Psychology 405: Psychometric Theory
Analysis of the midterm

William Revelle

Department of Psychology
Northwestern University
Evanston, Illinois USA

May 2025

Outline

total scores as sum scores

```
sort(mid$total)
[1] 16.5 18.0 18.5 18.5 20.0 21.0 21.5


> describe(mid$total)
   vars n  mean   sd median trimmed  mad  min  max range  skew kurt
X1    1 7 19.14 1.77   18.5   19.14 2.22 16.5 21.5     5 -0.01    -

round(sort(scale(mid$total)),2)
[1] -1.49 -0.64 -0.36 -0.36  0.48  1.05  1.33
```

Raw descriptives show some missing data

```
describe(mid)
     vars n  mean   sd median trimmed  mad  min  max range  skew kurtosis   se
q1      1 7  0.29 0.49    0.0    0.29 0.00  0.0  1.0   1.0  0.75    -1.60 0.18
q2      2 7  0.93 0.19    1.0    0.93 0.00  0.5  1.0   0.5 -1.62     0.80 0.07
q3      3 7  0.71 0.49    1.0    0.71 0.00  0.0  1.0   1.0 -0.75    -1.60 0.18
q4      4 7  0.50 0.29    0.5    0.50 0.00  0.0  1.0   1.0  0.00    -0.43 0.11
q5a     5 7  1.00 0.00    1.0    1.00 0.00  1.0  1.0   0.0   NaN      NaN 0.00
q5b     6 7  0.86 0.38    1.0    0.86 0.00  0.0  1.0   1.0 -1.62     0.80 0.14
q5c     7 7  0.57 0.53    1.0    0.57 0.00  0.0  1.0   1.0 -0.23    -2.20 0.20
q5d     8 7  0.71 0.49    1.0    0.71 0.00  0.0  1.0   1.0 -0.75    -1.60 0.18
q5e     9 7  0.86 0.38    1.0    0.86 0.00  0.0  1.0   1.0 -1.62     0.80 0.14
q5f    10 7  0.14 0.38    0.0    0.14 0.00  0.0  1.0   1.0  1.62     0.80 0.14
q5g    11 7  0.71 0.49    1.0    0.71 0.00  0.0  1.0   1.0 -0.75    -1.60 0.18
q5h    12 7  0.43 0.53    0.0    0.43 0.00  0.0  1.0   1.0  0.23    -2.20 0.20
q5i    13 7  0.71 0.49    1.0    0.71 0.00  0.0  1.0   1.0 -0.75    -1.60 0.18
q5j    14 7  0.29 0.49    0.0    0.29 0.00  0.0  1.0   1.0  0.75    -1.60 0.18
q5k    15 7  0.71 0.49    1.0    0.71 0.00  0.0  1.0   1.0 -0.75    -1.60 0.18
q5l    16 7  0.43 0.53    0.0    0.43 0.00  0.0  1.0   1.0  0.23    -2.20 0.20
q5m    17 7  0.57 0.53    1.0    0.57 0.00  0.0  1.0   1.0 -0.23    -2.20 0.20
q6     18 7  0.71 0.49    1.0    0.71 0.00  0.0  1.0   1.0 -0.75    -1.60 0.18
q7     19 7  0.43 0.53    0.0    0.43 0.00  0.0  1.0   1.0  0.23    -2.20 0.20
q8     20 7  0.71 0.49    1.0    0.71 0.00  0.0  1.0   1.0 -0.75    -1.60 0.18
q9     21 7  0.86 0.38    1.0    0.86 0.00  0.0  1.0   1.0 -1.62     0.80 0.14
q10    22 7  0.57 0.53    1.0    0.57 0.00  0.0  1.0   1.0 -0.23    -2.20 0.20
q11    23 7  0.71 0.49    1.0    0.71 0.00  0.0  1.0   1.0 -0.75    -1.60 0.18
q12    24 7  0.86 0.38    1.0    0.86 0.00  0.0  1.0   1.0 -1.62     0.80 0.14
q13    25 7  0.43 0.53    0.0    0.43 0.00  0.0  1.0   1.0  0.23    -2.20 0.20
q14    26 7  0.00 0.00    0.0    0.00 0.00  0.0  0.0   0.0   NaN      NaN 0.00
q15    27 7  0.00 0.00    0.0    0.00 0.00  0.0  0.0   0.0   NaN      NaN 0.00
q16    28 7  0.14 0.38    0.0    0.14 0.00  0.0  1.0   1.0  1.62     0.80 0.14
q17    29 7  0.14 0.38    0.0    0.14 0.00  0.0  1.0   1.0  1.62     0.80 0.14
q18    30 7  0.29 0.49    0.0    0.29 0.00  0.0  1.0   1.0  0.75    -1.60 0.18
q19    31 7  0.57 0.53    1.0    0.57 0.00  0.0  1.0   1.0 -0.23    -2.20 0.20
q20    32 7  0.14 0.38    0.0    0.14 0.00  0.0  1.0   1.0  1.62     0.80 0.14
```

