

Learning Centre

Factor Analysis



Content



What is factor analysis?



Types of factor analysis



Steps in conducting a factor analysis



Worked example

What is factor analysis



Factor analysis is a data reduction technique



Associate and group variables into factors



Consider the Big Five personality test:

How did researchers arrive at the 5 factors, from thousands of individual items and questions?

Factor analysis helped cluster those items into the 5 personality factors that we know today, while also identifying redundant items

Types of Factor Analyses

2 main branches: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA)

We will only focus on EFA

Exploratory Factor Analysis



Typically used in earlier stages of research, no concrete expectations of underlying factor structure



Thus, the number of factors will be determined by the data



For example:

A questionnaire of 100 items subjected to EFA suggested 70 redundant items

From the remaining 30 items, they loaded onto 3 factors, suggesting a 3-factor structure, with 10 items each

Exploratory Factor Analysis

EFA is also separated into different factor extraction types, with the more common ones being **Principal Axis Factoring (PAF)** and **Principal Component Analysis (PCA)**

PAF uses common variance to generate a factor solution, while PCA looks at total variance

Exploratory Factor Analysis

- Next, data is usually rotated after factor extraction
- This is done by rotating factor axes to best produce a simple structure
-
- This makes the factors easier to distinguish from one another
 - E.g., items on the Big Five measuring extraversion are clearly distinct from items measuring agreeableness
 - No item should be measuring both extraversion and agreeableness (complex structure)

Exploratory Factor Analysis

There are 2 common types of rotations:

- **Orthogonal** → assume factors are not correlated
 - Methods include Varimax, Equamax, Quartimax
- **Oblique** → assume factors are correlated
 - Methods include Direct Oblimin, Promax

Steps in Conducting Factor Analysis

Assess suitability of data to be subjected to factor analysis



Select factor extraction method (e.g., PCA, PAF)



Select rotation method (e.g., Varimax, Promax)



Interpreting the factor structure

An Example...

A team of researchers wants to develop a questionnaire that measures the characteristics of statistics teachers. They thought of 10 valid questions that is supposed to assess said characteristics, and sent the survey out to 300 students. After collecting the responses, they then carried out an exploratory factor analysis...

Location of the SPSS Data File for Practice

Example SPSS data for practice are available on LearnJCU:

Log in to LearnJCU -> Organisations -> Learning Centre JCU Singapore ->
Statistics Support -> Statistics Resources -> SPSS Data for Practice

Before we begin...

The following slides will go through each step of the factor analysis procedure separately for an explanation purpose, but *they can be conducted together in a single step*

Assessing Suitability of Data

1

Kaiser-Meyer-Olkin (KMO) for sampling adequacy (this indicates the proportion of variance that is explained by underlying factors)

2

Barlett's Test for Sphericity (this indicates if items are related)

