr>
</tbody>
</table>

Reliability = -0.49 0.53 0.79 0.83 0.81 0.61 0.45

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 1.91
The number of observations was 1525 with Chi Square = 2893.45 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 = 0.723
RMSEA index = 0.018 and the 90 % confidence intervals are 0.018 0.137
BIC = 2131.15

6.3 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 scoreIrt). 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
Figure 34: 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.
> plot(iq.irt,type='test')

Figure 35: A graphic analysis of 16 ability items sampled from the SAPA project. The total test information at all levels of difficulty may be shown by specifying the type='test' option in the plot function.
> om <- omega(iq.irt$rho, 4)

Figure 36: 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.
average of the three (classical test theory) fails to consider that the items might differ in either discrimination or in difficulty. The \texttt{scoreIrt} 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:

```r
> v9 <- sim.irt(9,1000,-2.,2.,mod="normal") \# dichotomous items
> items <- v9$items
> test <- irt.fa(items)
>
> total <- rowSums(items)
> ord <- order(total)
> items <- items[ord,]
> # now delete some of the data - note that they are ordered by score
> items[1:333,5:9] <- NA
> items[334:666,3:7] <- NA
> items[667:1000,1:4] <- NA
> scores <- scoreIrt(test,items)
> unitweighted <- scoreIrt(items=items,keys=rep(1,9))
> scores.df <- data.frame(true=v9$theta[ord],scores,unitweighted)
> colnames(scores.df) <- c("True theta","irt theta","total","fit","rasch","total","fit")
```

These results are seen in Figure 37.

### 6.3.1 1 versus 2 parameter IRT scoring

In Item Response Theory, items can be assumed to be equally discriminating but to differ in their difficulty (the Rasch model) or to vary in their discriminability. Two functions (\texttt{scoreIrt.1pl} and \texttt{scoreIrt.2pl}) are meant to find multiple IRT based scales using the Rasch model or the 2 parameter model. Both allow for negatively keyed as well as positively keyed items. Consider the \texttt{bfi} data set with scoring keys \texttt{key.list} and items listed as an \texttt{item.list}. (This is the same as the \texttt{key.list}, but with the negative signs removed.)

```r
> keys.list <- list(agree=c("-A1","A2","A3","A4","A5"),
+ conscientious=c("C1","C2","C3","-C4","-C5"),
+ extraversion=c("-E1","-E2","E3","E4","E5"),
+ neuroticism=c("N1","N2","N3","N4","N5"),
+ openness = c("O1","-O2","O3","O4","-O5"))
> item.list <- list(agree=c("A1","A2","A3","A4","A5"),
+ conscientious=c("C1","C2","C3","C4","C5"),
+ extraversion=c("E1","E2","E3","E4","E5"),
+ neuroticism=c("N1","N2","N3","N4","N5"),
+ openness = c("O1","O2","O3","O4","O5"))
> bfi.1pl <- scoreIrt.1pl(keys.list,bfi) \# the one parameter solution
> bfi.2pl <- scoreIrt.2pl(item.list,bfi) \# the two parameter solution
> bfi.ctt <- scoreFast(keys.list,bfi) \# fast scoring function
```
Figure 37: IRT based scoring and total test scores for 1000 simulated subjects. True theta values are reported and then the IRT and total scoring systems.

```r
> pairs.panels(scores.df,pch='.',gap=0)
> title('Comparing true theta for IRT, Rasch and classically based scoring',line=3)
```

Comparing true theta for IRT, Rasch and classically based scoring.

<table>
<thead>
<tr>
<th></th>
<th>True theta</th>
<th>IRT theta</th>
<th>Total</th>
<th>Fit</th>
<th>Rasch</th>
<th>Total</th>
<th>Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.83</td>
<td>0.50</td>
<td>0.10</td>
<td>-0.01</td>
<td></td>
<td>0.96</td>
<td>-0.14</td>
</tr>
<tr>
<td></td>
<td>0.40</td>
<td>0.08</td>
<td>0.96</td>
<td>0.10</td>
<td>0.96</td>
<td>1.00</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>0.06</td>
<td>0.68</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.55</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.19</td>
<td>-0.20</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


