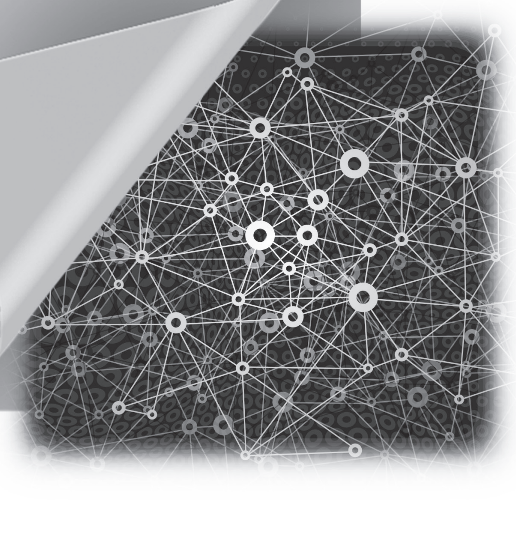
Exploratory Factor Analysis with SAS®

Jason W. Osborne · Erin S. Banjanovic

Exercise Solutions



support.sas.com/bookstore

Exploratory Factor Analysis with SAS®

ISBN 978-1-62960-064-2 (Hard copy)
ISBN 978-1-62960-241-7 (Epub)
ISBN 978-1-62960-242-4 (Mobi)
ISBN 978-1-62960-243-1 (PDF)

SAS Institute Inc., SAS Campus Drive, Cary, NC 27513-2414

March 2016

Exploratory Factor Analysis with SAS

End of Chapter Exercise Solutions

Please note, unless indicated otherwise, the syntax for each example is provided in the Exercise Solutions SAS syntax file.

Chapter 2

1. The results are presented in text and the syntax is captured in the SAS syntax file accompanying Chapter 2. Based on the results, all of the extraction methods yield similar results for the engineering and SDQ data. These data meet the assumptions of multivariate normality thus any method may be appropriate for these data (although the literature exhibits a preference for ML extraction under these conditions). The methods yield different results for the GDS data which violates assumptions of normality. The ULS or iterated PAF methods may be most appropriate for this data.

2. The data for this example is available on the book website and is called spq_osborne_1997.sas7bdat. Since the measure was designed to have three scales we extract three factors and compare the eigenvalues and communalities between the extraction methods. The results are presented in the tables below.

   Extracted Eigenvalues across Methods

Factor	Initial	ML	PAF	Iterated PAF	ULS	Alpha
Factor1	4.24	3.76	4.24	4.19	4.19	4.15
Factor2	1.47	1.97	1.47	1.55	1.56	1.28
Factor3	0.70	0.69	0.70	0.68	0.68	0.73
Factor4	0.60					
Factor5	0.51					
Factor6	0.49					
Factor7	0.41					
Factor8	0.33					
Factor9	0.20					
Factor10	0.12					

   Extracted Communalities across Methods

VAR	Initial	ML	PAF	Iterated PAF	ULS	Alpha
SPQ01	0.40	0.37	0.35	0.34	0.34	0.31
SPQ02	0.58	0.81	0.64	0.75	0.76	0.48
SPQ03	0.16	0.08	0.11	0.09	0.09	0.17
SPQ04	0.50	0.43	0.44	0.44	0.44	0.43
SPQ05	0.41	0.31	0.33	0.32	0.32	0.31
SPQ06	0.59	0.67	0.63	0.66	0.66	0.43
SPQ07	0.19	0.16	0.13	0.12	0.12	0.13
SPQ08	0.40	0.34	0.32	0.31	0.31	0.34
SPQ09	0.32	0.12	0.25	0.23	0.23	0.27

SPQ10	0.35	0.25	0.25	0.24	0.24	0.20
SPQ11	0.49	0.70	0.49	0.48	0.48	0.48
SPQ12	0.46	0.45	0.37	0.34	0.34	0.41
SPQ13	0.27	0.23	0.18	0.16	0.16	0.17
SPQ14	0.31	0.18	0.23	0.22	0.22	0.26
SPQ15	0.38	0.36	0.38	0.38	0.38	0.35
SPQ16	0.18	0.12	0.15	0.15	0.15	0.15
SPQ17	0.22	0.09	0.20	0.21	0.21	0.24
SPQ18	0.44	0.41	0.43	0.44	0.44	0.42
SPQ19	0.23	0.09	0.17	0.14	0.14	0.15
SPQ20	0.27	0.14	0.27	0.30	0.30	0.33
SPQ21	0.25	0.11	0.11	0.10	0.10	0.13

In general we can see there are differences in eigenvalues and communalities across the methods. If you run some descriptive statistics (e.g., means and frequencies) on the variables you will notice that the majority of the variables are highly skewed to one end of the seven point scale or the other. In addition, the full range of responses (i.e., 1-7) are not captured for a number of the variables. These data likely do not meet the assumptions of multivariate normality and thus the ULS or iterated PAF methods will provide the best extraction results.