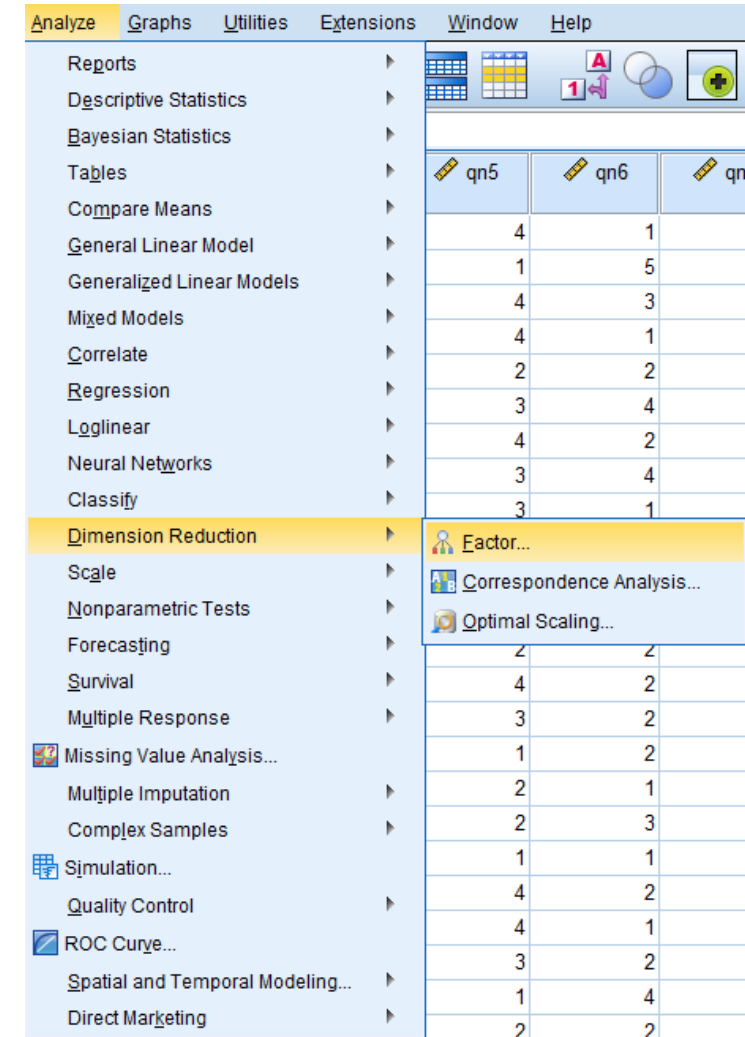
3

Anti-image matrices to further evaluate suitability of the data

KMO and Bartlett's test

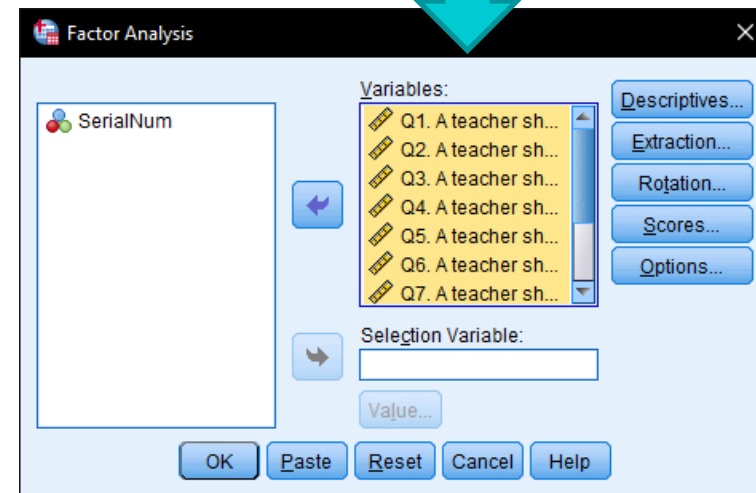
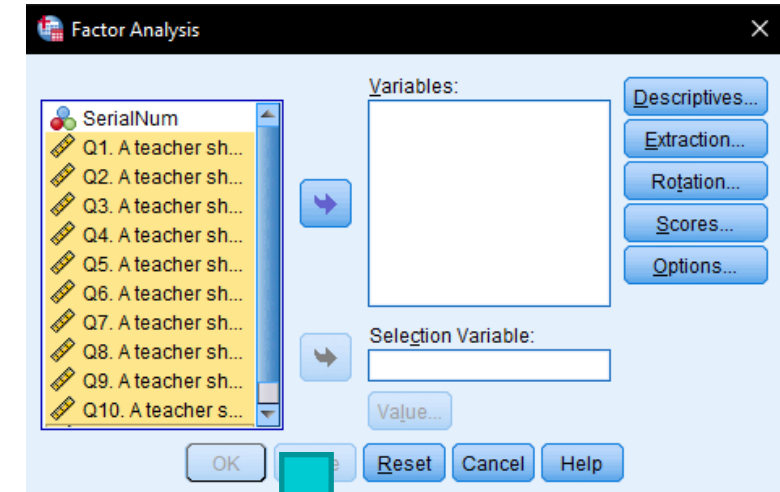
Analyze → Dimension
Reduction → Factor...

