

# Psychology 371: Personality Research Scale Construction

William Revelle

Department of Psychology  
Northwestern University  
Evanston, Illinois USA



May, 2019

10 Steps ○○○	The Problem ○○	Preliminaries ○ ○○	Scoring ○ ○○ ○○○○○○○	How many? ○○○○○ ○○○	Alternatives ○ ○○	Show the items ○○○○ ○○○○○	Empirical ○○○○○	Conclusion ○○○ ○
-----------------	-------------------	--------------------------	-------------------------------	---------------------------	-------------------------	---------------------------------	--------------------	------------------------

## Outline

Steps towards scale construction

The Problem

Preliminary steps

Data checking

Score the scales

Keying

Scoring using scoreItems

Determining how many constructs are in a set of items

Factor Analysis

Cluster Analysis

Scoring the alternative solutions

MultiTrait-MultiMethod validity

Multitrait-MultiMethod

Show the items

Factor analysis solution

Cluster analysis solution

Empirical scale construction

## Scale construction: A 10 steps program

1. Personality scales are not created in a theoretical vacuum. Perhaps the most important step in developing a new scale is a consideration of what is the construct of interest. What is it, what are manifestations of it, what is it not, and what should it not relate to.
2. Then, what is the population of interest? Are they young or old, highly literate, or somewhat challenged by literacy. Write items suitable for the population of interest.
3. Give the items to the participants. Make sure that they are engaged in the task.

## Scale construction: A 10 steps program (continued)

- To analyze the data, it is necessary to enter the data into a machine readable form.
  - This is a source of error. Double check for data entry errors.
  - Double entry (two different people enter the data and then the two files are automatically compared) is recommended.
  - Even better is automatic data entry (but then you need to check and double check the program).
  - `my.data <- read.file()` #go find the file on your computer
  - `my.data <- read.file(myfile)` #if you have the file name some
  - `my.data <- read.clipboard()` #if you have already copied the data to the clipboard
- Run basic descriptive statistics to do one more check for errors. Graphically check as well.
  - `describe(my.data)`
  - `pairs.panels(my.data)`
- Form the variance/covariance matrix from the items and examine the dimensionality of the resulting space.



## Scale construction: A 10 steps program (continued)

7. Apply various data reduction techniques (factor analysis, principal components analysis, cluster analysis).
  - fa
  - irt.fa # if you have polytomous or dichotomous items
  - principal
  - iclust
8. Form composite scales of the selected items. Check these scales for various measures of internal consistency.
  - make.keys
  - scoreItems
  - bestScales (For empirical scale construction)
9. Discriminant validity requires that the scales not correlate with other, unrelated traits.
10. Convergent validity requires that the scale do correlate with other, alternative measures of the same trait.

## Basic item development

As a demonstration of scale construction and validation, consider the following problem. N self report items are given to a number of people. This inventory has is composed of subsets of items that measure believed to measure different traits. In addition, each subject is rated by a friend on those same traits. There are several questions we can ask of these data:

1. Do the items form reliable scales?
2. What are the correlations of these scales?
3. Do the scales correlate with the peer ratings?
4. Can we empirically find a better structure of the items?
5. Do these revised scales show greater independence, reliability, and validity?

## Item writing

To show the procedures, 12 students in a personality research course spent several weeks learning about each of four personality dimensions. Each student then wrote five items to assess each of four constructs.

1. Need for Achievement
2. Anxiety
3. Sociability
4. Impulsivity

As a group they examined all of the items and formed the best 80 items into one questionnaire with 20 items believed to measure each of the constructs. An additional four items were the simple stem: "I think I am ... ". They administered this questionnaire to approximately ten friends each whom they also rated on these four constructs. Thus, we have a data set of about 110 participants assessed on 89 items (the 84 self report items and the 4 peer ratings + Gender).

These four sets of items can be seen as samples from four domains.

## Initial data reading

The data, item labels, and scoring keys are saved on a web server. They may be accessed by the `read.table(file.name)` command. We then use the `dim` command to find out the dimensions of the data file as well as the `names` command to find out what the names are. Unfortunately, given that the server is now an https server, it is necessary to read the data using a browser and then copy to the clipboard.

```
prq.data.name <- "http://personality-project.org/revelle/syllabi/371/prq.data"
prq.dictionary <- "http://personality-project.org/revelle/syllabi/371/prq.labels"
prq.data <- read.table(prq.data.name,header=TRUE)
pro <- read.clipboard.tab()
prq.dictionary <- read.clipboard.tab()
dim(prq.data)
names(prq.data)

> dim(prq)
[1] 110 89
names(prq)
 [1] "Q1" "Q2" "Q3" "Q4" "Q5" "Q6" "Q7" "Q8" "Q9" "Q10" "Q11" "Q12" "Q13" "Q14" "Q15" "Q16"
[17] "Q17" "Q18" "Q19" "Q20" "Q21" "Q22" "Q23" "Q24" "Q25" "Q26" "Q27" "Q28" "Q29" "Q30" "Q31" "Q32"
[33] "Q33" "Q34" "Q35" "Q36" "Q37" "Q38" "Q39" "Q40" "Q41" "Q42" "Q43" "Q44" "Q45" "Q46" "Q47" "Q48"
[49] "Q49" "Q50" "Q51" "Q52" "Q53" "Q54" "Q55" "Q56" "Q57" "Q58" "Q59" "Q60" "Q61" "Q62" "Q63" "Q64"
[65] "Q65" "Q66" "Q67" "Q68" "Q69" "Q70" "Q71" "Q72" "Q73" "Q74" "Q75" "Q76" "Q77" "Q78" "Q79" "Q80"
[81] "Q81" "Q82" "Q83" "Q84" "N" "A" "S" "I" "G"
```

## Data checking

Always check the data first. Use the describe function.

```
> describe(prq)
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
Q1	1	110	2.26	1.16	2.0	2.14	1.48	1	6	5	0.87	0.49	0.11
Q2	2	110	3.94	1.50	4.0	4.00	1.48	1	6	5	-0.34	-0.97	0.14
Q3	3	110	4.42	1.28	5.0	4.55	1.48	1	6	5	-0.65	-0.26	0.12
Q4	4	110	3.85	1.24	4.0	3.85	1.48	1	6	5	-0.04	-0.83	0.12
Q5	5	110	4.22	1.30	4.0	4.32	1.48	1	6	5	-0.43	-0.27	0.12
Q6	6	110	3.19	1.71	3.0	3.11	1.48	1	6	5	0.20	-1.20	0.16
Q7	7	110	3.62	1.31	4.0	3.58	1.48	1	6	5	0.07	-0.79	0.13
Q8	8	110	4.49	1.27	5.0	4.61	1.48	1	6	5	-0.73	0.24	0.12
...													
Q78	78	110	4.29	1.29	4.5	4.42	0.74	1	6	5	-0.75	0.09	0.12
Q79	79	110	3.95	1.28	4.0	3.99	1.48	1	6	5	-0.41	-0.39	0.12
Q80	80	110	3.01	1.44	3.0	2.93	1.48	1	6	5	0.39	-0.69	0.14
Q81	81	110	4.90	1.10	5.0	5.05	1.48	1	6	5	-1.03	0.98	0.10
Q82	82	110	3.25	1.53	3.0	3.23	1.48	1	6	5	-0.02	-0.98	0.15
Q83	83	110	4.27	1.24	4.0	4.35	1.48	1	6	5	-0.38	-0.39	0.12
Q84	84	110	3.08	1.40	3.0	3.02	1.48	1	6	5	0.36	-0.67	0.13
N	85	110	6.15	2.09	6.0	6.23	1.48	1	10	9	-0.29	-0.12	0.20
A	86	110	5.15	2.07	5.0	5.12	2.22	1	10	9	0.18	-0.64	0.20
S	87	110	5.64	2.10	6.0	5.60	2.97	2	10	8	0.07	-0.93	0.20
I	88	110	4.32	2.00	4.0	4.16	1.48	1	9	8	0.53	-0.36	0.19
G	89	110	1.58	0.50	2.0	1.60	0.00	1	2	1	-0.33	-1.91	0.05

## Data checking

In doing this, we discovered (on the first pass through the data) that one of the variables had a range of 46 rather than the 6 that was appropriate. Correcting the data, we can start over again. Even with well meaning, careful data entry, mistakes will happen in data entry. It is recommended that data be entered twice and then compared using software that compares the two files line by line and entry by entry. In all cases, make sure to describe the data and check that the ranges are appropriate for the data.

Thus, the data were edited and the prior steps were done again until there were no incorrectly entered subjects. One error that makes data checking complicated is a blank field in Excel is read improperly. However, if we copy the data file to the clipboard and then use the `read.clipboard.tab` function, this solves that problem. Note that the describe output shows that some variables do not have as many subjects as others.

## Score the scales

1. Forming scale scores as linear sums (or averages) of the items is easy to do in R.
2. One technique (not recommended) is to do a series of recodings, creating new variables for each scale.
3. A simpler technique, using the `scoreItems` function from the *psych* package does this for all scales defined in a matrix of keys (the keys matrix).
4. This is essentially a matrix of -1, 0, and 1s where 0 means don't include the item in the scale, and a 1 means to include it. -1 means to reverse key the item.

## Making up the scoring keys

```
> nach <- c(-1, 5, 9, 13, 17, 21, 25, 29, -33, 37, 41, 45, 49, 53, 57, 61, 65, 69, 73, -77, 81)
> anx <- c(2, 6, -10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50, 54, 58, 62, 66, -70, -74, -78, 82)
> soc <- c(3, 7, 11, -15, -19, 23, -27, 31, 35, 39, -43, 47, 51, 55, 59, 63, 67, 71, 75, 79, 83)
> imp <- c(4, -8, 12, -16, 20, -24, -28, 32, 36, 40, 44, 48, 52, 56, 60, 64, -68, -72, -76, 80, 84)
> prq.keys <- make.keys(prq,list(nach=nach,anx=anx,soc=soc,imp=imp,PeerNach=85,
                                PeerAnx=86, PeerSoc=87,PeerImp=88,gender=89))
```

By having the scoring key information in this form, we can always reproduce it.

We can also save it using dput

But the keys.list format is easiest to use.



## Saving the prq.keys so that we can use them later if we need to do SO.

