Reliability Analysis Example

SPSS

This example comes from a set of items my class developed to measure internet addiction. There were three items that were negatively keyed that needed to be rescored.

It is always a good idea to examine the distributions of individual variables, but I have omitted that output here to save paper.

Syntax

```
get file='c:\jason\spsswin\uvclass\ias.sav'.

*some items needed to be reverse coded--this may not be necessary on the homework problem.
recode q4 q5 q10 (1=7) (2=6) (3=5) (4=4) (5=3) (6=2) (7=1) into q4r q5r q10r.

*frequencies vars=all
  /histogram=normal.

descriptives vars=all
    /statistics=default variance skew kurtosis.

reliability vars=q1 q2 q3 q4r q5r q6 q7 q8 q9 q10r q11 q12
   /scale(ias)=q1 q2 q3 q4r q5r q6 q7 q8 q9 q10r q11 q12
  /statistics=correlations scale
  /summary=means corr total.
```

Menus

Analyze → Scale → Reliability Analysis

Drag over the desired variables. Click the **Statistics** button and check “Item,” “Scale,” and “Scale-if-item-deleted” under **Descriptives**; check “Correlations” under **Inter-Item**, and “Means and Correlations” under **Summaries**

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
q1 I stay up later than I intended to use the internet.	18	1.0	7.0	4.611	1.6852	2.840	-.714	.536	-.510	1.038
q2 I find myself using the internet while talking to someone in person.	18	1	7	4.06	1.798	3.232	-.161	.536	-.554	1.038
q3 I have more online friends than in person friends.	18	1	7	3.39	2.004	4.016	.331	.536	-.838	1.038
q4 I can go one full day without using the internet.	18	1	7	4.94	1.862	3.467	-.709	.536	-.555	1.038
q5 I don't often use social media	18	1	7	3.17	1.917	3.676	.579	.536	-.639	1.038
q6 I think about using the internet when I'm not using the internet.	18	1	7	4.50	1.689	2.853	-.082	.536	-.491	1.038
q7 I get bored when I don't have my phone when I use the bathroom	18	1	7	3.44	2.307	5.320	.355	.536	-1.370	1.038
q8 I like using social media.	18	2	7	4.72	1.526	2.330	-.141	.536	-1.171	1.038
q9 I have missed social events because of internet use.	18	1	6	1.94	1.626	2.644	1.578	.536	1.220	1.038
q10 I am comfortable staying in a place without internet access.	18	2	7	4.94	1.514	2.291	-.124	.536	-.724	1.038
q11 I have lost sleep because of internet use.	18	1	7	4.44	1.886	3.556	-.910	.536	-.274	1.038
q12 My friends or family have commented on my internet use.	18	1	7	2.78	1.896	3.595	.941	.536	-.123	1.038
q4r	18	1.00	7.00	3.0556	1.86207	3.467	.709	.536	-.555	1.038
q5r	18	1.00	7.00	4.8333	1.91741	3.676	-.579	.536	-.639	1.038
q10r	18	1.00	6.00	3.0556	1.51356	2.291	.124	.536	-.724	1.038
Valid N (listwise)	18									

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.724	.737	12

Inter-Item Correlation Matrix

	q1 I stay up later than I intended to use the internet.	q2 I find myself using the internet while talking to someone in person.	q3 I have more online friends than in person friends.	q4r	q5r	q6 I think about using the internet when I'm not using the internet.	q7 I get bored when I don't have my phone when I use the bathroom	q8 I like using social media.	q9 I have missed social events because of internet use.	q10r	q11 I have lost sleep because of internet use.	q12 My friends or family have commented on my internet use.
q1 I stay up later than I intended to use the internet.	1.000	.415	-.040	.026	-.112	.444	.092	-.182	.400	.124	.909	.266
q2 I find myself using the internet while talking to someone in person.	.415	1.000	-.039	.227	.617	.649	.121	.478	.243	.388	.322	.349
q3 I have more online friends than in person friends.	-.040	-.039	1.000	-.022	.048	.391	.177	-.001	.458	.012	.185	.365
q4r	.026	.227	-.022	1.000	.349	-.028	.309	.399	-.310	.416	-.141	-.263
q5r	-.112	.617	.048	.349	1.000	.354	-.022	.506	-.041	.470	-.141	.022
q6 I think about using the internet when I'm not using the internet.	.444	.649	.391	-.028	.354	1.000	.151	.308	.610	.426	.425	.386
q7 I get bored when I don't have my phone when I use the bathroom	.092	.121	.177	.309	-.022	.151	1.000	.405	-.087	-.092	.128	-.232
q8 I like using social media.	-.182	.478	-.001	.399	.506	.308	.405	1.000	-.078	.185	-.179	-.185
q9 I have missed social events because of internet use.	.400	.243	.458	-.310	-.041	.610	-.087	-.078	1.000	.001	.411	.625
q10r	.124	.388	.012	.416	.470	.426	-.092	.185	.001	1.000	-.009	-.200
q11 I have lost sleep because of internet use.	.909	.322	.185	-.141	-.141	.425	.128	-.179	.411	-.009	1.000	.309
q12 My friends or family have commented on my internet use.	.266	.349	.365	-.263	.022	.386	-.232	-.185	.625	-.200	.309	1.000