- **KMO:** A measure of sampling adequacy (MSA) measuring if the data are suitable for factor analysis. It indicates how much each item is predicted by the other items in the data
- **Bartlett's test:** A measure of overall correlation between the items. Items should be significantly correlated with each other for factor analysis

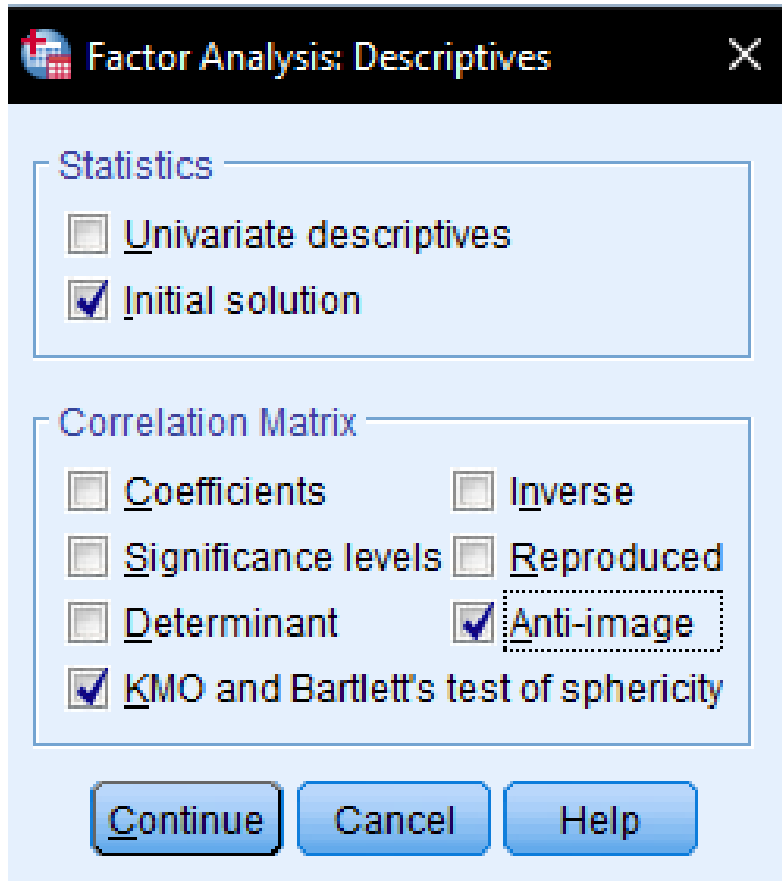


KMO and Bartlett's test

- Select all the items (except SerialNum) and shift them under Variables



KMO and Bartlett's test



- Ensure that KMO and Bartlett's test of sphericity, and Anti-image are selected
- (Optional) Coefficients can be selected to assess bivariate correlations between variables (if most of the correlations are small i.e., less than .3, it suggests that there is *no underlying structure*)
- Continue, and OK

KMO and Bartlett's test

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.796
Bartlett's Test of Sphericity	Approx. Chi-Square	1049.565
	df	55
	Sig.	.000

Based on $\alpha = .05$, the test statistic is significant at $p < .001$. This suggests that the data is suitable for factor analysis; there are correlations between all the items due to possible underlying factors

The KMO value is .796; these data are suitable for factor analysis. Generally, a value of $>.6$ is considered acceptable to conduct factor analysis

Anti-Image Matrices

- On the off-diagonal, the negatives of the partial correlation coefficients of all the items indicate parts of the items that cannot be predicted.
- On the diagonal, measure of sampling adequacy (MSAs) indicate how much individual item is predicted by the other items in the data. These can be used to determine whether individual items should be removed from the factor analysis.

data. These can be used to determine whether individual items should be removed from the factor analysis.