Chapter 3

1. The results are presented in text and the syntax is captured in the SAS syntax file accompanying Ch 3.

2. The spq_osborne_1997.sas7bdat data should be used for this example. We will use the ULS extraction method for our analysis as we determined the data likely do not meet the assumptions of multivariate normality in the examples at the end of Chapter 2 (note we could also use the iterated PAF extraction method). The results of each of the extraction criteria are reviewed below.

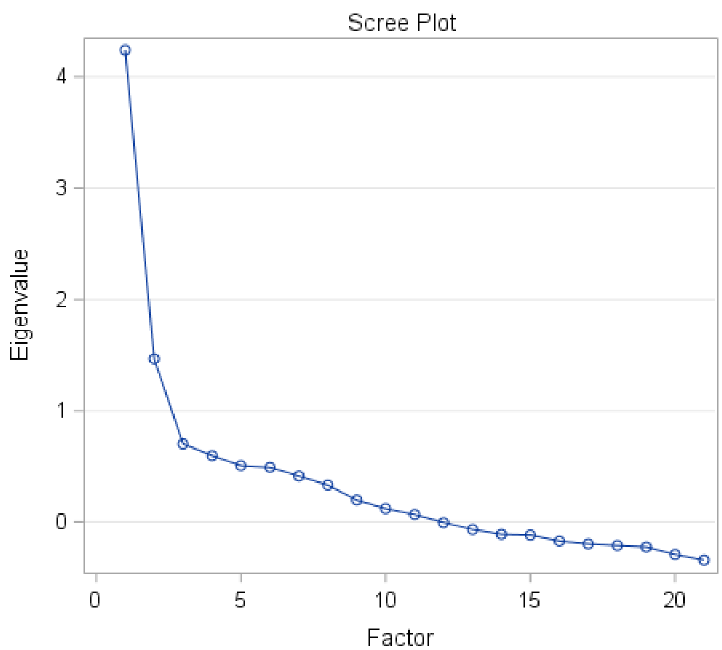
   Theory - The overall measure was designed to capture three scales thus theory would indicate there are three factors underlying the data.

   Kaiser Criterion, Minimum eigenvalue, & Proportion of variance – The table of preliminary eigenvalues are presented below. The Kaiser criterion would recommend two factors are extracted, the minimum eigenvalue would recommend seven factors are extracted, and the proportion of variance criterion would recommend five are extracted.

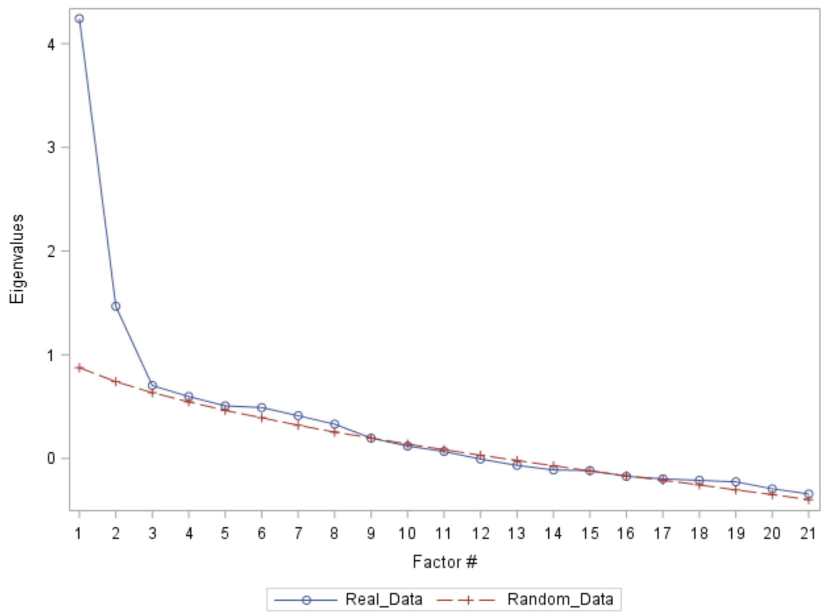
Preliminary Eigenvalues: Total = 7.3897546 Average = 0.35189308				
	Eigenvalue	Difference	Proportion	Cumulative
1	4.24211787	2.77498793	0.5741	0.5741
2	1.46712994	0.76386971	0.1985	0.7726
3	0.70326022	0.10656667	0.0952	0.8678
4	0.59669356	0.09013331	0.0807	0.9485
5	0.50656025	0.01594609	0.0685	1.0171
6	0.49061415	0.07807812	0.0664	1.0834
7	0.41253604	0.08192416	0.0558	1.1393
8	0.33061187	0.13456232	0.0447	1.1840
9	0.19604955	0.07729503	0.0265	1.2105
10	0.11875452	0.05248769	0.0161	1.2266
11	0.06626683	0.07196638	0.0090	1.2356
12	-.00569954	0.06106570	-0.0008	1.2348
13	-.06676524	0.04346906	-0.0090	1.2258
14	-.11023430	0.00672823	-0.0149	1.2109
15	-.11696253	0.05525257	-0.0158	1.1950
16	-.17221510	0.02466445	-0.0233	1.1717
17	-.19687955	0.01453062	-0.0266	1.1451
18	-.21141017	0.01451608	-0.0286	1.1165
19	-.22592626	0.06679000	-0.0306	1.0859
20	-.29271626	0.04931497	-0.0396	1.0463
21	-.34203123		-0.0463	1.0000

5 factors will be retained by the PROPORTION criterion.