Summary Item Statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	3.736	1.944	4.833	2.889	2.486	.854	12
Inter-Item Correlations	.189	-.310	.909	1.219	-2.935	.071	12

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
q1 I stay up later than I intended to use the internet.	40.2222	101.242	.422	.961	.698
q2 I find myself using the internet while talking to someone in person.	40.7778	91.477	.688	.839	.660
q3 I have more online friends than in person friends.	41.4444	103.203	.274	.783	.718
q4r	41.7778	108.536	.164	.757	.732
q5r	40.0000	101.647	.338	.657	.709
q6 I think about using the internet when I'm not using the internet.	40.3333	90.824	.767	.832	.652
q7 I get bored when I don't have my phone when I use the bathroom	41.3889	105.428	.161	.506	.740
q8 I like using social media.	40.1111	106.340	.309	.742	.712
q9 I have missed social events because of internet use.	42.8889	102.575	.400	.723	.701
q10r	41.7778	107.007	.290	.677	.714
q11 I have lost sleep because of internet use.	40.3889	100.016	.393	.943	.701
q12 My friends or family have commented on my internet use.	42.0556	105.350	.242	.735	.722

R

```
> cat("\014") #clear console
> rm(d) #clear active frame from previous analyses
> library(lessR)
> d = Read("c:/jason/spsswin/uvclass/ias.sav",quiet=TRUE)
> #reverse score items (lessR functions)
> #these were already recoded for my data set--can use lessR to recode if needed, e.g.,
> #d <- recode(q4,old=1:7, new=7:1)
> #d <- recode(q5,old=1:7, new=7:1)
> #d <- recode(q10,old=1:7, new=7:1)
> #caution alpha command here does all variables in the data set[1]
> library(psych)
> alpha(d)

Reliability analysis
Call: alpha(x = d)

  raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
      0.72      0.74    0.92      0.19 2.8 0.098  3.7 0.91     0.19

    95% confidence boundaries
         lower alpha upper
Feldt     0.49  0.72  0.88
Duhachek  0.53  0.72  0.92

 Reliability if an item is dropped:
     raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
q1        0.70      0.71    0.87      0.18 2.5    0.107 0.067  0.19
q2        0.66      0.67    0.89      0.16 2.1    0.122 0.073  0.13
q3        0.72      0.73    0.91      0.20 2.7    0.100 0.079  0.24
q4r       0.73      0.74    0.91      0.21 2.9    0.095 0.070  0.19
q5r       0.71      0.72    0.92      0.19 2.6    0.103 0.072  0.19
q6        0.65      0.66    0.89      0.15 2.0    0.125 0.071  0.12
q7        0.74      0.75    0.93      0.21 2.9    0.092 0.077  0.27
q8        0.71      0.73    0.91      0.20 2.7    0.102 0.071  0.19
q9        0.70      0.72    0.91      0.19 2.5    0.106 0.067  0.19
q10r      0.71      0.73    0.91      0.20 2.7    0.102 0.075  0.23
q11       0.70      0.72    0.88      0.19 2.5    0.106 0.066  0.19
q12       0.72      0.73    0.91      0.20 2.8    0.100 0.067  0.19

 Item statistics
      n raw.r std.r r.cor r.drop mean  sd
q1   18  0.54  0.55  0.57   0.42  4.6 1.7
q2   18  0.77  0.78  0.79   0.69  4.1 1.8
q3   18  0.44  0.42  0.40   0.27  3.4 2.0
q4r  18  0.33  0.32  0.29   0.16  3.1 1.9
q5r  18  0.49  0.50  0.46   0.34  4.8 1.9
q6   18  0.83  0.84  0.85   0.77  4.5 1.7
q7   18  0.36  0.32  0.25   0.16  3.4 2.3
q8   18  0.43  0.44  0.41   0.31  4.7 1.5
q9   18  0.52  0.53  0.51   0.40  1.9 1.6
q10r 18  0.41  0.45  0.41   0.29  3.1 1.5
q11  18  0.53  0.53  0.54   0.39  4.4 1.9
q12  18  0.40  0.40  0.37   0.24  2.8 1.9

Non missing response frequency for each item
        1    2    3    4    5    6    7 miss
q1   0.06 0.06 0.22 0.00 0.28 0.33 0.06    0
q2   0.11 0.11 0.11 0.22 0.28 0.06 0.11    0
q3   0.28 0.06 0.22 0.11 0.22 0.00 0.11    0
q4r  0.22 0.28 0.17 0.06 0.17 0.06 0.06    0
q5r  0.06 0.11 0.11 0.00 0.39 0.06 0.28    0
q6   0.06 0.00 0.28 0.17 0.22 0.11 0.17    0
q7   0.33 0.11 0.06 0.17 0.11 0.06 0.17    0
q8   0.00 0.06 0.22 0.17 0.17 0.28 0.11    0
q9   0.67 0.11 0.00 0.11 0.06 0.06 0.00    0
q10r 0.22 0.11 0.28 0.22 0.11 0.06 0.00    0
q11  0.17 0.00 0.11 0.06 0.33 0.28 0.06    0
q12  0.33 0.22 0.17 0.06 0.11 0.06 0.06    0
```