Anti-image Matrices												
		Q1. A teacher should be able to communicate effectively	Q2. A teacher should be entertaining	Q3. A teacher should be approachable	Q4. A teacher should be able to motivate me	Q5. A teacher should not display passion when teaching	Q6. I can explain statistical concepts well	Q7. I have the necessary expertise in statistics	Q8. I should be well versed in SPSS	Q9. I can explain how to read SPSS outputs	Q10. I am knowledgeable in research methodology	Q11. I am good looking and humorous
Anti-image Correlation	Q1. A teacher should be able to communicate effectively	.811 ^a	-.303	-.049	-.393	.157	.022	.021	.060	-.037	.079	.057
	Q2. A teacher should be entertaining	-.303	.712 ^a	-.430	.073	-.131	-.083	.068	-.037	.047	-.052	.009
	Q3. A teacher should be approachable	-.049	-.430	.777 ^a	-.359	.175	.080	-.035	.094	-.031	-.019	.088
	Q4. A teacher should be able to motivate me	-.393	.073	-.359	.778 ^a	.291	.011	-.026	-.045	.010	.018	.036
	Q5. A teacher should not display passion when	.157	-.131	.175	.291	.828 ^a	.075	-.058	.079	.036	-.008	-.058
	Q6. I can explain statistical concepts well	.022	-.083	.068	.011	.075	.809 ^a	-.236	-.291	-.076	-.083	.069
	Q7. I have the necessary expertise in statistics	.021	.068	-.035	-.026	-.058	-.236	.818 ^a	-.188	-.138	-.270	.017
	Q8. I should be well versed in SPSS	.060	-.037	.094	-.045	.079	-.291	.817 ^a	-.138	-.085	.058	.026
	Q9. I can explain how to read SPSS outputs	-.037	.047	-.031	.010	.036	-.076	-.138	.810 ^a	-.357	.787 ^a	.075
	Q10. A teacher should be knowledgeable in research methodology	.079	-.052	-.019	.018	-.008	-.083	-.270	-.085	-.357	.787 ^a	.075
	Q11. A teacher should be good looking and humorous	.057	.009	.088	.036	-.058	.069	.017	.058	.026	.075	.914 ^a

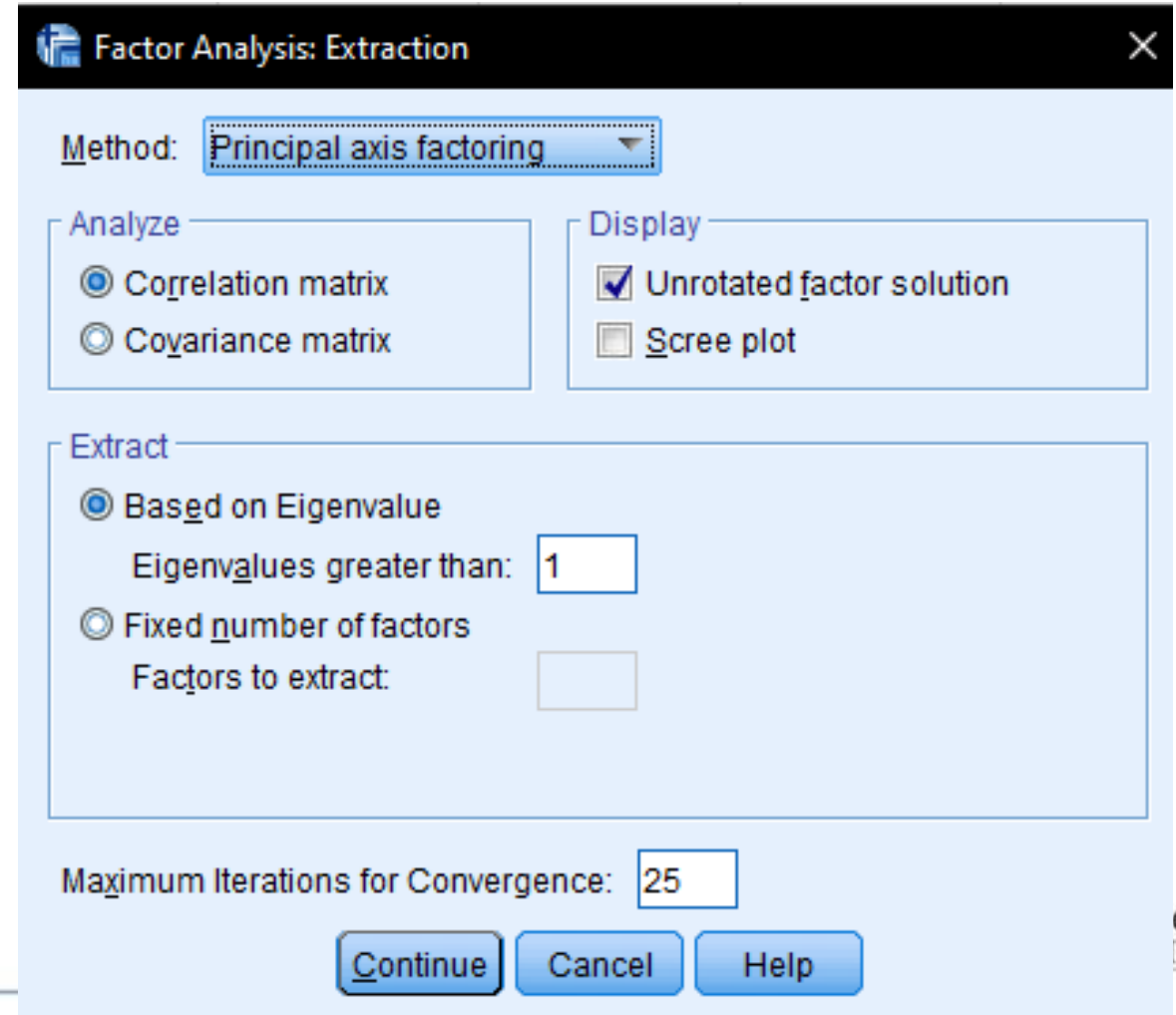
In the Anti-Image CORRELATION table, look at the diagonal values (those with the superscript 'a'). Values below .5 implies that the item has low partial correlations. In this example, all values are > .5. The off-diagonal values are also small.