```
dput(prq.keys)
dput(prq.keys)
structure(c(-1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0,
1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, -1, 0, 0, 0,
1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1,
0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, -1, 0, 0, 0, 1, 0, 0, 0, -1,
0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0,
-1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0,
1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1,
0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, -1, 0, 0, 0, -1,
0, 0, 0, -1, 0, 0, 0, -1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0,
0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, -1, 0, 0, 0, -1, 0, 0,
0, 1, 0, 0, 0, -1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0,
0, -1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0,
0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0,
1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, -1, 0, 0, 0, 1, 0,
0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0,
0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, -1, 0, 0, 0, -1, 0, 0,
0, -1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
```

## A keys matrix

```
> prq.keys
```

	nach	anx	soc	imp	PeerNach	PeerAnx	PeerSoc	PeerImp	gender
Q1	-1	0	0	0	0	0	0	0	0
Q2	0	1	0	0	0	0	0	0	0
Q3	0	0	1	0	0	0	0	0	0
Q4	0	0	0	1	0	0	0	0	0
Q5	1	0	0	0	0	0	0	0	0
Q6	0	1	0	0	0	0	0	0	0
Q7	0	0	1	0	0	0	0	0	0
Q8	0	0	0	-1	0	0	0	0	0
...									
Q81	1	0	0	0	0	0	0	0	0
Q82	0	1	0	0	0	0	0	0	0
Q83	0	0	1	0	0	0	0	0	0
Q84	0	0	0	1	0	0	0	0	0
N	0	0	0	0	1	0	0	0	0
A	0	0	0	0	0	1	0	0	0
S	0	0	0	0	0	0	1	0	0
I	0	0	0	0	0	0	0	1	0
G	0	0	0	0	0	0	0	0	1

## Score the items

We use the `scoreItems` function.

We first do this just for the items. The `item.scores` is a list of multiple values:

1. `scores` – the actual scores for each subject
2. `missing` – where there any missing values for any subject?
3. `alpha` – coefficient alpha for each scale
4. `av.r` – the average `r` within each scale
5. `n.items` – how many items in each scale?
6. `item.cor` – the correlation of each item with each scale
7. `cor` – the correlation matrix of the scales (based upon the correlations of the items - with SAPA data this will differ from correlating the scales)
8. `corrected` – the raw correlations of the scales (below the diagonal), the alpha reliabilities (on the diagonal), and the intercorrelations corrected for unreliability (above the diagonal).

## Using scoreItems

```
> prq.scores <- scoreItems(prq.keys, prq)
> prq.scores
```

(Unstandardized) Alpha:

	nach	anx	soc	imp	PeerNach	PeerAnx	PeerSoc	PeerImp	gender
alpha	0.8	0.85	0.85	0.86	1	1	1	1	1

Standard errors of unstandardized Alpha:

	nach	anx	soc	imp	PeerNach	PeerAnx	PeerSoc	PeerImp	gender
ASE	0.033	0.027	0.026	0.026	NaN	NaN	NaN	NaN	NaN

Average item correlation:

	nach	anx	soc	imp	PeerNach	PeerAnx	PeerSoc	PeerImp	gender
average.r	0.16	0.21	0.22	0.22	NaN	NaN	NaN	NaN	NaN

Guttman 6\* reliability:

	nach	anx	soc	imp	PeerNach	PeerAnx	PeerSoc	PeerImp	gender
Lambda.6	0.98	0.98	0.98	0.98	0.85	0.88	0.9	0.9	0.88

Signal/Noise based upon av.r :

	nach	anx	soc	imp	PeerNach	PeerAnx	PeerSoc	PeerImp	gender
Signal/Noise	4.1	5.6	5.8	5.9	NaN	NaN	NaN	NaN	NaN

## Show more of the output

```
> item.scores
```

Scale intercorrelations corrected for attenuation

raw correlations below the diagonal, alpha on the diagonal

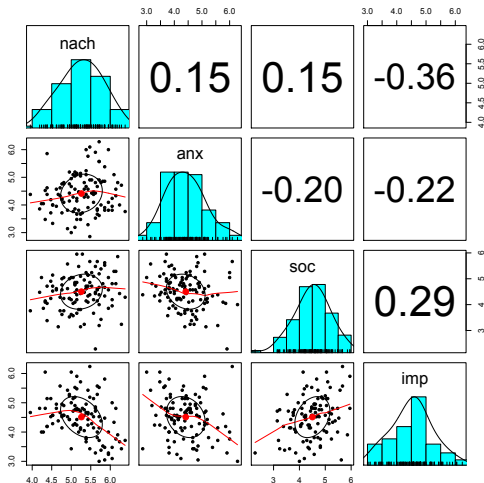
corrected correlations above the diagonal:

	nach	anx	soc	imp	PeerNach	PeerAnx	PeerSoc	PeerImp	gender
nach	0.803	0.178	0.182	-0.43	0.476	0.033	0.186	-0.041	0.149
anx	0.146	0.847	-0.231	-0.25	0.105	0.382	-0.192	-0.246	0.456
soc	0.151	-0.197	0.853	0.34	-0.161	-0.174	0.459	0.258	0.050
imp	-0.356	-0.217	0.294	0.86	-0.414	-0.331	0.135	0.477	-0.175
PeerNach	0.427	0.096	-0.149	-0.38	1.000	0.259	0.315	-0.016	-0.096
PeerAnx	0.029	0.351	-0.161	-0.31	0.259	1.000	-0.135	-0.161	0.028
PeerSoc	0.167	-0.176	0.424	0.12	0.315	-0.135	1.000	0.542	-0.024
PeerImp	-0.037	-0.227	0.238	0.44	-0.016	-0.161	0.542	1.000	-0.198
gender	0.134	0.419	0.046	-0.16	-0.096	0.028	-0.024	-0.198	1.000

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

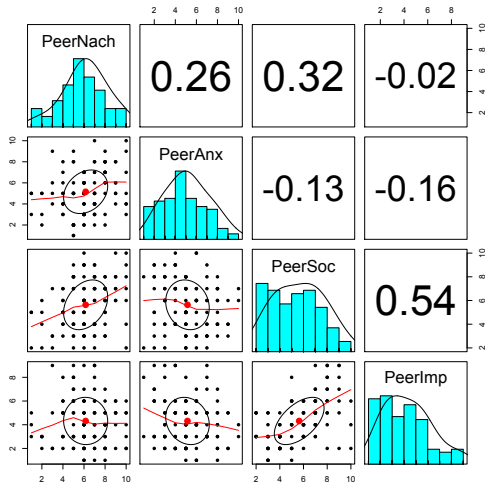
## Display the four self report dimensions

`pairs.panels(prq.scores$scores[,1:4])` # note that scores is an object in prq.scores



## Show the peer rating structure

`pairs.panels(prq.scores$scores[,5:8])`

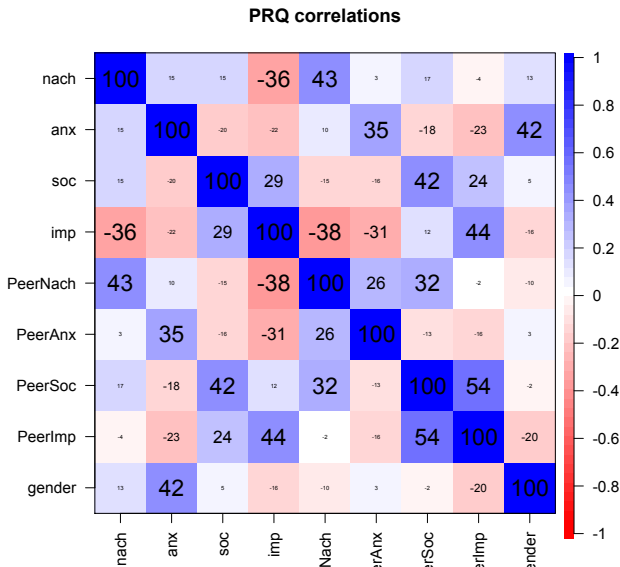


## The Multi-Trait- Multi- Method Matrix

- Correlations within method combine trait and method variance
  - What is the structure of NASI within self report
  - What is the structure of NASI within peer report
- Correlations across method show trait variance
  - Do the self report dimensions match the peer ratings?
  - Note the correlations of gender differ between self and peer report. What could account for this difference?



Show the MMTM matrix graphically – `cor.ci(prq.scores$scores)`



## Factor Analysis

The items analysed were meant to represent four constructs. Given the previous analysis, they probably do. But what if we did not know how many separate dimensions were in the data? Is it possible to find out? Three alternative procedure address this question.

1. Principal components analysis
2. Factor analysis
3. Cluster analysis

All three of these procedures are attempting to approximate the  $nvar * nvar$  correlation matrix  $R$  with a matrix of lesser rank, one that is  $nvar * nf$ . That is, can we find a Factor (Component or Cluster) such that

$$R \approx FF' + U^2 \quad (1)$$

or

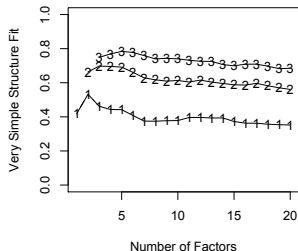
$$R \approx CC' \quad (2)$$

## Factor analysis of PRQ

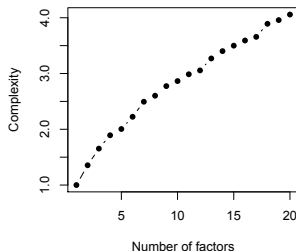
1. We need more people than items to make the matrix invertible
2. Can be solved in either case by using minimum residuals (OLS)
3. Can be solved by the `fa` function using `minres` option
4. How many factors to extract is a perpetual problem.
  - `nfactors(prq)`
  - Use VSS 2 ( complexity 1) or 3 (complexity 2)
  - Use MAPS 9
  - Empirical BIC 3 factors
5. Theory says 4

## VSS of prq

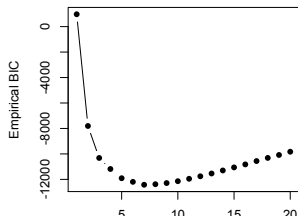
**Very Simple Structure**



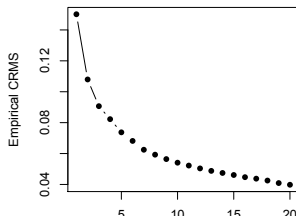
**Complexity**



**Empirical BIC**



**Root Mean Residual**



## Find a 4 factor as well as a 4 component solution – very similar