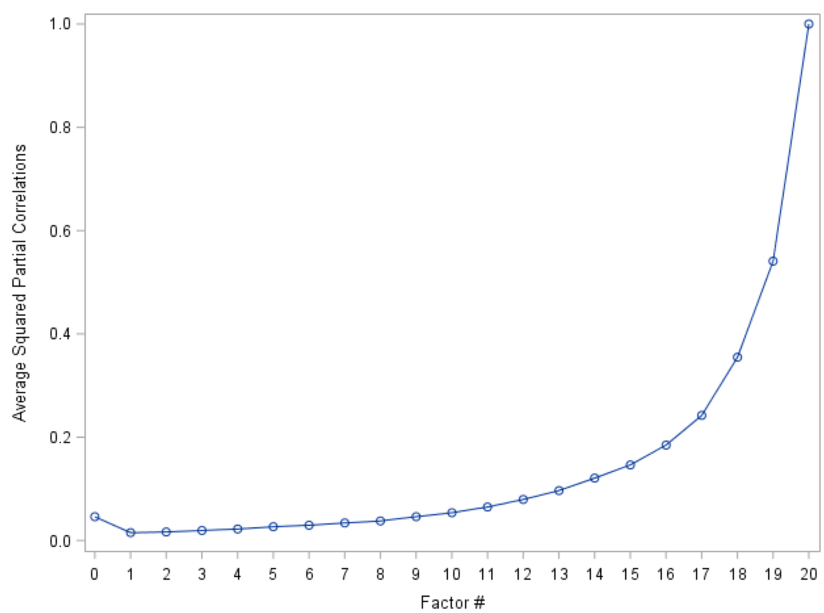
Scree Plot – The scree plot is presented below. The scree plot recommends extraction of 2, 3 or 5 factors.



Parallel Analysis - A plot of the parallel analysis results is presented below. These results are not very clear or useful as the random eigenvalues are very close but just below the real eigenvalues for factors three to eight. The real go below the random at factor nine, technically recommending eight factors are extracted.



MAP Analysis – A plot depicting the results of the MAP analysis is presented below. The MAP analysis recommends two factors are extracted.



Summary – The different criteria give us different recommendations for the number of factors to extract, suggesting we must explore several options. Two, three and five factor solutions are supported by more than one criterion thus these may be good starting points for exploration.

Chapter 4

1. A. The factor loading plots for the engineering data are presented in text and the syntax to produce them is captured in the SAS syntax file accompanying Ch 4. In general, as factors become correlated the oblique methods will produce clearer rotation results. If the factors are not correlated, the methods will yield nearly identical results.

   B. The three factor pattern matrix is presented below. These factors are found to account for 32.2% of the variance and the communalities range from 0.14-0.56, not too far below the variance and communalities of the five factor solution. Only one item (GDS27) is not found to have a loading greater than 0.30 on any of the factors. If we were to continue with this solution we may consider dropping that item. The three factors make conceptual sense: the first factor could be interpreted as depression and worry, the second as activity and mental clarity, and the third as positivity. These are similar to the three factors of the five factor model that made sense. Overall, this solution makes more sense than the five and eight factor model and represents more facets than the one factor model.

Rotated Factor Pattern

Item and Stem		Factor1	Factor2	Factor3
GDS16	Do you often feel downhearted and blue?	0.64	.	.
GDS25	Do you frequently feel like crying?	0.63	.	.
GDS06	Are you bothered by thoughts you can't get out of your head?	0.58	.	.
GDS03	Do you feel that your life is empty?	0.53	.	0.33
GDS18	Do you worry a lot about the past?	0.51	.	.
GDS23	Do you think that most people are better off than you are?	0.50	.	.
GDS22	Do you feel that your situation is hopeless?	0.48	.	.
GDS10	Do you often feel helpless?	0.47	.	.
GDS24	Do you frequently get upset over little things?	0.45	.	.
GDS04	Do you often get bored?	0.45	.	.
GDS08	Are you afraid that something bad is going to happen to you?	0.44	.	.
GDS11	Do you often get restless and fidgety?	0.43	.	.
GDS13	Do you frequently worry about the future?	0.41	.	.
GDS17	Do you feel pretty worthless the way you are now?	0.37	0.31	.
GDS21	Do you feel full of energy?	.	0.66	.
GDS20	Is it hard for you to get started on new projects?	.	0.56	.
GDS30	Is your mind as clear as it used to be?	.	0.52	.
GDS19	Do you find life very exciting?	.	0.51	.
GDS12	Do you prefer to stay at home, rather than going out and doing new things?	.	0.45	.
GDS02	Have you dropped many of your activities and interests?	.	0.45	.
GDS26	Do you have trouble concentrating?	.	0.44	.
GDS28	Do you prefer to avoid social gatherings?	.	0.36	.
GDS29	Is it easy for you to make decisions?	.	0.34	.
GDS09	Do you feel happy most of the time?	.	.	0.58
GDS15	Do you think it is wonderful to be alive now?	.	.	0.52
GDS07	Are you in good spirits most of the time?	0.32	.	0.48
GDS05	Are you hopeful about the future?	.	.	0.46
GDS01	Are you basically satisfied with your life?	0.34	.	0.41
GDS27	Do you enjoy getting up in the morning?	.	.	.
GDS14	Do you feel you have more problems with memory than most?	.	.	-0.33

Values less than 0.3 are not printed.

