

## Measurement Error 2: Scale Construction (Very Brief Overview)

Richard Williams, University of Notre Dame, <https://www3.nd.edu/~rwilliam/>

Last revised January 22, 2015

This handout draws heavily from Marija Norusis's *SPSS 14.0 Statistical Procedures Companion*. See her chapters 18 (Reliability Analysis) and 17 (Factor Analysis), as well as Hamilton's (2004) ch. 12 (Principal Components, Factor, and Cluster Analysis) for much more information. I've adapted Norusis's example from ch. 18 to show how the analyses could be done.

As we have seen, individual items often suffer from random measurement error; and when several items all tap the same underlying concept, including them all separately in a regression can lead to problems of multicollinearity and unnecessarily complicated results. Therefore, it is sometimes desirable to create scales out of items. When individual items all tap the same concept, a well constructed scale will be more reliable than each item individually and will also be more parsimonious. When constructing scales, naturally we want to know how good the scale is. How reliable is the scale? What items belong in the scale? Both theoretical concerns and empirical results should guide you in scale construction.

The General Social Survey asked 1,334 respondents questions concerning anomie, i.e. the breakdown of social norms. It consists of 9 statements to which people answer either agree (coded 1) or disagree (coded 0). In SPSS, the scale = the number of statements with which they agree; in Stata that score is divided by the number of items in the scale. Here are the descriptive statistics:

```
. use https://www3.nd.edu/~rwilliam/statafiles/anomia.dta, clear
. describe an*
```

variable name	storage type	display format	value label	variable label
anomia1	byte	%8.0g	anomia1	next to health, money is most important
anomia2	byte	%8.0g	anomia2	wonder if anything is worthwhile
anomia3	byte	%8.0g	anomia3	no right and wrong ways to make money
anomia4	byte	%8.0g	anomia4	live only for today
anomia5	byte	%8.0g	anomia5	lot of average man getting worse
anomia6	byte	%8.0g	anomia6	not fair to bring child into world
anomia7	byte	%8.0g	anomia7	officials not interested in average man
anomia8	byte	%8.0g	anomia8	don't know whom to trust
anomia9	byte	%8.0g	anomia9	most don't care what happens to others

```
. sum an*
```

Variable	Obs	Mean	Std. Dev.	Min	Max
anomia1	3975	.3116981	.4632455	0	1
anomia2	3975	.4125786	.4923601	0	1
anomia3	3975	.247044	.4313468	0	1
anomia4	3975	.4445283	.4969759	0	1
anomia5	3975	.5836478	.4930154	0	1
anomia6	3975	.3889308	.4875689	0	1
anomia7	3975	.6332075	.4819898	0	1
anomia8	3975	.7177358	.4501578	0	1
anomia9	3975	.5554717	.4969759	0	1

As you can see, only 25% agree with the statement “there are no right or wrong ways to make money” (anomia3) whereas 72% agree that they “don’t know whom to trust” (anomia8).

In Stata, the alpha command can be used to assess how well these 9 items form a single scale measuring the same concept.

```
. alpha an*, c
```

```
Test scale = mean(unstandardized items)
```

```
Average interitem covariance:    .0550546
Number of items in the scale:    9
Scale reliability coefficient:    0.7412
```

(The `c` option does casewise/listwise deletion when computing correlations; the default in Stata is pairwise.) The “Scale reliability coefficient” in the above is Cronbach’s Alpha. It can be interpreted in a couple of different ways:

- It is the correlation between the present scale and all other possible nine-item scales measuring the same thing
- It is the squared correlation between the score a person obtains on a particular scale (the observed score) and the score he or she would have obtained if questioned on all the possible items in the universe (the true score)

The higher Cronbach’s Alpha, the better, i.e. you want the correlation between the observed value and the true value to be as high as possible. A rule of thumb is that .80 and above is considered pretty good; in this case Cronbach’s alpha is .7412, which is a little below that.