```
f4 <- fa(prq,4)
p4 <- principal(prq,4)
> factor.congruence(f4,p4)
```

	RC1	RC2	RC4	RC3
MR1	0.97	0.08	-0.22	0.28
MR2	0.11	0.99	0.19	-0.17
MR3	-0.35	0.21	0.98	-0.07
MR4	0.02	-0.12	-0.15	0.97

## Summary of the 4 factor solution

```
> summary(f4)
```

```
Factor analysis with Call: fa(r = prq, nfactors = 4)
```

```
Test of the hypothesis that 4 factors are sufficient.
```

```
The degrees of freedom for the model is 3566 and the objective function was 65.08
```

```
The number of observations was 110 with Chi Square = 4935.07 with prob < 5e-48
```

```
The root mean square of the residuals (RMSA) is 0.08
```

```
The df corrected root mean square of the residuals is 0.08
```

```
Tucker Lewis Index of factoring reliability = 0.54
```

```
RMSEA index = 0.095 and the 90 % confidence intervals are 0.055 0.063
```

```
BIC = -11826.85
```

```
With factor correlations of
```

	MR1	MR2	MR3	MR4
MR1	1.00	0.11	-0.24	0.14
MR2	0.11	1.00	0.14	-0.15
MR3	-0.24	0.14	1.00	-0.06
MR4	0.14	-0.15	-0.06	1.00

## Also try a cluster analysis

```
> ic <- iclust(prq)
> summary(ic)
```

```
ICLUST (Item Cluster Analysis)Call: iclust(r.mat = prq)
ICLUST
```

```
Purified Alpha:
  C84 C82 C81 C77 C23
0.93 0.91 0.80 0.73 0.51
```

```
Guttman Lambda6*
  C84 C82 C81 C77 C23
0.99 0.99 0.97 0.96 0.91
```

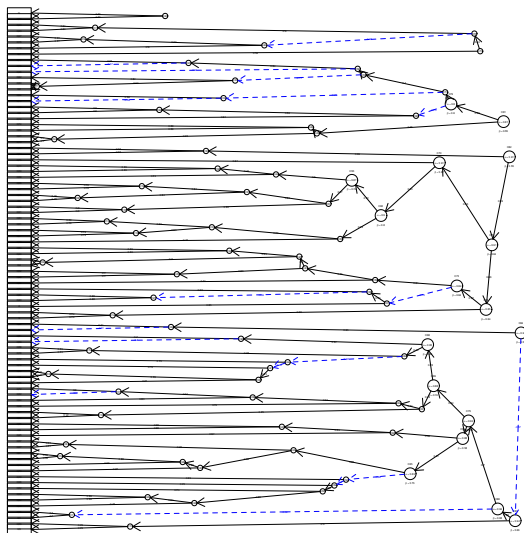
```
Original Beta:
  C84 C82 C81 C77 C23
0.53 0.55 0.55 0.39 0.51
```

```
Cluster size:
C84 C82 C81 C77 C23
 36 28 16 7 2
```

```
Purified scale intercorrelations
reliabilities on diagonal
correlations corrected for attenuation above diagonal:
      C84 C82 C81 C77 C23
C84  0.925 0.18 0.0559 0.3261 -0.34
C82  0.164 0.91 0.3016 0.2529 0.23
C81  0.048 0.26 0.8029 -0.0045 -0.33
C77  0.267 0.21 -0.0034 0.7271 0.10
```

## The cluster solution

ICLUST





## Compare the solutions

```
> factor.congruence(list(f4,p4,ic))
```

	MR1	MR2	MR3	MR4	RC1	RC2	RC4	RC3	C84	C82	C81	C77	C23
MR1	1.00	0.05	-0.17	0.07	0.97	0.08	-0.22	0.28	-0.86	0.01	0.07	-0.34	0.43
MR2	0.05	1.00	0.11	-0.07	0.11	0.99	0.19	-0.17	0.01	0.87	0.32	0.20	0.39
MR3	-0.17	0.11	1.00	-0.06	-0.35	0.21	0.98	-0.07	0.60	0.52	0.03	0.63	-0.08
MR4	0.07	-0.07	-0.06	1.00	0.02	-0.12	-0.15	0.97	-0.28	-0.31	-0.93	0.04	0.54
RC1	0.97	0.11	-0.35	0.02	1.00	0.12	-0.36	0.22	-0.91	0.00	0.12	-0.49	0.40
RC2	0.08	0.99	0.21	-0.12	0.12	1.00	0.26	-0.19	0.03	0.91	0.37	0.30	0.41
RC4	-0.22	0.19	0.98	-0.15	-0.36	0.26	1.00	-0.18	0.65	0.59	0.10	0.55	-0.14
RC3	0.28	-0.17	-0.07	0.97	0.22	-0.19	-0.18	1.00	-0.44	-0.36	-0.90	0.00	0.58
C84	-0.86	0.01	0.60	-0.28	-0.91	0.03	0.65	-0.44	1.00	0.27	0.14	0.52	-0.48
C82	0.01	0.87	0.52	-0.31	0.00	0.91	0.59	-0.36	0.27	1.00	0.47	0.38	0.19
C81	0.07	0.32	0.03	-0.93	0.12	0.37	0.10	-0.90	0.14	0.47	1.00	0.04	-0.31
C77	-0.34	0.20	0.63	0.04	-0.49	0.30	0.55	0.00	0.52	0.38	0.04	1.00	0.00
C23	0.43	0.39	-0.08	0.54	0.40	0.41	-0.14	0.58	-0.48	0.19	-0.31	0.00	1.00

