

Data analysis using sem

practical advice

Six steps

I. Theory construction

II. Data collection

III. Data entry

IV. Scale analysis

V. Structural analysis

VI. Theory evaluation

Theory construction

- I. What is the particular theoretical hypothesis
- II. What are plausible alternatives
- III. Is there a way to tease these apart

Theory simulation

- I. Can you generate data that match your theory?
- II. Can you generate data that match the alternatives?
- III. Can you tell them apart

Data collection

I. Choosing appropriate instruments

A. Do they have good psychometric properties?

B. Are you sure?

II. Need to justify the instruments given, don't rely on "that is what others do"

III. How were the items chosen

Data collection (cont)

I. Who are the participants?

A. How are they selected

B. To what population are you trying to generalize

II. Were they engaged in the task

Data entry and storage

- I. Keep a record of all instruments given
- II. Clearly identify data codes (for M/F, education, ...)
so that you can analyze the data 5 years from now.
- III. Use a missing value code that is distinctive
 - A. Different programs treat blanks differently
 - B. Tab delimited, comma delimited files are different
than text files
 - C. R does not like spaces in variable names

Data entry checks

I. Validate data entry

A. double enter the data with two different people

II. Find descriptive statistics that include range checks

III. Obsess

Data entry in R

- I. Data tables (data.frame) is typically organized with rows for subjects, columns for variables.
- II. Can be read from a file, from the clipboard, from SPSS, from SAS (using the foreign package).

A.my.data <- read.clipboard() #copy from clipboard

B.fn <- file.choose() #find the file on your disk

C.my.data <- read.table(fn, header=TRUE) #read data

Understand descriptive stats

```
> dim(hypo)
[1] 706 96
> describe(hypo)
```

	var	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
hyp1	1	700	0.54	0.50	1.00	0.55	0.00	0.00	1.00	1.00	-0.15	-1.98	0.02
hyp2	2	700	0.47	0.50	0.00	0.46	0.00	0.00	1.00	1.00	0.14	-1.98	0.02
hyp3	3	700	0.19	0.39	0.00	0.11	0.00	0.00	1.00	1.00	1.60	0.56	0.01
hyp4	4	700	0.18	0.39	0.00	0.10	0.00	0.00	1.00	1.00	1.65	0.72	0.01
hyp5	5	694	0.63	0.48	1.00	0.66	0.00	0.00	1.00	1.00	-0.52	-1.74	0.02
hyp6	6	698	0.56	0.50	1.00	0.57	0.00	0.00	1.00	1.00	-0.24	-1.95	0.02
hyp7	7	700	0.37	0.48	0.00	0.34	0.00	0.00	1.00	1.00	0.54	-1.71	0.02
...													
IPIPM5	65	424	2.31	1.17	2.00	2.21	1.48	1.00	5.00	4.00	0.58	-0.68	0.06
IPIPM6	66	423	3.04	1.31	3.00	3.06	1.48	1.00	5.00	4.00	-0.08	-1.16	0.06
IPIPM7	67	423	3.07	1.23	3.00	3.09	1.48	1.00	5.00	4.00	-0.23	-1.02	0.06
IPIPM8	68	424	2.77	1.25	3.00	2.73	1.48	1.00	5.00	4.00	0.13	-1.09	0.06
IPIPM9	69	422	3.30	1.17	4.00	3.36	1.48	1.00	5.00	4.00	-0.38	-0.77	0.06
...													
hyp1r	73	700	0.46	0.50	0.00	0.45	0.00	0.00	1.00	1.00	0.15	-1.98	0.02
hyp2r	74	700	0.53	0.50	1.00	0.54	0.00	0.00	1.00	1.00	-0.14	-1.98	0.02
hyp6r	75	698	0.44	0.50	0.00	0.43	0.00	0.00	1.00	1.00	0.24	-1.95	0.02
hyp48r	85	686	0.35	0.48	0.00	0.32	0.00	0.00	1.00	1.00	0.61	-1.62	0.02
hypmean	86	700	0.37	0.17	0.35	0.36	0.18	0.00	0.96	0.96	0.40	-0.07	0.01
IPIP_EX	87	428	2.86	0.68	2.92	2.86	0.74	1.17	4.75	3.58	0.05	-0.35	0.03
IPIP_MO	88	424	2.55	0.72	2.50	2.52	0.74	1.00	4.75	3.75	0.35	-0.26	0.03
ipipe10r	89	424	3.03	1.36	3.00	3.04	1.48	0.00	5.00	5.00	-0.11	-1.22	0.07

Scale analysis

- I. If items are available, analyze the measurement quality of each scale being used.
- II. Examine the reliabilities of the individual scales, the item characteristics for each scale
- III. Use `make.keys`, `(super.matrix)`, `score.items`

Scoring scales

```
mpq.keys <-  
make.keys(66,list(stress=c(1,4,6,9,11,14,-16,19,21,24,27,29,31,34,37,39,42,44,47,  
-49,52,54,57,60,62,64),aggr=c(2,5,8,12,15,18,22,25,-28,32,35,38,41,45,-48,51,55,5  
8,61,65),alia=c(3,7,10,13,17,-20,23,26,30,33,36,40,43,-46,50,53,56,59,63,66)))
```

```
iip.keys <-  
make.keys(64,list(vin=c(1,22,24,29,32,40,56,64),exp=c(2,10,25,34,38,42,53,61),soc  
=c(3,7,14,18,33,35,55,62),int=c(4,26,30,41,43,47,48,58),non=c(5,6,8,9,12,13,19,39  
) ,cold=c(11,15,16,20,23,27,36,60),dom=c(17,31,44,45,50,52,57,59),ove=c(21,28,37,4  
6,49,51,54,63)))
```

```
ipde.keys <-  
make.keys(77,list(schd=c(-1,12,-21,31,46,-57,-77),scht=c(-2,24,30,52,64,67,70,71,  
-77),para=c(-2,-14,36,38,58,66,72),obco=c(-3,19,23,32,41,48,54,59),bord=c(4,8,13,  
-25,40,43,53,60,75),hist=c(5,10,17,26,28,-35,44,45),depe=c(6,33,42,49,50,62,69,73  
) ,narc=c(-7,9,15,22,37,55,61,65,68),anti=c(-11,-18,20,29,47,56,74),avoi=c(16,27,3  
4,38,39,51,63,76)))
```

A keys matrix

```
> headtail(iip.keys)
```

	vin	exp	soc	int	non	cold	dom	ove
1	1	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0
4	0	0	0	1	0	0	0	0
...	...	...	...	...	...	...	...	...
61	0	1	0	0	0	0	0	0
62	0	0	1	0	0	0	0	0
63	0	0	0	0	0	0	0	1
64	1	0	0	0	0	0	0	0

Combining keys using super.matrix

```
> mpq.iip.keys <- super.matrix(mpq.keys,iip.keys)
> headtail(mpq.iip.keys,6,6)
```

	stress	aggr	alia	vin	exp	soc	int	non	cold	dom	ove
1	1	0	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	0
4	1	0	0	0	0	0	0	0	0	0	0
5	0	1	0	0	0	0	0	0	0	0	0
6	1	0	0	0	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...
125	0	0	0	0	0	0	0	0	0	1	0
126	0	0	0	0	0	0	0	0	1	0	0
127	0	0	0	0	1	0	0	0	0	0	0
128	0	0	0	0	0	1	0	0	0	0	0
129	0	0	0	0	0	0	0	0	0	0	1
130	0	0	0	1	0	0	0	0	0	0	0

