



Factor analysis

Sebastian Jentschke



Welcome to a combined lecture with PC exercise on factor analysis.



Agenda

- some theoretical background
- Exploratory Factor analysis (EFA)
- Principal Component Analysis (PCA)
- Confirmatory Factor Analysis (CFA)
- Reliability Analysis



In this lecture, I will introduce several concepts that stand back factor analysis. First, this is the distinction between latent variables (representing theoretical concepts, also called psychological constructs) and manifest variables that we measure.

After that, the different steps within a factor analysis will be introduced: factor extraction, factor selection, factor rotation and factor interpretation. Factor analysis is organized as a kind of a ping-pong: The software extracts the factors, you select them; the software rotates the factors to make interpretation easier for you. The latter, interpreting what the factors mean represents your major contribution to a factor analysis.



Agenda

- some theoretical background
- Exploratory Factor analysis (EFA)
- Principal Component Analysis (PCA)
- Confirmatory Factor Analysis (CFA)
- Reliability Analysis



Afterwards we will turn towards how four different factor-analytical approaches are calculated. The main focus will be on exploratory factor analysis. Principal component analysis is mathematically similar but conceptually different from EFA and is covered very brief. Confirmatory Factor Analysis is only for those who already know that they might use it (e.g., in their M.Sc. thesis to validate a questionnaire). The others may skip that part. Finally, Reliability analysis deals with one important aspect of reliability: internal consistency. It measures to what degree the items in a questionnaire or a scale within a questionnaire are assessing a common underlying concept.



Some theoretical background

- Factor analysis is a statistical method used to **describe** common variability among **observed, correlated variables** in terms of a potentially lower number of **unobserved variables** called factors.
- **Factors** are **latent variables** that are not observed but inferred (through a mathematical model). Factors are **derived from manifest variables** that are observed (directly measured).



Factor analysis is a statistical method used to describe common variability between correlated variables that we measured in terms of a potentially lower number of unobserved variables called factors.

With those characteristics, it serves two purposes: (1) It aims to account for different aspects of behaviour by hypothesizing that they are brought about by an underlying psychological construct (or theory). (2) It also serves as a data reduction technique by describing a larger number of observed variables by a smaller number of underlying factors.

Central for these two purposes is the concept of latent variables. Those latent variables (or factors) are initially hidden and then inferred through a statistical analysis based upon the correlation of the observed variables (which are directly measured).



Some theoretical background

- Factor analysis is a statistical method used to **describe** common variability among **observed, correlated variables** in terms of a potentially lower number of **unobserved variables** called factors.
- **Factors** are **latent variables** that are not observed but inferred (through a mathematical model). Factors are **derived from manifest variables** that are observed (directly measured).

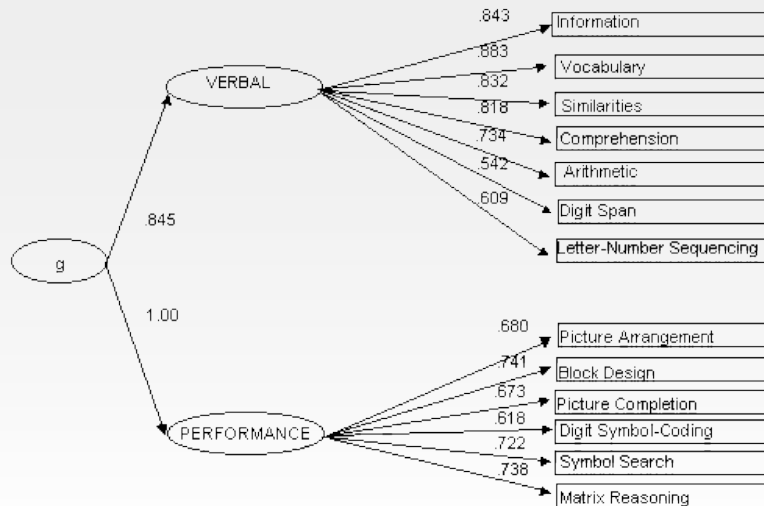


Latent variables represent psychological phenomena or constructs that are difficult to directly observe or measure. Examples for latent variables are personality dimensions, intelligence or certain attitudes or opinions.

An easy way to imagine a latent variable is that is a hidden characteristic of a person that we can't directly see, e.g., job satisfaction. We can, however, describe that characteristic by combining different aspect of the behaviour of that person. A person satisfied with their job could be expected to answer certain questions in a (relatively) predictable manner or show certain behaviour (e.g., work a lot, be dedicated to their employer).



Some theoretical background



Another classical example for latent variables (the circles in the figure) is intelligence. Psychological constructs of intelligence differ considerably with regard to into how many subdimensions intelligence is split. The figure represents the assumption is that one general intelligence factor is split into two large subdimensions for verbal and performance. We assume the position of a person on the latent variable «intelligence» and its subdimensions by measuring their behaviour (i.e., their performance) in certain tasks assumed to contribute to these subdimensions of intelligence. Somebody who is intelligent should, e.g., possess a large vocabulary, have good working memory (digit span), etc. Test results from the individual tasks are then “converted” into a score on the latent variable on the subdimensions and further on to general intelligence (expressed as IQ) by using a mathematical combination for (weighing) of these results.



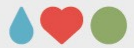
Some theoretical background

- Factor analysis is a statistical method used to **describe** common variability among **observed, correlated variables** in terms of a potentially lower number of **unobserved variables** called factors.
- **Factors** are **latent variables** that are not observed but inferred (through a mathematical model). Factors are **derived from manifest variables** that are observed (directly measured).



It is also assumed that certain behaviour – from performance in specific tasks to how somebody replies to questions in a questionnaire – can be assessed with manifest variables. These manifest variables, strictly speaking a combination of them, are “caused” by an individual's position on one or more underlying latent variable(s).

Once the position on the latent variables is determined (and controlled for), it is assumed that the manifest variables won't correlate any more and have nothing in common beyond these influences resulting from the latent variables.



A practical example

Participants responded to the following six statements with 1 = totally disagree to 5 = totally agree

- 1) My friends think I should use eco-friendly products.
- 2) My neighbors think I should use eco-friendly products
- 3) My colleagues think that I should use eco-friendly products.
- 4) I feel a moral obligation to buy eco-friendly product.
- 5) I feel a moral obligation to recycle household waste.
- 6) I feel a moral obligation to buy products that contain recycled ingredients.



Let's try this with an example: A group of participants was asked 6 questions that assessed their attitudes towards ecology. Already from the way the questions are formulated, you will see that the first three questions rather reflect expectations from other people towards the person answering the questionnaire, whereas the latter three questions rather assess what a person feels as a duty or an obligation.



A practical example

	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6
Var 1	–					
Var 2	0.82	–				
Var 3	0.74	0.86	–			
Var 4	-0.11	0.14	0.08	–		
Var 5	0.02	-0.12	-0.07	0.67	–	
Var 6	0.06	0.08	0.13	0.55	0.73	–

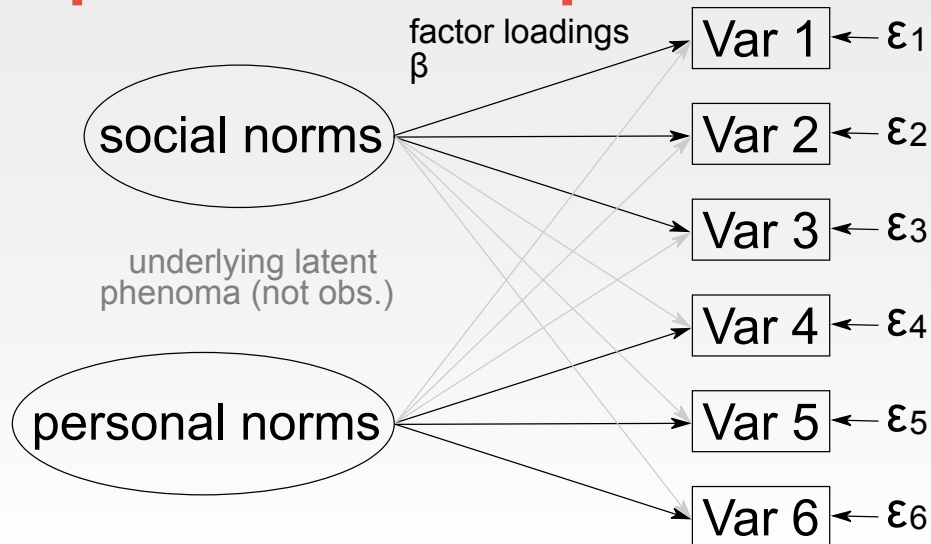


As a consequence of this pattern in the questions (i.e., that the questions fall in two “categories” or have two topics), the variables are correlated with each other in a predictable fashion. Those questions substantially correlated with each other fall into a common category. There must be something underlying, which “causes” this pattern of correlations. The underlying cause are our latent variables, and we aim to extract them based upon the pattern of correlations that we observed.

There is another thing to note. Outside the two groups of variables that correlate substantially due to an underlying latent variable (or factor), the remaining correlations are rather low.



A practical example



When visualizing the two presumed underlying latent variables and their consequences, we see that the first factor “social norms” has a relatively strong impact on variable 1 to 3, whereas the second factor “personal norms” has a relatively strong impact on variable 4 to 6. Apart from that the influence is rather minor. That is visualized by the grey arrows: The first factor has rather a negligible influence on variable 4 to 6 and vice versa has the second factor little influence on variable 1 to 3.

This is exactly the pattern that we saw in the correlation matrix on the previous slide.

Our aim is now to determine factor loadings or weights. Those loadings are “extracted” from the correlation matrix.



A practical example

similarity with a regression model:

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

for example:

$$\text{Var1}_i = \beta_1 \cdot \text{social norms}_i + \beta_2 \cdot \text{personal norms}_i + \varepsilon_i$$

$$\text{Var2}_i = \beta_1 \cdot \text{social norms}_i + \beta_2 \cdot \text{personal norms}_i + \varepsilon_i$$

... and so on for Var3 to Var6



These loadings can be described as similar to the β -weights in a regression model (one can use the General Linear Model for nearly everything...). If we would calculate the value of person i on variable Var 1, we do this by multiplying the first (relatively large) weight β_1 with the position of person i on the (hidden) latent variable “social norms” and adding this up with the result of multiplying the (relatively small) weight β_2 with the position of that person i on the other (hidden) latent variable “personal norms”. For variable 1 to 3 the first weight β_1 is relative large (think of the black arrow on the previous slide) for variable 4 to 6 the second weight β_2 .

For a person relatively high on “social norms” and low on “personal norms” we can expect high values / scores for variable 1 to 3 and rather low values / scores for variable 4 to 6. If a person where high on both social and personal norms, the values for all variables would be high.



A practical example

similarity with a regression model:

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

for example:

$$\text{Var1}_i = \beta_1 \cdot \text{social norms}_i + \beta_2 \cdot \text{personal norms}_i + \varepsilon_i$$

$$\text{Var2}_i = \beta_1 \cdot \text{social norms}_i + \beta_2 \cdot \text{personal norms}_i + \varepsilon_i$$

... and so on for Var3 to Var6

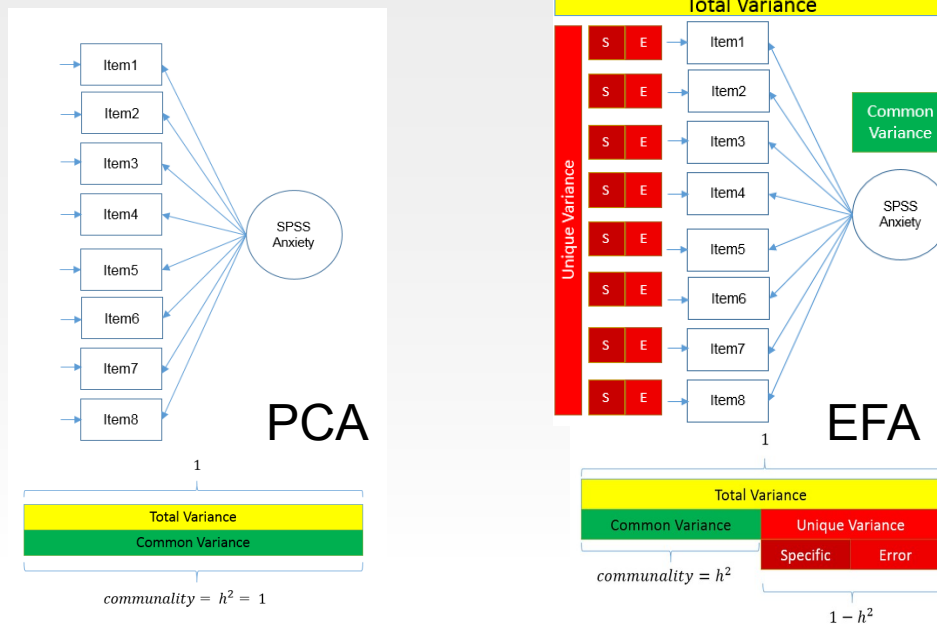


There is a last point in the regression equation, that I would like to mention, the error term ε_i . This is the part of the variance in the variable that we can not account for by our knowledge of a person's position on the two latent variables "social norms" and "personal norm" (in combination with the weights for those variables).

The error term is also important for a central distinction between two central factor-analytical methods: exploratory factor analysis (EFA) and principal component analysis (PCA). Quite some statistics programs (including SPSS) treat the PCA as a method within EFA. From a mathematical standpoint, the way they are calculated is (more or less) identical, however, the two methods have different underlying assumptions and serve different purposes.