```
>
```

## Combine the factor scores with the empirical scores

```
> scores.df <- data.frame(f4$scores, prq.scores$scores)
> lowerCor(scores.df)
```

	MR1	MR2	MR3	MR4	nach	anx	soc	imp	PrNch	PrAnx	PerSc	PrIm
MR1	1.00											
MR2	0.11	1.00										
MR3	-0.24	0.14	1.00									
MR4	0.14	-0.16	-0.06	1.00								
nach	0.89	0.25	-0.36	-0.06	1.00							
anx	0.32	-0.16	-0.17	0.91	0.15	1.00						
soc	0.06	0.91	0.27	-0.13	0.15	-0.20	1.00					
imp	-0.34	0.27	0.90	-0.16	-0.36	-0.22	0.29	1.00				
PeerNach	0.49	-0.14	-0.33	-0.06	0.43	0.10	-0.15	-0.38	1.00			
PeerAnx	0.16	-0.15	-0.27	0.41	0.03	0.35	-0.16	-0.31	0.26	1.00		
PeerSoc	0.19	0.41	0.17	-0.25	0.17	-0.18	0.42	0.12	0.32	-0.13	1.00	
PeerImp	0.00	0.22	0.47	-0.30	-0.04	-0.23	0.24	0.44	-0.02	-0.16	0.54	1.00
gender	0.21	0.09	-0.19	0.46	0.13	0.42	0.05	-0.16	-0.10	0.03	-0.02	-0.2

## Compare original, factors and clusters

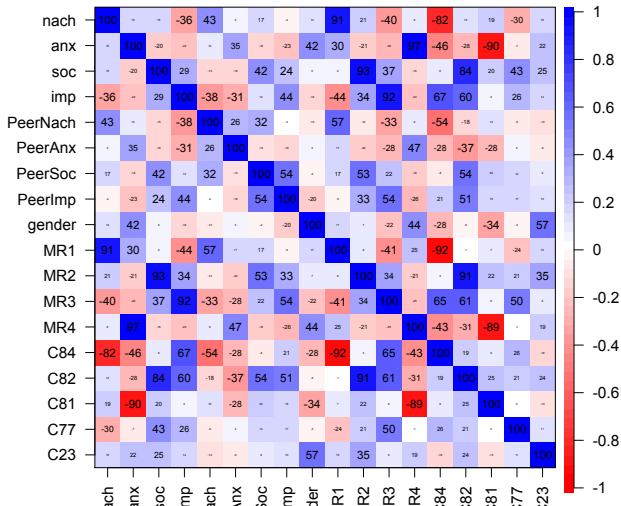
```
> fkeys <- factor2cluster(f4)
> ckeys <- cluster2keys(ic)
> all.keys <- cbind(prq.keys,fkeys,ckeys)
> all.scores <- scoreItems(all.keys,prq)
> lowerMat(all.scores$cor)
```

Coefficients and bootstrapped confidence intervals

	nach	anx	soc	imp	PrNch	PrAnx	PerSc	PrImp	gendr	MR1	MR2	MR3	MR4	C84
nach	1.00													
anx	0.15	1.00												
soc	0.15	-0.20	1.00											
imp	-0.36	-0.22	0.29	1.00										
PeerNach	0.43	0.10	-0.15	-0.38	1.00									
PeerAnx	0.03	0.35	-0.16	-0.31	0.26	1.00								
PeerSoc	0.17	-0.18	0.42	0.12	0.32	-0.13	1.00							
PeerImp	-0.04	-0.23	0.24	0.44	-0.02	-0.16	0.54	1.00						
gender	0.13	0.42	0.05	-0.16	-0.10	0.03	-0.02	-0.20	1.00					
MR1	0.91	0.30	0.02	-0.44	0.57	0.17	0.17	-0.05	0.17	1.00				
MR2	0.21	-0.21	0.93	0.34	-0.10	-0.20	0.53	0.33	0.07	0.08	1.00			
MR3	-0.40	-0.26	0.37	0.92	-0.33	-0.28	0.22	0.54	-0.22	-0.41	0.34	1.00		
MR4	0.08	0.97	-0.16	-0.23	0.09	0.47	-0.18	-0.26	0.44	0.25	-0.21	-0.25	1.00	
C84	-0.82	-0.46	0.06	0.67	-0.54	-0.28	-0.08	0.21	-0.28	-0.92	0.03	0.65	-0.43	1.00
C82	0.13	-0.28	0.84	0.60	-0.18	-0.37	0.54	0.51	-0.05	-0.01	0.91	0.61	-0.31	0.19
C81	0.19	-0.90	0.20	0.02	0.10	-0.28	0.20	0.16	-0.34	0.07	0.22	0.06	-0.89	0.09
C77	-0.30	-0.07	0.43	0.26	-0.09	0.02	0.14	0.17	-0.03	-0.24	0.21	0.50	-0.01	0.26
C23	0.16	0.22	0.25	0.14	-0.12	-0.06	0.09	0.10	0.57	0.15	0.35	0.06	0.19	-0.16
	C82	C81	C77	C23										
C82	1.00													
C81	0.25	1.00												
C77	0.21	0.00	1.00											
C23	0.24	-0.13	0.11	1.00										

## The correlations between rational keying, peer ratings factors and clusters

Correlation plot



## Best items sorted by factor loading Factor 1