Drop items with no variance

```
describe(mid[-c(5,26,27,33,36,39,43)])
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
q1	1	7	0.29	0.49	0.0	0.29	0	0.0	1	1.0	0.75	-1.60	0.18
q2	2	7	0.93	0.19	1.0	0.93	0	0.5	1	0.5	-1.62	0.80	0.07
q3	3	7	0.71	0.49	1.0	0.71	0	0.0	1	1.0	-0.75	-1.60	0.18
q4	4	7	0.50	0.29	0.5	0.50	0	0.0	1	1.0	0.00	-0.43	0.11
q5b	5	7	0.86	0.38	1.0	0.86	0	0.0	1	1.0	-1.62	0.80	0.14
q5c	6	7	0.57	0.53	1.0	0.57	0	0.0	1	1.0	-0.23	-2.20	0.20
q5d	7	7	0.71	0.49	1.0	0.71	0	0.0	1	1.0	-0.75	-1.60	0.18
q5e	8	7	0.86	0.38	1.0	0.86	0	0.0	1	1.0	-1.62	0.80	0.14
q5f	9	7	0.14	0.38	0.0	0.14	0	0.0	1	1.0	1.62	0.80	0.14
q5g	10	7	0.71	0.49	1.0	0.71	0	0.0	1	1.0	-0.75	-1.60	0.18
q5h	11	7	0.43	0.53	0.0	0.43	0	0.0	1	1.0	0.23	-2.20	0.20
q5i	12	7	0.71	0.49	1.0	0.71	0	0.0	1	1.0	-0.75	-1.60	0.18
q5j	13	7	0.29	0.49	0.0	0.29	0	0.0	1	1.0	0.75	-1.60	0.18
q5k	14	7	0.71	0.49	1.0	0.71	0	0.0	1	1.0	-0.75	-1.60	0.18
q5l	15	7	0.43	0.53	0.0	0.43	0	0.0	1	1.0	0.23	-2.20	0.20
q5m	16	7	0.57	0.53	1.0	0.57	0	0.0	1	1.0	-0.23	-2.20	0.20
q6	17	7	0.71	0.49	1.0	0.71	0	0.0	1	1.0	-0.75	-1.60	0.18
q7	18	7	0.43	0.53	0.0	0.43	0	0.0	1	1.0	0.23	-2.20	0.20
q8	19	7	0.71	0.49	1.0	0.71	0	0.0	1	1.0	-0.75	-1.60	0.18
q9	20	7	0.86	0.38	1.0	0.86	0	0.0	1	1.0	-1.62	0.80	0.14
q10	21	7	0.57	0.53	1.0	0.57	0	0.0	1	1.0	-0.23	-2.20	0.20
q11	22	7	0.71	0.49	1.0	0.71	0	0.0	1	1.0	-0.75	-1.60	0.18
q12	23	7	0.86	0.38	1.0	0.86	0	0.0	1	1.0	-1.62	0.80	0.14
q13	24	7	0.43	0.53	0.0	0.43	0	0.0	1	1.0	0.23	-2.20	0.20
q16	25	7	0.14	0.38	0.0	0.14	0	0.0	1	1.0	1.62	0.80	0.14
q17	26	7	0.14	0.38	0.0	0.14	0	0.0	1	1.0	1.62	0.80	0.14
q18	27	7	0.29	0.49	0.0	0.29	0	0.0	1	1.0	0.75	-1.60	0.18
q19	28	7	0.57	0.53	1.0	0.57	0	0.0	1	1.0	-0.23	-2.20	0.20
q20	29	7	0.14	0.38	0.0	0.14	0	0.0	1	1.0	1.62	0.80	0.14
q22	30	7	0.43	0.53	0.0	0.43	0	0.0	1	1.0	0.23	-2.20	0.20
q23	31	7	0.57	0.53	1.0	0.57	0	0.0	1	1.0	-0.23	-2.20	0.20
q26	32	7	0.14	0.38	0.0	0.14	0	0.0	1	1.0	1.62	0.80	0.14

alpha of the scores

R code