This can be repeated

```
> mpq.iip.ipde.keys <- super.matrix(mpq.iip.keys,ipde.keys)
```

```
> headtail(mpg.iip.ipde.keys,6,6)
```

	stress	aggr	alia	vin	exp	soc	int	non	cold	dom	ove	schd	scht	para	obco	bord	hist	depe	narc	anti	avoi
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
202	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
203	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
204	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
205	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
206	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
207	0	0	0	0	0	0	0	0	0	0	0	-1	-1	0	0	0	0	0	0	0	0

Score the scales

```
> sylvia.scores <- score.items(mpg.iip.ipde.keys,sylvia)
> sylvia.scores
```

Alpha:

```
      stress aggr alia  vin  exp  soc  int  non cold dom  ove schd scht para obco bord hist depe narc anti avoi
[1,]    0.92 0.73 0.84 0.79 0.82 0.84 0.76 0.89 0.81 0.8 0.79 0.35 0.21 0.57 0.6 0.66 0.42 0.53 0.42 0.6 0.72
```

Average item correlation:

```
      stress aggr alia  vin  exp soc  int non cold  dom  ove schd scht para obco bord hist depe narc anti avoi
[1,]    0.3 0.12 0.21 0.32 0.37 0.4 0.29 0.5 0.35 0.33 0.32 0.07 0.03 0.16 0.16 0.18 0.08 0.12 0.07 0.17 0.25
```

Scale intercorrelations:

	stress	aggr	alia	vin	exp	soc	int	non	cold	dom	ove	schd	scht	para	obco	bord	hist	depe	narc
stress	1.00	0.34	0.47	0.29	0.15	0.48	0.09	0.32	0.37	0.14	0.10	0.24	0.14	0.07	0.09	0.16	-0.10	0.02	0.05
aggr	0.34	1.00	0.46	0.05	0.22	0.24	0.13	0.25	0.06	0.05	0.11	0.20	0.10	0.03	0.09	0.11	0.05	0.14	0.05
alia	0.47	0.46	1.00	0.21	0.31	0.48	0.13	0.36	0.22	0.09	0.25	0.19	0.14	0.07	0.22	0.20	0.05	0.17	-0.03
vin	0.29	0.05	0.21	1.00	0.20	0.56	0.46	0.29	0.76	0.74	0.31	0.19	0.29	0.48	0.33	0.51	0.18	0.23	0.25
exp	0.15	0.22	0.31	0.20	1.00	0.50	0.46	0.76	0.27	0.15	0.76	0.07	0.09	0.05	0.19	0.13	0.19	0.41	0.08
soc	0.48	0.24	0.48	0.56	0.50	1.00	0.19	0.64	0.70	0.32	0.40	0.28	0.20	0.28	0.23	0.30	-0.01	0.27	0.08
int	0.09	0.13	0.13	0.46	0.46	0.19	1.00	0.33	0.30	0.56	0.60	0.00	0.20	0.14	0.29	0.35	0.44	0.37	0.23
non	0.32	0.25	0.36	0.29	0.76	0.64	0.33	1.00	0.41	0.15	0.55	0.09	0.08	0.04	0.12	0.12	0.09	0.33	-0.05
cold	0.37	0.06	0.22	0.76	0.27	0.70	0.30	0.41	1.00	0.55	0.32	0.22	0.27	0.35	0.25	0.35	0.06	0.19	0.12
dom	0.14	0.05	0.09	0.74	0.15	0.32	0.56	0.15	0.55	1.00	0.39	0.19	0.30	0.42	0.32	0.49	0.30	0.23	0.29
ove	0.10	0.11	0.25	0.31	0.76	0.40	0.60	0.55	0.32	0.39	1.00	0.06	0.25	0.23	0.36	0.27	0.29	0.49	0.16
schd	0.24	0.20	0.19	0.19	0.07	0.28	0.00	0.09	0.22	0.19	0.06	1.00	0.36	0.26	0.19	0.24	0.04	0.13	0.01
scht	0.14	0.10	0.14	0.29	0.09	0.20	0.20	0.08	0.27	0.30	0.25	0.36	1.00	0.49	0.29	0.51	0.37	0.30	0.16
para	0.07	0.03	0.07	0.48	0.05	0.28	0.14	0.04	0.35	0.42	0.23	0.26	0.49	1.00	0.38	0.55	0.18	0.33	0.19
obco	0.09	0.09	0.22	0.33	0.19	0.23	0.29	0.12	0.25	0.32	0.36	0.19	0.29	0.38	1.00	0.42	0.26	0.36	0.30
bord	0.16	0.11	0.20	0.51	0.13	0.30	0.35	0.12	0.35	0.49	0.27	0.24	0.51	0.55	0.42	1.00	0.48	0.49	0.24
hist	-0.10	0.05	0.05	0.18	0.19	-0.01	0.44	0.09	0.06	0.30	0.29	0.04	0.37	0.18	0.26	0.48	1.00	0.38	0.29
depe	0.02	0.14	0.17	0.23	0.41	0.27	0.37	0.33	0.19	0.23	0.49	0.13	0.30	0.33	0.36	0.49	0.38	1.00	0.23
narc	0.05	0.05	-0.03	0.25	0.08	0.08	0.23	-0.05	0.12	0.29	0.16	0.01	0.16	0.19	0.30	0.24	0.29	0.23	1.00
anti	0.09	0.01	-0.01	0.35	-0.08	0.05	0.22	-0.10	0.23	0.38	0.03	0.21	0.34	0.40	0.15	0.50	0.28	0.22	0.23
avoi	0.25	0.11	0.32	0.41	0.26	0.61	0.08	0.35	0.38	0.23	0.25	0.34	0.29	0.48	0.36	0.44	0.05	0.42	0.11

The default option is find mean response rather than sum response

More scoring

Scale intercorrelations corrected for attenuation

raw correlations below the diagonal, alpha on the diagonal

corrected correlations above the diagonal:

	stress	aggr	alia	vin	exp	soc	int	non	cold	dom	ove	schd	scht	para	obco	bord	hist	depe	narc
stress	0.92	0.41	0.54	0.34	0.17	0.54	0.11	0.36	0.43	0.16	0.12	0.43	0.32	0.09	0.12	0.20	-0.16	0.03	0.08
aggr	0.34	0.73	0.58	0.06	0.29	0.30	0.17	0.31	0.08	0.06	0.14	0.39	0.25	0.05	0.13	0.16	0.10	0.23	0.08
alia	0.47	0.46	0.84	0.26	0.38	0.57	0.17	0.41	0.26	0.11	0.31	0.35	0.33	0.10	0.31	0.26	0.08	0.26	-0.04
vin	0.29	0.05	0.21	0.79	0.25	0.69	0.59	0.35	0.95	0.93	0.39	0.36	0.70	0.71	0.47	0.70	0.31	0.36	0.43
exp	0.15	0.22	0.31	0.20	0.82	0.60	0.58	0.89	0.33	0.18	0.94	0.13	0.21	0.07	0.26	0.18	0.32	0.62	0.13
soc	0.48	0.24	0.48	0.56	0.50	0.84	0.24	0.74	0.84	0.39	0.49	0.51	0.47	0.40	0.33	0.40	-0.01	0.40	0.13
int	0.09	0.13	0.13	0.46	0.46	0.19	0.76	0.40	0.38	0.72	0.77	0.00	0.50	0.21	0.43	0.49	0.78	0.58	0.40
non	0.32	0.25	0.36	0.29	0.76	0.64	0.33	0.89	0.48	0.18	0.66	0.17	0.20	0.06	0.17	0.16	0.14	0.49	-0.08
cold	0.37	0.06	0.22	0.76	0.27	0.70	0.30	0.41	0.81	0.69	0.39	0.41	0.64	0.51	0.35	0.48	0.10	0.29	0.20
dom	0.14	0.05	0.09	0.74	0.15	0.32	0.56	0.15	0.55	0.80	0.50	0.36	0.73	0.63	0.46	0.68	0.52	0.36	0.50
ove	0.10	0.11	0.25	0.31	0.76	0.40	0.60	0.55	0.32	0.39	0.79	0.11	0.61	0.34	0.52	0.38	0.50	0.76	0.28
schd	0.24	0.20	0.19	0.19	0.07	0.28	0.00	0.09	0.22	0.19	0.06	0.35	1.33	0.58	0.41	0.51	0.12	0.29	0.03
scht	0.14	0.10	0.14	0.29	0.09	0.20	0.20	0.08	0.27	0.30	0.25	0.36	0.21	1.42	0.82	1.35	1.22	0.89	0.53
para	0.07	0.03	0.07	0.48	0.05	0.28	0.14	0.04	0.35	0.42	0.23	0.26	0.49	0.57	0.65	0.89	0.38	0.60	0.40
obco	0.09	0.09	0.22	0.33	0.19	0.23	0.29	0.12	0.25	0.32	0.36	0.19	0.29	0.38	0.60	0.66	0.52	0.64	0.60
bord	0.16	0.11	0.20	0.51	0.13	0.30	0.35	0.12	0.35	0.49	0.27	0.24	0.51	0.55	0.42	0.66	0.90	0.83	0.45
hist	-0.10	0.05	0.05	0.18	0.19	-0.01	0.44	0.09	0.06	0.30	0.29	0.04	0.37	0.18	0.26	0.48	0.42	0.81	0.69
depe	0.02	0.14	0.17	0.23	0.41	0.27	0.37	0.33	0.19	0.23	0.49	0.13	0.30	0.33	0.36	0.49	0.38	0.53	0.50
narc	0.05	0.05	-0.03	0.25	0.08	0.08	0.23	-0.05	0.12	0.29	0.16	0.01	0.16	0.19	0.30	0.24	0.29	0.23	0.42
anti	0.09	0.01	-0.01	0.35	-0.08	0.05	0.22	-0.10	0.23	0.38	0.03	0.21	0.34	0.40	0.15	0.50	0.28	0.22	0.23
avoi	0.25	0.11	0.32	0.41	0.26	0.61	0.08	0.35	0.38	0.23	0.25	0.34	0.29	0.48	0.36	0.44	0.05	0.42	0.11

More than enough information

Item by scale correlations:

	stress	aggr	alia	vin	exp	soc	int	non	cold	dom	ove	schd
MPQ_often_worry	0.62	0.09	0.18	0.13	-0.01	0.16	-0.03	0.11	0.12	0.06	-0.02	0.13
MPQ_point_out_weakness	0.29	0.49	0.26	-0.04	0.11	0.21	-0.08	0.16	0.07	-0.05	-0.03	0.16
MPQ_people_keep_from_getting_ahead	0.10	0.14	0.48	0.15	0.15	0.22	0.13	0.16	0.11	0.11	0.20	0.12
MPQ_feelings_hurt_easily	0.57	0.10	0.22	0.10	0.13	0.28	0.02	0.21	0.21	-0.01	0.12	0.15
MPQ_retaliate	0.19	0.61	0.24	-0.03	0.04	0.04	0.00	0.09	-0.01	-0.08	-0.02	0.07
MPQ_easily_rattled	0.62	0.09	0.36	0.17	0.08	0.32	0.11	0.19	0.17	0.08	0.06	0.09
MPQ_people_try_to_push	0.22	0.03	0.44	0.12	0.05	0.14	0.07	0.07	0.00	0.00	0.06	-0.06
MPQ_enjoy_violent_movies	0.02	0.46	0.22	-0.05	0.08	0.06	0.03	0.02	-0.03	-0.01	0.05	0.10
MPQ_often_irritated	0.57	0.35	0.27	0.16	0.11	0.26	0.06	0.19	0.23	0.08	0.10	0.15
MPQ_people_friends_for_useful	0.22	0.18	0.45	0.17	0.10	0.21	0.14	0.11	0.19	0.10	0.05	0.01
MPQ_suffer_from_nervousness	0.47	0.18	0.15	0.09	0.07	0.10	0.13	0.13	0.06	0.08	0.02	0.11
MPQ_get_ahead_in_line	-0.06	0.01	-0.17	-0.08	0.00	-0.08	0.09	-0.01	-0.06	-0.09	0.03	-0.02
MPQ_people_take_advantage	0.25	0.11	0.52	0.20	0.11	0.22	0.14	0.16	0.17	0.13	0.19	0.13
MPQ_often_lonely	0.55	0.21	0.41	0.19	0.09	0.22	0.10	0.14	0.22	0.11	0.04	0.22
MPQ_step_on_toes	0.08	0.42	0.30	-0.11	0.06	0.10	-0.05	0.06	-0.08	-0.12	-0.01	-0.01
MPQ_can_put_fears_out_of_mind	-0.46	-0.03	-0.15	-0.09	0.00	-0.11	-0.08	-0.07	-0.16	-0.08	0.01	-0.16

Just the high points

```
> summary(sylia.scores)
```

Alpha:

```
      stress aggr alia  vin  exp  soc  int  non cold dom  ove schd scht para obco bord hist depe narc anti avoi  
[1,]  0.92 0.73 0.84 0.79 0.82 0.84 0.76 0.89 0.81 0.8 0.79 0.35 0.21 0.57  0.6 0.66 0.42 0.53 0.42  0.6 0.72
```

Scale intercorrelations:

	stress	aggr	alia	vin	exp	soc	int	non	cold	dom	ove	schd	scht	para	obco	bord	hist	depe	narc
stress	1.00	0.34	0.47	0.29	0.15	0.48	0.09	0.32	0.37	0.14	0.10	0.24	0.14	0.07	0.09	0.16	-0.10	0.02	0.05
aggr	0.34	1.00	0.46	0.05	0.22	0.24	0.13	0.25	0.06	0.05	0.11	0.20	0.10	0.03	0.09	0.11	0.05	0.14	0.05
alia	0.47	0.46	1.00	0.21	0.31	0.48	0.13	0.36	0.22	0.09	0.25	0.19	0.14	0.07	0.22	0.20	0.05	0.17	-0.03
vin	0.29	0.05	0.21	1.00	0.20	0.56	0.46	0.29	0.76	0.74	0.31	0.19	0.29	0.48	0.33	0.51	0.18	0.23	0.25
exp	0.15	0.22	0.31	0.20	1.00	0.50	0.46	0.76	0.27	0.15	0.76	0.07	0.09	0.05	0.19	0.13	0.19	0.41	0.08
soc	0.48	0.24	0.48	0.56	0.50	1.00	0.19	0.64	0.70	0.32	0.40	0.28	0.20	0.28	0.23	0.30	-0.01	0.27	0.08
int	0.09	0.13	0.13	0.46	0.46	0.19	1.00	0.33	0.30	0.56	0.60	0.00	0.20	0.14	0.29	0.35	0.44	0.37	0.23
non	0.32	0.25	0.36	0.29	0.76	0.64	0.33	1.00	0.41	0.15	0.55	0.09	0.08	0.04	0.12	0.12	0.09	0.33	-0.05
cold	0.37	0.06	0.22	0.76	0.27	0.70	0.30	0.41	1.00	0.55	0.32	0.22	0.27	0.35	0.25	0.35	0.06	0.19	0.12
dom	0.14	0.05	0.09	0.74	0.15	0.32	0.56	0.15	0.55	1.00	0.39	0.19	0.30	0.42	0.32	0.49	0.30	0.23	0.29
ove	0.10	0.11	0.25	0.31	0.76	0.40	0.60	0.55	0.32	0.39	1.00	0.06	0.25	0.23	0.36	0.27	0.29	0.49	0.16
schd	0.24	0.20	0.19	0.19	0.07	0.28	0.00	0.09	0.22	0.19	0.06	1.00	0.36	0.26	0.19	0.24	0.04	0.13	0.01
scht	0.14	0.10	0.14	0.29	0.09	0.20	0.20	0.08	0.27	0.30	0.25	0.36	1.00	0.49	0.29	0.51	0.37	0.30	0.16
para	0.07	0.03	0.07	0.48	0.05	0.28	0.14	0.04	0.35	0.42	0.23	0.26	0.49	1.00	0.38	0.55	0.18	0.33	0.19
obco	0.09	0.09	0.22	0.33	0.19	0.23	0.29	0.12	0.25	0.32	0.36	0.19	0.29	0.38	1.00	0.42	0.26	0.36	0.30
bord	0.16	0.11	0.20	0.51	0.13	0.30	0.35	0.12	0.35	0.49	0.27	0.24	0.51	0.55	0.42	1.00	0.48	0.49	0.24
hist	-0.10	0.05	0.05	0.18	0.19	-0.01	0.44	0.09	0.06	0.30	0.29	0.04	0.37	0.18	0.26	0.48	1.00	0.38	0.29
depe	0.02	0.14	0.17	0.23	0.41	0.27	0.37	0.33	0.19	0.23	0.49	0.13	0.30	0.33	0.36	0.49	0.38	1.00	0.23
narc	0.05	0.05	-0.03	0.25	0.08	0.08	0.23	-0.05	0.12	0.29	0.16	0.01	0.16	0.19	0.30	0.24	0.29	0.23	1.00
anti	0.09	0.01	-0.01	0.35	-0.08	0.05	0.22	-0.10	0.23	0.38	0.03	0.21	0.34	0.40	0.15	0.50	0.28	0.22	0.23
avoi	0.25	0.11	0.32	0.41	0.26	0.61	0.08	0.35	0.38	0.23	0.25	0.34	0.29	0.48	0.36	0.44	0.05	0.42	0.11

But where are the scores?

```
> str(sylia.scores)
List of 7
 $ scores      : num [1:216, 1:21] 1.038 0.962 0.308 0.654 0.462 ...
  ..- attr(*, "dimnames")=List of 2
  .. ..$ : NULL
  .. ..$ : chr [1:21] "stress" "aggr" "alia" "vin" ...
 $ alpha       : num [1, 1:21] 0.92 0.73 0.84 0.79 0.82 0.84 0.76 0.89 0.81 0.8 ...
  ..- attr(*, "dimnames")=List of 2
  .. ..$ : NULL
  .. ..$ : chr [1:21] "stress" "aggr" "alia" "vin" ...
 $ av.r        : num [1, 1:21] 0.3 0.12 0.21 0.32 0.37 0.4 0.29 0.5 0.35 0.33 ...
  ..- attr(*, "dimnames")=List of 2
  .. ..$ : NULL
  .. ..$ : chr [1:21] "stress" "aggr" "alia" "vin" ...
 $ n.items     : Named num [1:21] 26 20 20 8 8 8 8 8 8 8 ...
  ..- attr(*, "names")= chr [1:21] "stress" "aggr" "alia" "vin" ...
 $ item.cor    : num [1:207, 1:21] 0.62 0.29 0.1 0.57 0.19 0.62 0.22 0.02 0.57 0.22 ...
  ..- attr(*, "dimnames")=List of 2
  .. ..$ : chr [1:207] "MPQ_often_worry" "MPQ_point_out_weakness" "MPQ_people_keep_from_getting_ahead"
"MPQ_feelings_hurt_easily" ...
  .. ..$ : chr [1:21] "stress" "aggr" "alia" "vin" ...
 $ cor         : num [1:21, 1:21] 1 0.34 0.47 0.29 0.15 0.48 0.09 0.32 0.37 0.14 ...
  ..- attr(*, "dimnames")=List of 2
  .. ..$ : chr [1:21] "stress" "aggr" "alia" "vin" ...
  .. ..$ : chr [1:21] "stress" "aggr" "alia" "vin" ...
 $ corrected: num [1:21, 1:21] 0.92 0.34 0.47 0.29 0.15 0.48 0.09 0.32 0.37 0.14 ...
  ..- attr(*, "dimnames")=List of 2
  .. ..$ : chr [1:21] "stress" "aggr" "alia" "vin" ...
  .. ..$ : chr [1:21] "stress" "aggr" "alia" "vin" ...
- attr(*, "class")= chr "psych"
```

Describe the results

```
> describe(sylian.scores$scores)
```

	var	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
stress	1	216	0.62	0.27	0.60	0.60	0.31	0.23	1.23	1.00	0.43	-0.83	0.02
aggr	2	216	0.56	0.16	0.51	0.55	0.14	0.35	1.15	0.80	0.99	0.51	0.01
alia	3	216	0.42	0.16	0.35	0.39	0.07	0.30	1.15	0.85	1.68	2.71	0.01
vin	4	216	0.75	0.61	0.62	0.68	0.56	0.00	3.62	3.62	1.26	2.21	0.04
exp	5	216	1.27	0.71	1.12	1.23	0.74	0.00	3.50	3.50	0.53	-0.10	0.05
soc	6	216	1.01	0.73	0.88	0.95	0.74	0.00	3.25	3.25	0.64	-0.36	0.05
int	7	216	0.97	0.64	0.88	0.93	0.74	0.00	3.25	3.25	0.64	0.18	0.04
non	8	216	1.33	0.84	1.25	1.31	0.93	0.00	3.75	3.75	0.32	-0.63	0.06
cold	9	216	0.74	0.64	0.62	0.66	0.56	0.00	3.38	3.38	1.33	2.02	0.04
dom	10	216	0.81	0.64	0.75	0.75	0.74	0.00	3.50	3.50	0.98	1.04	0.04
ove	11	216	1.38	0.70	1.38	1.36	0.74	0.00	3.75	3.75	0.35	-0.25	0.05
schd	12	216	1.89	0.16	1.86	1.87	0.21	1.71	2.57	0.86	0.92	0.97	0.01
scht	13	216	0.86	0.13	0.78	0.84	0.16	0.67	1.33	0.67	1.16	1.76	0.01
para	14	216	1.06	0.19	1.00	1.03	0.21	0.86	1.71	0.86	1.22	1.28	0.01
obco	15	216	0.70	0.24	0.62	0.69	0.19	0.38	1.38	1.00	0.57	-0.30	0.02
bord	16	216	0.52	0.19	0.44	0.49	0.16	0.33	1.11	0.78	1.27	1.14	0.01
hist	17	216	0.65	0.19	0.62	0.63	0.19	0.38	1.25	0.88	0.72	0.15	0.01
depe	18	216	0.29	0.18	0.25	0.27	0.19	0.00	0.88	0.88	0.90	0.51	0.01
narc	19	216	0.67	0.16	0.67	0.66	0.16	0.44	1.22	0.78	0.81	0.69	0.01
anti	20	216	0.95	0.15	0.86	0.91	0.00	0.86	1.71	0.86	2.31	5.99	0.01
avoi	21	216	0.32	0.26	0.25	0.29	0.19	0.00	1.00	1.00	0.67	-0.43	0.02

But these are peculiar. What happened?