2. These three scales are hypothesized to be related and thus an oblique method should be used with the data. The promax and oblimin rotation methods provide similar, but slightly different results (presented below). For both solutions, the first factor generally be interpreted as importance of good grades (note a number of these items with negative items should be recoded), the second factor in the promax solution and the third in the oblimin could be interpreted as feelings about failure, and the last could be feelings of discrimination. The promax solution may be slightly easier to interpret as it has fewer items with loadings greater than .3 on multiple factors.

		Rotated Factor Pattern					
		Promax			Oblimin		
Item and Stem		Factor1	Factor2	Factor3	Factor1	Factor2	Factor3
SPQ18	If I were to list all the things that are important in my life, school would be near the top of the list.	0.55	.	.	0.65	.	.
SPQ09	I always feel good about myself when I do well on a test.	0.55	.	.	0.48	.	.
SPQ16	There is no better feeling in the world than when I get an A on an assignment.	0.44	.	.	0.36	.	.
SPQ19	Sometimes I get depressed or “bummed out” when I don’t do as well on an assignment as I wanted to.	0.43	.	.	0.37	.	.
SPQ04	Being a good student is an important part of who I am	0.40	-0.33	.	0.58	.	.
SPQ01	Doing well in school is very important to me.	0.35	-0.31	.	0.52	.	.
SPQ05	I feel that the grades I get are an accurate reflection of my abilities.	0.32	.	.	0.48	.	.
SPQ07	It is very important to do well on tests.	0.30	.	.	0.34	.	.
SPQ21	If I don’t get at least a B, I feel like I’ve failed.	.	.	.	.	.	.
SPQ13	No test will ever change my opinion of how smart I am.	-0.35	.	.	-0.39	.	.
SPQ08	My grades do not tell me anything about my academic potential.	-0.40	.	.	-0.52	.	.
SPQ10	School has no relationship to how I do in life.	-0.40	.	.	-0.48	.	.
SPQ12	I really don't care what tests say about my intelligence	-0.42	.	.	-0.55	.	.
SPQ20	I am often relieved if I just pass a course.	.	0.63	.	.	.	0.50
SPQ15	If I were to fail out of school I would still feel okay about myself, because school is not really important to me.	.	0.57	.	-0.37	.	0.40
SPQ11	It usually doesn't matter to me one way or the other how I do in school.	.	0.56	.	-0.51	.	0.37
SPQ17	If the tests we take were fair, I would be doing much better in school.	.	0.45	.	.	.	0.39
SPQ14	Doing well in school is a matter of luck, not intelligence.	.	0.36	.	-0.31	.	.
SPQ02	I feel that my education has suffered because of prejudice or discrimination.	.	.	0.87	.	0.86	.
SPQ06	I believe that I have been personally discriminated against because of my race, age, or sex, and that this discrimination has hurt my grades.	.	.	0.77	.	0.76	.
SPQ03	In general, whites have an advantage in school.	.	.	0.30	.	0.30	.
Values less than 0.3 are not printed.							

Chapter 5

1. Using a seed value of 5, 7, 9, and 11 for ratios A-D, ULS extraction and oblimin rotation, the following results were produced. Please note if you would like to replicate these results the seed values should be used when using proc surveyselect. Overall, the results of the smaller samples are not too bad in this example. They do generally tend to replicate, although the results in the larger samples may be slightly closer to the full sample results. These results could be much different in another, more volatile dataset.

Pattern Matrix Results

		Full Sample			**2:1 Ratio**		
Item and Description		**Factor1**	**Factor2**	**Factor3**	**Factor1**	**Factor2**	**Factor3**
Math1	MATHEMATICS IS ONE OF R^S BEST SUBJECTS	0.90	.	.	.	0.85	.
Math2	R HAS ALWAYS DONE WELL IN MATHEMATICS	0.86	.	.	.	0.81	.
Math3	R GETS GOOD MARKS IN MATHEMATICS	0.89	.	.	.	0.84	.
Math4	R DOES BADLY IN TESTS OF MATHEMATICS	-0.59	.	.	.	-0.69	.
Par1	R^S PARENTS TREAT R FAIRLY	.	0.71	.	.	.	0.66
Par2	R DOES NOT LIKE HIS PARENTS VERY MUCH	.	-0.69	.	.	.	-0.52
Par3	R GETS ALONG WELL WITH HIS/HER PARENTS	.	0.82	.	.	.	0.77
Par4	PARENTS DISAPPOINTED WITH WHAT R DOES	.	-0.60	.	-0.32	.	-0.55
Par5	R^S PARENTS UNDERSTAND HIM/HER	.	0.74	.	.	.	0.89
Eng1	LEARN THINGS QUICKLY IN ENGLISH CLASSES	.	.	0.78	0.94	.	.
Eng2	ENGLISH IS ONE OF R^S BEST SUBJECTS	.	.	0.84	0.67	.	.
Eng3	R GETS GOOD MARKS IN ENGLISH	.	.	0.85	0.77	.	.
Eng4	R HOPELESS IN ENGLISH CLASSES	.	.	-0.61	-0.79	.	.
Values less than 0.3 are not printed.							