We can compare these three ways of doing the analysis using the `cor2` function which correlates two separate data frames. All three models produce very similar results for the case of almost complete data. It is when we have massively missing completely at random data (MMCAR) that the results show the superiority of the irt scoring.

```r
> # compare the solutions using the cor2 function
> cor2(bfi.1pl,bfi.ctt)

<table>
<thead>
<tr>
<th></th>
<th>agree</th>
<th>conscientious</th>
<th>extraversion</th>
<th>neuroticism</th>
<th>openness</th>
</tr>
</thead>
<tbody>
<tr>
<td>agree</td>
<td>0.93</td>
<td>0.27</td>
<td>0.44</td>
<td>-0.20</td>
<td>0.19</td>
</tr>
<tr>
<td>conscientious</td>
<td>0.25</td>
<td>0.97</td>
<td>0.25</td>
<td>-0.22</td>
<td>0.17</td>
</tr>
<tr>
<td>extraversion</td>
<td>0.43</td>
<td>0.25</td>
<td>0.94</td>
<td>-0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>neuroticism</td>
<td>-0.19</td>
<td>-0.23</td>
<td>-0.22</td>
<td>1.00</td>
<td>-0.09</td>
</tr>
<tr>
<td>openness</td>
<td>0.12</td>
<td>0.19</td>
<td>0.19</td>
<td>-0.07</td>
<td>0.95</td>
</tr>
</tbody>
</table>

> cor2(bfi.2pl,bfi.ctt)

<table>
<thead>
<tr>
<th></th>
<th>agree</th>
<th>conscientious</th>
<th>extraversion</th>
<th>neuroticism</th>
<th>openness</th>
</tr>
</thead>
<tbody>
<tr>
<td>agree</td>
<td>0.98</td>
<td>0.25</td>
<td>0.49</td>
<td>-0.17</td>
<td>0.15</td>
</tr>
<tr>
<td>conscientious</td>
<td>0.26</td>
<td>0.97</td>
<td>0.25</td>
<td>-0.23</td>
<td>0.18</td>
</tr>
<tr>
<td>extraversion</td>
<td>0.43</td>
<td>0.24</td>
<td>0.94</td>
<td>-0.25</td>
<td>0.21</td>
</tr>
<tr>
<td>neuroticism</td>
<td>-0.19</td>
<td>-0.22</td>
<td>-0.18</td>
<td>0.98</td>
<td>-0.08</td>
</tr>
<tr>
<td>openness</td>
<td>0.17</td>
<td>0.21</td>
<td>0.27</td>
<td>-0.11</td>
<td>0.98</td>
</tr>
</tbody>
</table>

> cor2(bfi.2pl,bfi.1pl)

<table>
<thead>
<tr>
<th></th>
<th>agree</th>
<th>conscientious</th>
<th>extraversion</th>
<th>neuroticism</th>
<th>openness</th>
</tr>
</thead>
<tbody>
<tr>
<td>agree</td>
<td>0.88</td>
<td>0.25</td>
<td>0.45</td>
<td>-0.17</td>
<td>0.12</td>
</tr>
<tr>
<td>conscientious</td>
<td>0.26</td>
<td>1.00</td>
<td>0.24</td>
<td>-0.23</td>
<td>0.16</td>
</tr>
<tr>
<td>extraversion</td>
<td>0.43</td>
<td>0.23</td>
<td>0.99</td>
<td>-0.25</td>
<td>0.20</td>
</tr>
<tr>
<td>neuroticism</td>
<td>-0.21</td>
<td>-0.21</td>
<td>-0.19</td>
<td>0.98</td>
<td>-0.07</td>
</tr>
<tr>
<td>openness</td>
<td>0.20</td>
<td>0.19</td>
<td>0.28</td>
<td>-0.11</td>
<td>0.92</td>
</tr>
</tbody>
</table>

7 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.
7.1 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.

\[ r_{xy} = \eta_{wg} * \eta_{wg} * r_{xywg} + \eta_{bg} * \eta_{bg} * r_{xybg} \]  

(3)

where \( r_{xy} \) is the normal correlation which may be decomposed into a within group and between group correlations \( r_{xywg} \) and \( r_{xybg} \) and \( \eta \) (eta) is the correlation of the data with the within group values, or the group means.

7.2 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 trials 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 affect manipulation (statsBy(affect)).
7.3 Factor analysis by groups

Confirmatory factor analysis comparing the structures in multiple groups can be done in the lavaan package. However, for exploratory analyses of the structure within each of multiple groups, the faBy function may be used in combination with the statsBy function. First run pfunstatsBy with the correlation option set to TRUE, and then run faBy on the resulting output.