Cronbach’s Alpha will go up as you add more “good” items that measure the same underlying concept; however it can go down if you add other items that do not belong in the scale, e.g. items that measure some other concept. You should therefore examine how individual items are related to the overall scale. You do this by adding the `i` (short for `item`) option.

```
. alpha an*, i c
```

```
Test scale = mean(unstandardized items)
```

Item	Obs	Sign	item-test correlation	item-rest correlation	average inter-item covariance	alpha
anomia1	3975	+	0.4752	0.3094	.0591653	0.7351
anomia2	3975	+	0.5814	0.4238	.0543645	0.7167
anomia3	3975	+	0.5262	0.3809	.0575461	0.7237
anomia4	3975	+	0.5010	0.3259	.0577938	0.7336
anomia5	3975	+	0.6197	0.4706	.0527024	0.7084
anomia6	3975	+	0.6540	0.5156	.0513415	0.7006
anomia7	3975	+	0.5304	0.3664	.0566869	0.7263
anomia8	3975	+	0.6009	0.4629	.0543247	0.7107
anomia9	3975	+	0.6440	0.4998	.0515661	0.7031
Test scale					.0550546	0.7412

The item-test correlation shows how highly correlated each item is with the overall scale. The item-rest correlation (which may be more helpful; SPSS calls it the Corrected Item-Total Correlation) shows how the item is correlated with a scale computed from only the other 8 items. You want individual items that are correlated with the scale as a whole. You may want to drop items that do not correlate well with the scale, as they may not be measuring the same construct as the other variables.

The last column, labeled alpha (which SPSS labels as Cronbach's Alpha if Item Deleted) shows how the alpha for the scale would change if the item was deleted from the scale. In this particular case, removing any of the items would cause alpha to go down, i.e. the scale would become less reliable. That is an argument for keeping all the current items in the scale. If eliminating an item would substantially increase alpha, then you should consider removing that item from your scale. For example, let's see what happens when we add another variable, age, to the scale:

```
. alpha an* age, c i
```

```
Test scale = mean(unstandardized items)
```

Item	Obs	Sign	item-test correlation	item-rest correlation	average inter-item covariance	alpha
anomia1	3965	+	0.2519	0.2268	.0756966	0.0205
anomia2	3965	+	0.0717	0.0435	.1158421	0.0310
anomia3	3965	+	0.0795	0.0548	.1147608	0.0307
anomia4	3965	+	0.1099	0.0816	.1066084	0.0286
anomia5	3965	+	0.0695	0.0413	.116339	0.0312
anomia6	3965	+	0.1407	0.1131	.0995986	0.0268
anomia7	3965	+	0.0876	0.0600	.1122049	0.0301
anomia8	3965	+	0.1272	0.1016	.1040996	0.0280
anomia9	3965	+	0.0990	0.0706	.109236	0.0293
age	3965	+	0.9901	0.0611	.0550302	0.7411
Test scale					.1009416	0.0333

Clearly, age (at least as it is currently measured) would be a terrible addition to this scale. Alpha is much lower now. Further, the last column shows you that the alpha would be dramatically higher if age was dropped from the scale.

Of course, age is measured on a totally different scale than the other items, i.e. it is measured in years while the other items are just 0-1 dichotomies. If, for some reason, you believed it should be part of the scale, you'd probably want to standardize the variables first so all had a mean of 0 and sd of 1. You can do that with the `s` (short for `std`) option on `alpha`:

```
. alpha an* age, c i s
```

```
Test scale = mean(standardized items)
```

Item	Obs	Sign	item-test correlation	item-rest correlation	average inter-item correlation	alpha
anomia1	3965	+	0.5023	0.3406	0.2038	0.6974
anomia2	3965	+	0.5581	0.4064	0.1956	0.6864
anomia3	3965	+	0.5254	0.3676	0.2004	0.6929
anomia4	3965	+	0.4876	0.3236	0.2060	0.7001
anomia5	3965	+	0.5928	0.4482	0.1905	0.6793
anomia6	3965	+	0.6398	0.5059	0.1836	0.6694
anomia7	3965	+	0.5153	0.3559	0.2019	0.6948
anomia8	3965	+	0.5967	0.4530	0.1900	0.6785
anomia9	3965	+	0.6230	0.4852	0.1861	0.6730
age	3965	+	0.2495	0.0624	0.2410	0.7408
Test scale					0.1999	0.7142