The problem of item reversal

I. Some items have negative content, we want to reverse key the item.

II. $1 \rightarrow 4, 2 \rightarrow 3, 3 \rightarrow 2, 1 \rightarrow 4$

III. Or, $x_r \rightarrow x_{\max} + x_{\min} - x$

IV. Need to know $x_{\max}$ and $x_{\min}$

V. If items have different possible $x_{\max}$ and $x_{\min}$, then need to score parts at a time

Score the first and last

```
> syl.mpq.ipde <- score.items(mpq.ipde.keys,sy lia,min=0,max=1)
```

```
> describe(syl.mpq.ipde$scores)
```

	var	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
stress	1	216	0.39	0.27	0.37	0.37	0.31	0.00	1.00	1.00	0.43	-0.83	0.02
aggr	2	216	0.26	0.16	0.21	0.25	0.14	0.05	0.85	0.80	0.99	0.51	0.01
alia	3	216	0.12	0.16	0.05	0.09	0.07	0.00	0.85	0.85	1.68	2.71	0.01
schd	4	216	0.17	0.16	0.14	0.15	0.21	0.00	0.86	0.86	0.92	0.97	0.01
scht	5	216	0.19	0.13	0.11	0.18	0.16	0.00	0.67	0.67	1.16	1.76	0.01
para	6	216	0.20	0.19	0.14	0.17	0.21	0.00	0.86	0.86	1.22	1.28	0.01
obco	7	216	0.33	0.24	0.25	0.31	0.19	0.00	1.00	1.00	0.57	-0.30	0.02
bord	8	216	0.18	0.19	0.11	0.15	0.16	0.00	0.78	0.78	1.27	1.14	0.01
hist	9	216	0.27	0.19	0.25	0.26	0.19	0.00	0.88	0.88	0.72	0.15	0.01
depe	10	216	0.29	0.18	0.25	0.27	0.19	0.00	0.88	0.88	0.90	0.51	0.01
narc	11	216	0.34	0.16	0.33	0.33	0.16	0.11	0.89	0.78	0.81	0.69	0.01
anti	12	216	0.09	0.15	0.00	0.05	0.00	0.00	0.86	0.86	2.31	5.99	0.01
avoi	13	216	0.32	0.26	0.25	0.29	0.19	0.00	1.00	1.00	0.67	-0.43	0.02

Score the middle

```
> syl.iip <- score.items(iip.keys,sylia[,67:130])  
> syl.iip
```

Alpha:

	vin	exp	soc	int	non	cold	dom	ove
[1,]	0.79	0.82	0.84	0.76	0.89	0.81	0.8	0.79

Average item correlation:

	vin	exp	soc	int	non	cold	dom	ove
[1,]	0.32	0.37	0.4	0.29	0.5	0.35	0.33	0.32

Scale intercorrelations:

	vin	exp	soc	int	non	cold	dom	ove
vin	1.00	0.20	0.56	0.46	0.29	0.76	0.74	0.31
exp	0.20	1.00	0.50	0.46	0.76	0.27	0.15	0.76
soc	0.56	0.50	1.00	0.19	0.64	0.70	0.32	0.40
int	0.46	0.46	0.19	1.00	0.33	0.30	0.56	0.60
non	0.29	0.76	0.64	0.33	1.00	0.41	0.15	0.55
cold	0.76	0.27	0.70	0.30	0.41	1.00	0.55	0.32
dom	0.74	0.15	0.32	0.56	0.15	0.55	1.00	0.39
ove	0.31	0.76	0.40	0.60	0.55	0.32	0.39	1.00

Combine the columns

```
> syl.all <- data.frame(syl.mpq.ipde$scores, syl.iip$scores)
> describe(syl.all)
```

	var	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
stress	1	216	0.39	0.27	0.37	0.37	0.31	0.00	1.00	1.00	0.43	-0.83	0.02
aggr	2	216	0.26	0.16	0.21	0.25	0.14	0.05	0.85	0.80	0.99	0.51	0.01
alia	3	216	0.12	0.16	0.05	0.09	0.07	0.00	0.85	0.85	1.68	2.71	0.01
schd	4	216	0.17	0.16	0.14	0.15	0.21	0.00	0.86	0.86	0.92	0.97	0.01
scht	5	216	0.19	0.13	0.11	0.18	0.16	0.00	0.67	0.67	1.16	1.76	0.01
para	6	216	0.20	0.19	0.14	0.17	0.21	0.00	0.86	0.86	1.22	1.28	0.01
obco	7	216	0.33	0.24	0.25	0.31	0.19	0.00	1.00	1.00	0.57	-0.30	0.02
bord	8	216	0.18	0.19	0.11	0.15	0.16	0.00	0.78	0.78	1.27	1.14	0.01
hist	9	216	0.27	0.19	0.25	0.26	0.19	0.00	0.88	0.88	0.72	0.15	0.01
depe	10	216	0.29	0.18	0.25	0.27	0.19	0.00	0.88	0.88	0.90	0.51	0.01
narc	11	216	0.34	0.16	0.33	0.33	0.16	0.11	0.89	0.78	0.81	0.69	0.01
anti	12	216	0.09	0.15	0.00	0.05	0.00	0.00	0.86	0.86	2.31	5.99	0.01
avoi	13	216	0.32	0.26	0.25	0.29	0.19	0.00	1.00	1.00	0.67	-0.43	0.02
vin	14	216	0.75	0.61	0.62	0.68	0.56	0.00	3.62	3.62	1.26	2.21	0.04
exp	15	216	1.27	0.71	1.12	1.23	0.74	0.00	3.50	3.50	0.53	-0.10	0.05
soc	16	216	1.01	0.73	0.88	0.95	0.74	0.00	3.25	3.25	0.64	-0.36	0.05
int	17	216	0.97	0.64	0.88	0.93	0.74	0.00	3.25	3.25	0.64	0.18	0.04
non	18	216	1.33	0.84	1.25	1.31	0.93	0.00	3.75	3.75	0.32	-0.63	0.06
cold	19	216	0.74	0.64	0.62	0.66	0.56	0.00	3.38	3.38	1.33	2.02	0.04
dom	20	216	0.81	0.64	0.75	0.75	0.74	0.00	3.50	3.50	0.98	1.04	0.04
ove	21	216	1.38	0.70	1.38	1.36	0.74	0.00	3.75	3.75	0.35	-0.25	0.05