EFA vs. PCA



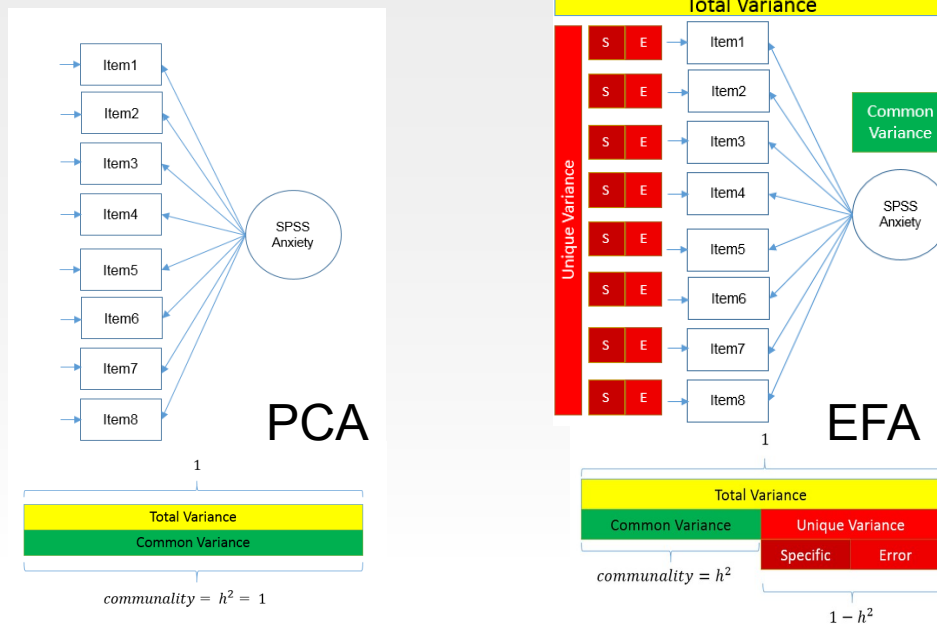
SLIDE 13



The fundamental difference between the two methods is that the PCA (left) assumes or aims to assign all available variance to underlying components. The main function of the PCA is therefore not to test any theoretical assumptions about some underlying latent variables but to reduce the data (strictly speaking, to try to describe the originally k dimensions – one per variable – with a lower number of dimensions). The aim is to describe as much as possible of the available information with as less as possible dimensions. The PCA further assumes that all variation in the variables can be accounted for by the extracted components.



EFA vs. PCA



SLIDE 14

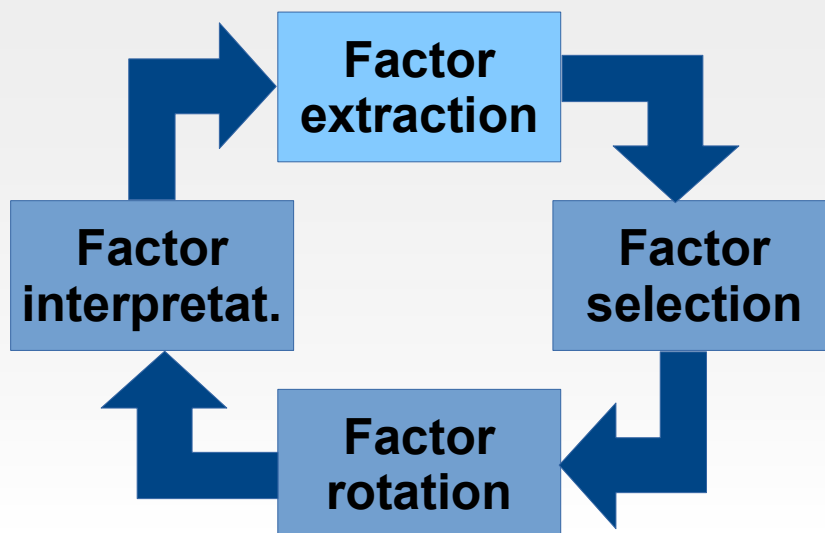


The EFA assumes instead that only a certain proportion of the variation in the variable can be accounted for by the proposed underlying latent variable(s). Given this assumption, the total variance is split in two components. The common variance, and the unique variance. The common variance is how much of the variation in one variable can be accounted for by the remaining variables. In principle is the common variance in the EFA what results from a linear regression where one variable is the dependent and the remaining variables are the independent ones, which are used to predict the dependent variable.

The unique variance (i.e., the part which can't be predicted by the other variables) is again split in a part that is specific to a particular item or variable and another part which is error variance (e.g., caused by measurement errors).



Steps of a factor analysis



When conducting a factor analysis, we have four steps. As I said, it works a little like ping-pong. In a first step, the software extracts factors. In the second step, you carry out a selection of how many of these extracted factors are to be considered (taken further) and how many are deselected.

Afterwards follows the factor rotation. The aim is here to make the extracted factors you selected easier interpretable. This is done by shifting or rotating the factor loadings (β) with the aim that each variable loads as highly as possible on one factor and as lowly as possible on other factors.

The interpretation is primarily based upon these high loadings. Variables that load high on one factor are important to the factor. What the factor may represent can be inferred from what the variable means. This may sound a bit cryptic. I hope it gets a bit clearer once we have a look at each of these steps one after another.



Factor extraction

	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6
Var 1	0.70					
Var 2	0.82	0.79				
Var 3	0.74	0.86	0.71			
Var 4	-0.11	0.14	0.08	0.61		
Var 5	0.02	-0.12	-0.07	0.67	0.79	
Var 6	0.06	0.08	0.13	0.55	0.73	0.75



The first part, the factor extraction, is starting with the correlation matrix (such as the one we saw some slides ago for the items where attitudes to the environment were assessed).

The basic idea is to “shift” the dimensions in the data (6 in this case, one for each variable) in a way such that one factor “takes care” of the correlations between variable 1 to 3, another factor of the correlations between variable 4 to 6 whereas the remaining factors “take care” of what variation is left over. Those interested in a more “mathematical” account, please ask in the discussion.

Here is also the place where the difference between EFA and PCA is decisive for what happens in the calculation. The values that can be found in the main diagonal are either the results of the prediction of one variables by all others (for the EFA) or 1 (for the PCA). That is, the example you see in the table represents an EFA.



Factor extraction

	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6
Var 1	0.70					
Var 2	0.82	0.79				
Var 3	0.74	0.86	0.71			
Var 4	-0.11	0.14	0.08	0.61		
Var 5	0.02	-0.12	-0.07	0.67	0.79	
Var 6	0.06	0.08	0.13	0.55	0.73	0.75



Remember the slide with the regression equation.

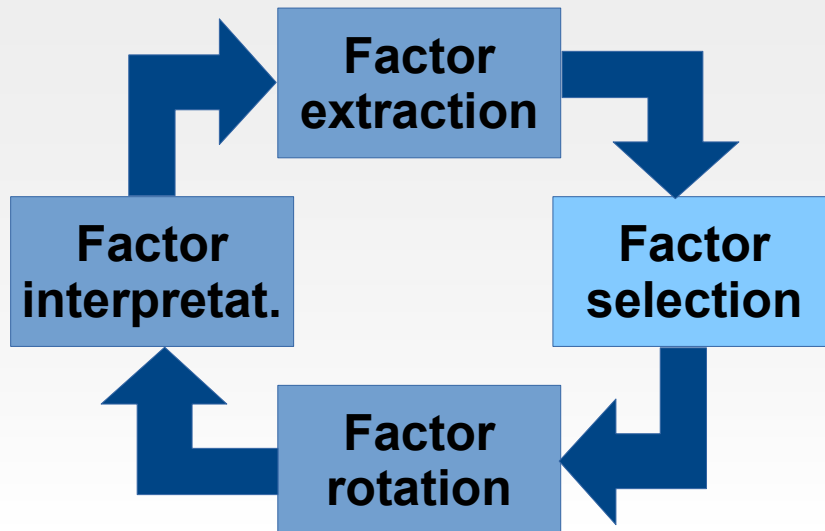
What basically happens during what I described as “shifting” on the previous slide is that the β -weights are iteratively (step-by-step) adjusted in a way that β_1 best “represents” the first latent variable and β_2 the second latent variable. Of course, there are also β_3 to β_6 but these are during the process adjusted to rather low loadings as they represent the remaining small correlations that we also find in the matrix / table on the slide.

There is one last concept I need to introduce:

Eigenvalues. Let's assume that the variation in each variable gets assigned the value 1 (mathematically, this is simply a z-transformation). For our 6 variables, we have a total eigenvalue of 6 to distribute. β_1 and β_2 having relative large influence will get assigned relatively large eigenvalues, the remaining β_3 to β_6 rather small ones.



Steps of a factor analysis



FACTOR ANALYSIS

SLIDE 18



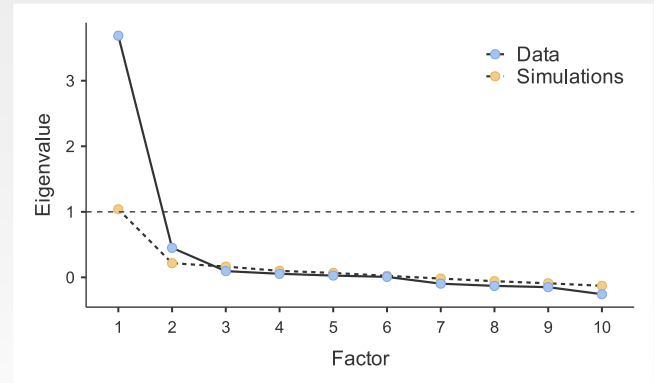
The eigenvalues are central for how many of the extracted factors we are going to select.



Factor selection

four possible methods:

- parallel analysis
- eigenvalue
- scree plot
- theoretical assumptions



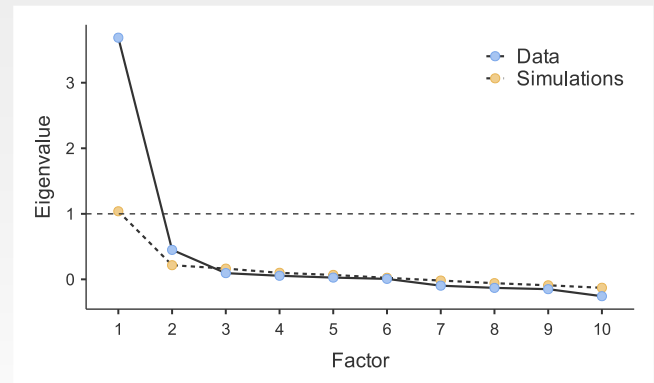
We have four possible choices how to make the selection, two mathematical ones (where we just choose the method and max. set one parameter and the rest is done automatically) and two that require our interaction. Let's begin with the latter.



Factor selection

four possible methods:

- parallel analysis
- eigenvalue
- scree plot
- theoretical assumptions



Here we define how many factors we want to select by entering that into our statistics software.

The decision can either be based upon theoretical assumptions (i.e., that we claimed, our construct contains of a certain number of factors; the Big Five personality model proposes 5 factors or dimensions with a certain “meaning”, e.g., Neuroticism).

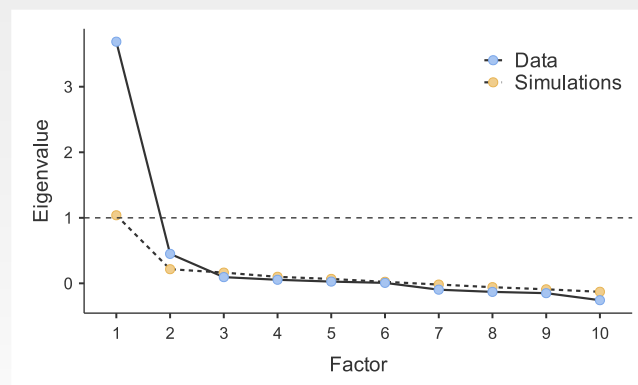
The second opportunity would be based upon the scree plot (shown in the figure on the right). The eigenvalue of each factor is visualized in descending order (the line with the blue dots). At one point of that line, the descent goes from steep to flat (like an elbow). Everything left from that (i.e., the first factor for that screeplot) is considered meaningful.



Factor selection

four possible methods:

- parallel analysis
- eigenvalue
- scree plot
- theoretical assumptions



The other two criteria are based on thresholds (and don't need our interaction after choosing that method). The first of those is parallel analysis. Here, the eigenvalues for a random matrix (with the same number of variables and participants as in our data matrix) are determined (by running a parallel factor extraction on these random data). The results are shown as line with yellow dots. Whenever the eigenvalue of the real data exceeds that of the random data (i.e., if the blue dot is above the yellow dot) the factor is selected. In the figure, factor 1 and 2 would be selected.

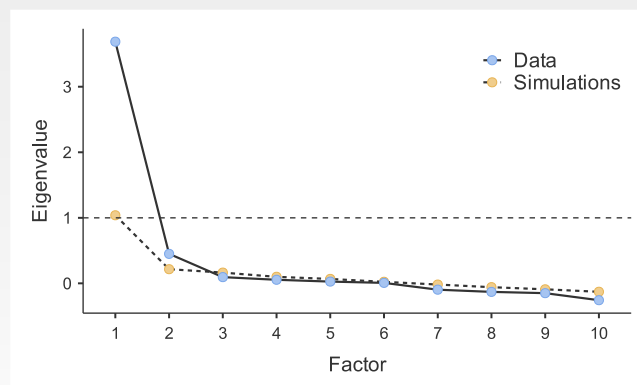
Finally, we have to so called Kaiser-criterium, an eigenvalue of 1 as threshold (represented by the dotted line). Using that criterium, factor 1 would be selected. Initially, each variable has an eigenvalue of 1. Factors that are selected would therefore explain a higher amount of variance than any of the original variables.