```
sb <- statsBy(bfi[c(1:25,27)], group="education",cors=TRUE)
faBy(sb,nfactors=5)  #find the 5 factor solution for each education level
```

8 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 setCor function. Set correlation is

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

where \( \lambda_i \) is the ith eigen value of the eigen value decomposition of the matrix

\[ R = R_{xx}^{-1}R_{xy}R_{xy}^{-1}R_{yy}^{-1}. \]

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.

setCor 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 setCor. Two things to notice. setCor works on the correlation or covariance or raw data matrix, and thus if using the correlation matrix, will report standardized or raw \( \hat{\beta} \) 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.

```r
> C <- cov(sat.act, use="pairwise")
> model1 <- lm(ACT ~ gender + education + age, data=sat.act)
> summary(model1)

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 `setCor`.

```r
> # compare with sector
> setCor(c(4:6), c(1:3), C, n.obs=700)

Call: setCor(y = c(4:6), x = c(1:3), data = C, n.obs = 700)

Multiple Regression from matrix input

Beta weights

<table>
<thead>
<tr>
<th></th>
<th>ACT</th>
<th>SATV</th>
<th>SATQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>gender</td>
<td>-0.05</td>
<td>-0.03</td>
<td>-0.18</td>
</tr>
<tr>
<td>education</td>
<td>0.14</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>age</td>
<td>0.03</td>
<td>-0.10</td>
<td>-0.09</td>
</tr>
</tbody>
</table>

Multiple R

<table>
<thead>
<tr>
<th></th>
<th>ACT</th>
<th>SATV</th>
<th>SATQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.16</td>
<td>0.10</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Multiple R2

<table>
<thead>
<tr>
<th></th>
<th>ACT</th>
<th>SATV</th>
<th>SATQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0272</td>
<td>0.0096</td>
<td>0.0359</td>
</tr>
</tbody>
</table>

Unweighted multiple R

<table>
<thead>
<tr>
<th></th>
<th>ACT</th>
<th>SATV</th>
<th>SATQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.15</td>
<td>0.05</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Unweighted multiple R2

<table>
<thead>
<tr>
<th></th>
<th>ACT</th>
<th>SATV</th>
<th>SATQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.02</td>
<td>0.00</td>
<td>0.01</td>
</tr>
</tbody>
</table>
```

105
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 setCor 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.
9 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.

107
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 a major factor structure with some small correlated residuals.

Although coefficient $\omega_h$ 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

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

where $\gamma$ is the lower asymptote or guessing parameter, $\zeta$ is the upper asymptote (normally 1), $\alpha_j$ is item discrimination and $\delta_j$ 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.