```
alpha(mid[-c(5,26,27,33,36,39,43)])
Some items ( q1 q3 q5c q5f q5h q5l q5m q8 q10 q11 q12 q13 q19 q23
      q27 q29 q30 ) were negatively correlated with the first principa
probably should be reversed.
To do this, run the function again with the 'check.keys=TRUE' option
```

```
Reliability analysis
Call: alpha(x = mid[-c(5, 26, 27, 33, 36, 39, 43)])

  raw_alpha std.alpha G6(smc) average_r   S/N ase mean    sd median_r
      -1.5     -0.95   -0.93    -0.014 -0.49 1.2  0.5 0.049        0

    95% confidence boundaries
         lower alpha upper
Feldt    -5.07 -1.46  0.50
Duhachek -3.90 -1.46  0.98

   raw_alpha std.alpha G6(smc) average_r   S/N var.r  med.r
q1      -0.88     -0.55   -0.54   -0.0103 -0.36  0.17  0.000
q2      -1.20     -0.58   -0.57   -0.0106 -0.37  0.17  0.000
q3      -0.89     -0.55   -0.54   -0.0103 -0.36  0.17  0.000
q4      -1.64     -1.12   -1.09   -0.0154 -0.53  0.17 -0.091
q5b     -2.16     -1.63   -1.59   -0.0180 -0.62  0.16  0.000
q5c     -1.05     -0.63   -0.62   -0.0112 -0.39  0.16  0.000
q5d     -1.06     -0.78   -0.76   -0.0127 -0.44  0.17  0.000
q5e     -2.16     -1.63   -1.59   -0.0180 -0.62  0.16  0.000
...
```

Allow items to be reversed – but is this meaningful for an exam?

R code

```
alpha(mid[-c(5,26,27,33,36,39,43)],check.keys=TRUE)
```

```
Reliability analysis
Call: alpha(x = mid[-c(5, 26, 27, 33, 36, 39, 43)], check.keys = TRUE)

  raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
       0.9       0.91    0.88      0.21 9.7 0.053 0.53 0.22     0.26

    95\% confidence boundaries
         lower alpha upper
Feldt     0.76   0.9  0.98
Duhachek  0.80   0.9  1.01

 Reliability if an item is dropped:
     raw_alpha std.alpha G6(smc) average_r  S/N var.r med.r
q1-       0.90      0.90    0.88      0.21  9.5  0.12  0.26
q2        0.90      0.90    0.88      0.21  9.4  0.12  0.26
q3-       0.90      0.90    0.88      0.21  9.5  0.12  0.26
q4        0.90      0.91    0.88      0.22  9.9  0.12  0.26
q5b       0.90      0.90    0.87      0.21  9.1  0.12  0.26
q5c-      0.90      0.90    0.88      0.21  9.2  0.12  0.26
q5d       0.90      0.90    0.88      0.21  9.5  0.12  0.26
q5e       0.90      0.90    0.87      0.21  9.1  0.12  0.26
q5f-      0.91      0.91    0.88      0.22 10.1  0.12  0.26
q5g       0.89      0.90    0.87      0.20  8.7  0.12  0.26
q5h-      0.91      0.91    0.88      0.22 10.1  0.12  0.26
q5i       0.90      0.90    0.88      0.21  9.2  0.12  0.26