Factor selection

four possible methods:

- parallel analysis
- eigenvalue
- scree plot
- theoretical assumptions

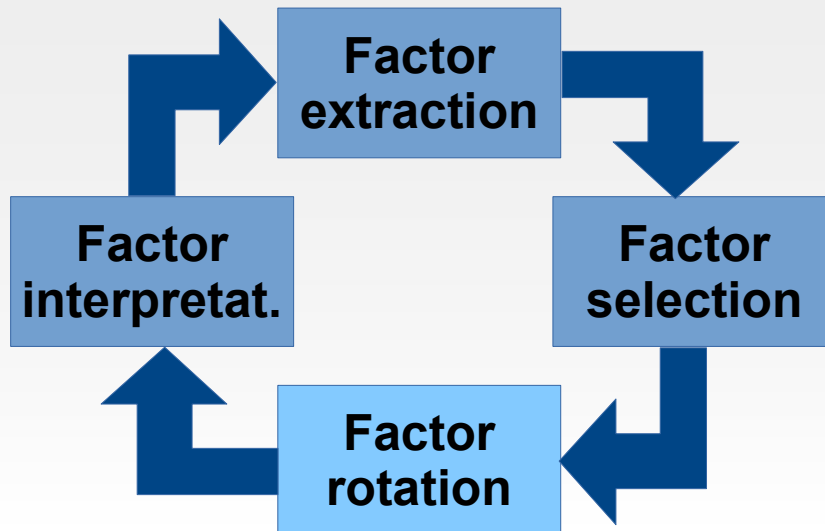


It is possible and common to consider these criteria in combination. For example would both the scree-plot and the eigenvalue-criterium have suggested to keep one factor whereas the parallel analysis would have suggested two.

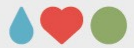
In any case, if we have any theoretical assumption regarding how many factors we expect, this should take precedence. Otherwise, for the current case, I would recommend to only consider one factor. Both scree plot and eigenvalue-criterium suggest that. In the end, it is always our responsibility to decide what we deem best, and to properly describe and justify this decision.



Steps of a factor analysis



After we decided about how many factors we want to included the procedure continues with rotating the selected factors.



Factor rotation

- aim: simplify the solution, help interpretability
- **orthogonal:**
Varimax, Quartimax
- **oblique**
Oblimin, Promax, Simplimax

unrotated

Oblimin

Factor Loadings						Factor Loadings						
	Factor						Factor					
	1	2	3	4	5	Uniqueness	1	2	3	4	5	Uniqueness
A1						0.719						0.719
A2	0.374	0.404	0.322			0.480				-0.482		0.712
A3	0.499	0.345	0.370			0.442				0.683		0.442
A4	0.429			0.306		0.666				0.435		0.666
A5	0.580					0.531				0.525		0.531
C1	0.465		-0.435		0.330	0.480		0.717				0.480
C2	0.395		-0.398		0.323	0.524		0.676				0.524
C3			-0.369			0.747		0.596				0.747
C4	-0.584		0.356			0.475		-0.635				0.475
C5	-0.548		0.380			0.519		-0.579				0.519
E1	-0.410				0.339	0.696						0.696
E2	-0.700				0.385	0.353				-0.531		0.353
E3	0.538	0.308				0.478				-0.705		0.478
E4	0.605					0.425				0.493	0.362	0.425
E5	0.542					0.575				0.592		0.575
N1	-0.414	0.615				0.347	0.794			0.421		0.347
N2		0.627				0.399	0.789					0.399
N3	-0.394	0.588				0.474	0.693					0.474
N4	-0.505	0.461				0.449	0.504		-0.453			0.449
N5		0.473				0.632	0.472					0.632
O1	0.391			-0.432		0.626						0.626
O2				0.443		0.662					0.566	0.662
O3				-0.487		0.414					-0.490	0.682
O4			0.302			0.746				0.326		0.746
O5				0.447		0.679					-0.515	0.679

A situation that is quite typical after factor extraction and selection is that variables have several rather small loadings on many or all extracted factors. In the example in the left table, the majority of variables has loadings on the first factor whereas other factors receive few loadings.

This has to do with the way factors are extracted. The first factor has the largest amount of explained variance, the other factors less (according to their order). This can also be seen at the scree plot. If you try to conceptualize this in terms of our example so are both social and personal norms related to environmental-responsible behaviour (the first factor).



Factor rotation

- aim: simplify the solution, help interpretability
- **orthogonal:**
Varimax, Quartimax
- **oblique**
Oblimin, Promax, Simplimax

unrotated

Oblimin

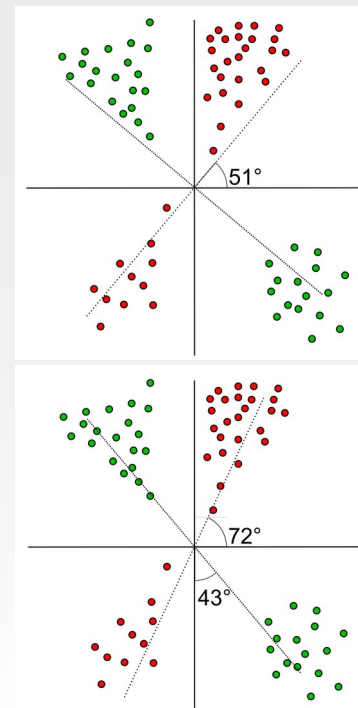
Factor Loadings						Factor Loadings							
	Factor					Uniqueness		Factor					Uniqueness
	1	2	3	4	5		1	2	3	4	5		
A1			-0.447			0.719	A1			-0.482		0.719	
A2	0.374	0.404	0.322			0.480	A2			0.712		0.480	
A3	0.499	0.345	0.370			0.442	A3			0.683		0.442	
A4	0.429			0.306		0.666	A4			0.435		0.666	
A5	0.580					0.531	A5			0.525		0.531	
C1	0.465		-0.435		0.330	0.480	C1		0.717			0.480	
C2	0.395		-0.398		0.323	0.524	C2		0.676			0.524	
C3			-0.369			0.747	C3		0.596			0.747	
C4	-0.584		0.356			0.475	C4		-0.635			0.475	
C5	-0.548		0.380			0.519	C5		-0.579			0.519	
E1	-0.410				0.339	0.696	E1			-0.531		0.696	
E2	-0.700				0.385	0.353	E2			-0.705		0.353	
E3	0.538	0.308				0.478	E3			0.493	0.386	0.478	
E4	0.605					0.425	E4			0.592	0.362	0.425	
E5	0.542					0.575	E5			0.421		0.575	
N1	-0.414	0.615				0.347	N1	0.794				0.347	
N2		0.627				0.399	N2	0.789				0.399	
N3	-0.394	0.588				0.474	N3	0.693				0.474	
N4	-0.505	0.461				0.449	N4	0.504	-0.453			0.449	
N5		0.473				0.632	N5	0.472				0.632	
O1	0.391			-0.432		0.626	O1				0.566	0.626	
O2				0.443		0.682	O2				-0.490	0.682	
O3				-0.487		0.414	O3				0.685	0.414	
O4			0.302			0.746	O4			0.326		0.746	
O5				0.447		0.679	O5				-0.515	0.679	

What we would like to achieve is that one class of variables has rather high loadings on one factor and rather low loadings on others like in the example in the right table. Or in terms of our example: We would like that instead of the variables loading very strongly on one factor (environmental-responsible behaviour) the loadings are «redistributed» so that some of the original variables have high loadings on social norms and low loadings on personal norms and that others variables show the opposite pattern, i.e. low loadings on social norms and high loadings on personal norms.

I mentioned before that determining which variables have high loadings on the factor (i.e., are especially representative for it) is essential for our interpretation what a factor means.

Factor rotation

- aim: simplify the solution, help interpretability
- **orthogonal** (top):
Varimax, *Quartimax*
- **oblique** (bottom)
Oblimin, *Promax*, *Simplimax*



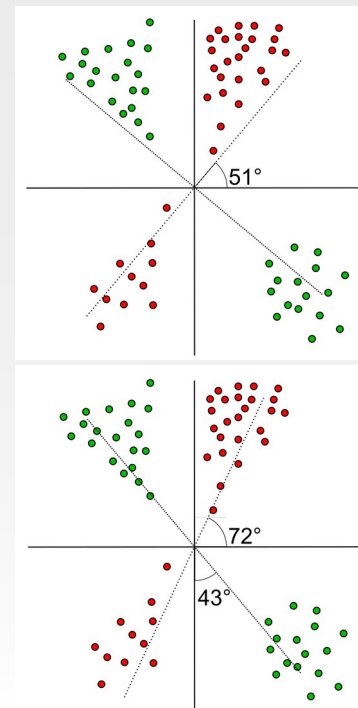
In order to achieve that, we rotate the factors. The dots in the original (unrotated) situation – that is relative to the coordinate system and the axis with solid lines – have relatively considerable distances from both axes / factors.

Through the rotation, we achieve something similar to fitting a regression line: The distance to one factor is minimized and to the other factor maximized.

In the top example, the whole coordinate system is rotated by a certain angle, from the original position where the lines are solid and arranged like a normal coordinate system, to the position of the dotted lines. The 90°-relationship of the two axes to each other is kept. Therefore, this is called an *orthogonal rotation*. It means means that the two factors are independent of each other and not correlated.

Factor rotation

- aim: simplify the solution, help interpretability
- **orthogonal** (top):
Varimax, *Quartimax*
- **oblique** (bottom)
Oblimin, *Promax*, *Simplimax*

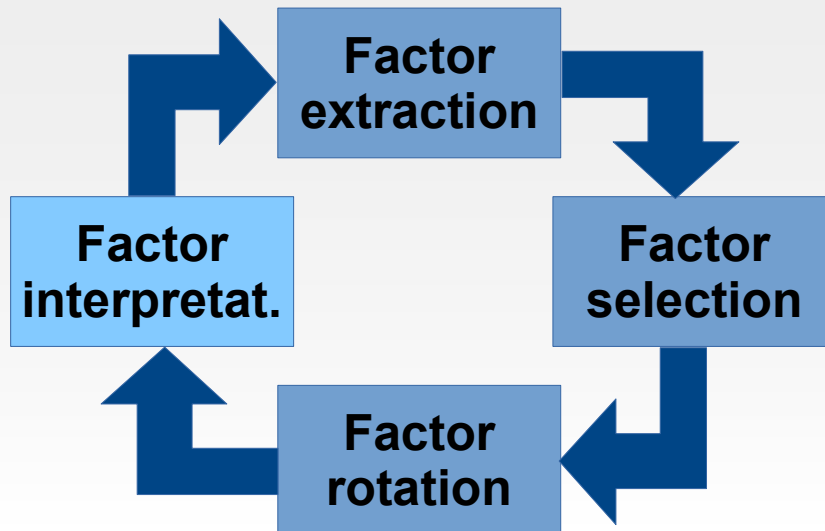


The bottom example is called *oblique rotation*: Each axis is rotated separately, not the coordinate system as a whole. They are therefore not orthogonal any longer. This has the advantage that they can be fitted more easily right through the clouds of dots. But not being orthogonal also means that the two factors are not independent of each other any more. Once they are correlated, changing the position on one factor also means that the position on other factors is changed as well (though to a lesser degree).

That psychological constructs such as personality dimensions or different aspects of intelligence are correlated is the typical case. Therefore, an oblique rotation also often bears more ecological validity than a orthogonal rotation.



Steps of a factor analysis



The rotation has the aim of making interpreting the factors easier. The interpretation is the last step and also the steps that needs our contribution most. The reason that an arrow is shown from factor interpretation to a new round of factor extraction indicates that we might decide to adjust certain decisions we made in the first round in order to refine or improve our model.



Factor interpretation

- assess the factor loading
- start with variables with high loadings and assess their meaning (e.g., variable label, questionnaire item)
- should be in accordance with your theoretical construct

	Factor					Uniqueness
	1	2	3	4	5	
A1				-0.482		0.719
A2				0.712		0.480
A3				0.683		0.442
A4				0.435		0.666
A5				0.520		0.531
C1		0.717				0.480
C2		0.676				0.524
C3		0.506				0.747
C4		-0.635				0.475
C5		-0.579				0.519
E1			-0.581			0.696
E2			0.705			0.353
E3			0.495		0.386	0.478
E4			0.592	0.362		0.425
E5			0.421			0.575
N1	0.794					0.347
N2	0.789					0.399
N3	0.693					0.474
N4	0.504		-0.453			0.449
N5	0.473					0.632
O1					0.566	0.626
O2					-0.490	0.682
O3					0.685	0.414
O4				0.326		0.746
O5					-0.515	0.679



Factor interpretation means to assign meaning to the results we obtained. The main part of our results are the so called factor loadings, i.e., information about how much a certain variable contributes to a factor.

Typically, we would – within each factor – look for variables that have especially high loadings and are therefore especially representative for the factor. A sensible approach would be to assess the two, three variables with the highest loadings and trying to figure out what they have in common. This reasoning is typically based upon what the variable was supposed to measure.