108
With the $a = \alpha$ parameter $= 1.702$ in the logistic model the two models are practically identical.

10 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)

```r
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)
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.

11 Converting output to APA style tables using \LaTeX

Although for most purposes, using the *Sweave* or *KnitR* packages produces clean output, some prefer output pre formatted for APA style tables. This can be done using the *xtable* package for almost anything, but there are a few simple functions in *psych* for the most common tables. *fa2latex* will convert a factor analysis or components analysis output to a \LaTeX table, *cor2latex* will take a correlation matrix and show the lower (or upper diagonal), *irt2latex* converts the item statistics from the *irt.fa* function to more convenient \LaTeX output, and finally, *df2latex* converts a generic data frame to \LaTeX.

An example of converting the output from *fa* to \LaTeX appears in Table 11.
> 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.rect(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=ul$right,labels ="right to left")
> dia.curved.arrow(to=ur,from=ul$right,labels ="left to right")
> dia.arrow(from=ul,to=ul,right,labels ="right to left")
> dia.curve(ll$top,ul$bottom,"double",-1) #for rectangles, specify where to point
> dia.arrow(mr,ur,"up") #but for ellipses, just point to it.
> dia.arrow(ml,mr,"across")
> dia.arrow(ur,lr,"top down")
> dia.curved.arrow(bl$top,lr$bottom,"up")
> dia.curved.arrow(bl,lr$bottom)
> dia.curved.arrow(ml,ll$top,scale=-1)
> dia.curved.arrow(mr,lr$top)

Demonstration of dia functions

Figure 38: The basic graphic capabilities of the dia functions are shown in this figure.
Table 11: fa2latex  
A factor analysis table from the psych package in R

<table>
<thead>
<tr>
<th>Variable</th>
<th>MR1</th>
<th>MR2</th>
<th>MR3</th>
<th>h2</th>
<th>u2</th>
<th>com</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentences</td>
<td>0.91</td>
<td>-0.04</td>
<td>0.04</td>
<td>0.82</td>
<td>0.18</td>
<td>1.01</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.89</td>
<td>0.06</td>
<td>-0.03</td>
<td>0.84</td>
<td>0.16</td>
<td>1.01</td>
</tr>
<tr>
<td>Sent.Completion</td>
<td>0.83</td>
<td>0.04</td>
<td>0.00</td>
<td>0.73</td>
<td>0.27</td>
<td>1.00</td>
</tr>
<tr>
<td>First.Leters</td>
<td>0.00</td>
<td>0.86</td>
<td>0.00</td>
<td>0.73</td>
<td>0.27</td>
<td>1.00</td>
</tr>
<tr>
<td>4.Letter.Words</td>
<td>-0.01</td>
<td>0.74</td>
<td>0.10</td>
<td>0.63</td>
<td>0.37</td>
<td>1.04</td>
</tr>
<tr>
<td>Suffixes</td>
<td>0.18</td>
<td>0.63</td>
<td>-0.08</td>
<td>0.50</td>
<td>0.50</td>
<td>1.20</td>
</tr>
<tr>
<td>Letter.Series</td>
<td>0.03</td>
<td>-0.01</td>
<td>0.84</td>
<td>0.72</td>
<td>0.28</td>
<td>1.00</td>
</tr>
<tr>
<td>Pedigrees</td>
<td>0.37</td>
<td>-0.05</td>
<td>0.47</td>
<td>0.50</td>
<td>0.50</td>
<td>1.93</td>
</tr>
<tr>
<td>Letter.Group</td>
<td>-0.06</td>
<td>0.21</td>
<td>0.64</td>
<td>0.53</td>
<td>0.47</td>
<td>1.23</td>
</tr>
</tbody>
</table>

SS loadings 2.64 1.86 1.5

<table>
<thead>
<tr>
<th>MR1</th>
<th>1.00</th>
<th>0.59</th>
<th>0.54</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR2</td>