```
> fa.lookup(f4,prq.dictionary)
```

	MR1	MR2	MR3	MR4	com	h2	
Q25	0.87	-0.02	-0.04	-0.02	1.01	0.76	content
Q73	0.81	0.01	0.00	-0.01	1.00	0.65	I am goal oriented
Q61	0.79	-0.05	-0.04	0.12	1.06	0.68	Success motivates me
Q17	0.77	-0.01	0.06	-0.01	1.01	0.58	I set goals for myself
Q37	0.65	0.07	-0.11	-0.04	1.09	0.47	I strive to be the best
Q53	0.60	0.04	0.17	-0.07	1.20	0.35	I push myself to succeed
Q62	0.58	-0.05	0.09	0.29	1.55	0.46	I am a competitive person
Q57	0.57	0.09	-0.23	-0.04	1.39	0.45	I feel uncomfortable when I do not have control over a situation
Q29	0.53	-0.23	0.13	0.25	1.94	0.40	I am a motivated person
Q33	0.51	-0.18	0.03	0.38	2.13	0.48	I get upset when I lose or do poorly
Q76	0.51	0.03	-0.29	-0.18	1.89	0.41	I am afraid of failure
Q21	0.51	0.13	0.04	-0.44	2.12	0.43	When I start a task I always finish it
Q42	0.49	-0.20	0.07	0.27	1.97	0.37	I am not afraid of difficult tasks
N	0.48	-0.19	-0.20	-0.17	1.96	0.33	I have trouble letting go of things
Q65	0.48	0.16	0.00	0.12	1.34	0.29	Nach
Q45	0.44	0.23	0.03	0.05	1.54	0.27	The reward often justifies the effort
Q81	0.43	0.20	-0.30	-0.04	2.30	0.37	I feel accomplished when I reach my goals
Q13	0.39	0.14	-0.03	-0.29	2.16	0.25	I believe that if something is worth doing, it is worth doing well.
Q68	0.38	0.15	-0.26	0.01	2.13	0.28	I like challenging tasks
Q36	-0.32	0.15	0.26	0.06	2.48	0.22	I know what I am doing next week
Q5	0.32	-0.16	-0.05	0.11	1.80	0.16	I sometimes switch goals with no real reason
Q66	0.29	0.19	-0.20	-0.27	3.58	0.25	It is important for me to outperform my peers
Q9	0.24	0.14	-0.07	0.01	1.82	0.09	I feel like I have control over my life
							I would prefer a moderately difficult task over an easy or hard one

## Factor 2

Q83	0.06	0.83	0.00	0.11	1.05	0.69	I am a very sociable person.
Q47	-0.07	0.78	0.07	0.05	1.04	0.61	I like to meet new people.
Q55	-0.10	0.76	-0.08	0.01	1.06	0.56	I can easily hold a conversation with a stranger.
Q35	-0.02	0.73	-0.15	0.02	1.09	0.52	I find it easy to make new friends.
Q51	-0.01	0.69	0.27	0.03	1.30	0.60	I like to be around groups of people.
Q79	0.00	0.66	-0.06	-0.34	1.50	0.62	I am relaxed and confident around others.
Q31	-0.17	0.65	-0.14	-0.25	1.55	0.54	I am relaxed when meeting new people.
Q39	0.00	0.65	-0.14	-0.10	1.14	0.45	Expressing myself to others comes naturally.
Q67	0.13	0.64	0.19	0.21	1.49	0.51	I often crave interaction with other people.
Q3	0.00	0.62	-0.21	-0.05	1.23	0.40	I say hello to acquaintances on the street.
Q11	0.14	0.59	0.16	0.03	1.29	0.42	I am talkative.
Q71	-0.02	0.58	0.27	0.14	1.55	0.45	I prefer interacting with others to spending time by myself.
Q75	-0.16	0.43	0.17	0.04	1.63	0.26	I prefer working with others to working alone.
Q59	0.30	0.38	0.15	0.08	2.32	0.28	I am often the first person to speak during a conversation.
S	0.22	0.33	0.17	-0.22	3.13	0.26	Sociability
Q60	0.15	0.30	0.14	0.21	2.81	0.17	When shopping I find my spending money on things I never planned to
Q80	-0.16	0.28	0.25	-0.25	3.57	0.30	I like not knowing what comes next
Q64	0.01	0.21	0.06	-0.18	2.16	0.10	Do you often switch lanes when you are driving?

## Factor 3

Q84	0.17	0.12	0.73	0.00	1.17	0.55	I am an impulsive person.
Q72	0.19	0.19	-0.64	-0.07	1.40	0.51	I think before I act
Q41	0.18	0.16	-0.63	-0.03	1.31	0.48	I consider myself to be a perfectionist
Q32	-0.22	-0.01	0.62	0.13	1.34	0.50	I often act without thinking
Q4	0.16	0.37	0.62	-0.06	1.79	0.58	I like to do things spur of the moment
Q44	0.29	-0.10	0.57	-0.17	1.77	0.35	I want to try sky-diving
Q16	0.14	0.20	-0.56	0.02	1.38	0.38	I plan out my actions in detail
Q20	0.10	0.34	0.52	-0.04	1.82	0.44	I like making decisions on the spur of the moment
Q34	0.21	-0.07	-0.51	0.30	2.08	0.50	Others would describe me as uptight or high strung
Q48	0.02	0.24	0.50	-0.05	1.47	0.35	I find myself doing things I had not planned to do that day
Q24	0.27	0.24	-0.50	0.02	2.05	0.43	I consider all of my options before making a decision
Q40	0.09	0.25	0.49	-0.22	2.00	0.40	I like to take risks
Q56	-0.17	0.21	0.49	0.03	1.65	0.36	Do you go on unplanned trips or excursions
I	0.13	0.10	0.47	-0.27	1.89	0.32	Impulsivity
Q8	0.30	0.07	-0.45	-0.06	1.84	0.36	I like to plan out my day
Q28	0.00	-0.05	-0.42	0.02	1.03	0.19	I analyze my thoughts before saying them out loud?
Q63	-0.05	0.38	0.41	0.06	2.07	0.36	I enjoy being in a crowded area.
Q52	0.08	0.26	0.39	-0.18	2.28	0.30	I enjoy surprises?
Q69	-0.09	-0.12	0.39	-0.04	1.32	0.19	I frequently cheat to succeed
Q12	-0.31	-0.17	0.39	-0.10	2.51	0.35	I rarely plan for the future
Q43	-0.12	0.17	0.39	0.21	2.21	0.25	I would prefer to have many friends rather than a few close ones.
Q23	0.12	0.00	0.32	-0.01	1.29	0.10	I like to be the center of attention.
Q1	-0.26	0.01	0.31	0.20	2.71	0.22	I give up easily
Q15	0.14	-0.04	-0.21	-0.14	2.65	0.09	I like quiet time alone.
Q27	0.17	-0.15	-0.19	0.12	3.62	0.13	I need time to recharge after spending time with others.

## Factor 4

Q2	0.08	0.13	-0.06	0.70	1.11	0.52	I have difficulty stopping myself from worrying
Q22	0.14	-0.03	-0.03	0.68	1.09	0.52	I often feel nervous or on edge
Q30	0.03	0.10	-0.15	0.63	1.18	0.42	Sometimes, I am so worried, I can not focus
Q6	-0.05	0.05	0.21	0.53	1.34	0.30	I often fidget or bite my nails
Q54	0.06	-0.05	-0.06	0.51	1.08	0.30	I often feel threatened or judged by other people
Q70	0.31	0.27	0.08	-0.50	2.37	0.44	I am confident in my abilities
Q50	-0.21	0.17	0.19	0.48	2.00	0.30	I have trouble concentrating on difficult tasks
G	0.09	0.18	-0.16	0.47	1.64	0.28	Gender
Q26	0.38	-0.09	0.15	0.44	2.30	0.38	I often worry that my life will not turn out as planned
Q18	0.39	0.05	-0.37	0.42	3.00	0.60	I often fret over details for future plans
Q14	0.17	-0.10	0.21	0.42	1.97	0.26	I anticipate the worst outcome of a situation
Q78	0.19	0.14	0.02	-0.42	1.63	0.23	I usually think things will work out
Q58	0.21	0.25	-0.34	0.42	3.16	0.43	I overthink details
Q74	0.20	0.05	-0.09	-0.40	1.66	0.20	I work well under pressure
Q77	0.34	-0.12	-0.09	0.38	2.32	0.35	Avoiding failure motivates me
Q82	-0.16	0.31	0.07	0.38	2.43	0.22	I am more emotional than my friends
A	0.07	-0.07	-0.22	0.37	1.81	0.23	Anxiety
Q10	-0.31	0.19	0.17	-0.35	3.08	0.36	It is easy for me to relax
Q46	-0.04	-0.05	-0.20	0.32	1.75	0.15	I get nervous before speaking in public
Q49	0.15	0.25	-0.06	-0.31	2.49	0.20	Failure is a sign to try again
Q38	0.22	-0.06	-0.02	0.27	2.03	0.15	I often feel restlessness or insomnia
Q7	0.12	0.16	0.12	0.25	2.69	0.11	Many of my goals involve other people.



## Show the items for the clusters

*fa.lookup(ic,prq.dictionary)*

	C84	C82	C81	C77	C23		content	scal
Q25	-0.77	-0.01	0.09	-0.13	0.19		I am goal oriented	
Q61	-0.75	-0.06	-0.04	-0.14	0.24		I set goals for myself	
Q73	-0.71	0.05	0.06	-0.23	0.24		Success motivates me	
Q18	-0.70	-0.20	-0.34	-0.21	0.33		I often fret over details for future plans	
Q17	-0.66	0.07	0.10	-0.02	0.18		I strive to be the best	
Q57	-0.63	-0.02	0.14	-0.20	0.13		I am a motivated person	
Q37	-0.63	0.05	0.07	-0.27	0.01		I push myself to succeed	
Q76	-0.61	0.00	0.22	-0.03	0.13		When I start a task I always finish it	
Q34	-0.58	-0.37	-0.20	-0.07	0.32		Others would describe me as uptight or high strung	
Q62	-0.58	-0.05	-0.21	-0.02	0.26		I feel uncomfortable when I do not have control over a situation	
Q33	-0.55	-0.22	-0.28	-0.03	0.22		I am afraid of failure	
Q12	0.54	0.02	0.00	0.19	-0.25		I rarely plan for the future	