	5:1 Ratio			**10:1 Ratio**			**20:1 Ratio**		
Var	**Factor1**	**Factor2**	**Factor3**	**Factor1**	**Factor2**	**Factor3**	**Factor1**	**Factor2**	**Factor3**
Math1	0.90	.	.	0.88	.	.	.	0.88	.
Math2	0.85	.	.	0.89	.	.	.	0.78	.
Math3	0.86	.	.	0.94	.	.	.	0.90	.
Math4	-0.52	.	.	-0.58	.	.	.	-0.59	.
Par1	.	0.52	.	.	0.70	.	0.74	.	.
Par2	.	-0.51	.	.	-0.72	.	-0.70	.	.
Par3	.	0.95	.	.	0.80	.	0.83	.	.
Par4	.	-0.66	.	.	-0.52	.	-0.69	.	.
Par5	.	0.70	.	.	0.78	.	0.71	.	.
Eng1	.	.	0.72	.	.	0.65	.	.	0.79
Eng2	.	.	0.83	.	.	0.90	.	.	0.78
Eng3	.	.	0.76	.	.	0.85	.	.	0.84
Eng4	.	.	-0.57	.	.	-0.53	.	.	-0.65
Values less than 0.3 are not printed.									

2. Use online research in your field of study to address this question. In general, we support the recommendation of at least 10 participants per parameter estimated but strongly encourage researchers to try to get closer to 20. Keep in mind these are just recommendations and you may need more or less depending on the clarity of the factor structure.

Chapter 6

1. The results for this analysis are presented in text and the syntax to produce them is captured in the SAS syntax file accompanying Ch 6.

2. The pattern matrix results for the two samples (using iterated PAF extraction and oblimin rotation) are presented below, along with the corresponding factor loading squared differences. The seed value of 2 was used to obtain these exact groupings. The two solutions were found to have identical factor structures and the squared differences are fairly small. The largest difference is .02, which corresponds to a difference in factor loadings of .14. This falls below our value of .04 (corresponding to a .20 difference in factor loadings) that suggests the loadings may be volatile. Overall, we find the strong factor structure previously observed does indeed replicate.

	Subsample1		**Subsample2**		**Squared Differences**
Var	**Factor1**	**Factor2**	**Factor1**	**Factor2**	
EngProbSolv1	0.82	.	0.88	.	0.0033
EngProbSolv2	0.76	.	0.91	.	0.0226
EngProbSolv3	0.82	.	0.92	.	0.0116
EngProbSolv4	0.88	.	0.93	.	0.0019
EngProbSolv5	0.87	.	0.90	.	0.0014
EngProbSolv6	0.85	.	0.88	.	0.0009
EngProbSolv7	0.85	.	0.89	.	0.0016
EngProbSolv8	0.78	.	0.79	.	0.0001
INTERESTeng1	.	0.78	.	0.82	0.0020
INTERESTeng2	.	0.88	.	0.96	0.0051
INTERESTeng3	.	0.92	.	0.93	0.0001
INTERESTeng4	.	0.88	.	0.93	0.0026
INTERESTeng5	.	0.86	.	0.93	0.0047
INTERESTeng6	.	0.83	.	0.89	0.0030
Values less than 0.3 are not printed.					

3. The pattern matrix results for the two samples (using ULS extraction and oblimin rotation) are presented below, along with the corresponding factor loading squared differences. The seed value of 26 was used to obtain these exact groupings. Ten of the 21 items failed to replicate the basic structure in these samples. Among the remaining items with a replicated structure the squared differences in item loadings ranged from .0003 to .0673. Some of these differences exceed what we consider reasonable and thus we would conclude these factor loadings to be volatile and the solution to not replicate. Larger samples may be necessary to further explore the scale.