<td>0.59</td>
<td>1.00</td>
<td>0.52</td>
</tr>
<tr>
<td>MR3</td>
<td>0.54</td>
<td>0.52</td>
<td>1.00</td>
</tr>
</tbody>
</table>

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 \LaTeX{} table. May be used when Sweave is not convenient.

Cor2latex Will format a correlation matrix in APA style in a \LaTeX{} 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.

111
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 Bechtholdt 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 (1984), 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 download that version, go to the repository http://personality-project.org/r/src/contrib/ and wander around. For a Mac, this version can be installed directly using the “other repository” option in the package installer. For a PC, the zip file for the most recent release has been created using the win-builder facility at CRAN. The development release for the Mac is usually several weeks ahead of the PC development version.

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 > "1.5.0", 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) (2017-01-04 r71889)
Platform: x86_64-apple-darwin13.4.0 (64-bit)
Running under: macOS Sierra 10.12.2
locale:
[1] C
attached base packages:
[1] stats graphics grDevices utils datasets methods base
other attached packages:
loaded via a namespace (and not attached):
[1] Rcpp_0.12.4 lattice_0.20-34 matrixcalc_1.0-3 MASS_7.3-45
[5] grid_3.4.0 arm_1.8-6 nleq_3.1-128 stats4_3.4.0
[9] coda_0.18-1 knn_1.2-4 mi_1.0 nlme_1.0-10
[13] Matrix_1.2-7.1 boot_1.3-18 splines_3.4.0 lme4_1.1-12
[17] sem_3.1-7 tools_3.4.0 foreign_0.8-67 abind_1.4-3
[21] parallel_3.4.0 compiler_3.4.0 mnormt_1.6-4
References


Bliese, P. D. (2009). Multilevel modeling in r (2.3) a brief introduction to r, the multilevel package and the nlme package.


<table>
<thead>
<tr>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>affect, 13, 22</td>
</tr>
<tr>
<td>agnes, 52</td>
</tr>
<tr>
<td>alpha, 5, 6, 71–74, 87</td>
</tr>
<tr>
<td>Bechtoldt.1, 112</td>
</tr>
<tr>
<td>Bechtoldt.2, 112</td>
</tr>
<tr>
<td>bestScales, 89</td>
</tr>
<tr>
<td>bfi, 23, 81–83, 89, 100, 112</td>
</tr>
<tr>
<td>bi.bars, 6, 23</td>
</tr>
<tr>
<td>bifactor, 6, 42, 52</td>
</tr>
<tr>
<td>biquartimin, 42</td>
</tr>
<tr>
<td>biserial, 12, 31</td>
</tr>
<tr>
<td>block.random, 111</td>
</tr>
<tr>
<td>burt, 31</td>
</tr>
<tr>
<td>circadian.cor, 111</td>
</tr>
<tr>
<td>circadian.linear.cor, 111</td>
</tr>
<tr>
<td>circadian.mean, 111</td>
</tr>
<tr>
<td>circular statistics, 111</td>
</tr>
<tr>
<td>cities, 113</td>
</tr>
<tr>
<td>cluster analysis, 39, 52</td>
</tr>
<tr>
<td>cluster scores, 39</td>
</tr>
<tr>
<td>cluster.cor, 71, 81, 83</td>
</tr>
<tr>
<td>cluster.loadings, 81, 85</td>
</tr>
<tr>
<td>cohen.kappa, 112</td>
</tr>
<tr>
<td>common variance, 48</td>
</tr>
<tr>
<td>component scores, 39</td>
</tr>
<tr>
<td>congruence coefficient, 55</td>
</tr>
<tr>
<td>cor, 24</td>
</tr>
<tr>
<td>cor.smooth, 31</td>
</tr>
<tr>
<td>cor.test, 25</td>
</tr>
<tr>
<td>cor2, 102</td>
</tr>
<tr>
<td>cor2latex, 109, 111</td>
</tr>
<tr>
<td>corPlot, 6</td>
</tr>
<tr>
<td>corr.p, 25, 29</td>
</tr>
<tr>
<td>corr.test, 25, 29</td>
</tr>
<tr>