Q81	-0.54	0.08	0.02	-0.39	0.16	I believe that if something is worth doing, it is worth doing well.		
Q72	-0.54	-0.11	0.06	-0.20	0.07		I think before I act	
Q32	0.54	0.27	-0.12	0.33	0.09		I often act without thinking	
Q24	-0.53	-0.03	0.07	-0.26	0.08		I consider all of my options before making a decision	
Q8	-0.53	-0.09	0.10	-0.18	0.29		I like to plan out my day	
Q41	-0.53	-0.14	0.12	-0.18	0.24		I consider myself to be a perfectionist	
N	-0.52	-0.18	0.12	-0.09	-0.13		Nach	
Q58	-0.51	-0.05	-0.28	-0.36	0.25		I overthink details	
Q68	-0.51	0.01	0.15	-0.13	0.17		I know what I am doing next week	
Q10	0.50	0.33	0.35	-0.04	-0.15		It is easy for me to relax	
Q42	-0.50	-0.18	-0.28	-0.09	0.08		I have trouble letting go of things	
Q16	-0.49	-0.07	0.03	-0.29	0.13		I plan out my actions in detail	
Q65	-0.48	0.13	0.05	-0.07	0.32		The reward often justifies the effort.	
Q77	-0.48	-0.26	-0.22	-0.02	0.20		Avoiding failure motivates me	
Q29	-0.47	-0.21	-0.19	-0.01	0.13		I get upset when I lose or do poorly	
Q36	0.46	0.20	-0.08	0.11	0.00		I sometimes switch goals with no real reason	
Q56	0.46	0.37	0.01	0.21	-0.06		Do you go on unplanned trips or excursions	

## Cluster 2

IQ83	-0.12	0.71	0.08	0.07	0.24
Q47	0.05	0.70	0.15	0.17	0.15
Q51	0.16	0.70	0.13	0.30	0.25
Q4	0.25	0.66	0.08	0.02	0.02
Q79	0.04	0.66	0.40	-0.08	0.03
Q55	0.02	0.65	0.10	0.00	0.13
Q31	0.09	0.62	0.31	0.14	0.01
Q11	-0.04	0.62	0.17	0.21	0.27
Q35	-0.08	0.59	0.09	-0.03	0.03
Q20	0.22	0.58	0.06	0.10	0.08
Q71	0.10	0.58	0.08	0.45	0.13
Q40	0.24	0.58	0.15	0.23	-0.06
Q67	-0.08	0.56	0.05	0.18	0.29
Q39	-0.07	0.55	0.25	0.02	0.04
Q84	0.28	0.53	-0.04	0.16	0.16
Q52	0.17	0.53	0.13	0.00	-0.13
Q63	0.23	0.52	0.06	0.38	0.07
S	-0.07	0.50	0.20	0.14	0.05
Q48	0.28	0.49	0.05	0.09	0.05
Q3	-0.12	0.48	0.17	-0.12	0.10
Q80	0.36	0.47	0.19	0.01	-0.13
I	0.21	0.46	0.15	0.18	-0.01
Q75	0.22	0.46	0.07	0.25	0.18
Q59	-0.23	0.44	0.07	0.35	0.19
Q64	0.06	0.36	0.14	0.04	0.15
Q49	-0.08	0.34	0.28	-0.12	0.10
A	-0.28	-0.34	-0.27	0.02	-0.03
Q44	0.12	0.30	0.10	0.06	0.03

I am a very sociable person.  
 I like to meet new people.  
 I like to be around groups of people.  
 I like to do things spur of the moment  
 I am relaxed and confident around others.  
 I can easily hold a conversation with a stranger.  
 I am relaxed when meeting new people.  
 I am talkative.  
 I find it easy to make new friends.  
 I like making decisions on the spur of the moment  
 I prefer interacting with others to spending time by myself.  
 I like to take risks  
 I often crave interaction with other people.  
 Expressing myself to others comes naturally.  
 I am an impulsive person.  
 I enjoy surprises?  
 I enjoy being in a crowded area.  
 Sociability  
 I find myself doing things I had not planned to do that day  
 I say hello to acquaintances on the street.  
 I like not knowing what comes next  
 Impulsivity  
 I prefer working with others to working alone.  
 I am often the first person to speak during a conversation.  
 Do you often switch lanes when you are driving?  
 Failure is a sign to try again  
 Anxiety  
 I want to try sky-diving

## Cluster 3

Q2	-0.31	-0.11	-0.66	-0.06	0.24
Q22	-0.32	-0.21	-0.65	0.05	0.19
Q30	-0.31	-0.14	-0.63	-0.10	0.27
Q70	-0.10	0.39	0.62	0.03	-0.16
Q78	-0.02	0.22	0.56	0.01	0.10
Q6	0.03	-0.01	-0.52	0.01	0.22
Q21	-0.30	0.26	0.51	0.00	0.00
Q54	-0.24	-0.23	-0.49	-0.04	0.14
Q74	-0.11	0.10	0.48	-0.23	-0.07
Q14	-0.13	-0.07	-0.47	-0.07	-0.09
Q50	0.15	0.14	-0.47	0.14	0.39
Q66	-0.30	0.14	0.45	-0.01	0.03
Q46	-0.14	-0.20	-0.41	-0.17	0.26
Q13	-0.24	0.20	0.38	-0.08	0.13
Q38	-0.28	-0.09	-0.34	-0.05	0.14
Q82	0.09	0.16	-0.29	0.03	0.07

I have difficulty stopping myself from worrying  
     I often feel nervous or on edge  
 Sometimes, I am so worried, I can not focus  
     I am confident in my abilities  
     I usually think things will work out  
     I often fidget or bite my nails  
     I am not afraid of difficult tasks  
 I often feel threatened or judged by other people  
     I work well under pressure  
     I anticipate the worst outcome of a situation  
 I have trouble concentrating on difficult tasks  
     I feel like I have control over my life  
     I get nervous before speaking in public  
     I like challenging tasks  
     I often feel restlessness or insomnia  
     I am more emotional than my friends

## Cluster 4

Q15	-0.17	-0.08	0.00	-0.72	-0.06	I like quiet time alone.
Q43	0.26	0.23	-0.08	0.69	0.16	I would prefer to have many friends rather than a few close ones.
Q19	-0.21	0.04	0.04	-0.68	-0.07	I can have fun alone.
Q27	-0.27	-0.25	-0.17	-0.65	0.10	I need time to recharge after spending time with others.
Q69	0.31	0.09	0.01	0.55	-0.06	I frequently cheat to succeed
Q23	0.06	0.15	0.06	0.53	0.00	I like to be the center of attention.
Q7	-0.12	0.12	-0.15	0.39	0.14	Many of my goals involve other people.

## Cluster 5

G	-0.28	-0.04	-0.32	-0.03	0.78		Gender
Q60	-0.10	0.29	-0.03	0.13	0.78	When shopping I find my spending money on things I never planned to	

## Empirical scale construction

1. Identify those items that most correlate with the criteria
  - Form item composites based upon those items
2. `best.scales` will do this
  - `bs <- best.scales(prq,c("N","A","I","S"),dictionary=prq.dictionary)`

## Empirical 1

The items most correlated with the criteria yield r's of  
correlation n.items

N	0.58	9
A	0.55	10
I	0.53	9
S	0.48	8

The best items, their correlations and content are  
\$N

	N		content scale
Q25	0.42	I am goal oriented	N
Q73	0.41	Success motivates me	N
Q76	0.38	When I start a task I always finish it	I
Q72	0.37	I think before I act	I
Q82	-0.37	I am more emotional than my friends	A
Q17	0.35	I strive to be the best	N
Q34	0.35	Others would describe me as uptight or high strung	A
Q61	0.34	I set goals for myself	N
Q10	-0.32	It is easy for me to relax	A

## Empirical 2

\$A

	A		content scale
Q22	0.35	I often feel nervous or on edge	A
Q40	-0.33	I like to take risks	I
Q34	0.33	Others would describe me as uptight or high strung	A
Q58	0.32	I overthink details	A
Q18	0.31	I often fret over details for future plans	A
Q2	0.31	I have difficulty stopping myself from worrying	A
Q84	-0.31	I am an impulsive person.	I
Q33	0.30	I am afraid of failure	N
Q49	-0.29	Failure is a sign to try again	N
Q62	0.29	I feel uncomfortable when I do not have control over a situation	A



## Empirical 3

\$I		I		content scale
Q84	0.41		I am an impulsive person.	I
Q4	0.37		I like to do things spur of the moment	I
Q44	0.36		I want to try sky-diving	I
Q40	0.36		I like to take risks	I
Q48	0.34	I find myself doing things I had not planned to do that day		I
Q32	0.33		I often act without thinking	I
Q52	0.31		I enjoy surprises?	I
Q20	0.31	I like making decisions on the spur of the moment		I
Q77	-0.29		Avoiding failure motivates me	N

## Empirical 4

\$S

S content scale

Q79	0.42	I am relaxed and confident around others.	S
Q31	0.36	I am relaxed when meeting new people.	S
Q47	0.34	I like to meet new people.	S
Q11	0.33	I am talkative.	S
Q35	0.33	I find it easy to make new friends.	S
Q51	0.33	I like to be around groups of people.	S
Q52	0.30	I enjoy surprises?	I
Q83	0.30	I am a very sociable person.	S

## Multiple ways to construct scales

### 1. Rational/Theoretical

- Learn Theory
- Write good items

### 2. Homogeneous keying

- Write good items
- Factor/Cluster analyze

### 3. Empirical Keys

- Write good items
- Select those items that correlate with the criteria

## Reliability of various ways of scoring