	Subsample1			Subsample2			Squared Differences
	Factor1	Factor2	Factor3	Factor1	Factor2	Factor3	
SPQ01	0.61	.	.	-0.49	.	.	0.0142
SPQ02	.	.	0.87	.	0.93	.	0.0091
SPQ03	.	.	0.43	.	.	.	failed
SPQ04	0.56	.	.	-0.62	.	.	0.0041
SPQ05	0.73	.	.	-0.47	-0.33	.	0.0673
SPQ06	.	.	0.79	.	0.87	.	0.0336
SPQ07	.	-0.36	.	-0.31	.	.	failed
SPQ08	-0.44	.	.	0.60	.	.	0.0252
SPQ09	0.36	-0.35	.	.	.	0.77	failed
SPQ10	.	.	.	0.57	.	.	failed
SPQ11	-0.58	0.31	.	0.59	.	.	0.0003
SPQ12	.	0.77	.	0.70	.	.	failed
SPQ13	.	0.36	.	0.53	.	.	failed
SPQ14	-0.53	.	.	0.50	.	.	0.0011
SPQ15	-0.47	.	.	0.61	.	.	0.0193
SPQ16	0.39	.	.	.	.	.	failed
SPQ17	.	.	.	.	.	.	failed
SPQ18	0.60	.	.	-0.47	.	.	0.0176
SPQ19	.	-0.44	.	.	.	0.54	0.0107
SPQ20	.	.	.	0.38	.	.	failed
SPQ21	.	-0.39	.	.	.	.	failed
Values less than 0.3 are not printed.							

Chapter 7

1. The results for this analysis are presented in text and the syntax to produce them is captured in the SAS syntax file accompanying Ch 7.

2. The results for this analysis are presented in text and the syntax to produce them is captured in the SAS syntax file accompanying Ch 7. Note the CI results will be different before and after running the %alignedFactors macro (NOTE TYPO in text -- %checkFactorOrder should be %alignedFactors) because the macro checks and occasionally switches the factor a set of variables loaded on and the direction of the loading. These changes are made in order to maintain consistency in the order and direction of the loadings so CI can be computed. When this is not done then you will get CI that might be computed with incorrect loadings or loadings that were extracted in the opposite direction (e.g., -.5 vs .5).

3. The original eigenvalues along with their bootstrapped 95% CI are presented below. Please note we used a seed value of 16 to select the bootstrap resamples and we had to increase the default number of iterations for rotation to 500 (from 300; RITER=500) to allow a few resamples to converge on a solution. Since the purpose behind bootstrap resampling is to get an idea of the range in a potential estimate and

a limited occurrence of outliers in the distribution is almost expected, a few estimates that do not converge should not be too problematic. However, with these changes we must be careful to interpret the results and quick to question them.

These results do not really help to clarify the factor structure. The 95% CI around the eigenvalues are rather large with the third factor having a 95% CI from .98 to 1.43. If we were to evaluate these results using the Kaiser criterion, these results do generally support a three factor structure. If the 1 was further centered in the 95% CI these results would not support that conclusion. If we examine the results further we notice that some of the eigenvalues are not captured within the corresponding 95% CI. This makes us question the precision of our estimates at larger factor numbers.

Factor	Eigenvalue	95% CI
1	7.26	(6.34,8.33)
2	1.46	(1.37,1.96)
3	1.04	(0.98,1.43)
4	0.67	(0.73,1.07)
5	0.55	(0.62,0.86)
6	0.48	(0.52,0.73)
7	0.43	(0.44,0.62)
8	0.33	(0.36,0.53)
9	0.30	(0.29,0.44)
10	0.23	(0.23,0.37)
11	0.17	(0.18,0.3)
12	0.09	(0.14,0.24)
13	0.07	(0.09,0.18)
14	0.03	(0.06,0.14)
15	0.02	(0.03,0.1)
16	0.00	(-0.01,0.07)
17	-0.02	(-0.03,0.03)
18	-0.05	(-0.06,0)
19	-0.07	(-0.08,-0.03)
20	-0.10	(-0.11,-0.05)
21	-0.12	(-0.13,-0.08)
22	-0.13	(-0.15,-0.11)
23	-0.14	(-0.17,-0.13)
24	-0.16	(-0.2,-0.15)
25	-0.17	(-0.22,-0.17)
26	-0.19	(-0.24,-0.19)
27	-0.24	(-0.26,-0.21)
28	-0.26	(-0.28,-0.24)
29	-0.27	(-0.31,-0.26)
30	-0.30	(-0.34,-0.28)

Chapter 8

1. We used proc freq and proc mi to examine some descriptives and to examine patterns in the data. We identified 58 cases with missing data for all of the variables and an additional 36 that were missing data for at least half of the items. These cases were removed. Next we imputed missing data using the EM algorithm. This left us with a final sample of 562 cases. We then ran the EFA with the reduced sample using the covariance matrix as our input dataset to look for potential items that may not belong in the scale. The results of the pattern matrix are presented below. Three items (GDS08, GDS18, and GDS29) are found to have loadings below .3 on the factors and should be considered for removal. Two additional items are found to have relatively weak loadings on two separate factors (GDS17 and GDS25) and may also be considered for removal after reviewing the content.