In addition, we should base our reasoning on our theoretical construct. What we decide regarding the meaning of a factor has to fall within the limits of our construct.



Assumption checks

before starting a factor analysis:

- ***N***: at least 300 respondents (Tabachnick & Fidell, 2007); may be smaller ($N \sim 150$) with high factor loadings (above .80)
- ***correlations per variable***: at least some above .30 and none with .90 or higher



For being on the safe side when conducting a factor analysis, one should have a dataset consisting of at least 300 respondents (Tabachnick & Fidell, 2007). The number of participants can be lower and a smaller dataset (of approximately 150 respondents) may also be suitable when factor loadings are high (above .80).

Regarding the strength of the correlations between the variables: Each variable should be correlated with a couple of other variables with above .30. Variables that don't show any correlations of that magnitude should possibly be excluded as they contribute little to the understanding of the structure of the data.

We also have to check for the other extreme:

Variables that have a correlation of .90 or higher, cause multicollinearity, and one of the two variables in such a pair of highly correlating variables should be removed.



Assumption checks

within the factor analysis:

- **Bartlett's test of sphericity**: deviation from an identity matrix → variables are correlated with one another → can be “summarized” into a factor
- **Kaiser-Meyer-Olkin (KMO)**: < 0.6 not suitable; 0.6 – 0.7 mediocre; 0.7 – 0.9 good; > 0.9 excellent



Within the factor analysis procedure, we can request two further statistical measures:

Bartlett's test of sphericity assesses to what degree our correlation matrix deviates from an identity matrix. In an identity matrix (ones in the main diagonal, zeros outside the main diagonal), variables are only correlated with themselves and not with other variables. The more our variable structure deviates from an identity matrix, i.e., the more variables in your dataset are correlated with each other, the better these variables can potentially be summarised with a smaller set of factors. Bartlett's test must be significant ($p < .05$).



Assumption checks

within the factor analysis:

- **Bartlett's test of sphericity**: deviation from an identity matrix → variables are correlated with one another → can be “summarized” into a factor
- **Kaiser-Meyer-Olkin (KMO)**: < 0.6 not suitable; 0.6 – 0.7 mediocre; 0.7 – 0.9 good; > 0.9 excellent



The Kaiser-Meyer-Olkin (KMO) criterion of sampling adequacy is conducted for each possible pair of variables and then averaged. It is tested whether only that pair of variables is correlated (without any further correlation with other variables). The correlation for each variable pair is set into relation to the partial correlation for that variable pair (while the influence of all other variables is controlled for). If the correlation is equally large as the partial correlation, the two variables don't further correlate with other variables. This means, a factor involving this variable pair would likely be just based on these two variables. Typically, we aim for to extract factors based on several variables.

KMO assesses this danger: KMO values smaller than 0.6 indicates that EFA is not suitable. Values between 0.6 and 0.7 are considered mediocre, values between 0.7 and 0.9 are good and values between 0.9 and 1.0 are excellent.



Exploratory factor analysis in jamovi

Now shall we take our theoretical reasoning into praxis by conducting an exploratory factor analysis in jamovi.

All analyses in jamovi are using a data set called `bfi_sample.omv` (which is available on MittUiB).



Getting started...

- choose «Factor» → «Exploratory Factor Analysis»
- assign the variable to the variable box
- choose extraction method: typically «Minimum Residuals» works well

Then you start the factor analysis by choosing «Factor» from the icon bar and then selecting «Exploratory Factor Analysis». In the window that opens, you assign the variables you want to include to the variable box.

Afterwards, you decide about the extraction method you want to choose.

“Minimum residuals” is the most stable option and therefore the default. It works like a linear regression where you with an Ordinary Least Squares try to find the solution where the sum of the squared residuals is minimal (i.e., ε from some slides ago is minimized).



Getting started...

- choose «Factor» → «Exploratory Factor Analysis»
- assign the variable to the variable box
- choose extraction method: typically «Minimum Residuals» works well

“Maximum likelihood” is probably the choice with the potential to lead to the best extraction (i.e., the clearest factor structure). It would therefore be preferable. However, it also bears a risk to be unstable for data sets that are not “well behaved”, i.e. deviating from assumptions such as minimum (0.3) and maximum correlations (0.9) between variables.

The method aims to obtain “optimal” estimates of factor loadings and unique variances as to maximize the multivariate normal likelihood function summarizing the similarity between observed covariances and the covariances accounted for by our factor structure.



Getting started...

- choose «Factor» → «Exploratory Factor Analysis»
- assign the variable to the variable box
- choose extraction method: typically «Minimum Residuals» works well

“Principal axis” is the most conventional technique. It is based upon eigenvalue decomposition of a correlation matrix. This was shown in the example in the theoretical part (slide 16/17).

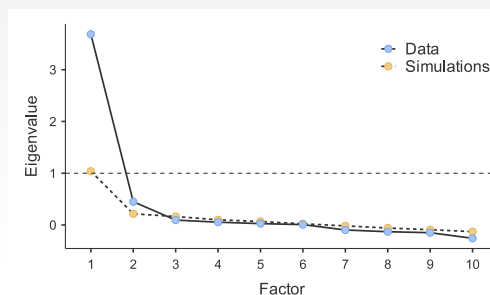
The communalities represent the proportion of the variation of one variable that can be accounted for by the remaining variables. These communalities are entered into the main diagonal.

What the eigenvalue decomposition does is to “describe” the communality (in the main diagonal) by the correlations (outside the main diagonal). It is an so-called iteration, and the procedure is repeated until the weights in the diagonal converge and do not vary any more.



Determine number of factors

- automatic: parallel analysis
- semi-automatic: eigenvalue $> x$
- manual: determine factor number



FACTOR ANALYSIS

Exploratory Factor Analysis ➔

ID

age

gender

Variables

- A3
- C1
- C2
- C3
- C4
- C5
- E1
- E2

Method

Extraction: Minimum residuals

Rotation: Oblimin

Number of Factors

Based on parallel analysis

Based on eigenvalue

Eigenvalues greater than

Fixed number

component(s)

Assumption Checks

Bartlett's test of sphericity

KMO measure of sampling adequacy

Factor Loadings

Hide loadings below

Sort loadings by size

Additional Output

Factor summary

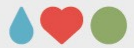
Factor correlations

Model fit measures

Initial eigenvalues

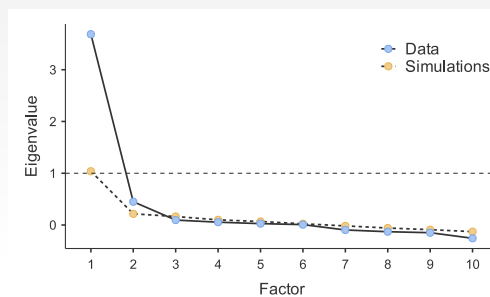
Scree plot

When determining the number of factors to be selected within our analysis, we can distinguish methods based upon their “automaticity” vs. the degree to which they require user interaction. Parallel analysis is a fully automatic method to determine the number of parameters. It determines eigenvalues to be compared with those in our observed data by using a random matrix of the same size (i.e., same number of participants and variables). If the eigenvalue obtained by the real data (blue dots) lies above the eigenvalue in the random data (yellow dots), the factor is included.



Determine number of factors

- automatic: parallel analysis
- semi-automatic: eigenvalue $> x$
- manual: determine factor number



FACTOR ANALYSIS

Exploratory Factor Analysis ➔

ID

age

gender

Variables

- R3
- C1
- C2
- C3
- C4
- C5
- E1
- E2

Method

Extraction Minimum residuals

Rotation Oblimin

Number of Factors

Based on parallel analysis

Based on eigenvalue

Eigenvalues greater than

Fixed number

component(s)

Assumption Checks

Bartlett's test of sphericity

KMO measure of sampling adequacy

Factor Loadings

Hide loadings below

Sort loadings by size

Additional Output

Factor summary

Factor correlations

Model fit measures

Initial eigenvalues

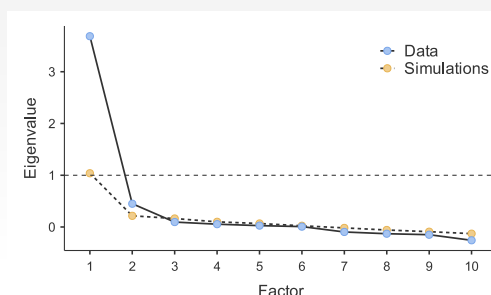
Scree plot

A semi-automatic method is to set a threshold to a certain eigenvalue. If the eigenvalue of a factor obtained from the data exceeds this threshold, the factor is included. A very common threshold is the so-called Kaiser-criterium where the threshold is set to 1. With this threshold, a factor explains at least as much variance as it contributed to the original data.



Determine number of factors

- automatic: parallel analysis
- semi-automatic: eigenvalue $> x$
- manual: determine factor number



FACTOR ANALYSIS

Exploratory Factor Analysis

Variables: age, gender, C1, C2, C3, C4, C5, E1, E2

Method: Extraction: Minimum residuals, Rotation: Oblimin

Number of Factors:

- Based on parallel analysis
- Based on eigenvalue (Eigenvalues greater than 0)
- Fixed number (1 component(s))

Assumption Checks:

- Bartlett's test of sphericity
- KMO measure of sampling adequacy

Factor Loadings:

- Hide loadings below 0.3
- Sort loadings by size

Additional Output:

- Factor summary
- Factor correlations
- Model fit measures
- Initial eigenvalues
- Scree plot

At the manual option, the number of factors is set by the user. The user may be guided by two sources of information when making that decision: (1) Theoretical assumptions – e.g., for the Big Five personality model, it is logical to assume that 5 factors should be extracted. (2) Based upon the scree plot. In the scree plot shown, there is a relative steep decline between first and second factor. Afterwards, the remaining factors have relatively low eigenvalues. There is a characteristic change in the plot that looks like an elbow. The factors before the elbow turns direction («1» in the example on the figure) are manually chosen. Typically, our decision regarding the number of factors is made up by weighing or combining the different metrics. If, e.g., scree plot and eigenvalue = 1-criterion suggest one factor and parallel analysis two, one may choose one.



Choosing rotation

- typically «Varimax» (orthogonal) or «Oblimin» (oblique) are good choices

Exploratory Factor Analysis

Variables: C1, C2, C3, C4, C5, E1, E2

Method

Extraction: Minimum residuals

Rotation: **Oblimin**

Number of Factors

Based on parallel analysis

Based on eigenvalue

Eigenvalues greater than: 0

Fixed number

1 component(s)

Assumption Checks

Bartlett's test of sphericity

KMO measure of sampling adequacy

Factor Loadings

Hide loadings below: 0.3

Sort loadings by size

Additional Output

Factor summary

Factor correlations

Model fit measures

Initial eigenvalues

Scree plot

Rotations are performed iteratively. Iteratively means that the angles for rotation are step-wise increased or decreased to reach an optimal stage. The optimization either maximizes (*Varimax*) or minimizes (*Oblimin*) an objective criterion simultaneously for all the factors.

The most common classes of rotations are either orthogonal and oblique. For an orthogonal rotation, the coordinate system in which all measurements and the factors lie remains unchanged. For an oblique rotation, the axes of the factors don't have to be orthogonal and the factors can be correlated.

Varimax is an orthogonal rotation that tries to maximize the sum of the squared loadings in each factor.

Quartimax is also an orthogonal rotation that seeks to maximize the sum of all loadings raised to power 4 (i.e. squared and squared again).



Choosing rotation

- typically «Varimax» (orthogonal) or «Oblimin» (oblique) are good choices

Exploratory Factor Analysis

Variables: C1, C2, C3, C4, C5, E1, E2

Method

Extraction: Minimum residuals

Rotation: **Oblimin**

Number of Factors

Based on parallel analysis

Based on eigenvalue

Eigenvalues greater than 0

Fixed number

1 component(s)

Assumption Checks

Bartlett's test of sphericity

KMO measure of sampling adequacy

Factor Loadings

Hide loadings below 0.3

Sort loadings by size

Additional Output

Factor summary

Factor correlations

Model fit measures

Initial eigenvalues

Scree plot

Promax is the oblique equivalent of Varimax: It keeps trying to maximize the variance of the squared loadings in each factor while permitting that the factors become correlated.

Oblimin is an oblique rotation which aims to minimize the “obliqueness”, i.e., the amount with which factors are correlated, while also maximizing the variance.

Simplimax is an oblique rotation which minimizes the small loadings in order to make the factor structure as simple as possible.