[1] If you only want to get reliability statistics for a subset of items, create a new data frame with a subset of the items (base R):
`newd <- subset(d, select=c(q1,q2,q3,q4r,q5r))`

Write-up

Internal reliability of the 12-item Internet Addiction Scale was investigated using Cronbach's alpha.[2] Results indicated that the alpha for the total scale was equal to .72. Examination of individual item statistics suggested that elimination of several items would increase the reliability of the scale. Subsequent analyses indicated that alpha could be improved after eliminating each of the following items individually: "I can go one full day without using the internet" (reverse-scored), "I get bored when I don't have my phone when I use the bathroom," "I like using social media," "I have more online friends than in-person friends," "I don't often use social media" (reverse-scored), and "I am comfortable staying in a place without internet access" (reverse-scored). The final reliability for the resulting six-item scale was considered acceptable, α = .82.

Comment: There are few important points to keep in mind. What is an acceptable alpha values is debatable. Many sources give .7 as minimally acceptable,[3] but as we saw here, the reliability of a measure can often be improved for a scale with an alpha of approximately this value. The above process is potentially problematic, because it involves multiple changes in the scale. The more changes that are made, the greater the risk that the final result will not be replicated in another sample. Moreover, all of the changes to the scale were based on empirical findings (*post hoc*) rather than guided by theory (*a priori*), and reviewers may criticize too many *post hoc* changes to a scale. It may be unwise to eliminate items without some theoretical explanation for why the item does not perform well, and it may be quite reasonable to decide to retain items purely for theoretical reasons (i.e., even though the results suggest the item may be dropped). Preferably, a process like this one would be used with a smaller pilot data set that would later be replicated with a larger study intended to be published. It is also important to keep in mind that items may perform poorly, because they are related to another construct rather than because they are necessarily "bad" items. Factor analytic approaches are better suited to examining whether there are multiple constructs underlying a set of items (a topic I will address briefly next term).

References

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334.

[2] It is not necessary any longer to cite Cronbach's paper, as the statistic is widely known, but the usual citation is to his seminal paper published in 1951.
[3] The .7 standard seems to be traced to Nunnally (1978), who was actually stating .7 may not be high enough: "what a satisfactory level of reliability is depends on how a measure is being used. In the early stages of research . . . one saves time and energy by working with instruments that have only modest reliability, for which purpose reliabilities of .70 or higher will suffice. . . . In contrast to the standards in basic research, in many applied settings a reliability of .80 is not nearly high enough." (pp. 245-246).