COOK
UNIVERSITY
BIRMINGHAM
CORE

In the Anti-Image CORRELATION table, look at the diagonal values (those with the superscript 'a'). Values below .5 implies that the item has low partial correlations. In this example, all values are > .5. The off-diagonal values are also small.

Extracting Factors

- Analyze → Dimension Reduction → Factor → Extraction
- Select 'Principal axis factoring' or 'Principal component analysis' as a Method
- Select Scree plot if you want to determine a number of extracted factors using a visual method
- Continue, and OK



The image shows the 'Factor Analysis: Extraction' dialog box in SPSS. The 'Method' dropdown is set to 'Principal axis factoring'. Under the 'Analyze' section, 'Correlation matrix' is selected. Under the 'Display' section, 'Unrotated factor solution' is checked, and 'Scree plot' is unchecked. Under the 'Extract' section, 'Based on Eigenvalue' is selected, with 'Eigenvalues greater than' set to 1. 'Fixed number of factors' is also an option. At the bottom, 'Maximum iterations for Convergence' is set to 25. The 'Continue', 'Cancel', and 'Help' buttons are at the bottom right.

Factor Analysis: Extraction

Method: Principal axis factoring

Analyze

- ☒ Correlation matrix
- ☐ Covariance matrix

Display

- ☒ Unrotated factor solution
- ☐ Scree plot

Extract

- ☒ Based on Eigenvalue
 - Eigenvalues greater than: 1
- ☐ Fixed number of factors
 - Factors to extract:

Maximum iterations for Convergence: 25

Continue Cancel Help

Extracting Factors

Total Variance Explained						
Factor	Total	Initial Eigenvalues		Extraction Sums of Squared Loadings		
		% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.112	28.294	28.294	2.645	24.047	24.047
2	2.814	25.581	53.874	2.245	20.413	44.460
3	.858	7.796	61.670			
4	.832	7.565	69.235			
5	.760	6.912	76.147			
6	.554	5.032	81.179			
7	.507	4.605	85.784			
8	.481	4.372	90.156			
9	.430	3.909	94.065			
10	.388	3.523	97.588			
11	.265	2.412	100.000			