```
> mixed.key <- cbind(bs$key, prq.keys)
> mixed <- scoreItems(mixed.key, prq)
> mixed
```

```
> mixed
Call: scoreItems(keys = mixed.key, items = prq)
```

(Unstandardized) Alpha:

	N	A	S	I	nach	anx	soc	imp	PeerNach	PeerAnx	PeerSoc	PeerImp	gender
alpha	0.8	0.77	0.86	0.82	0.8	0.85	0.85	0.86	1	1	1	1	1

Standard errors of unstandardized Alpha:

	N	A	S	I	nach	anx	soc	imp	PeerNach	PeerAnx	PeerSoc	PeerImp	gender
ASE	0.044	0.046	0.038	0.041	0.033	0.027	0.026	0.026	NaN	NaN	NaN	NaN	NaN

Average item correlation:

	N	A	S	I	nach	anx	soc	imp	PeerNach	PeerAnx	PeerSoc	PeerImp	gender
average.r	0.3	0.25	0.42	0.34	0.16	0.21	0.22	0.22	NaN	NaN	NaN	NaN	NaN

Guttman 6\* reliability:

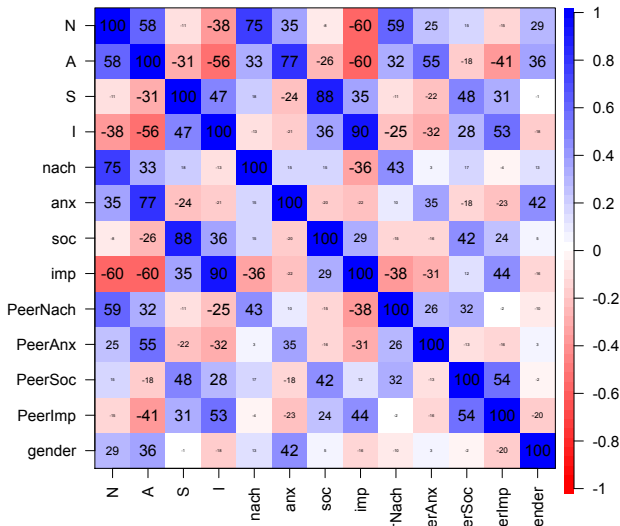
	N	A	S	I	nach	anx	soc	imp	PeerNach	PeerAnx	PeerSoc	PeerImp	gender
Lambda.6	0.98	0.97	0.98	0.98	0.98	0.98	0.98	0.98	0.85	0.88	0.9	0.9	0.88

Signal/Noise based upon av.r :

	N	A	S	I	nach	anx	soc	imp	PeerNach	PeerAnx	PeerSoc	PeerImp	gender
Signal/Noise	3.9	3.4	5.9	4.5	4.1	5.6	5.8	5.9	NaN	NaN	NaN	NaN	NaN

## Show the MMTM matrix graphically – cor.ci(mixed\$scores)

Empirical, rational and peer ratings



## 10 steps: Reprise

1. Specify your theory of relevant constructs
2. Define the population of interest
3. Give items to engaged subjects
4. Enter the data (carefully)
5. Descriptives to double check data entry and subject engagement
6. Find the variance/covariance matrix
7. Reduce its dimensionality through FA, PC, or clustering
8. Score composites (classical or IRT based)
9. Discriminant validity versus other constructs
10. Convergent validity with similar constructs and different methods



## Methods of scale construction

1. Empirical
  - MMPI
  - Strong Vocational Interest Blank
2. Rational
  - California Psychological Inventory
3. Theoretical
  - Measures of Need Achievement (e.g., Jackson PI)
4. Homegeneous keying
  - Eysenck Personality Inventory
  - NEO
  - BFI
  - TIPI



## Empirical

1. Ask items that discriminate known groups
  - People in general versus specific group
  - Choose items that are maximally independent and that have highest validities
2. Example:
  - MMPI
  - Strong-Campbell
  - sex and ethnic differences in personality and music
3. Problem:
  - What is the meaning of the scale?
  - Need to develop new scale for every new group





○○○○○

○○○○○○

## Sex differences at item level

Item	effect size
Get overwhelmed by emotions.	0.59
Sympathize with others' feelings.	0.45
Worry about things.	0.43
Feel others' emotions.	0.39
Get stressed out easily.	0.51
Have a soft heart.	0.38
Panic easily	0.50
Inquire about others' well-being.	0.41
Get upset by unpleasant thoughts that come into my mind.	0.38
Get upset easily.	0.37
Am indifferent to the feelings of others.	-0.33
Am not interested in other people's problems.	-0.35
Feel little concern for others.	-0.35
Am not easily bothered by things	0.34
Love to help others.	-0.32
Am not really interested in others.	0.30
Think of others first.	0.29
Take offense easily.	0.33
Take time out for others.	



## Sex differences and music preference

effect size	Item
0.9	Broadway Musicals (e.g. Rent, Cats, Phantom of the Opera)
0.68	Top 40/Pop Vocal Music (e.g. Kelly Clarkson, Madonna, The Black Eyed Peas)
0.65	Broadway, Movie and TV Soundtrack Music in General
0.59	Contemporary Rhythm and Blues (e. g. Whitney Houston, Usher, Alicia Keys)
0.59	Modern Country Music (e.g. Garth Brooks, Dixie Chicks, Tim McGraw)
0.37	Country Music in General
0.37	Movie Soundtracks (e.g. Starwars, Good Will Hunting, Garden State)
0.36	Top 40 Music/Pop in General
0.32	Pop Rock (e.g. Maroon 5, Counting Crows, John Mayer)
0.31	Modern Religious Music (e.g. 4Him, Casting Crowns)
0.3	Soul Rock (e.g. Stevie Wonder, Earth Wind and Fire)
-0.3	Acid Rock (e.g. Pink Floyd, The Doors, Jefferson Airplane)
-0.4	Heavy Metal (e.g. Metallica, Marilyn Manson, System of a Down)



○○○○○

○○○○○○

## Ethnic differences and music preference

effect size	Item
1.26	Acid Rock (e.g. Pink Floyd, The Doors, Jefferson Airplane)
1	Alternative (e.g. Pearl Jam, Incubus, Radiohead)
0.97	Electronic Music in General
0.91	Rock Music In General
0.87	Jam Bands (e.g. The Grateful Dead, Phish, String Cheese Incident)
0.87	Classic Rock (e.g. The Beatles, The Rolling Stones, Led Zeppelin)
0.85	Country Rock (e.g. The Allman Brothers, Lynyrd Skynyrd)
0.61	Electronic Dance Music (e.g. DJ Tiesto, Paul Van Dyk, Keoki)
0.59	Folk Music in General (e.g. Bob Dylan, Iron and Wine, Simon and Garfunkel)
0.57	Pop Rock (e.g. Maroon 5, Counting Crows, John Mayer)
0.56	Country Music in General
0.51	Bluegrass (e.g. Alison Krauss, Lester Flatt, Nickel Creek)
-0.56	Contemporary Rhythm and Blues (e. g. Whitney Houston, Usher, Alicia Keys)
-0.6	Blues in General (e.g. Ray Charles, Stevie Ray Vaughn, B.B. King)
-0.63	Instrumental Hip-Hop (e.g. DJ Hi-Tek, RJD2, Prefuse 73)
-0.64	Gospel Soul (e.g. Aretha Franklin, Solomon Burke)
-0.67	Soul in General (e.g. Otis Redding, Marvin Gaye)
-0.84	Religious Music in General
-1.04	Soul Rock (e.g. Stevie Wonder, Earth Wind and Fire)
-1.11	Rhythm and Blues in General
-1.43	Religious Gospel (e.g. Andre Crouch, Gospel Quartet)



## Rational Keying

1. Ask items with direct content relevance
2. Example: California Psychological Inventory
3. Problems
  - Not all items predict in obvious way
  - Need evidence for validity
  - Easy to fake



## Theoretical Keying

1. Ask items with theoretical relevance
2. Example: Jackson Personality Research Form
3. Problems:
  - Theoretical circularity
  - Need evidence for validity



## Homogeneous Keying

1. Select items to represent single domain
2. Exclude items based upon internal consistency
3. Examples:
  - 16PF
  - EPI/EPQ,
  - NEO/NEO-PIR
4. Problems
  - Garbage In, Garbage Out
  - Need evidence for validity



## Methods of Homogeneous keying

1. Cluster analysis (e.g. iclust)
2. Principal Components analysis (e.g., pca)
3. Factor analysis (e.g., fa)



## The Hase and Goldberg and Goldberg studies

1. Hase and Goldberg: a direct comparison of different techniques
  - Differential validity of scale construction
  - Factor analytic
  - Empirical Group discrimination
  - Intuitive theoretical
  - Intuitive rational
  - Stylistic-psychometric
  - Random
2. 200 University Freshman women
3. CPI items and 13 criteria



## Hase and Goldberg: 13 criteria

1. Sorority Membership
2. An experimental measure of conformity
3. Peer ratings of
  - Dominance
  - Sociability
  - Responsibility
  - Psychological Mindedness
  - Femininity
4. Peer ratings of how well known the person is
5. Average number of dates per month
6. College Grade Point Average
7. College Achievement relative to ability
8. College Major
9. College Dropout



## Does it make a difference?

1. Hase and Goldberg ([Hase & Goldberg, 1967](#)) No
2. [Goldberg \(1972\)](#) YES



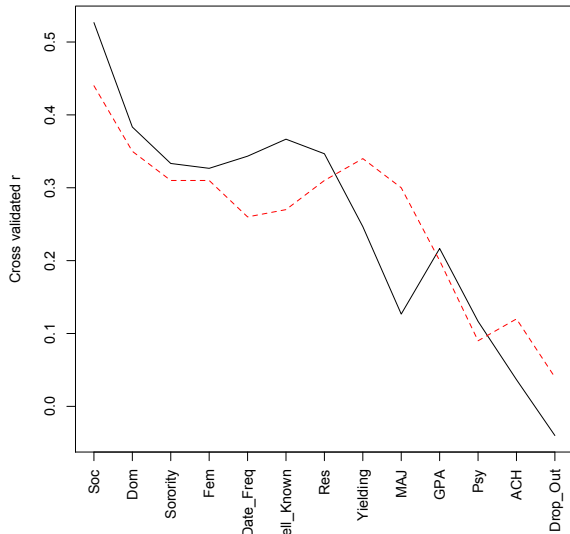
## Hase and Goldberg; mean values)

Original Hase and Goldberg showed no difference between methods, except that stylistic and random were much worse.

	var	n	mean	sd	median	trimmed	mad	min	max	range	se
Factor	1	13	0.25	0.18	0.27	0.25	0.13	-0.05	0.57	0.62	0.05
Theoretical	2	13	0.25	0.16	0.26	0.25	0.18	0.01	0.52	0.51	0.04
Rational	3	13	0.26	0.16	0.32	0.27	0.09	-0.08	0.49	0.57	0.04
Empirical	4	13	0.26	0.11	0.30	0.26	0.06	0.04	0.44	0.40	0.03
Stylistic	5	13	0.13	0.12	0.11	0.13	0.12	-0.07	0.35	0.42	0.03
Random	6	13	0.10	0.12	0.11	0.10	0.13	-0.08	0.30	0.38	0.03

## Prediction depends upon criteria: Goldberg: 72

Hase and Goldberg



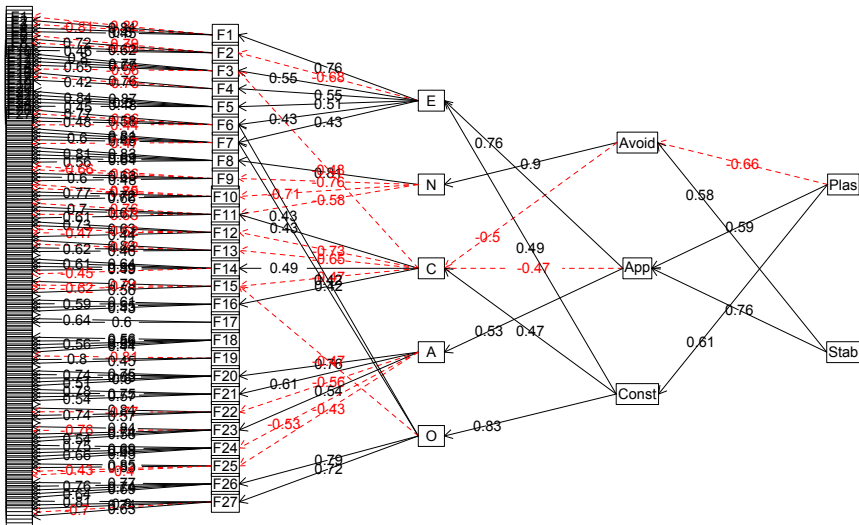


## Another factorial versus empirical example

1. Sapa Personality Inventory best 135 item ([Condon \(2017\)](#))
  - From 1800 IPIP items, found that 696 were most common
  - Factor structure of these 696 showed 135 very clear items
  - 5/27 factors, but not hierarchically organized
2. 4,000 subjects on spi 135 in the *psych* package
3. 135 items plus 10 criteria variables

# Applying the 'Bass Ackward' function

**BassAckward**



○  
○○○○  
○○○○

○○○○○

○○●○○○

## Score the Big 5 and predict the criteria

R code