Find the correlations

```
> round(cor(syl.all),2)
```

	stress	aggr	alia	schd	scht	para	obco	bord	hist	depe	narc	anti	avoi	vin	exp	soc	int	non	cold
stress	1.00	0.34	0.47	0.24	0.14	0.07	0.09	0.16	-0.10	0.02	0.05	0.09	0.25	0.29	0.15	0.48	0.09	0.32	0.37
aggr	0.34	1.00	0.46	0.20	0.10	0.03	0.09	0.11	0.05	0.14	0.05	0.01	0.11	0.05	0.22	0.24	0.13	0.25	0.06
alia	0.47	0.46	1.00	0.19	0.14	0.07	0.22	0.20	0.05	0.17	-0.03	-0.01	0.32	0.21	0.31	0.48	0.13	0.36	0.22
schd	0.24	0.20	0.19	1.00	0.36	0.26	0.19	0.24	0.04	0.13	0.01	0.21	0.34	0.19	0.07	0.28	0.00	0.09	0.22
scht	0.14	0.10	0.14	0.36	1.00	0.49	0.29	0.51	0.37	0.30	0.16	0.34	0.29	0.29	0.09	0.20	0.20	0.08	0.27
para	0.07	0.03	0.07	0.26	0.49	1.00	0.38	0.55	0.18	0.33	0.19	0.40	0.48	0.48	0.05	0.28	0.14	0.04	0.35
obco	0.09	0.09	0.22	0.19	0.29	0.38	1.00	0.42	0.26	0.36	0.30	0.15	0.36	0.33	0.19	0.23	0.29	0.12	0.25
bord	0.16	0.11	0.20	0.24	0.51	0.55	0.42	1.00	0.48	0.49	0.24	0.50	0.44	0.51	0.13	0.30	0.35	0.12	0.35
hist	-0.10	0.05	0.05	0.04	0.37	0.18	0.26	0.48	1.00	0.38	0.29	0.28	0.05	0.18	0.19	-0.01	0.44	0.09	0.06
depe	0.02	0.14	0.17	0.13	0.30	0.33	0.36	0.49	0.38	1.00	0.23	0.22	0.42	0.23	0.41	0.27	0.37	0.33	0.19
narc	0.05	0.05	-0.03	0.01	0.16	0.19	0.30	0.24	0.29	0.23	1.00	0.23	0.11	0.25	0.08	0.08	0.23	-0.05	0.12
anti	0.09	0.01	-0.01	0.21	0.34	0.40	0.15	0.50	0.28	0.22	0.23	1.00	0.08	0.35	-0.08	0.05	0.22	-0.10	0.23
avoi	0.25	0.11	0.32	0.34	0.29	0.48	0.36	0.44	0.05	0.42	0.11	0.08	1.00	0.41	0.26	0.61	0.08	0.35	0.38
vin	0.29	0.05	0.21	0.19	0.29	0.48	0.33	0.51	0.18	0.23	0.25	0.35	0.41	1.00	0.20	0.56	0.46	0.29	0.76
exp	0.15	0.22	0.31	0.07	0.09	0.05	0.19	0.13	0.19	0.41	0.08	-0.08	0.26	0.20	1.00	0.50	0.46	0.76	0.27
soc	0.48	0.24	0.48	0.28	0.20	0.28	0.23	0.30	-0.01	0.27	0.08	0.05	0.61	0.56	0.50	1.00	0.19	0.64	0.70
int	0.09	0.13	0.13	0.00	0.20	0.14	0.29	0.35	0.44	0.37	0.23	0.22	0.08	0.46	0.46	0.19	1.00	0.33	0.30
non	0.32	0.25	0.36	0.09	0.08	0.04	0.12	0.12	0.09	0.33	-0.05	-0.10	0.35	0.29	0.76	0.64	0.33	1.00	0.41
cold	0.37	0.06	0.22	0.22	0.27	0.35	0.25	0.35	0.06	0.19	0.12	0.23	0.38	0.76	0.27	0.70	0.30	0.41	1.00
dom	0.14	0.05	0.09	0.19	0.30	0.42	0.32	0.49	0.30	0.23	0.29	0.38	0.23	0.74	0.15	0.32	0.56	0.15	0.55
ove	0.10	0.11	0.25	0.06	0.25	0.23	0.36	0.27	0.29	0.49	0.16	0.03	0.25	0.31	0.76	0.40	0.60	0.55	0.32

More scale analysis

I. Dimensionality of the inventories done for each inventory

II. Factor analysis, Cluster analysis

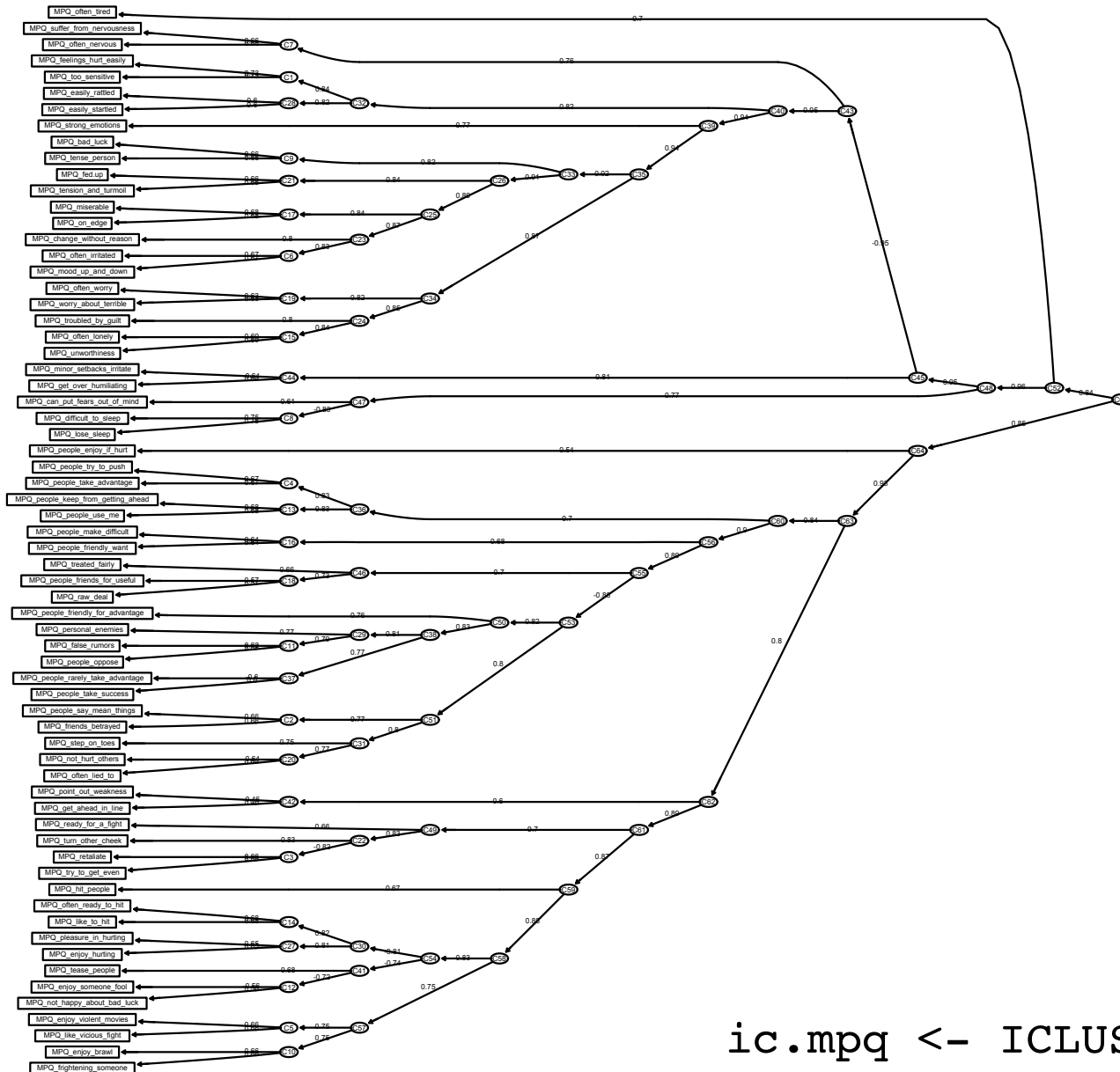
III. But what is the expectation?