....
```

Try IRT analysis

R code

```
f1 <- irt.fa(mid[-c(5,26,27,33,36,39,43)])
f1
```

```
 f1
Item Response Analysis using Factor Analysis

Call: irt.fa(x = mid[-c(5, 26, 27, 33, 36, 39, 43)])
Item Response Analysis using Factor Analysis

 Summary information by factor and item
 Factor =  1
              -3   -2   -1    0    1    2    3
q1          0.07 0.09 0.11 0.11 0.08 0.06 0.03
q3          0.02 0.02 0.02 0.03 0.03 0.02 0.02
q5b         0.32 0.57 0.36 0.11 0.03 0.01 0.00
q5c         0.04 0.08 0.12 0.15 0.14 0.10 0.06
q5e         0.32 0.57 0.36 0.11 0.03 0.01 0.00
q5g         0.11 0.40 0.66 0.33 0.08 0.02 0.00
q5i         0.07 0.11 0.14 0.13 0.10 0.06 0.03
q5j         0.02 0.05 0.12 0.26 0.32 0.22 0.10
q5k         0.05 0.06 0.07 0.07 0.06 0.05 0.03
q5l         0.06 0.10 0.15 0.16 0.13 0.08 0.04
q5m         0.04 0.06 0.08 0.09 0.08 0.07 0.05
q7          0.02 0.02 0.03 0.03 0.03 0.02 0.02
q8          0.03 0.05 0.08 0.10 0.11 0.09 0.06
q9          0.32 0.57 0.36 0.11 0.03 0.01 0.00
q10         0.04 0.09 0.18 0.25 0.22 0.13 0.06
q12         0.02 0.04 0.07 0.11 0.14 0.14 0.10
q13         0.02 0.02 0.03 0.03 0.03 0.02 0.02
q16         0.02 0.04 0.07 0.13 0.18 0.18 0.12
q17         0.02 0.02 0.02 0.03 0.03 0.03 0.02
q18         0.02 0.05 0.12 0.26 0.32 0.22 0.10
q19         0.03 0.05 0.05 0.06 0.06 0.05 0.04
```

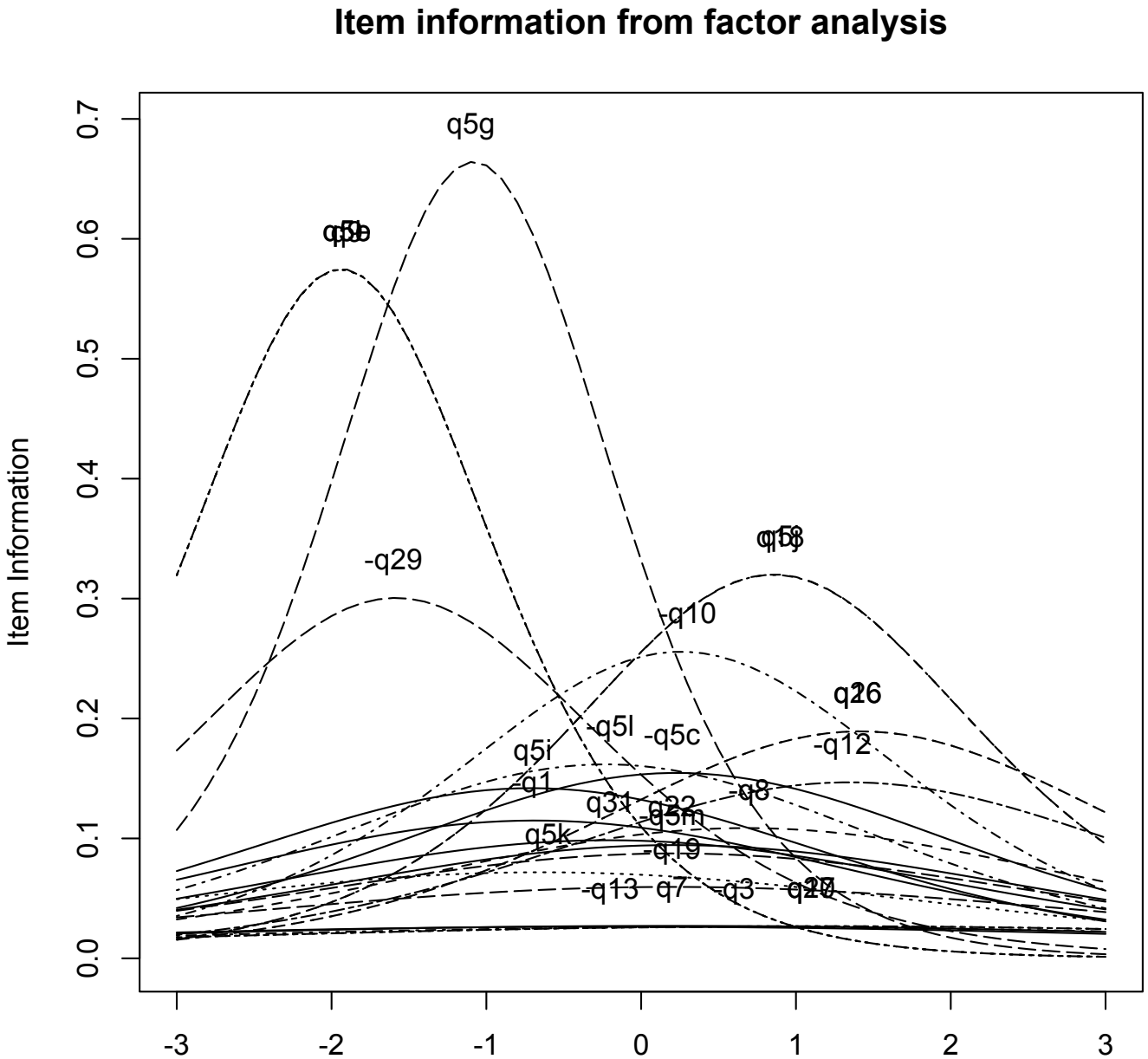
irt continued

R code

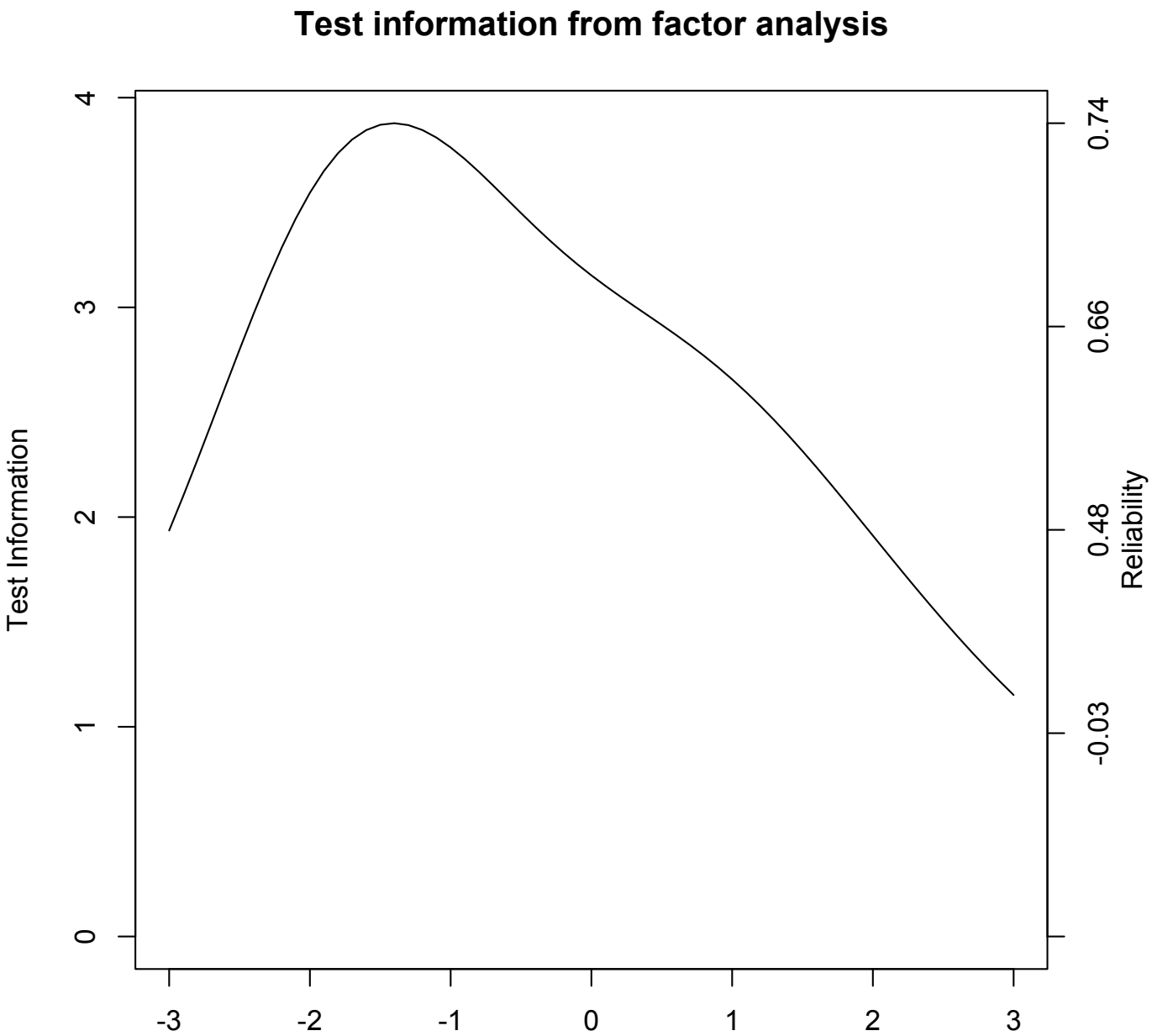
```
plot(f1)
plot(f1,type="test")
```