Choosing rotation

Factor Loadings	Factor					Uniqueness
	1	2	3	4	5	
A1			-0.447			0.719
A2	0.374	0.404	0.322			0.480
A3	0.499	0.345	0.370			0.442
A4	0.429			0.306		0.666
A5	0.580					0.531
C1	0.465		-0.435		0.330	0.480
C2	0.395		-0.398		0.323	0.524
C3			-0.369			0.747
C4	-0.584		0.356			0.475
C5	-0.548		0.380			0.519
E1	-0.410				0.339	0.696
E2	-0.700				0.385	0.353
E3	0.538	0.308				0.478
E4	0.605					0.425
E5	0.542					0.575
N1	-0.414	0.615				0.347
N2		0.627				0.399
N3	-0.394	0.588				0.474
N4	-0.505	0.461				0.449
N5		0.473				0.632
O1	0.391			-0.432		0.626
O2				0.443		0.682
O3	0.511			-0.487		0.414
O4			0.302			0.746
O5				0.447		0.679

Unrotated

FACTOR ANALYSIS

Factor Loadings	Factor					Uniqueness
	1	2	3	4	5	
A1				-0.473		0.719
A2				0.693		0.480
A3				0.690		0.442
A4				0.424		0.666
A5			0.325	0.549		0.531
C1				0.683		0.480
C2				0.656		0.524
C3				0.483		0.747
C4				-0.628		0.475
C5				-0.585		0.519
E1			-0.536			0.696
E2			-0.725			0.353
E3			0.602		0.327	0.478
E4			0.631		0.380	0.425
E5			0.521			0.575
N1		0.776				0.347
N2		0.761				0.399
N3		0.714				0.474
N4		0.607				0.449
N5		0.518				0.632
O1				0.534		0.626
O2				-0.504		0.682
O3				0.647		0.414
O4			0.324			0.746
O5				-0.531		0.679

Varimax

Correlation Matrix					
	1	2	3	4	5
1	1	-0.024	-0.027	-4.747e-4	-0.024
2		1	0.082	0.080	0.055
3			1	0.012	0.029
4				1	0.042
5					1

Factor Loadings	Factor					Uniqueness
	1	2	3	4	5	
A1				-0.482		0.719
A2				0.712		0.480
A3				0.683		0.442
A4				0.435		0.666
A5				0.520		0.531
C1		0.717				0.480
C2		0.676				0.524
C3		0.506				0.747
C4		-0.635				0.475
C5		-0.579				0.519
E1				-0.531		0.696
E2				-0.705		0.353
E3				0.499	0.386	0.478
E4				0.592		0.425
E5				0.421		0.575
N1		0.794				0.347
N2		0.789				0.399
N3		0.693				0.474
N4		0.504		-0.453		0.449
N5		0.473				0.632
O1					0.566	0.626
O2					-0.490	0.682
O3					0.685	0.414
O4				0.326		0.746
O5					-0.515	0.679

Oblimin

Correlation Matrix					
	1	2	3	4	5
1	1	-0.193	-0.203	0.020	-0.130
2		1	0.399	0.172	0.245
3			1	0.303	0.304
4				1	0.272
5					1

SLIDE 44



Typically, oblique rotations result in a factor structure that is easier to interpret: Factor loadings are a little larger while small, spurious loadings rather have disappeared: In the table, the factor related to extraversion (2 for Varimax, 3 for Oblimin) has more small loadings in Varimax (A5, O3) whereas the loading for N4 appears in either rotation method.

Please note that the remaining loadings are not 0 or non-existent, but rather not shown because of the threshold we set under «Hide loadings below».



Further settings

- **Assumption Checks:**
both options should be ticked and evaluated
- **Factor Loadings:**
leave the defaults (0.3)
- **Additional Output:**
tick all options except «Initial eigenvalues»

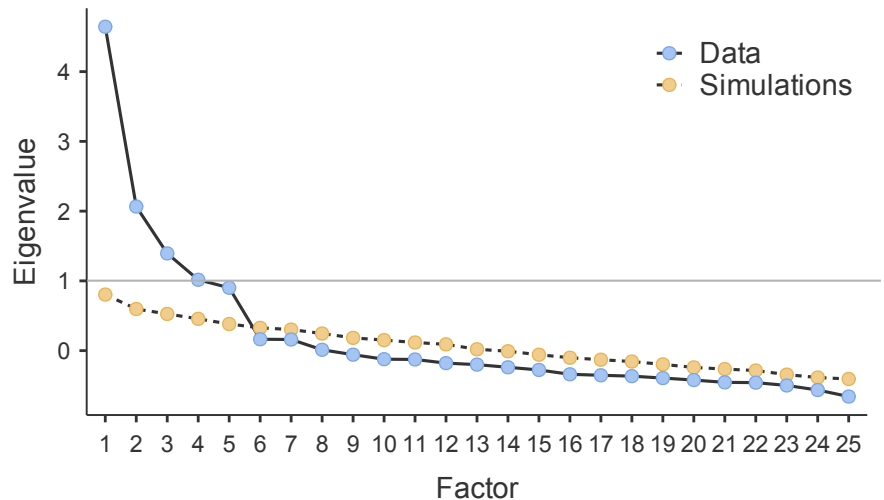
Finally, there are three categories of input:

- 1) Both assumption checks should be ticked and evaluated.
- 2) Setting the threshold from which factor loadings are shown in the table makes it easier to hold the overview when assessing the factor loadings.
- 3) From the «Additional Output» all options except «Initial eigenvalues» should be ticked.
 - «Factor summary» gives an overview of what proportion of variance each factor explains and for much variance all factors up to that factor account for in total.
 - «Factor correlations» is relevant if we choose an oblique rotation because the table reports how much the factors are correlated.
 - «Model Fit measures» allows us to assess the quality of our factor analysis.
 - «Scree plot» should always be chosen to help assess how many factors should be extracted.



Assessing the output

- **Scree plot**



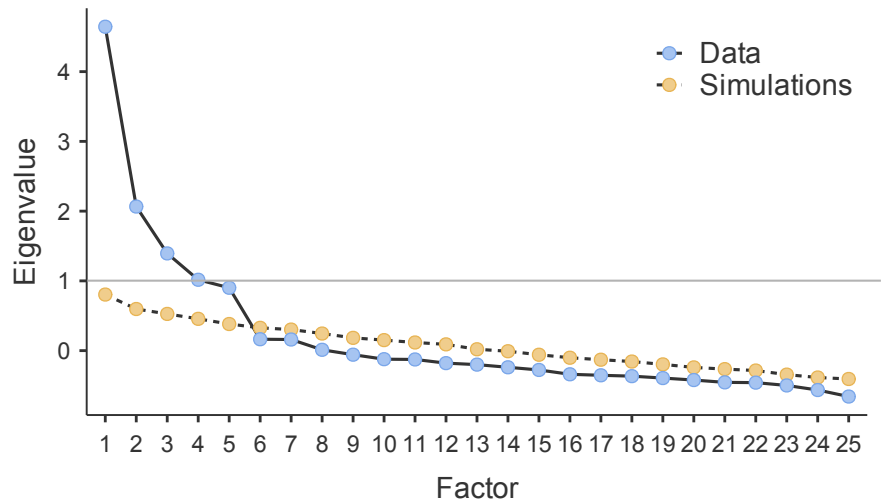
In the results, the first thing to look at should be the scree plot (at the end). Given that the automatic option – parallel analysis – is the default, you should check whether you would like to accept the number of extracted factors proposed by the parallel analysis. In our case, the parallel analysis proposes five factors. This is in accordance with our theoretical assumptions that personality can be described by five dimensions. It is also in accordance with that in the scree plot, the «elbow» is lying left from the 6th factor: After that criterium we also would extract 5 factors.

For the eigenvalues (please note that the grey line in the figure is added manually): This criterium above «1» would extract four factors. Whether the fourth is inside or not can be found out in two ways. Either just switch to the eigenvalue-criterium within «Number of extracted factors» or tick «Initial eigenvalues» to get them in a table.



Assessing the output

- **Scree plot**



When being required to decide, how you should weigh the information? Clearly, the theoretical assumptions you had before or a theoretical model that your test is based upon always should take precedence. For how the weigh the remaining information it comes down to preference.

I would weigh the scree plot next: It gives you a (relatively clear) picture of the structure of the data. In the scree plot above you see a decrease from factor 1 to 5, then a little of a drop and before the rest is rather flat and low (asymptotic). Especially such a drop as the one we have between factor 5 and 6 is rare to see. More often, we will have a less clear picture. In such cases, the scree plot might be not as informative as we have here. My preference for the scree plot primarily rests on the information we can extract from it. If there is little information in it (e.g., if it is unclear where the «elbow» are), other measures should be weighed higher.

Assessing the output

- Bartlett's Test of Sphericity has to be significant
- KMO for «Overall» should be > 0.6
check KMO for individual variables
- model fit: RMSEA < 0.10

χ^2	df	p
2204.278	300	<.0001

RMSEA	RMSEA 90% CI		TLI	BIC	Model Test		
	Lower	Upper			χ^2	df	p
0.054	0.044	0.064	0.882	-699.981	321.489	185	<.0001

FACTOR ANALYSIS

SEBASTIAN.JENTSCHKE@UIB.NO

KMO Measure of Sampling Adequacy

	MSA
Overall	0.808
A1	0.594
A2	0.838
A3	0.820
A4	0.835
A5	0.864
C1	0.800
C2	0.807
C3	0.746
C4	0.793
C5	0.842
E1	0.828
E2	0.845
E3	0.834
E4	0.869
E5	0.910
N1	0.736
N2	0.712
N3	0.768
N4	0.825
N5	0.804
O1	0.840
O2	0.698
O3	0.816
O4	0.743
O5	0.760

Next, we turn to checking the assumptions. The Bartlett's Test of Sphericity should be significant. This indicates that the correlation matrix we used for determining our factors significantly deviate from an identity matrix. Which again means that the variables in our dataset were correlated with each other, and could be summarised with a smaller set of factors.

Assessing the output

- Bartlett's Test of Sphericity has to be significant
- KMO for «Overall» should be > 0.6
check KMO for individual variables
- model fit: RMSEA < 0.10

χ^2	df	p
2204.278	300	<.0001

RMSEA	RMSEA 90% CI		TLI	BIC	Model Test		
	Lower	Upper			χ^2	df	p
0.054	0.044	0.064	0.882	-699.981	321.489	185	<.0001

FACTOR ANALYSIS

SEBASTIAN.JENTSCHKE@UIB.NO

KMO Measure of Sampling Adequacy

	MSA
Overall	0.808
A1	0.594
A2	0.838
A3	0.820
A4	0.835
A5	0.864
C1	0.800
C2	0.807
C3	0.746
C4	0.793
C5	0.842
E1	0.828
E2	0.845
E3	0.834
E4	0.869
E5	0.910
N1	0.736
N2	0.712
N3	0.768
N4	0.825
N5	0.804
O1	0.840
O2	0.698
O3	0.816
O4	0.743
O5	0.760

The KMO-measure for the whole test is a little over 0.8 which counts as “good”. It might be worthwhile to check the KMO-values for the individual variables in order to see whether it should be considered not to include individual variables. These values are generally in the same range as the KMO-value for “Overall”. “A1” is a bit of an exception with 0.594. Unfortunately, there are no firm recommendations, regarding when to exclude a variable. I would start considering that if the value were under 0.5 and if the “Overall”-value is much higher (difference over 0.3).

Assessing the output

- Bartlett's Test of Sphericity has to be significant
- KMO for «Overall» should be > 0.6
check KMO for individual variables
- model fit: RMSEA < 0.10

χ^2	df	p
2204.278	300	<.0001

RMSEA	RMSEA 90% CI		TLI	BIC	Model Test		
	Lower	Upper			χ^2	df	p
0.054	0.044	0.064	0.882	-699.981	321.489	185	<.0001

FACTOR ANALYSIS

SEBASTIAN.JENTSCHKE@UIB.NO

KMO Measure of Sampling Adequacy

	MSA
Overall	0.808
A1	0.594
A2	0.838
A3	0.820
A4	0.835
A5	0.864
C1	0.800
C2	0.807
C3	0.746
C4	0.793
C5	0.842
E1	0.828
E2	0.845
E3	0.834
E4	0.869
E5	0.910
N1	0.736
N2	0.712
N3	0.768
N4	0.825
N5	0.804
O1	0.840
O2	0.698
O3	0.816
O4	0.743
O5	0.760

The final bit of checking is evaluating the model fit.

Typically, to use model fit criteria isn't too common in the context of an EFA. These criteria are much more needed and appropriate when conducting a confirmatory factor analysis (CFA). Whereas we have a hypothetical model in the CFA, in the EFA we are (kind of) comparing our factor structure to a "model" without (i.e., 0) factors.

The measure which should be checked is the root mean square error of approximation (RMSEA). This fit index is an estimate of the discrepancy between the model and the data. Values less than .05 constitute good fit, values between 0.05 and 0.08 constitute acceptable fit, a values between 0.08 and 0.10 constitute marginal fit and values greater than 0.10 indicate poor fit. Our model lies therefore within the margin of an acceptable fit, just above the border to "good".



Assessing the output

- Factor loadings, summary and correlations

Summary			
Factor	SS Loadings	% of Variance	Cumulative %
1	2.613	10.452	10.452
2	2.379	9.516	19.968
3	2.407	9.628	29.597
4	2.207	8.826	38.423
5	1.836	7.343	45.766

Inter-Factor Correlations					
	1	2	3	4	5
1	—	-0.158	-0.162	0.023	-0.104
2		—	0.315	0.133	0.187
3			—	0.229	0.223
4				—	0.203
5					—

	Factor					Uniqueness
	1	2	3	4	5	
A1				-0.482		0.719
A2				0.712		0.480
A3				0.683		0.442
A4				0.435		0.666
A5				0.520		0.531
C1		0.717				0.480
C2		0.676				0.524
C3		0.506				0.747
C4		-0.635				0.475
C5		-0.579				0.519
E1			-0.531			0.696
E2			-0.705			0.353
E3			0.493		0.386	0.478
E4			0.592	0.362		0.425
E5			0.421			0.575
N1	0.794					0.347
N2	0.789					0.399
N3	0.693					0.474
N4	0.504		-0.453			0.449
N5	0.473					0.632
O1					0.566	0.626
O2					-0.490	0.682
O3					0.685	0.414
O4				0.326		0.746
O5					-0.515	0.679

Note: 'Minimum residual' extraction method was used in combination with a 'oblimin' rotation



Once we decided about how many of the extracted factors we would like to include and after checking assumptions and model fit, we can turn to the main results.