A. one, two, many factors

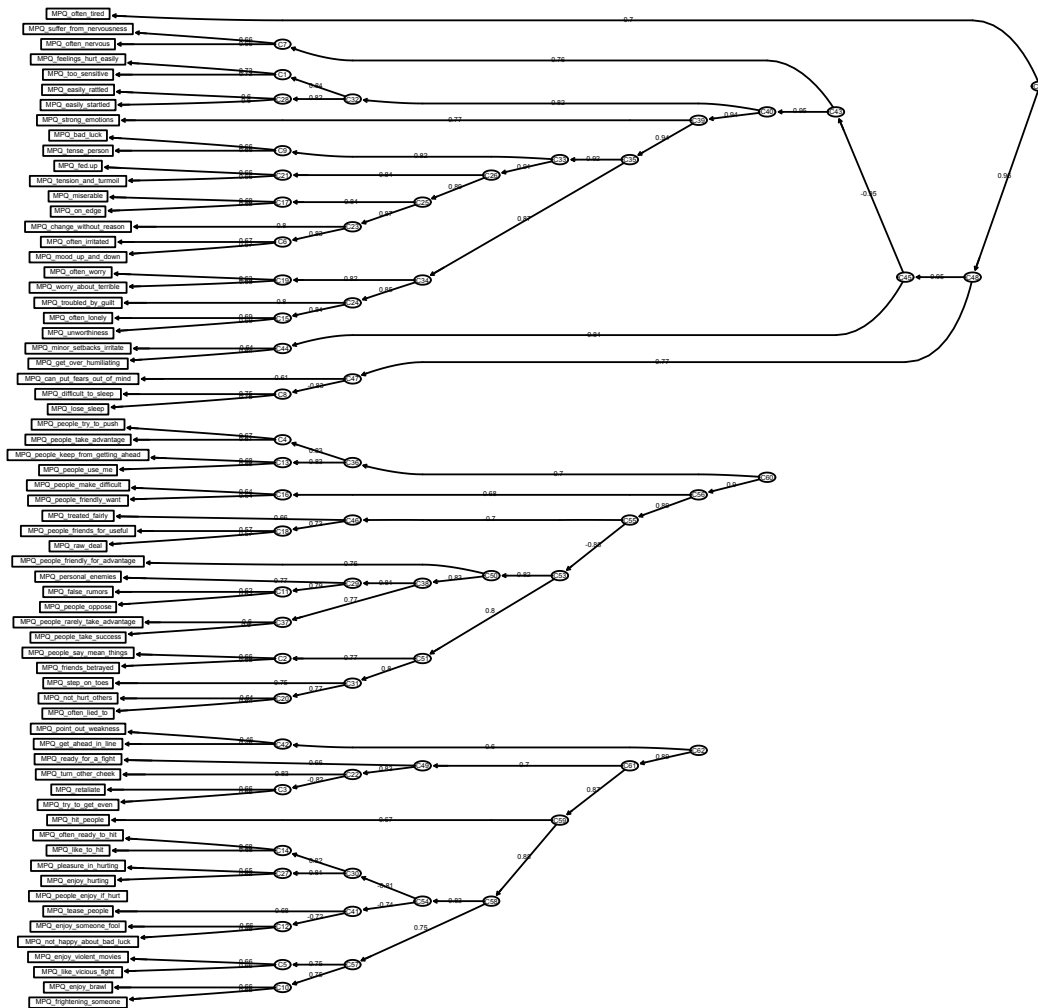
B. what should the structure look like

ICLUST

ICLUST of mpq



ICLUST



3 large
but
correlated
clusters

```
ic.mpq4 <- ICLUST(sylia[,1:66],4)
```

```
> summary(ic.mpq4)
```

```
ICLUST (Item Cluster Analysis)Call: ICLUST(r.mat = sylvia[,  
1:66], nclusters = 4)
```

```
ICLUST
```

```
Purified Alpha:
```

C52	C60	C62	V59
0.92	0.84	0.81	1.00

```
Original Beta:
```

C52	C60	C62	V59
0.63	0.63	0.47	NA

```
Cluster size:
```

C52	C60	C62	V59
26	21	18	1

```
Purified scale intercorrelations
```

```
reliabilities on diagonal
```

```
correlations corrected for attenuation above diagonal:
```