<td>cortest, 30</td>
</tr>
<tr>
<td>cosinor, 111</td>
</tr>
<tr>
<td>cvt, 7</td>
</tr>
<tr>
<td>cubes, 112</td>
</tr>
<tr>
<td>densityBy, 13</td>
</tr>
<tr>
<td>describe, 6, 9, 87, 111</td>
</tr>
<tr>
<td>describeBy, 9, 10</td>
</tr>
<tr>
<td>df2latex, 109, 111</td>
</tr>
<tr>
<td>dia.rect, 109</td>
</tr>
<tr>
<td>diagram, 7</td>
</tr>
<tr>
<td>draw.cor, 31</td>
</tr>
<tr>
<td>draw.tetra, 31</td>
</tr>
<tr>
<td>dummy.code, 12</td>
</tr>
<tr>
<td>edit, 4</td>
</tr>
<tr>
<td>epi.bfi, 112</td>
</tr>
<tr>
<td>eRm, 95</td>
</tr>
<tr>
<td>error bars, 17</td>
</tr>
<tr>
<td>error.bars, 6, 13, 17</td>
</tr>
<tr>
<td>error.bars.by, 9, 13, 17</td>
</tr>
<tr>
<td>error.bars.tab, 17</td>
</tr>
<tr>
<td>error.crosses, 17</td>
</tr>
<tr>
<td>errorCircles, 22</td>
</tr>
<tr>
<td>errorCrosses, 21</td>
</tr>
<tr>
<td>esem, 68</td>
</tr>
<tr>
<td>esem.diagram, 68</td>
</tr>
<tr>
<td>FA, 39, 48</td>
</tr>
<tr>
<td>fa, 6, 7, 39, 41, 42, 48, 49, 55, 87, 109</td>
</tr>
<tr>
<td>fa.diagram, 6, 41, 47, 67, 109</td>
</tr>
<tr>
<td>fa.extension, 67, 113</td>
</tr>
<tr>
<td>fa.graph, 41</td>
</tr>
<tr>
<td>fa.irt, 95</td>
</tr>
<tr>
<td>fa.parallel, 4, 6, 40, 62, 64, 87</td>
</tr>
<tr>
<td>fa.parallel.poly, 64</td>
</tr>
<tr>
<td>fa.plot, 40, 42</td>
</tr>
<tr>
<td>fa.poly, 39, 55</td>
</tr>
<tr>
<td>fa2latex, 109, 111</td>
</tr>
<tr>
<td>faBy, 104</td>
</tr>
<tr>
<td>factanal, 40, 41</td>
</tr>
</tbody>
</table>
factor analysis, 6, 39, 48
factor extension, 66
factor scores, 39
factor.congruence, 40, 61
factor.minres, 7, 40–42
factor.pa, 7, 40–42, 75
factor.wls, 7, 40
factors, 39
fisherz, 111
galton, 113
generalized least squares, 6
generalized least squares factor analysis, 39
geometric.mean, 112
GPArotation, 7, 42, 52
guttman, 6, 76, 80
harmonic.mean, 112
hclust, 52
head, 112
headtail, 112
heights, 112
het.diagram, 6
hierarchical cluster analysis, 52
Hmisc, 25
Holzinger, 112
ICC, 6, 112
iclust, 6, 41, 52, 53, 59, 70, 75, 87, 109
iclust.diagram, 6, 109
Index, 111
introduction to psychometric theory with applications in R, 7
iqitems, 112
irt.fa, 6, 90, 95, 99, 109
irt.npl, 108
irt.npn, 108
irt.responses, 87, 88
irt2latex, 109, 111
item.lookup, 87
keys, 73
KnitR, 109
lavaan, 66, 104
library, 8
lm, 31
logistic, 91
lowerCor, 24
lowerMat, 24
lowerUpper, 24
lowess, 13
ltm, 95
make.keys, 82
MAP, 6, 40, 62
mardia, 112
maximum likelihood, 6
maximum likelihood factor analysis, 39
mediate, 36, 37
mediate.diagram, 36
min.res, 41
minimum residual, 6
minres factor analysis, 39
mixed.cor, 31
multi.hist, 6
multilevel, 103
multiple regression, 31
nfactors, 6, 40
nlme, 103
oblimin, 42
omega, 6, 7, 48, 55, 70, 71, 75, 95, 109
omegaSem, 75, 79
outlier, 4, 10, 11
p.adjust, 25
p.rep, 112
pairs, 13
pairs.panels, 4, 6, 11–15, 83
parallels, 95
paran, 64
partial.r, 112
PCA, 39, 48
peas, 112, 113
plot, 42, 46
plot.irt, 6
plot.poly, 6
plot.psych, 109
pnorm, 108
polychoric, 6, 31, 55, 64, 90
polyserial, 31
principal, 5–7, 40, 42, 48, 75
principal axis, 6
principal axis factor analysis, 39
principal components, 39
principal components analysis, 39, 48
princomp, 48
print, 42
print.psych, 45
progressBar, 55
Promax, 42, 48
psych, 4–6, 8, 25, 36, 39, 41, 52, 82, 95, 109, 111, 112, 114
quartimax, 42
quartimin, 42
circadian.mean, 111