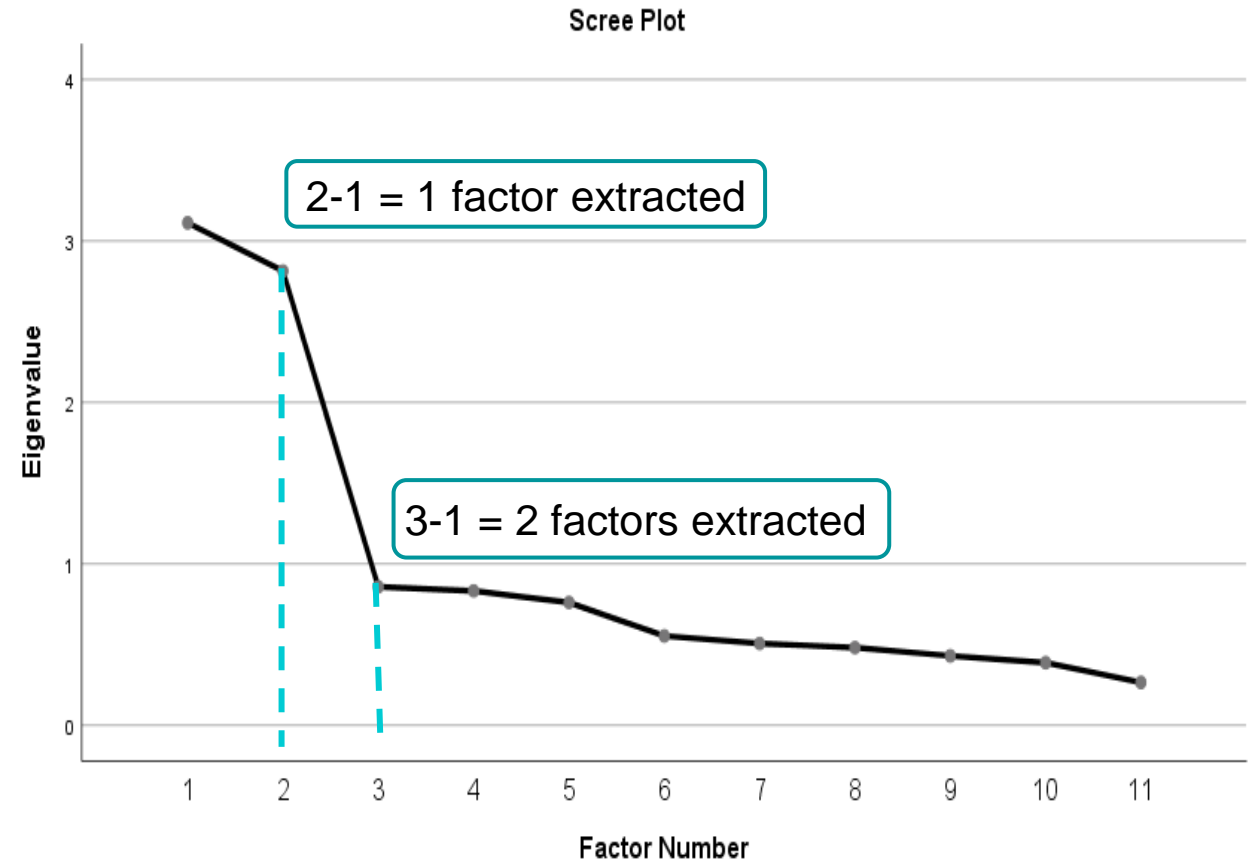
Extraction Method: Principal Axis Factoring.

Any factor with eigenvalues > 1 will be retained. In this case, there are 2.

This shows the percentage of variance explained by each factor, as well as the cumulative percentage. The 2 factors extracted explained 44.46% of total variance.

Extraction Via Scree Plot

- The number of factors to be extracted is based on the inflexion point (point where the graph drops off drastically) and subtracting 1 from that point
- This scree plot suggests either a 1 or 2 factor solution