Rotated Factor Pattern (Standardized Regression Coefficients)					
	Factor1	**Factor2**	**Factor3**	**Factor4**	**Factor5**
GDS09	0.79	.	.	.	.
GDS07	0.72	.	.	.	.
GDS15	0.67	.	.	.	.
GDS01	0.49	.	.	.	.
GDS05	0.48	.	.	.	.
GDS27	0.31	.	.	.	.
GDS08	.	.	.	.	.
GDS04	.	0.64	.	.	.
GDS11	.	0.57	.	.	.
GDS24	.	0.47	.	.	.
GDS16	.	0.44	.	.	.
GDS06	.	0.35	.	.	.
GDS13	.	0.33	.	.	.
GDS18	.	.	.	.	.
GDS21	.	.	0.50	.	.
GDS12	.	.	0.48	.	.
GDS20	.	.	0.45	.	.
GDS28	.	.	0.45	.	.
GDS02	.	.	0.44	.	.
GDS19	.	.	0.44	.	.
GDS17	.	.	0.34	0.34	.
GDS22	.	.	.	0.60	.
GDS03	.	.	.	0.41	.
GDS25	.	0.34	.	0.36	.
GDS23	.	.	.	0.33	.
GDS10	.	.	.	0.32	.
GDS14	.	.	.	.	0.62
GDS26	.	.	.	.	0.53
GDS30	.	.	.	.	0.52
GDS29	.	.	.	.	.
Values less than 0.3 are not printed.					

2. The results for this analysis are presented in text and the syntax to produce them is captured in the SAS syntax file accompanying Ch 8.

3. This exercise does repeat some of the steps that were incorporated into our solution for exercise 1. Further direction may be required if these exercises are used in a class to encourage less repetition across the exercises. Unfortunately these solutions were written sometime after the exercises. We address this exercise by conducting the same data cleaning as done in exercise 1, and then comparing the results after using EM imputation to those using the default of missing case deletion. The resulting pattern matrices are presented below. In general, the results are fairly similar; however the imputed sample has fewer items exhibiting cross-loadings on factors. The imputed sample results are thus slightly easier to interpret.

Rotated Factor Pattern (Standardized Regression Coefficients)						
	Non-Imputed Sample			Imputed Sample		
	Factor1	Factor2	Factor3	Factor1	Factor2	Factor3
GDS16	0.64	.	.	0.61	.	.
GDS25	0.63	.	.	0.65	.	.
GDS06	0.58	.	.	0.51	.	.
GDS03	0.53	.	0.33	0.48	.	0.34
GDS18	0.51	.	.	0.48	.	.
GDS23	0.50	.	.	0.52	.	.
GDS22	0.48	.	.	0.47	.	.
GDS10	0.47	.	.	0.43	.	.
GDS24	0.45	.	.	0.44	.	.
GDS04	0.45	.	.	0.46	.	.
GDS08	0.44	.	.	0.35	.	.
GDS11	0.43	.	.	0.48	.	.
GDS13	0.41	.	.	0.42	.	.
GDS17	0.37	0.31	.	0.40	.	.
GDS21	.	0.66	.	.	0.66	.
GDS20	.	0.56	.	.	0.57	.
GDS30	.	0.52	.	.	0.48	.
GDS19	.	0.51	.	.	0.50	.
GDS12	.	0.45	.	.	0.45	.
GDS02	.	0.45	.	.	0.46	.
GDS26	.	0.44	.	.	0.42	.
GDS28	.	0.36	.	.	0.34	.
GDS29	.	0.34	.	.	0.32	.
GDS09	.	.	0.58	.	.	0.65
GDS15	.	.	0.52	.	.	0.58
GDS07	0.32	.	0.48	.	.	0.58
GDS05	.	.	0.46	.	.	0.47
GDS01	0.34	.	0.41	.	.	0.45
GDS27	.	.	.	.	.	.
GDS14	.	.	-0.33	.	.	.
Values less than 0.3 are not printed.						

Chapter 9

1. The solutions and syntax for A and B are discussed in the chapter and reviewed in the SAS syntax file accompanying the chapter. The solution for C is a straightforward mean (or it could also be a sum!) of the 8 items in the problem solving scale and the 6 in the interest scale.

2. The distribution of the respective factor scores are as follows and their corresponding correlations are as follows. Notice the factor scores from the scoring coefficients are centered at zero and the improper factor scores have a larger mean and generally a larger range of values. These differences may make one method more useful than another in a particular context. Additionally, the factor scores that represent the same factor are highly correlated across methods and exhibit similar correlations with the other factor score.

Variable	N	Mean	Std Dev	Minimum	Maximum
scoreCoeffF1	372	0.00	0.98	-3.35	1.63
scoreCoeffF2	372	0.00	0.98	-5.26	1.12
patternF1	372	2.48	0.59	0.49	3.46
patternF2	372	2.28	0.36	0.38	2.69
improperF1	372	5.02	1.20	1.00	7.00
improperF2	372	5.98	0.93	1.00	7.00

Prob > |r| under H0: Rho=0						
	scoreCoeffF1	**scoreCoeffF2**	**patternF1**	**patternF2**	**improperF1**	**improperF2**
scoreCoeffF1	1	0.38253	0.999	0.37276	0.99864	0.36856
		<.0001	<.0001	<.0001	<.0001	<.0001
scoreCoeffF2	0.38253	1	0.37139	0.99831	0.37155	0.99755
	<.0001		<.0001	<.0001	<.0001	<.0001
patternF1	0.999	0.37139	1	0.3617	0.99996	0.35749
	<.0001	<.0001		<.0001	<.0001	<.0001
patternF2	0.37276	0.99831	0.3617	1	0.36171	0.99973
	<.0001	<.0001	<.0001		<.0001	<.0001
improperF1	0.99864	0.37155	0.99996	0.36171	1	0.35739
	<.0001	<.0001	<.0001	<.0001		<.0001
improperF2	0.36856	0.99755	0.35749	0.99973	0.35739	1
	<.0001	<.0001	<.0001	<.0001	<.0001	

3. The average differences in the factor scores are summarized below. The same subjects were assigned factor scores that differed -.23 to .001 on factor 1 and -.25 to .35 on factor 2. These differences are attributed to the differences in the samples with which the models were estimated.