```
q19          0.03 0.05 0.05 0.06 0.06 0.05 0.04
q20          0.02 0.02 0.02 0.03 0.03 0.03 0.02
q22          0.04 0.06 0.08 0.09 0.09 0.07 0.05
q26          0.02 0.04 0.07 0.13 0.18 0.18 0.12
q29          0.17 0.29 0.27 0.15 0.06 0.02 0.01
q31          0.05 0.07 0.09 0.10 0.09 0.06 0.04
Test Info    1.94 3.55 3.76 3.15 2.66 1.91 1.15
SEM          0.72 0.53 0.52 0.56 0.61 0.72 0.93
Reliability 0.48 0.72 0.73 0.68 0.62 0.48 0.13
```

Item information



Test information



But some of those items seem to be reversed scored!

R code

```
f1$fa$loadings
```

```
'f1$fa$loadings

Loadings:
    MR1
q1  -0.561
q2  -0.123
q3  -0.307
q4  -0.166
q5b  0.835
q5c -0.618
q5d  0.184
q5e  0.835
q5f
q5g  0.852
q5h
q5i  0.602
q5j  0.749
q5k  0.472
q5l -0.627
q5m -0.509
q6   0.132
q7   0.312
q8  -0.550
q9   0.835
q10 -0.711
q11
q12 -0.608
q13 -0.309
q16  0.656
q17  0.311
q18  0.749
```

Tests require meaningful variability and good items

1. Although IRT does not require variance of the total score, the FA approach to IRT is based on the correlations, and does.
2. Need to try other IRT approaches