Extracting Factors

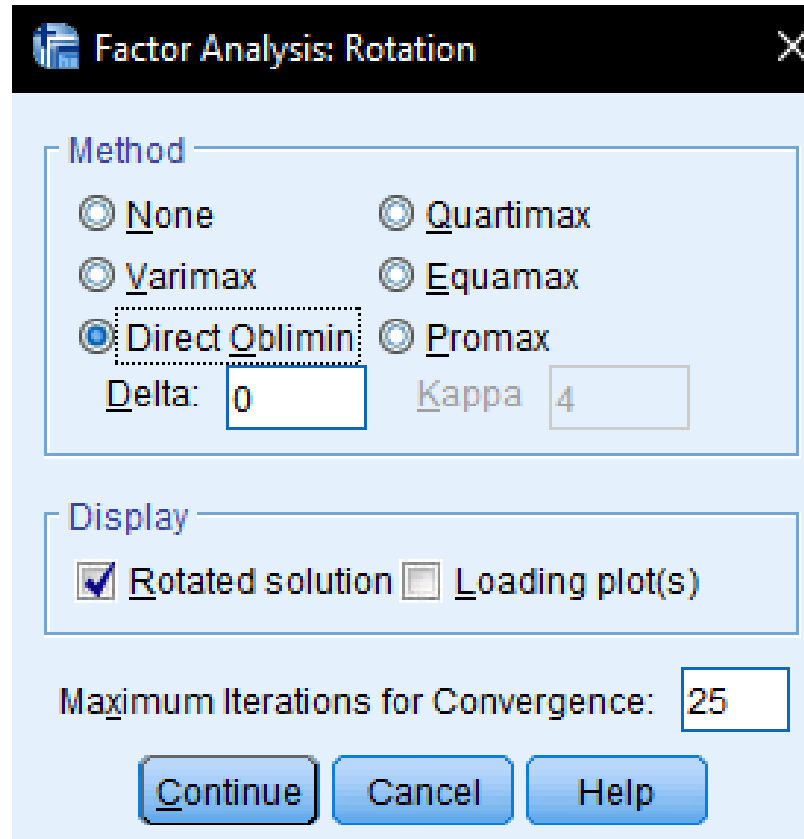
- We decide to stick to the two-factor solution provided by SPSS
- The factor matrix shows the correlations between items and the 2 factors BEFORE rotation

Factor Matrix ^a		
	Factor	
	1	2
Q1. A teacher should be able to communicate effectively	.771	-.089
Q2. A teacher should be entertaining	.563	-.025
Q3. A teacher should be approachable	.795	-.035
Q4. A teacher should be able to motivate me	.811	.006
Q5. A teacher should not display passion when teaching	-.598	-.087
Q6. A teacher should be able to explain statistical concepts well	.004	.631
Q7. A teacher should have the necessary expertise in statistics	-.041	.703
Q8. A teacher should be well versed in SPSS	-.010	.642
Q9. A teacher should be able to clearly explain how to read SPSS outputs	.048	.634
Q10. A teacher should be knowledgeable in research methodology	-.017	.683
Q11. A teacher should be good looking and humorous	-.287	-.233

Extraction Method: Principal Axis Factoring.
a. 2 factors extracted. 6 iterations required.

Rotation

- Analyze → Dimension Reduction → Factor → Rotation
- Select Direct Oblimin as a Method
- Continue
- Select Options → Suppress small coefficients → .3 (this is done to aid interpretation of the matrix table)
- Continue, and OK