Variable	N	Mean	Std Dev	Minimum	Maximum
diff_F1	22	-0.0745	0.0557	-0.2272	0.0011
diff_F2	22	-0.0118	0.1190	-0.2457	0.3516

Chapter 10

1. The results for this analysis are presented in text and the syntax to produce them is captured in the SAS syntax file accompanying Ch 10.

2. The inter-factor correlations are presented below. In general these factors exhibit fairly weak correlations and do not support the notion of a higher-order factor. We do however still provide the eigenvalues and pattern matrix of the high-order factor analysis below. Notice the eigenvalues are small and the factor loadings indicate unequal contributions to the overall construct. It appears the second factor, parent relationships, primarily drives the overall construct.

   Inter-factor correlations

	Factor1	Factor2	Factor3
Factor1	1.00	0.23	0.14
Factor2	0.23	1.00	0.27
Factor3	0.14	0.27	1.00

Preliminary Eigenvalues: Total = 0.24461698				
Average = 0.08153899				
	Eigenvalue	Difference	Proportion	Cumulative
1	0.51	0.58	2.10	2.10
2	-0.07	0.13	-0.29	1.81
3	-0.20		-0.81	1.00

Factor Pattern	
	Factor1
Factor1	0.34
Factor2	0.66
Factor3	0.41

3. Explore with your own data!

Chapter 11

1. Cronbach’s alpha is an indicator of the internal consistency of the scale. It can conceptually be thought of as the correlation between one test and an alternate form of the test that contains the same number of items. An alpha of .85 suggests there is a correlation of .92 between the actual score and the true score that would be captured through perfect measurement. This estimate includes approximately 15% error variance.

2. The results for this analysis are presented in text and the syntax to produce them is captured in the SAS syntax file accompanying Ch 11.

3. The item-total correlations and alphas for the two scales are presented below. We review the item-total correlations to make sure all items are coded in the same direction and find that no recoding is necessary. The raw item-totals and Cronbach alpha should be interpreted for these results since the raw data is input and not standardized data (e.g., z-scores). In examining the item-total correlations and alpha with deleted variables, it appears that the scale reliability cannot be further improved by removing any of the variables.

Problem Solving Scale

Cronbach Coefficient Alpha	
Variables	**Alpha**
Raw	0.9588
Standardized	0.959151

Cronbach Coefficient Alpha with Deleted Variable				
Deleted Variable	**Raw Variables**		**Standardized Variables**	
	Correlation with Total	**Alpha**	**Correlation with Total**	**Alpha**
EngProbSolv1	0.84	0.95	0.83	0.95
EngProbSolv2	0.80	0.96	0.80	0.96
EngProbSolv3	0.86	0.95	0.86	0.95
EngProbSolv4	0.88	0.95	0.88	0.95
EngProbSolv5	0.87	0.95	0.87	0.95
EngProbSolv6	0.86	0.95	0.86	0.95
EngProbSolv7	0.86	0.95	0.86	0.95
EngProbSolv8	0.80	0.96	0.80	0.96

Interest Scale

Cronbach Coefficient Alpha	
Variables	**Alpha**
Raw	0.955323
Standardized	0.955923

Cronbach Coefficient Alpha with Deleted Variable				
Deleted Variable	**Raw Variables**		**Standardized Variables**	
	Correlation with Total	**Alpha**	**Correlation with Total**	**Alpha**
INTERESTeng1	0.80	0.95	0.80	0.95
INTERESTeng2	0.89	0.94	0.89	0.94
INTERESTeng3	0.89	0.94	0.89	0.94
INTERESTeng4	0.88	0.94	0.88	0.95
INTERESTeng5	0.87	0.95	0.87	0.95
INTERESTeng6	0.84	0.95	0.84	0.95

Using a seed of 1 to select a bootstrap sample containing 2000 resamples, a 95% confidence interval of .95 to .97 was found for the Problem Solving scale and a 95% confidence interval of .94 to .97 was found for the Interest scale. These ranges are relatively narrow and suggest these finding are likely to replicate in similar samples.

Overall, the alphas for the two engineering scales are relatively high, capturing only about 5% error variance. They suggest we would find a correlation of approximately .98 between a score produced by these measures and the “true” score. Furthermore, the confidence intervals are narrow, suggesting these relatively high alphas are likely to replicate. These results support good quality of measurement for these scales.