The factor loadings give you an impression about which items (or more generally variables) contribute to a certain factor (which again also indicates what the factor might mean). Please note that the parts which are blank typically still contain factor loadings, but small ones that are not shown based upon the setting not to show loadings below 0.3 (set this to 0 to see the table with all loadings). The loadings reveal a relatively clear factor structure with factor 1 mainly receiving contributions (i.e., the highest loadings) from N-items, factor 2 from C-items, factor 3 mainly from E-items (with an additional contribution from N4), factor 4 from A-items and factor 5 from O-items (with O4 having an odd loading pattern within factor 4 instead of 5).



Assessing the output

- Factor loadings, summary and correlations

Summary

Factor	SS Loadings	% of Variance	Cumulative %
1	2.613	10.452	10.452
2	2.379	9.516	19.968
3	2.407	9.628	29.597
4	2.207	8.826	38.423
5	1.836	7.343	45.766

Inter-Factor Correlations

	1	2	3	4	5
1	—	-0.158	-0.162	0.023	-0.104
2		—	0.315	0.133	0.187
3			—	0.229	0.223
4				—	0.203
5					—

Factor Loadings

	Factor					Uniqueness
	1	2	3	4	5	
A1				-0.482		0.719
A2				0.712		0.480
A3				0.683		0.442
A4				0.435		0.666
A5				0.520		0.531
C1		0.717				0.480
C2		0.676				0.524
C3		0.506				0.747
C4		-0.635				0.475
C5		-0.579				0.519
E1			-0.531			0.696
E2			-0.705			0.353
E3			0.493		0.386	0.478
E4			0.592	0.362		0.425
E5			0.421			0.575
N1	0.794					0.347
N2	0.789					0.399
N3	0.693					0.474
N4	0.504		-0.453			0.449
N5	0.473					0.632
O1					0.566	0.626
O2					-0.490	0.682
O3					0.685	0.414
O4				0.326		0.746
O5					-0.515	0.679

Note: 'Minimum residual' extraction method was used in combination with a 'oblimin' rotation



Regardless of that “outlier” the factor structure (i.e., how the factor loadings are distributed between factors) is still quite clear and we can turn to the next part of the results, the factor summary and their correlations.

In the factor summary, the most important information is how much of the variation in the original data is explained by our factors. After the rotation, the proportion to which each factor explains variance is quite equally distributed, all explain around 10%. This together accounts for about 46% of the original variation. Ideally, this could be a little higher (80% would be excellent).



Assessing the output

- Factor loadings, summary and correlations

Summary			
Factor	SS Loadings	% of Variance	Cumulative %
1	2.613	10.452	10.452
2	2.379	9.516	19.968
3	2.407	9.628	29.597
4	2.207	8.826	38.423
5	1.836	7.343	45.766

	Factor					Uniqueness
	1	2	3	4	5	
A1				-0.482		0.719
A2				0.712		0.480
A3				0.683		0.442
A4				0.435		0.666
A5				0.520		0.531
C1		0.717				0.480
C2		0.676				0.524
C3		0.506				0.747
C4		-0.635				0.475
C5		-0.579				0.519
E1			-0.531			0.696
E2			-0.705			0.353
E3			0.493		0.386	0.478
E4			0.592	0.362		0.425
E5			0.421			0.575
N1	0.794					0.347
N2	0.789					0.399
N3	0.693					0.474
N4	0.504		-0.453			0.449
N5	0.473					0.632
O1					0.566	0.626
O2					-0.490	0.682
O3					0.685	0.414
O4				0.326		0.746
O5					-0.515	0.679

Note: 'Minimum residual' extraction method was used in combination with a 'oblimin' rotation



Correlations between factors are typically in the range of 0.0 to 0.4 – which would be regarded marginal to small in terms of Cohen's effect sizes. This is desirable because the higher the correlations are, the more difficult it gets to explain what the factor means. If we consider, e.g., factor 1 we would like it to be as clearly as possible “connected” to one underlying cause (in this case neuroticism) without having too much influence or contributions from the other factor.



Assessing the output

- Factor loadings, summary and correlations

Summary			
Factor	SS Loadings	% of Variance	Cumulative %
1	2.613	10.452	10.452
2	2.379	9.516	19.968
3	2.407	9.628	29.597
4	2.207	8.826	38.423
5	1.836	7.343	45.766

	Factor					Uniqueness
	1	2	3	4	5	
A1				-0.482		0.719
A2				0.712		0.480
A3				0.683		0.442
A4				0.435		0.666
A5				0.520		0.531
C1		0.717				0.480
C2		0.676				0.524
C3		0.506				0.747
C4		-0.635				0.475
C5		-0.579				0.519
E1			-0.531			0.696
E2			-0.705			0.353
E3			0.493		0.386	0.478
E4			0.592	0.362		0.425
E5			0.421			0.575
N1	0.794					0.347
N2	0.789					0.399
N3	0.693					0.474
N4	0.504		-0.453			0.449
N5	0.473					0.632
O1					0.566	0.626
O2					-0.490	0.682
O3					0.685	0.414
O4				0.326		0.746
O5					-0.515	0.679

Note: 'Minimum residual' extraction method was used in combination with a 'oblimin' rotation



The most important task when calculating a factor analysis is the interpretation of what the factors mean.

The quite clear factor structure makes it relatively easy for the current data set. A description what the variables mean can be found here:

<https://personality-project.org/r/html/bfi.html>

We can mostly rely on which items load on which factor and how high. For factor 1 all items loading on the factor are those designed to assess behaviour which is regarded neurotic (e.g., is the item N1 with the highest loading “Get angry easily.”).

For factors where the loadings are a little more spread (such as factor 4) we have to integrate the meaning from the items assessing agreeableness with other ones such as O4 (“Spend time reflecting on things.”) where “things” possibly includes the own behaviour towards others.



Reporting an EFA

describe the:

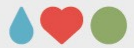
- 1) theoretical underpinnings and constructs
- 2) sample (e.g., demographics, sample size, etc.)
- 3) type of data (e.g., nominal, continuous) and descriptive statistics
- 4) test of assumpt., incl. details regd. sphericity, sampling adequacy
- 5) extraction method (e.g., maximum likelihood)
- 6) criteria and process for the number of factors
- 7) rotation methods, and the results of the rotation
- 8) final factor loadings in a table, incl. uniqueness
- 9) meaningful names for the extracted factors



Finally, we have to describe what we did in our factor analysis in order to report it in a manuscript.

Unfortunately, there is not a formal standard way to write up an EFA, and different disciplines and researchers seem vary considerably. However, some fairly standard pieces of information should be included in your report:

- 1) What are the theoretical underpinnings for the area you are studying, and specifically for the constructs that you are interested in uncovering through EFA.
- 2) A description of the sample (e.g., demographic information, sample size, sampling method).
- 3) Describe of the type of data used (e.g., nominal, continuous) and provide descriptive statistics.
- 4) Describe the testing of assumptions for the EFA. Details regarding sphericity checks and measures of sampling adequacy should be reported.



Reporting an EFA

describe the:

- 1) theoretical underpinnings and constructs
- 2) sample (e.g., demographics, sample size, etc.)
- 3) type of data (e.g., nominal, continuous) and descriptive statistics
- 4) test of assumpt., incl. details regd. sphericity, sampling adequacy
- 5) extraction method (e.g., maximum likelihood)
- 6) criteria and process for the number of factors
- 7) rotation methods, and the results of the rotation
- 8) final factor loadings in a table, incl. uniqueness
- 9) meaningful names for the extracted factors



- 5) Explain what FA extraction method (e.g., «minimum residuals») was used.
- 6) Explain the criteria and process used for deciding how many factors were extracted in the final solution. Clearly explain the rationale for key decisions.
- 7) Explain what rotation methods were attempted, the reasons why (at least for orthogonal vs. oblique), and the results.
- 8) Final factor loadings should be reported in the results, in a table. This table should also report the uniqueness (or communality) for each variable (in the final column). Factor loadings should be reported with descriptive labels in addition to item numbers. Correlations between the factors should also be included, either at the bottom of this table, in a separate table.



Reporting an EFA

describe the:

- 1) theoretical underpinnings and constructs
- 2) sample (e.g., demographics, sample size, etc.)
- 3) type of data (e.g., nominal, continuous) and descriptive statistics
- 4) test of assumpt., incl. details regd. sphericity, sampling adequacy
- 5) extraction method (e.g., maximum likelihood)
- 6) criteria and process for the number of factors
- 7) rotation methods, and the results of the rotation
- 8) final factor loadings in a table, incl. uniqueness
- 9) meaningful names for the extracted factors



- 9) Meaningful names for the extracted factors should be provided. You may like to use previously selected factor names (e.g., from a theoretical model you study was based upon, if such exist). It also might be, that on examining the actual items and factors you may think a different name is more appropriate. Generally, this process should be guided by the variables showing the highest loadings on the factor as those are most indicative.
- Remember that making sense of the factors is your main contribution. It is worth spending time on finding the most appropriate description for these factors, and how you arrived at that description.



Principal component analysis in jamovi

The introduction to the principal component analysis will be rather brief. As I said in the introduction: While being mathematically similar in the way the calculations are carried out for “Principal components” in the PCA and “Principal axis” in the EFA; the conceptual difference is considerable.

Whereas the EFA aims to describe the variation in our measured variables by an underlying factor (latent variable), the main aim of the PCA is data reduction (i.e., the components that are extracted don't necessarily have to have a connection to a theoretical concept as long as they do a reasonable job in compressing the original amount of data).



Principal Component Analysis

Principal Component Analysis

Method: Rotation **Oblimin**

Assumption Checks: Bartlett's test of sphericity, KMO measure of sampling adequacy

Number of Components: Based on parallel analysis, Based on eigenvalue (Eigenvalues greater than), Fixed number (component(s))

Factor Loadings: Hide loadings below Sort loadings by size

Additional Output: Component summary, Component correlations, Initial eigenvalues, Scree plot

Exploratory Factor Analysis

Method: Extraction **Minimum residuals**, Rotation **Oblimin**

Assumption Checks: Bartlett's test of sphericity, KMO measure of sampling adequacy


Number of Factors: Based on parallel analysis, Based on eigenvalue (Eigenvalues greater than), Fixed number (component(s))

Factor Loadings: Hide loadings below Sort loadings by size

Additional Output: Factor summary, Factor correlations, Initial eigenvalues, Scree plot, Model fit measures

FACTOR ANALYSIS

SLIDE 59



When looking at the input window for Principal Components Analysis and Exploratory Factor Analysis, they look more or less identical. There are small differences such as that there is no choice of methods for extraction (there is only one in PCA) and that model fit measures are missing. Model fit measures are not commonly used in EFA and for the PCA they conceptually don't make much sense: Why would I check a fit with an underlying concept or model if there isn't one, just the aim to reduce the data. Finally, there is a slight differences in wording with "component" being replaced by "factor" in Factor summary etc.



Principal Component Analysis

Component Loadings						Factor Loadings							
	Component					Uniqueness		Factor					Uniqueness
	1	2	3	4	5			1	2	3	4	5	
A1				-0.628		0.520	A1				-0.482		0.719
A2				0.763		0.387	A2				0.712		0.480
A3				0.725		0.374	A3				0.683		0.442
A4				0.497		0.547	A4				0.435		0.666
A5				0.574		0.456	A5				0.520		0.531
C1			0.767			0.386	C1		0.717				0.480
C2			0.754			0.405	C2		0.676				0.524
C3			0.653			0.576	C3		0.506				0.747
C4			-0.674			0.402	C4		-0.635				0.475
C5			-0.633			0.431	C5		-0.579				0.519
E1		-0.672				0.559	E1			-0.531			0.696
E2		-0.734				0.320	E2			-0.705			0.353
E3		0.637			0.336	0.383	E3			0.493		0.386	0.478
E4		0.648		0.363		0.360	E4			0.592	0.362		0.425
E5		0.542				0.480	E5			0.421			0.575
N1	0.815					0.302	N1	0.794					0.347
N2	0.830					0.324	N2	0.789					0.399
N3	0.767					0.384	N3	0.693					0.474
N4	0.599	-0.426				0.376	N4	0.504		-0.453			0.449
N5	0.565					0.504	N5	0.473					0.632
O1				0.640		0.486	O1				0.566		0.626
O2				-0.626		0.511	O2				-0.490		0.682
O3				0.672		0.360	O3				0.685		0.414
O4			0.430	0.347		0.573	O4			0.326			0.746
O5				-0.663		0.497	O5				-0.515		0.679

Note: 'oblimin' rotation was used

Note: 'Minimum residual' extraction method was used in combination with a 'oblimin' rotation



Given that the underlying mathematical calculations are very similar, it doesn't come too much as a surprise that the pattern of loadings for these components vs. factors looks very much alike. That factor 2 and 3 are swapped and that the values for uniqueness for the components in the PCA are generally lower comes from the different assumptions and what goes into the calculations. Whereas for the PCA all variance in the variables is considered, the EFA only considers the variance that is common with (one could also say: explained by) other variables.