```
spi.scales <- scoreItems(spi.keys[1:5],spi)

cor2(spi[1:10],spi.scales$scores)
```

	Agree	Consc	Neuro	Extra	Open
age	0.18	0.19	-0.17	-0.02	0.13
sex	0.17	0.09	0.24	0.06	-0.15
health	0.11	0.23	-0.34	0.21	0.07
p1edu	0.02	-0.02	-0.05	0.06	0.07
p2edu	0.02	-0.04	-0.04	0.08	0.07
education	0.13	0.12	-0.17	-0.01	0.15
wellness	0.11	0.12	-0.02	0.11	0.01
exer	0.07	0.19	-0.18	0.13	0.10
smoke	-0.09	-0.11	0.06	0.06	0.09
ER	-0.03	-0.01	0.12	0.02	-0.02

○  
○○○○  
○○○○

○○○○○

○○○●○○

## What about multiple R

R code

```
summary(setCor(1:10,11:15,data=spi.scores.df,plot=FALSE))
```

```
summary(setCor(1:10,11:15,data=spi.scores.df,plot=FALSE))
```

Multiple Regression from raw data

```
setCor(y = 1:10, x = 11:15, data = spi.scores.df, plot = FALSE)
```

Multiple Regression from matrix input

Beta weights

	age	sex	health	p1edu	p2edu	education	wellness	exer	smoke	ER
Agree	0.16	0.162	0.0063	0.015	0.014	0.116	0.0631	-0.0053	-0.083	-0.025
Consc	0.13	0.103	0.1715	-0.034	-0.049	0.065	0.1053	0.1613	-0.082	0.016
Neuro	-0.14	0.286	-0.2721	-0.036	-0.033	-0.147	0.0302	-0.1247	0.058	0.131
Extra	-0.11	0.086	0.1436	0.047	0.061	-0.086	0.0918	0.0876	0.084	0.050
Open	0.12	-0.122	0.0126	0.058	0.057	0.142	0.0031	0.0675	0.090	-0.012

Multiple R

	age	sex	health	p1edu	p2edu	education	wellness	exer	smoke	ER
	0.306	0.360	0.405	0.098	0.109	0.264	0.170	0.267	0.181	0.133

Multiple R2

	age	sex	health	p1edu	p2edu	education	wellness	exer	smoke	ER
	0.0939	0.1296	0.1642	0.0096	0.0118	0.0699	0.0288	0.0711	0.0329	0.0176

Cohen's set correlation R2

```
[1] 0.4
```

Squared Canonical Correlations

```
[1] 0.2394 0.1332 0.0620 0.0298 0.0079
```



## Compare with best scales

R code

```
bs <- bestScales(spi[11:145], spi[1:10], dictionary=spi.dictionary, n.iter=20)
```

```
Call = bestScales(x = spi[11:145], criteria = spi[1:10], dictionary = spi.dictionary,  
  n.iter = 20)
```

	derivation.mean	derivation.sd	validation.m	validation.sd	final.valid
age	0.37	0.014	0.360	0.021	0.35
sex	0.36	0.014	0.354	0.021	0.35
health	0.44	0.016	0.432	0.017	0.43
p1edu	0.15	0.030	0.124	0.026	NA
p2edu	0.17	0.027	0.098	0.024	NA
education	0.32	0.022	0.285	0.026	0.18
wellness	0.25	0.014	0.213	0.026	0.22
exer	0.32	0.018	0.283	0.023	0.30
smoke	0.28	0.016	0.255	0.024	0.27
ER	0.17	0.025	0.127	0.025	0.12

Repeat from setCor:

Multiple R

age	sex	health	p1edu	p2edu	education	wellness	exer	smoke	ER
0.306	0.360	0.405	0.098	0.109	0.264	0.170	0.267	0.181	0.133

o  
oooo  
oooo

ooooo

ooooo●

## What are the items?

Criterion = age

	Freq	mean.r	sd.r	item_id	item	item_scale	resp_type
q_4296	20	-0.23	0.01	q_4296	Tell a lot of lies.	EPQ:P	reg
q_4249	20	-0.21	0.02	q_4249	Would call myself a nervous person.	EPQ:N	reg
q_501	20	-0.21	0.01	q_501	Cheat to get ahead.	IPIP	reg
q_1024	18	-0.21	0.01	q_1024	Hang around doing nothing.	IPIP	reg
q_803	19	0.20	0.02	q_803	Express myself easily.	IPIP	reg
q_1081	18	-0.20	0.01	q_1081	Have difficulty expressing my feelings.	IPIP	reg

Criterion = sex

	Freq	mean.r	sd.r	item_id	item	item_scale	resp_type
q_1505	20	0.29	0.01	q_1505	Panic easily.	IPIP	re
q_979	20	0.29	0.01	q_979	Get overwhelmed by emotions.	IPIP	re
q_793	20	0.25	0.01	q_793	Experience my emotions intensely.	IPIP	re
q_174	20	-0.24	0.01	q_174	Am not easily affected by my emotions.	IPIP	re
q_1989	18	0.21	0.01	q_1989	Worry about things.	IPIP	re
q_851	19	0.21	0.01	q_851	Feel sympathy for those who are worse off than myself.	IPIP	re
q_1763	18	0.21	0.02	q_1763	Sympathize with others feelings.	IPIP	re
q_4252	18	0.20	0.01	q_4252	Am a worrier.	EPQ:N	re

Criterion = health

	Freq	mean.r	sd.r	item_id	item	item_scale	resp_type
q_820	20	0.35	0.02	q_820	Feel comfortable with myself.	IPIP	reg
q_2765	20	0.35	0.01	q_2765	Am happy with my life.	IPIP	reg
q_811	20	-0.34	0.01	q_811	Feel a sense of worthlessness or hopelessness.	IPIP	reg
q_578	20	-0.34	0.02	q_578	Dislike myself.	IPIP	reg
q_1371	20	0.32	0.02	q_1371	Love life.	IPIP	reg
q_56	20	0.28	0.01	q_56	Am able to control my cravings.	IPIP	reg
q_1505	20	-0.27	0.01	q_1505	Panic easily.	IPIP	reg
q_808	18	-0.26	0.02	q_808	Fear for the worst.	IPIP	reg



- Condon, D. M. (2017). The SAPA Personality Inventory: An empirically-derived, hierarchically-organized self-report personality assessment model. Technical report, Northwestern University.
- Goldberg, L. R. (1972). Parameters of personality inventory construction and utilization: A comparison of prediction strategies and tactics. *Multivariate Behavioral Research Monographs*. No 72-2, 7.
- Hase, H. D. & Goldberg, L. R. (1967). Comparative validity of different strategies of constructing personality inventory scales. *Psychological Bulletin*, 67(4), 231–248.