Factor Analysis: Rotation

Method

☐ None ☐ Quartimax
☐ Varimax ☐ Equamax
☒ Direct Oblimin ☐ Promax

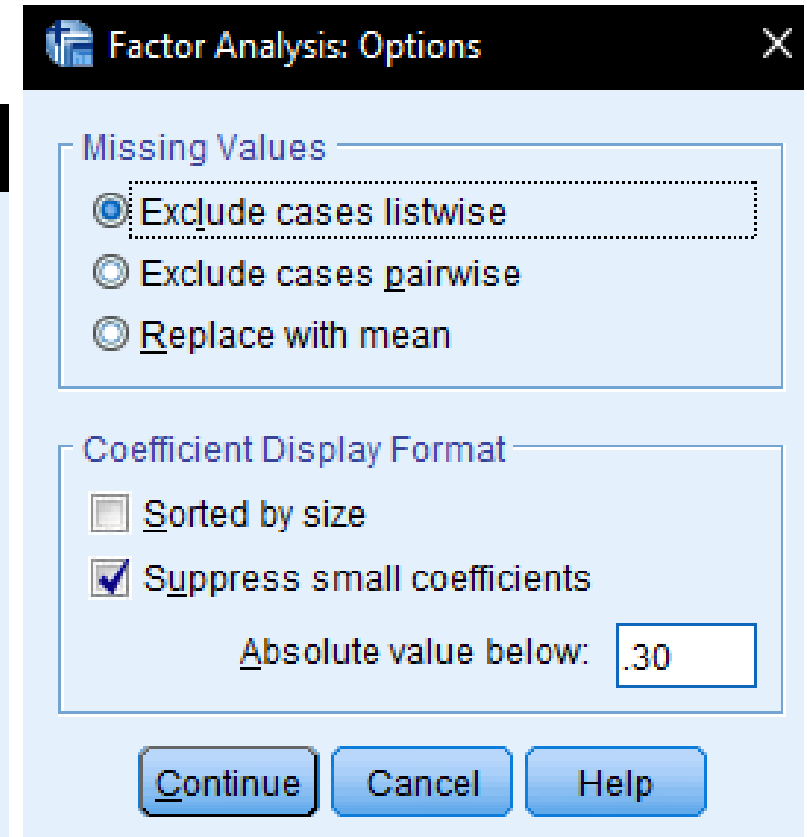
Delta: 0 Kappa: 4

Display

☒ Rotated solution ☐ Loading plot(s)

Maximum Iterations for Convergence: 25

Continue Cancel Help



Factor Analysis: Options

Missing Values

☒ Exclude cases listwise
☐ Exclude cases pairwise
☐ Replace with mean

Coefficient Display Format

☐ Sorted by size
☒ Suppress small coefficients

Absolute value below: .30

Continue Cancel Help

Rotation

- Look at Pattern Matrix for factor loadings AFTER rotation (if an orthogonal rotation was chosen, look at Rotated Factor Matrix instead)
- The empty fields represent suppressed coefficients ($<.3$)
- We can see that Q1-Q5 load heavily on factor 1, while Q6-Q10 load on factor 2
- Q11 appears to not load on either factor, so this item can be dropped

	Factor	
	1	2
Q1. A teacher should be able to communicate effectively	.774	
Q2. A teacher should be entertaining	.565	
Q3. A teacher should be approachable	.797	
Q4. A teacher should be able to motivate me	.812	
Q5. A teacher should not display passion when teaching	-.597	
Q6. A teacher should be able to explain statistical concepts well		.632
Q7. A teacher should have the necessary expertise in statistics		.706
Q8. A teacher should be well versed in SPSS		.643
Q9. A teacher should be able to clearly explain how to read SPSS outputs		.633
Q10. A teacher should be knowledgeable in research methodology		.685
Q11. A teacher should be good looking and humorous		

Extraction Method: Principal Axis Factoring.
Rotation Method: Oblimin with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Interpreting the Factors

- Now that we know which factor each item loads on, we can label the factors
- Looking at the items, it appears that Factor 1 refers to a teacher's personal characteristics (entertaining, approachable, etc.)
- Factor 2 seems to be subject oriented (statistical concepts, SPSS, etc.)
- Therefore, in this questionnaire, it appears that there are 2 dimensions: person-oriented, and subject-oriented characteristics.

Pattern Matrix^a

	Factor	
	1	2
Q1. A teacher should be able to communicate effectively	.774	
Q2. A teacher should be entertaining	.565	
Q3. A teacher should be approachable	.797	
Q4. A teacher should be able to motivate me	.812	
Q5. A teacher should not display passion when teaching	-.597	
Q6. A teacher should be able to explain statistical concepts well		.632
Q7. A teacher should have the necessary expertise in statistics		.706
Q8. A teacher should be well versed in SPSS		.643
Q9. A teacher should be able to clearly explain how to read SPSS outputs		.633
Q10. A teacher should be knowledgeable in research methodology		.685

Extraction Method: Principal Axis Factoring.
Rotation Method: Oblimin with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Write-Up

An example write-up can be found on page 231 in

Allen, P., Bennett, K., & Heritage, B. (2019). *SPSS Statistics: A Practical Guide* (4th ed.). Cengage Learning.

Questions?

learningcentre-singapore@jcu.edu.au