cities, 113
cluster.cor, 71, 81, 83
cluster.loadings, 81, 85
cohen.kappa, 112
cor, 24
cor.smooth, 31
cor.test, 25
cor2, 102
cor2latex, 109, 111
corPlot, 6
corr.p, 25, 29
corr.test, 25, 29
cortest, 30
cosinor, 111
cubits, 112
densityBy, 13
describe, 6, 9, 87, 111
describeBy, 9, 10
df2latex, 109, 111
dia.rect, 109
draw.cor, 31
draw.tetra, 31
dummy.code, 12
edit, 4
epi.bfi, 112
ero.bars, 6, 13, 17
error.bars.by, 9, 13, 17, 18
error.bars.tab, 17
error.circles, 17
errorCircles, 22
errorCrosses, 21
esem, 68
esem.diagram, 68
fa, 6, 7, 39, 41, 42, 48, 49, 55, 87, 109
fa.diagram, 6, 41, 47, 67, 109
fa.extension, 67, 113
fa.graph, 41
fa.irt, 95
fa.parallel, 4, 6, 40, 62, 64, 87
fa.parallel.poly, 64

R function
affect, 13
agines, 52
alpha, 5, 6, 71–74, 87
Bechtoldt.1, 112
Bechtoldt.2, 112
bestScales, 89
bfi, 23, 81–83, 89, 100, 112
bi.bars, 6, 23
bifactor, 42, 52
biquartimin, 42
biserial, 12, 31
block.random, 111
burt, 31
circadian.cor, 111
circadian.linear.cor, 111

122
<table>
<thead>
<tr>
<th>Function</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rangeCorrection</td>
<td>112</td>
</tr>
<tr>
<td>rcorr</td>
<td>25</td>
</tr>
<tr>
<td>read.clipboard</td>
<td>6, 8</td>
</tr>
<tr>
<td>read.clipboard.csv</td>
<td>8</td>
</tr>
<tr>
<td>read.clipboard.fwf</td>
<td>8</td>
</tr>
<tr>
<td>read.clipboard.lower</td>
<td>8</td>
</tr>
<tr>
<td>read.clipboard.tab</td>
<td>4, 8, 9</td>
</tr>
<tr>
<td>read.clipboard.upper</td>
<td>8</td>
</tr>
<tr>
<td>read.table</td>
<td>8</td>
</tr>
<tr>
<td>Reise</td>
<td>112</td>
</tr>
<tr>
<td>response.frequencies</td>
<td>87</td>
</tr>
<tr>
<td>reverse.code</td>
<td>112</td>
</tr>
<tr>
<td>Rgraphviz</td>
<td>41</td>
</tr>
<tr>
<td>sat.act</td>
<td>9, 30, 104</td>
</tr>
<tr>
<td>scatter.hist</td>
<td>6</td>
</tr>
<tr>
<td>schmid</td>
<td>6, 7, 75, 76</td>
</tr>
<tr>
<td>score.multiple.choice</td>
<td>6, 85, 87, 95</td>
</tr>
<tr>
<td>scoreIrt</td>
<td>96, 100</td>
</tr>
<tr>
<td>scoreIrt.1pl</td>
<td>100</td>
</tr>
<tr>
<td>scoreIrt.2pl</td>
<td>100</td>
</tr>
<tr>
<td>scoreItems</td>
<td>5, 6, 71, 74, 75, 81–83, 85–87</td>
</tr>
<tr>
<td>scrub</td>
<td>4, 12</td>
</tr>
<tr>
<td>sector</td>
<td>38</td>
</tr>
<tr>
<td>set.cor</td>
<td>66, 87, 112</td>
</tr>
<tr>
<td>setCor</td>
<td>31, 104–106</td>
</tr>
<tr>
<td>sim</td>
<td>107</td>
</tr>
<tr>
<td>sim.anova</td>
<td>107</td>
</tr>
<tr>
<td>sim.circ</td>
<td>107</td>
</tr>
<tr>
<td>sim.congeneric</td>
<td>107</td>
</tr>
<tr>
<td>sim.dichot</td>
<td>107</td>
</tr>
<tr>
<td>sim.hierarchical</td>
<td>107</td>
</tr>
<tr>
<td>sim.irt</td>
<td>107, 108</td>
</tr>
<tr>
<td>sim.item</td>
<td>107</td>
</tr>
<tr>
<td>sim.minor</td>
<td>62, 107, 108</td>
</tr>
<tr>
<td>sim.multilevel</td>
<td>103, 108</td>
</tr>
<tr>
<td>sim.npl</td>
<td>107, 108</td>
</tr>
<tr>
<td>sim.npn</td>
<td>107, 108</td>
</tr>
<tr>
<td>sim.omega</td>
<td>107, 108</td>
</tr>
<tr>
<td>sim.parallel</td>
<td>107</td>
</tr>
<tr>
<td>sim.rasch</td>
<td>107, 108</td>
</tr>
<tr>
<td>sim.simple</td>
<td>107</td>
</tr>
<tr>
<td>sim.simplex</td>
<td>107, 108</td>
</tr>
<tr>
<td>sim.structural</td>
<td>107</td>
</tr>
<tr>
<td>sim.structure</td>
<td>62, 107, 108</td>
</tr>
<tr>
<td>sim.vss</td>
<td>107</td>
</tr>
<tr>
<td>smc</td>
<td>40</td>
</tr>