Confirmatory factor analysis in jamovi

The next part of the lecture you may skip. Please read the possible use cases on the next two slides and decide whether it might be relevant for you. Otherwise, continue at “Reliability analysis”.



Introduction

- tests the presumed relationship between observed and latent variables
- several purposes:
 - confirm the “measurement model”
 - test a theoretically developed model of a psychological construct
 - get estimate of the measurement errors in our sample



Often, we have the case that the presumed relationship between observed and latent variables is known. This knowledge can be based upon theory (e.g., that we claim a relationship – factor structure – with certain properties), empirical (we developed a questionnaire in an earlier study, did a factor analysis and know this factor structure), or both.

One purpose of such confirmatory factor analysis might be that we want to confirm what is called the “measurement model”: the relationship between the observed and their proposed or known underlying factors. One reason to be interested in this question might be if you used an existing questionnaire in your study, developed through an earlier study and possibly in a different lab. You would like to ensure that the theoretical assumptions, the questionnaire was based upon, still hold for the sample from which you acquired your data.



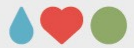
Introduction

- tests the presumed relationship between observed and latent variables
- several purposes:
 - confirm the “measurement model”
 - test a theoretically developed model of a psychological construct
 - get estimate of the measurement errors in our sample



The other common use case is that you theoretically developed a model where you proposed a psychological construct and described which aspects (i.e., factors) contribute in what way to that construct. Based upon that construct, you develop a questionnaire and collect data with that questionnaire. For the data collected, you would like to know, to what degree these data are in accordance with the psychological construct you proposed.

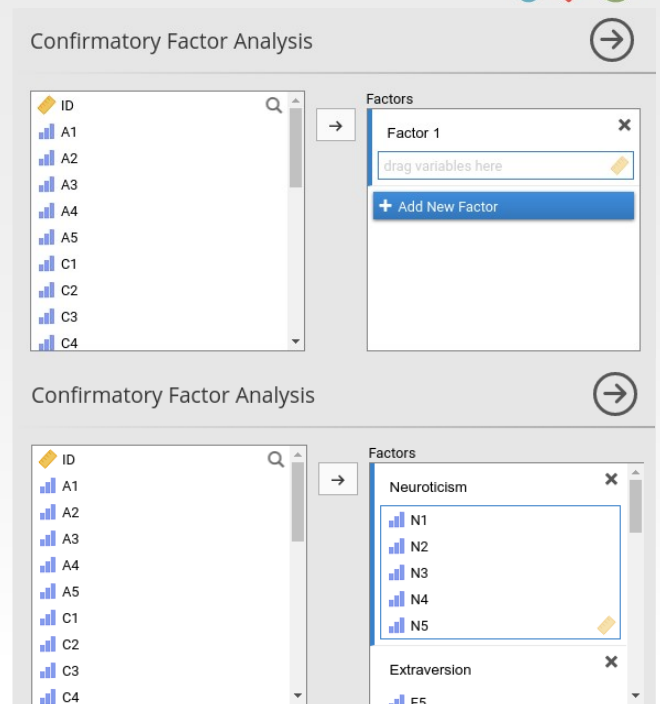
Finally, one further purpose is – also by comparing the proposed or known factor structure with our empirical data – to get an estimate of the measurement errors in our sample.



Getting started...

- choose «Factor» → «Confirmatory Factor Analysis»
- edit «Factor 1» in your variable box into a recognizable name
assign variables belonging to that factor
- add as many further factors as needed and assign vars.

FACTOR ANALYSIS



We start a confirmatory factor analysis (CFA) before we open out statistics software by specifying our hypothetical model: That is, we describe which variables belong to which factor and decide whether these factors are supposed to be independent of each other (orthogonal) or correlated (oblique).

One word before we start: I prefer a different order of factors than the one in the jamovi-book: Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. Don't be confused.



Getting started...

- choose «Factor» → «Confirmatory Factor Analysis»
- edit «Factor 1» in your variable box into a recognizable name
assign variables belonging to that factor
- add as many further factors as needed and assign vars.

FACTOR ANALYSIS

The image displays two screenshots of the Jamove software interface for Confirmatory Factor Analysis. The top screenshot shows the initial state with a list of variables (A1-A5, C1-C4) on the left and a 'Factors' panel on the right containing 'Factor 1' and an 'Add New Factor' button. The bottom screenshot shows the same interface after editing 'Factor 1' to 'Neuroticism' and assigning variables N1-N5 to it. A second factor, 'Extraversion', is also visible at the bottom.

To conduct a CFA in jamove, we choose «Factor» from the icon bar and choose «Confirmatory Factor Analysis». Then we assemble the factors as they are laid out in our hypothetical model. The default name «Factor 1» should be edited into a recognizable name in accordance with your model. For our dataset, we choose Neuroticism. We then assign the variables belonging to that factor (N1 to N5 in our example).

Afterwards add as many further factors as needed and assign the respective variables. For our example that are: Extraversion (E1 to E5), Openness to experience (O1 to O5), Agreeableness (A1 to A5), Conscientiousness (C1 to C5). The order in which the factors are added doesn't make a difference.



Assessing the output

- **model fit:**

Test for Exact Fit			Fit Measures				
χ^2	df	p	CFI	TLI	RMSEA	RMSEA 90% CI	
						Lower	Upper
739.726	265	< .0001	0.762	0.731	0.085	0.077	0.092



In order to quantify, how well our hypothetical model fits our empirical data, we can use several model fit criteria, including the χ^2 test for exact fit, the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI), the Root mean square error of approximation (RMSEA) as well as the (Standardized) Root Mean Square Residual (SRMR) and Akaike's and Bayes's information criteria (AIC, BIC).

A general remark regarding the χ^2 test first: It is most appropriate for models with about 75 to 200 cases. If the sample size exceeds 400, it has a tendency to always get significant. Useful rules of thumb for interpreting these measures are that a satisfactory fit is indicated by $CFI > 0.9$, $TLI > 0.9$, and RMSEA of about 0.05 to 0.08. A good fit is $CFI > 0.95$, $TLI > 0.95$, and RMSEA with an upper CI < 0.05 .



Assessing the output

- **model fit:**

Test for Exact Fit			Fit Measures				
χ^2	df	p	CFI	TLI	RMSEA	RMSEA 90% CI	
						Lower	Upper
739.726	265	< .0001	0.762	0.731	0.085	0.077	0.092



When looking at the output, we can see that the χ^2 value is large and highly significant. Given that our sample size is not too large ($N = 250$), this indicates a poor fit. The other indices point in the same direction: A CFI of 0.762, a TLI of 0.731, and an RMSEA of 0.085 with a 90% confidence interval from 0.077 to 0.092 all indicate a poor fit between the model and the data.

A possible explanation for those results is that we had only about half of the variance in the data accounted for by the five factor model in our earlier EFA.



Assessing the output

- factor loadings and covariances:

Factor Covariances						
		Estimate	SE	Z	p	Stand. Estimate
Neuroticism	Neuroticism	1.000 ^a				
	Extraversion	-0.238	0.076	-3.122	0.0018	-0.238
	Openness to experience	-0.182	0.078	-2.320	0.0203	-0.182
	Agreeableness	0.164	0.077	2.140	0.0323	0.164
	Conscientiousness	-0.289	0.074	-3.917	<.0001	-0.289
Extraversion	Extraversion	1.000 ^a				
	Openness to experience	0.532	0.067	7.949	<.0001	0.532
	Agreeableness	-0.581	0.063	-9.213	<.0001	-0.581
	Conscientiousness	0.504	0.065	7.756	<.0001	0.504
Openness to experience	Openness to experience	1.000 ^a				
	Agreeableness	-0.424	0.071	-5.948	<.0001	-0.424
	Conscientiousness	0.277	0.081	3.440	0.0006	0.277
Agreeableness	Agreeableness	1.000 ^a				
	Conscientiousness	-0.335	0.074	-4.507	<.0001	-0.335
Conscientiousness	Conscientiousness	1.000 ^a				

^a fixed parameter

FACTOR ANALYSIS

Factor Loadings						
Factor	Indicator	Estimate	SE	Z	p	Stand. Estimate
Neuroticism	N1	1.339	0.094	14.174	<.0001	0.821
	N2	1.215	0.093	13.067	<.0001	0.765
	N3	1.217	0.097	12.487	<.0001	0.747
	N4	0.956	0.110	8.715	<.0001	0.574
	N5	0.842	0.110	7.631	<.0001	0.494
Extraversion	E5	0.778	0.083	9.370	<.0001	0.596
	E4	1.040	0.092	11.278	<.0001	0.692
	E3	0.911	0.092	9.889	<.0001	0.627
Openness to experience	E2	-1.261	0.100	-12.667	<.0001	-0.761
	E1	-0.898	0.106	-8.512	<.0001	-0.550
	O1	0.665	0.082	8.119	<.0001	0.580
	O2	-0.653	0.115	-5.660	<.0001	-0.403
	O3	0.999	0.088	11.301	<.0001	0.831
Agreeableness	O4	0.233	0.085	2.735	0.0062	0.197
	O5	-0.514	0.093	-5.521	<.0001	-0.405
	A1	0.468	0.099	4.727	<.0001	0.328
Conscientiousness	A2	-0.771	0.079	-9.761	<.0001	-0.621
	A3	-1.061	0.082	-13.002	<.0001	-0.788
	A4	-0.749	0.096	-7.830	<.0001	-0.516
	A5	-0.941	0.082	-11.443	<.0001	-0.711
Conscientiousness	C1	0.802	0.076	10.531	<.0001	0.670
	C2	0.747	0.087	8.558	<.0001	0.569
	C3	0.539	0.087	6.191	<.0001	0.421
	C4	-1.067	0.086	-12.414	<.0001	-0.760
	C5	-1.121	0.104	-10.775	<.0001	-0.678



Within the factor loadings, the standardized estimates for each of the parameters as well as their z-statistic and p-value indicate that each variable makes a reasonable contribution to the model. There doesn't appear to be any reason to remove any of the specified variable-factor paths. The factor covariance estimates further indicate that our assumption the factors in the model might be correlated is supported by the data. The factor-factor correlations from the model as well as their respective z-statistic and p-value are also all significant.



Assessing the output

- modification indices:

	Neuroticism	Extraversion	Openness to experience	Agreeableness	Conscientiousness
N1		1.082	0.147	1.147	0.591
N2		7.580	3.178	1.034	9.833
N3		0.013	0.009	0.207	0.007
N4		28.786	0.652	1.653	14.031
N5		0.843	2.896	1.129	0.028
E5	2.507		2.894	2.400	8.775
E4	0.302		9.370	22.831	1.511
E3	6.201		28.017	4.395	5.853
E2	19.013		2.334	5.360	1.670
E1	1.357		3.140	13.545	0.501
O1	0.280	1.889		0.069	0.190
O2	6.849	8.060		6.429	2.426
O3	1.037	7.699		2.711	2.093
O4	8.873	10.541		1.851	13.387
O5	2.188	2.974		1.584	4.494
A1	1.532	14.214	0.391		11.138
A2	3.645	1.018	0.534		1.203
A3	2.952	4.695	0.001		1.760
A4	0.060	4.152	2.299		6.992
A5	8.179	11.412	3.501		3.843
C1	0.545	1.962	0.147	4.216	
C2	12.077	0.009	0.007	1.399	
C3	12.774	1.542	0.863	0.849	
C4	2.708	1.053	0.214	0.960	
C5	4.746	1.028	0.206	0.226	



To improve the model, there are two choices: One option is to go back and think again about the items / measures we are using and how they might be improved or changed. Another option is to make some post hoc tweaks to the model to improve the fit. One way of doing this is to select (tick) “Modification indices”, specified in the drop-down menu “Additional output”.



Assessing the output

- modification indices:

	Neuroticism	Extraversion	Openness to experience	Agreeableness	Conscientiousness
N1		1.082	0.147	1.147	0.591
N2		7.580	3.178	1.034	9.833
N3		0.013	0.009	0.207	0.007
N4		28.786	0.652	1.653	14.031
N5		0.843	2.896	1.129	0.028
E5	2.507		2.894	2.400	8.775
E4	0.302		9.370	22.831	1.511
E3	6.201		28.017	4.395	5.853
E2	19.013		2.334	5.360	1.670
E1	1.357		3.140	13.545	0.501
O1	0.280	1.889		0.069	0.190
O2	6.849	8.060		6.429	2.426
O3	1.037	7.699		2.711	2.093
O4	8.873	10.541		1.851	13.387
O5	2.188	2.974		1.584	4.494
A1	1.532	14.214	0.391		11.138
A2	3.645	1.018	0.534		1.203
A3	2.952	4.695	0.001		1.760
A4	0.060	4.152	2.299		6.992
A5	8.179	11.412	3.501		3.843
C1	0.545	1.962	0.147	4.216	
C2	12.077	0.009	0.007	1.399	
C3	12.774	1.542	0.863	0.849	
C4	2.708	1.053	0.214	0.960	
C5	4.746	1.028	0.206	0.226	



When checking the items / variables with the highest modification index value (MI), we could decide whether it would make sense to add additional terms to the model. Variable N4 (“Often feel blue”) with an MI of 28.786 does not only load on the factor “Neuroticism” it was assigned to but also on factor “Extraversion”. If we add this path into the model (i.e., adding N4 as a variable inside “Extraversion”) then the χ^2 value will reduce by around the amount of the MI.