This works a lot better, although you still see that the scale is better with age excluded than with it in. However, even if the inclusion of age had caused alpha to go up, you wouldn't want to include it unless you had good theoretical reasons for doing so. (In other words, you don't just use mindless empiricism when constructing your scales; theory should be guiding you as well.)

Finally, you can have alpha generate the scale as a new variable for you:

```
. quietly alpha an* , c gen(anscale)
. gen anscale2 = anscale*9
. sum anscale*
```

Variable	Obs	Mean	Std. Dev.	Min	Max
anscale	3975	.4772048	.2725445	0	1
anscale2	3975	4.294843	2.4529	0	9

As you see, when you multiply the scale by 9, you get the same results that Norusis reports using SPSS.

**Closing Comments.** This is a relatively simple problem. All items are measured the same way, and there are theoretical reasons for believing that all 9 measure a single underlying concept. In other situations, variables may be scaled in very different ways, they may reflect more than one underlying factor, and theory may not be clear as to what those factors are. In such cases, exploratory factor analysis may be more appropriate. See the texts for more details.

Appendix: SPSS Analysis (Optional) See ch. 18 (reliability) of Norusis's book for a detailed explanation. Using the menus, you click on Analyze \ Scale\ Reliability Analysis.

RELIABILITY

```

/VARIABLES=anomia1 anomia2 anomia3 anomia4 anomia5 anomia6 anomia7
anomia8 anomia9
/SCALE('ALL VARIABLES') ALL/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE CORR
/SUMMARY=TOTAL MEANS VARIANCE CORR .

```

## Reliability

[DataSet2] D:\SOC63993\SpssFiles\anomia.sav

### Scale: ALL VARIABLES

Case Processing Summary

		N	%
Cases	Valid	3975	100.0
	Excluded <sup>a</sup>	0	.0
	Total	3975	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.741	.741	9

Item Statistics

	Mean	Std. Deviation	N
anomia1 Next to health, money is most important	.31	.463	3975
anomia2 Wonder if anything is worthwhile	.41	.492	3975
anomia3 No right and wrong ways to make money	.25	.431	3975
anomia4 Live only for today	.44	.497	3975
anomia5 Lot of average man getting worse	.58	.493	3975
anomia6 Not fair to bring child into world	.39	.488	3975
anomia7 Officials not interested in average man	.63	.482	3975
anomia8 Don't know whom to trust	.72	.450	3975
anomia9 Most don't care what happens to others	.56	.497	3975

**Summary Item Statistics**

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	.477	.247	.718	.471	2.905	.024	9
Item Variances	.228	.186	.247	.061	1.327	.000	9
Inter-Item Correlations	.241	.105	.438	.332	4.157	.006	9

**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
anomia1 Next to health, money is most important	3.98	5.151	.309	.111	.735
anomia2 Wonder if anything is worthwhile	3.88	4.855	.424	.193	.717
anomia3 No right and wrong ways to make money	4.05	5.089	.381	.160	.724
anomia4 Live only for today	3.85	5.042	.326	.115	.734
anomia5 Lot of average man getting worse	3.71	4.761	.471	.254	.708
anomia6 Not fair to bring child into world	3.91	4.690	.516	.293	.701
anomia7 Officials not interested in average man	3.66	4.995	.366	.176	.726
anomia8 Don't know whom to trust	3.58	4.892	.463	.264	.711
anomia9 Most don't care what happens to others	3.74	4.694	.500	.294	.703

**Scale Statistics**

Mean	Variance	Std. Deviation	N of Items
4.29	6.017	2.453	9