<tr>
<td>spider</td>
<td>12</td>
</tr>
<tr>
<td>splitHalf</td>
<td>80</td>
</tr>
<tr>
<td>statsBy</td>
<td>103, 104</td>
</tr>
<tr>
<td>statsBy.boot</td>
<td>103</td>
</tr>
<tr>
<td>statsBy.boot.summary</td>
<td>103</td>
</tr>
<tr>
<td>structure.diagram</td>
<td>6, 109</td>
</tr>
<tr>
<td>superMatrix</td>
<td>112</td>
</tr>
<tr>
<td>table</td>
<td>17</td>
</tr>
<tr>
<td>tail</td>
<td>112</td>
</tr>
<tr>
<td>target.rot</td>
<td>42</td>
</tr>
<tr>
<td>tetrachoric</td>
<td>6, 31, 55, 90, 91</td>
</tr>
<tr>
<td>Thurstone</td>
<td>25, 112</td>
</tr>
<tr>
<td>Thurstone.33</td>
<td>112</td>
</tr>
<tr>
<td>topBottom</td>
<td>112</td>
</tr>
<tr>
<td>varimax</td>
<td>42</td>
</tr>
<tr>
<td>variminc</td>
<td>42</td>
</tr>
<tr>
<td>vegetables</td>
<td>112, 113</td>
</tr>
<tr>
<td>violinBy</td>
<td>13, 16</td>
</tr>
<tr>
<td>vss</td>
<td>4, 6, 40, 62, 63, 87</td>
</tr>
<tr>
<td>withinBetween</td>
<td>103</td>
</tr>
</tbody>
</table>

R package

<table>
<thead>
<tr>
<th>Package</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ctv</td>
<td>7</td>
</tr>
<tr>
<td>eRm</td>
<td>95</td>
</tr>
<tr>
<td>GPArotation</td>
<td>4, 42, 52</td>
</tr>
<tr>
<td>Hmisc</td>
<td>25</td>
</tr>
<tr>
<td>KnitR</td>
<td>109</td>
</tr>
<tr>
<td>lavaan</td>
<td>66, 104</td>
</tr>
<tr>
<td>ltm</td>
<td>95</td>
</tr>
<tr>
<td>multilevel</td>
<td>103</td>
</tr>
<tr>
<td>nlme</td>
<td>103</td>
</tr>
<tr>
<td>parallels</td>
<td>95</td>
</tr>
<tr>
<td>paran</td>
<td>64</td>
</tr>
<tr>
<td>psych</td>
<td>4–6, 8, 25, 36, 39, 41, 52, 82, 95, 109, 111, 112, 114</td>
</tr>
<tr>
<td>Rgraphviz</td>
<td>7</td>
</tr>
</tbody>
</table>

126
sem, 7, 66, 75, 79, 112
stats, 25, 41, 48
Sweave, 109
xtable, 109
r.test, 25
rangeCorrection, 112
rcorr, 25
read.clipboard, 6, 8
read.clipboard.csv, 8
read.clipboard.fwf, 8
read.clipboard.lower, 8
read.clipboard.tab, 4, 8, 9
read.clipboard.upper, 8
read.table, 8
Reise, 112
response.frequencies, 87
reverse.code, 112
Rgraphviz, 7, 41
rooted dendritic structure, 52
SAPA, 23, 85, 95–98, 113
sat.act, 9, 30, 104
scatter.hist, 6
schmid, 6, 7, 75, 76
Schmid-Leiman, 48
score.multiple.choice, 6, 85, 87, 95
scoreIrt, 96, 100
scoreIrt.1pl, 100
scoreIrt.2pl, 100
scoreItems, 5, 6, 71, 74, 75, 81–83, 85–87
scrub, 4, 12
sector, 38
sem, 7, 66, 75, 79, 112
set correlation, 66, 104
set.cor, 66, 87, 112
setCor, 31, 104–106
sim, 107
sim.anova, 107
sim.circ, 107
sim.congeneric, 107
sim.dichot, 107
sim.hierarchical, 107
sim.irt, 107, 108
sim.item, 107
sim.minor, 62, 107, 108
sim.multilevel, 103, 108
sim.npl, 107, 108
sim.npn, 107, 108
sim.omega, 107, 108
sim.parallel, 107
sim.rasch, 107, 108
sim.simple, 107
sim.simplex, 107, 108
sim.structural, 107
sim.structure, 62, 107, 108
sim.vss, 107
Singular Value Decomposition, 39
smc, 40
spider, 12
splitHalf, 80
stats, 25, 41, 48
statsBy, 103, 104
statsBy.boot, 103
statsBy.boot.summary, 103
structure.diagram, 6, 109
superMatrix, 112
Sweave, 109
table, 17
tail, 112
target.rot, 42
tetrachoric, 6, 31, 55, 90, 91
Thurstone, 25, 31, 112
Thurstone.33, 112
topBottom, 112
tree diagram, 52
varimax, 42
varimin, 42
vegetables, 112, 113
violinBy, 13, 16
vss, 4, 6, 40, 62, 63, 87
weighted least squares, 6
weighted least squares factor analysis, 39
withinBetween, 103
xtable, 109