However, given the underlying theoretical construct (where the variable is part of the factor “Neuroticism”) this path arguably doesn’t really make any theoretical or methodological sense, so it’s not a good idea (unless you can come up with a persuasive argument).



Reporting a CFA

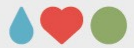
describe the:

- 1) theoretical and empirical justification for the hypothesized model
- 2) model specification (detailed: each latent factor with its indicator variables)
- 3) sample (e.g., demographics, sample size, etc.)
- 4) type of data (e.g., nominal, continuous) and descriptive statistics
- 5) estimation method used (and, possibly, assumption checks)
- 6) (possibly) missing data and how they were handled
- 7) software and version used to fit the model
- 8) criteria used, to assess the model fit
- 9) alterations made to the original model
- 10) parameter estimates (i.e., loadings, error variances, latent (co)variances) and their standard errors



When reporting a CFA it is recommended to include the following information in your report:

- 1) A theoretical and empirical justification for the hypothesized model.
- 2) A complete description of how the model was specified (e.g., each latent factor with its indicator variables, possible covariances between latent variables, and any correlations between error terms). Possibly, a path diagram would be good to summarize that.
- 3) A description of the sample (e.g., demographic information, sample size, sampling method).
- 4) A description of the type of data used (e.g., nominal, continuous) and descriptive statistics.
- 5) The estimation method used (and, possibly, assumption checks you conducted).



Reporting a CFA

describe the:

- 1) theoretical and empirical justification for the hypothesized model
- 2) model specification (detailed: each latent factor with its indicator variables)
- 3) sample (e.g., demographics, sample size, etc.)
- 4) type of data (e.g., nominal, continuous) and descriptive statistics
- 5) estimation method used (and, possibly, assumption checks)
- 6) (possibly) missing data and how they were handled
- 7) software and version used to fit the model
- 8) criteria used, to assess the model fit
- 9) alterations made to the original model
- 10) parameter estimates (i.e., loadings, error variances, latent (co)variances) and their standard errors



- 6) A description of whether there were missing data and how the missing data were handled.
- 7) The software and version used to fit the model.
- 8) Measures, and the criteria used, to judge model fit (χ^2 , CFI, TLI, RMSEA, etc.).
- 9) Any alterations made to the original model based on model fit or modification indices.
- 10) All parameter estimates (i.e., loadings, error variances, latent (co)variances) and their standard errors, probably in a table.



Reliability analysis

Finally, questionnaires (also called scales) are often developed using EFA. The Big Five in personality research constitute a known example for that. It started with submitting responses to short statements describing personality characteristics and mental states to an EFA. Afterwards, the derived factor structure – Neuroticism, Extraversion, Openness to experience, Agreeableness and Conscientiousness was further assessed using confirmatory factor analyses. This served to assess whether the proposed structure was in accordance with a theoretically proposed, hypothetical model with different samples.



Introduction

- how well measures a scale that combines the observed variables the proposed factors?
- reliability → internal consistency → Cronbach's α
- $\alpha = 1 - \text{error variance}$
- not a measure of unidimensionality
- $\alpha > 0.70$ – acceptable; $\alpha > 0.80$ – good; $\alpha > 0.95$ – to high



There is another aspect of interest: How well are the proposed factors measured using a scale that combines the observed variables? In psychometrics we use reliability analysis to provide information about how consistently a scale measures a psychological construct. Internal consistency is one aspect of reliability: It refers to the consistency across the individual items that make up a measurement scale. So, if we have V_1 , V_2 , V_3 , V_4 and V_5 as observed item variables, we can calculate a statistic that tells us how internally consistent these items are in measuring the underlying construct.



Introduction

- how well measures a scale that combines the observed variables the proposed factors?
- reliability → internal consistency → Cronbach's α
- $\alpha = 1 - \text{error variance}$
- not a measure of unidimensionality
- $\alpha > 0.70$ – acceptable; $\alpha > 0.80$ – good; $\alpha > 0.95$ – to high



Often Cronbach's α is used to assess the internal consistency of a scale (Cronbach, 1951). It is a measure of equivalence (whether different sets of scale items would give the same measurement outcomes). Equivalence is tested by dividing the scale items into two groups (a "split-half") and seeing whether the analysis of the two parts gives comparable results. Cronbach's α is a function of all the split-half coefficients for a scale. If a set of items that measures a construct (e.g., an Extraversion scale) has an alpha of 0.80, then the proportion of error variance in the scale is 0.20.



Introduction

- how well measures a scale that combines the observed variables the proposed factors?
- reliability → internal consistency → Cronbach's α
- $\alpha = 1 - \text{error variance}$
- not a measure of unidimensionality
- $\alpha > 0.70$ – acceptable; $\alpha > 0.80$ – good; $\alpha > 0.95$ – to high



However, Cronbach's α is not a measure of unidimensionality (i.e. an indicator whether a scale is measuring a single factor or construct rather than multiple related constructs). An α of 0.80 does not mean that 80% of a single underlying construct is accounted for. It could be that the 80% comes from more than one underlying construct. That's why it is useful to do an EFA and a CFA first.

Furthermore, α tends to be sample specific: it is not alone a characteristic of the scale, but also reflects the sample in which the scale has been used. A biased, unrepresentative, or small sample could produce a very different α than a large, representative sample.



Introduction

- how well measures a scale that combines the observed variables the proposed factors?
- reliability → internal consistency → Cronbach's α
- $\alpha = 1 - \text{error variance}$
- not a measure of unidimensionality
- $\alpha > 0.70$ – acceptable; $\alpha > 0.80$ – good; $\alpha > 0.95$ – to high



Nevertheless, despite these limitations, Cronbach's α has been popular for estimating reliability using internal consistency (it's pretty easy to calculate, understand and interpret, and therefore an useful initial check on scale performance, e.g., when you administer a scale with a different sample, from a different setting or population).

An alternative is McDonald's ω . Whereas alpha makes the assumptions: (a) no residual correlations, (b) items have identical loadings (i.e., the variables are summed up / averaged), and (c) the scale is unidimensional, ω does not and is therefore a more robust reliability statistic. If the assumptions are not violated then α and ω will be similar, otherwise ω is to be preferred.



Introduction

- how well measures a scale that combines the observed variables the proposed factors?
- reliability → internal consistency → Cronbach's α
- $\alpha = 1 - \text{error variance}$
- not a measure of unidimensionality
- $\alpha > 0.70$ – acceptable; $\alpha > 0.80$ – good; $\alpha > 0.95$ – to high



Typically, an α of 0.70 represents “acceptable” whereas 0.80 represents “good” reliability. Thresholds like this should, however, be used cautiously, as they also depend on what exactly the scale is supposed to be measuring.

An α of above 0.95 indicates high inter-correlations between the items. This indicates too much overlap in what the items assess and poses a risk that the construct being measured is perhaps overly narrow and redundant.



Getting started...

- «Factor» → «Reliability analysis»
- assign variables to «Items»
- check «Scale Reliability Statistics» for warnings regd. inverted items
- revert those items

	mean	sd	Cronbach's α	McDonald's ω
scale	4.207	0.771	0.469	0.622

Note. item 'A1' correlates negatively with the total scale and probably should be reversed

Reverse Scaled Items

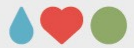
Normal Scaled Items	Reverse Scaled Items
A2	A1
A3	
A4	
A5	

FACTOR ANALYSIS

To conduct a Reliability Analysis in jamovi, click on “Factor” in the icon bar and select “Reliability analysis”. In the window that opens, assign the variables for which you want to assess the reliability to the variable box (i.e., the variables measuring / contributing to one factor).

Check the table “Scale Reliability” in the output first. It will tell you whether an item correlates negatively and has to be inverted. Inverted items are supposed to counter response tendencies (e.g., participants always saying “strongly agree” or “strongly disagree”).

In the case of the Agreeableness-scale, the inverted item is A1 (“Am indifferent to the feelings of others” – indicating an attitude or an behaviour which is opposite of agreeable).



Getting started...

- tick «Cronbach's α » and «McDonald's ω » for both Item and Scale Statistics
- tick «Mean» and «Stand. Deviation» for Scale

Reliability Analysis

ID: C1, C2, C3, C4, C5, E1, E2

Items: A1, A2, A3, A4, A5

Scale Statistics

- Cronbach's α
- McDonald's ω
- Mean
- Standard deviation

Item Statistics

- Cronbach's α (if item is dropped)
- McDonald's ω (if item is dropped)
- Mean
- Standard deviation
- Item-rest correlation

Additional Options

- Correlation heatmap



After we did that, we should tick other output options we would like to use. I chose all options within “Scale Statistics” (i.e., Cronbach's α , McDonald's ω , Mean and Standard deviation) and within “Item Statistics”, Cronbach's α and McDonald's ω .



Getting started...

- tick «Cronbach's α » and «McDonald's ω » for both Item and Scale Statistics
- tick «Mean» and «Stand. Deviation» for Scale

Reliability Analysis

ID: C1, C2, C3, C4, C5, E1, E2

Items: A1, A2, A3, A4, A5

Scale Statistics

- Cronbach's α
- McDonald's ω
- Mean
- Standard deviation

Item Statistics

- Cronbach's α (if item is dropped)
- McDonald's ω (if item is dropped)
- Mean
- Standard deviation
- Item-rest correlation

Additional Options

- Correlation heatmap



Once we are done with the assembling the first scale (Agreeableness), we will set up the analyses for the other scales. A time-saving way of doing that is to duplicate the first analysis four times (so that we have five analyses altogether). This is done by right-clicking within the output of our Reliability Analysis, and to choose “Analysis” → “Duplicate” in the menu that opens.

Afterwards, we go from analysis to analysis and change the variables from A1 to A5 to C1 to C5 for the second analysis, E1 to E5 for the third, N1 to N5 for the fourth, and O1 to O5 for the fifth. Repeat assigning the items to “Reversed Scale Items” where you got a warning that they should be reversed within each of these analyses.



Assessing the output

- first assess Cronbach's α and McDonald's ω
- check scale mean and std. deviation, think in terms of a normal distribution
- any item that would raise α or ω considerable (> 0.1)

Scale Reliability Statistics				
	mean	sd	Cronbach's α	McDonald's ω
scale	4.612	0.939	0.722	0.741

Item Reliability Statistics		
	if item dropped	
	Cronbach's α	McDonald's ω
A1 ^a	0.744	0.754
A2	0.646	0.684
A3	0.607	0.631
A4	0.705	0.725
A5	0.663	0.685

^a reverse scaled item



Then we can assess the analyses. All reveal generally acceptable to good α 's, (0.722, 0.759, 0.780, 0.811) except for openness (0.630). The latter might in part be due to our observation that the O4-item had that little strange pattern of correlations when we did the assumption check at the very beginning: It was the only item that correlated with any of the remaining variables with less than 0.3.

The mean and the standard deviation of a scale can be used to check whether the scale might be subject to floor or ceiling effects. In the example, the mean is 4.612. If we use a 6-point-scale, this mean is already relatively high. Given that the standard deviation is 0.939, that leaves about 1.5 standard deviations at top. Therefore, about 6.7% of participants have to choose the highest available rating category even though they might have chosen an even more extreme answer.



Assessing the output

- first assess Cronbach's α and McDonald's ω
- check scale mean and std. deviation, think in terms of a normal distribution
- any item that would raise α or ω considerable (> 0.1)

Scale Reliability Statistics				
	mean	sd	Cronbach's α	McDonald's ω
scale	4.612	0.939	0.722	0.741

Item Reliability Statistics		
	if item dropped	
	Cronbach's α	McDonald's ω
A1 ^a	0.744	0.754
A2	0.646	0.684
A3	0.607	0.631
A4	0.705	0.725
A5	0.663	0.685

^a reverse scaled item



The last thing to check is the item statistics: If there were any item where dropping the item would lead to a considerable raise in α or ω (i.e., a difference of more than 0.1).

In our example this threshold would be $\alpha > 0.822$ or $\omega > 0.841$. There is no item where this would be the case.



Summary and literature



Summary

- some theoretical background
- Exploratory Factor analysis (EFA)
- Principal Component Analysis (PCA)
- Confirmatory Factor Analysis (CFA)
- Reliability Analysis



Let's summarize today's lecture: We started with some background for the factor analysis such as latent variables, and how these latent variables influence the correlation patterns among manifest variables. Afterwards, we introduced the steps of a factor analysis: Factor extraction, factor selection, factor rotation and factor interpretation.

We continued with introducing how to conduct an exploratory factor analysis (EFA) in jamovi, incl. how to read the output.

Then there was a short demonstration about the principal component analysis being mathematically similar to the EFA while being conceptually different.

The part on confirmatory factor analysis was likely skipped by the majority of you.

Reliability analysis explained one aspect of reliability – internal consistency – and how it can be checked within jamovi.



Literature

- Navarro, D. J., & Foxcroft, D. R. (2022). *Learning statistics with jamovi*.
<https://doi.org/10.24384/hgc3-7p15> (Ch. 15.1 – 15.2, [15.3], 15.5; p. 419 – 439, [439 – 450], 458 – 464)



Most of what was covered in today's lecture can be found in chapter 15 of the jamovi-book (even a bit more... ;-). Relevant are the sub-chapters 15.1 (EFA), 15.2 (PCA) and 15.5 (Reliability analysis). For those who would like to use a CFA, this is covered in sub-chapter 15.3.



**Thank you for your
attention!**

Thank you for bearing with me!



UNIVERSITY OF BERGEN