	C52	C60	C62	V59
C52	0.92	0.51	-0.40	0.11
C60	0.45	0.84	-0.57	0.26
C62	-0.35	-0.47	0.81	-0.24
V59	0.10	0.24	-0.22	1.00

Sort the clusters

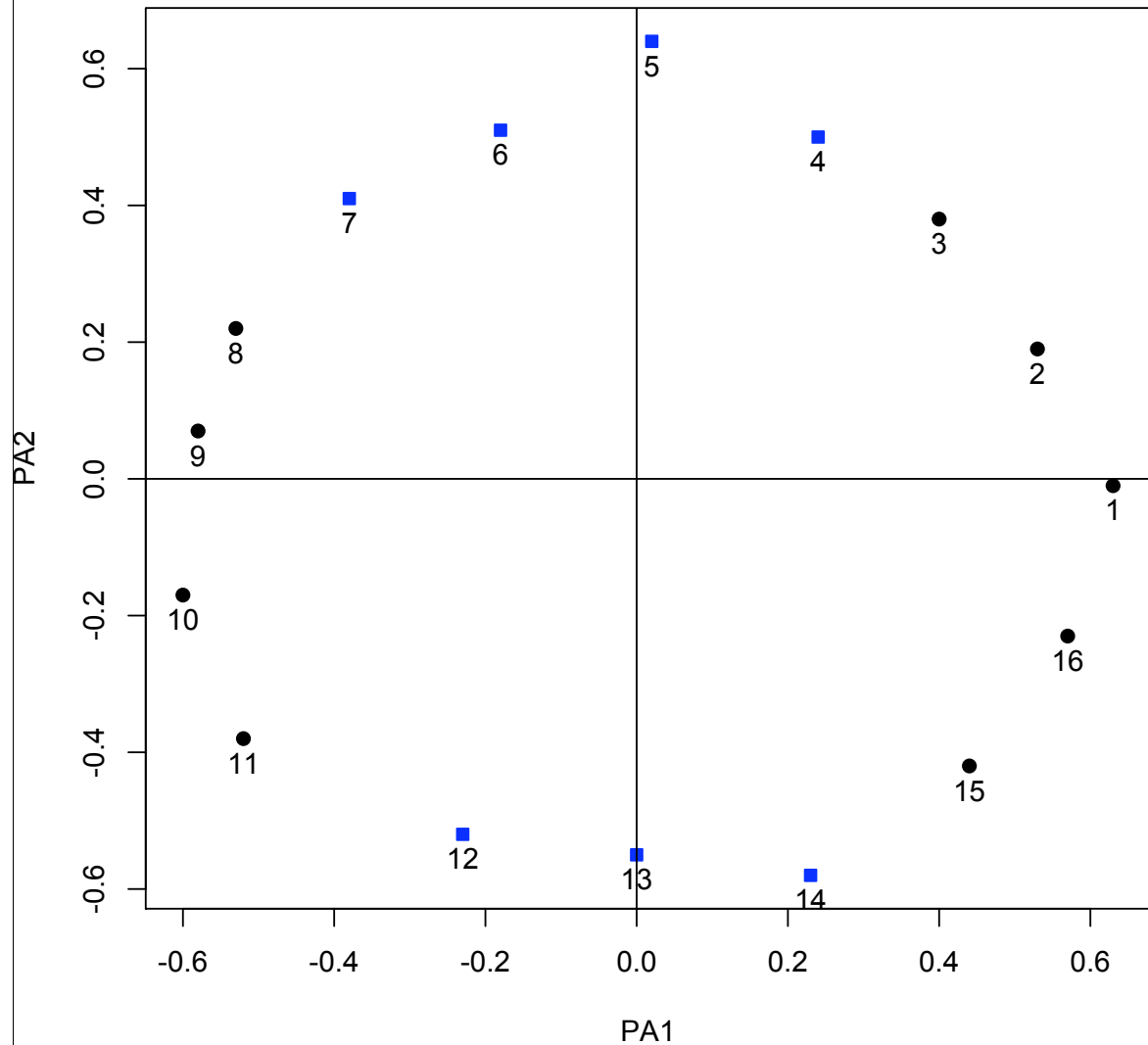
```
> ICLUST.sort(ic.mpq4)
$sorted
```

	item	content	cluster	C52	C60	C62	V59
MPQ_miserable	24	MPQ_miserable	1	0.60	0.33	-0.36	0.06
MPQ_too_sensitive	62	MPQ_too_sensitive	1	0.59	0.28	-0.13	0.09
MPQ_unworthiness	27	MPQ_unworthiness	1	0.58	0.26	-0.20	0.01
MPQ_lose_sleep	52	MPQ_lose_sleep	1	0.58	0.39	-0.08	0.07
MPQ_often_worry	1	MPQ_often_worry	1	0.57	0.17	-0.05	0.07
MPQ_easily_rattled	6	MPQ_easily_rattled	1	0.57	0.35	-0.11	0.17
MPQ_mood_up_and_down	21	MPQ_mood_up_and_down	1	0.57	0.31	-0.31	0.07
MPQ_change_without_reason	64	MPQ_change_without_reason	1	0.57	0.33	-0.34	0.04
...							
MPQ_often_lied_to	56	MPQ_often_lied_to	2	0.27	0.52	-0.34	0.10
MPQ_false_rumors	40	MPQ_false_rumors	2	0.21	0.51	-0.28	0.17
MPQ_people_oppose	66	MPQ_people_oppose	2	0.28	0.51	-0.28	0.18
MPQ_people_say_mean_things	33	MPQ_people_say_mean_things	2	0.28	0.50	-0.30	0.21
MPQ_bad_luck	23	MPQ_bad_luck	2	0.48	0.49	-0.45	0.14
MPQ_people_use_me	17	MPQ_people_use_me	2	0.26	0.47	-0.22	0.02
MPQ_people_make_difficult	30	MPQ_people_make_difficult	2	0.19	0.47	-0.23	0.16
MPQ_people_take_advantage	13	MPQ_people_take_advantage	2	0.25	0.45	-0.12	0.03
...							
MPQ_like_vicious_fight	51	MPQ_like_vicious_fight	3	0.03	0.21	-0.50	0.08
MPQ_often_ready_to_hit	18	MPQ_often_ready_to_hit	3	0.40	0.26	-0.49	0.08
MPQ_like_to_hit	65	MPQ_like_to_hit	3	0.26	0.42	-0.46	0.17
MPQ_try_to_get_even	58	MPQ_try_to_get_even	3	0.22	0.27	-0.44	0.18
MPQ_retaliate	5	MPQ_retaliate	3	0.19	0.23	-0.43	0.11
MPQ_enjoy_someone_fool	25	MPQ_enjoy_someone_fool	3	0.26	0.27	-0.41	0.12
MPQ_turn_other_cheek	28	MPQ_turn_other_cheek	3	-0.07	-0.17	0.38	-0.16
MPQ_point_out_weakness	2	MPQ_point_out_weakness	3	0.29	0.26	-0.36	0.18
...							

Structural analysis

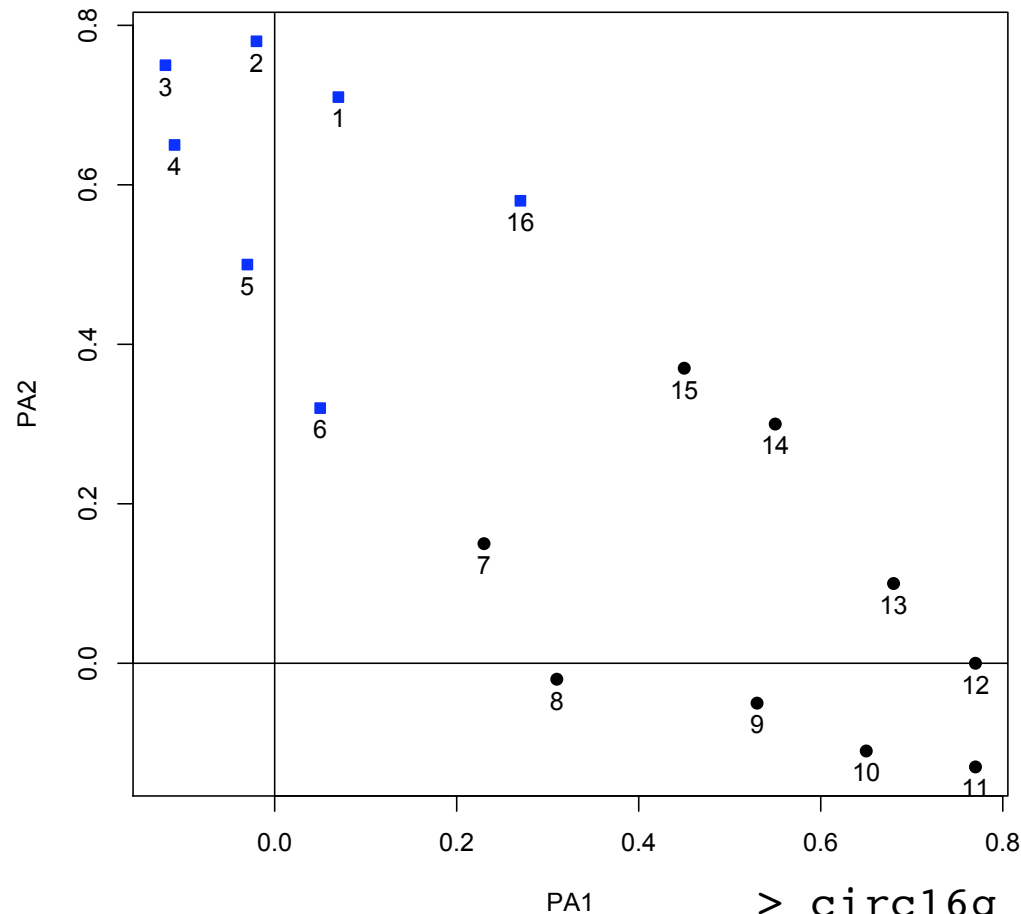
- I. Do on covariance matrix
- II. Consider theory (do we expect simple structure or do we expect more complicated structure)
- III. Compare with simulations

Consider circumplex items



```
> circ16 <- sim.circ(16)
> f2c <- factor.pa(circ16,2)
> plot(f2c)
```

Circumplex with a g factor



```
> circ16g <- sim.circ(16,gloading=.45)
> f2cg <- factor.pa(circ16g,2)
> plot(f2cg)
```

Theory Evaluation

- I. What do our data tell us about the theory?
- II. What needs to be improved
- III. What can be rejected
- IV. What do we know that we did not before?
- V. What do we know is wrong that we thought we knew before?