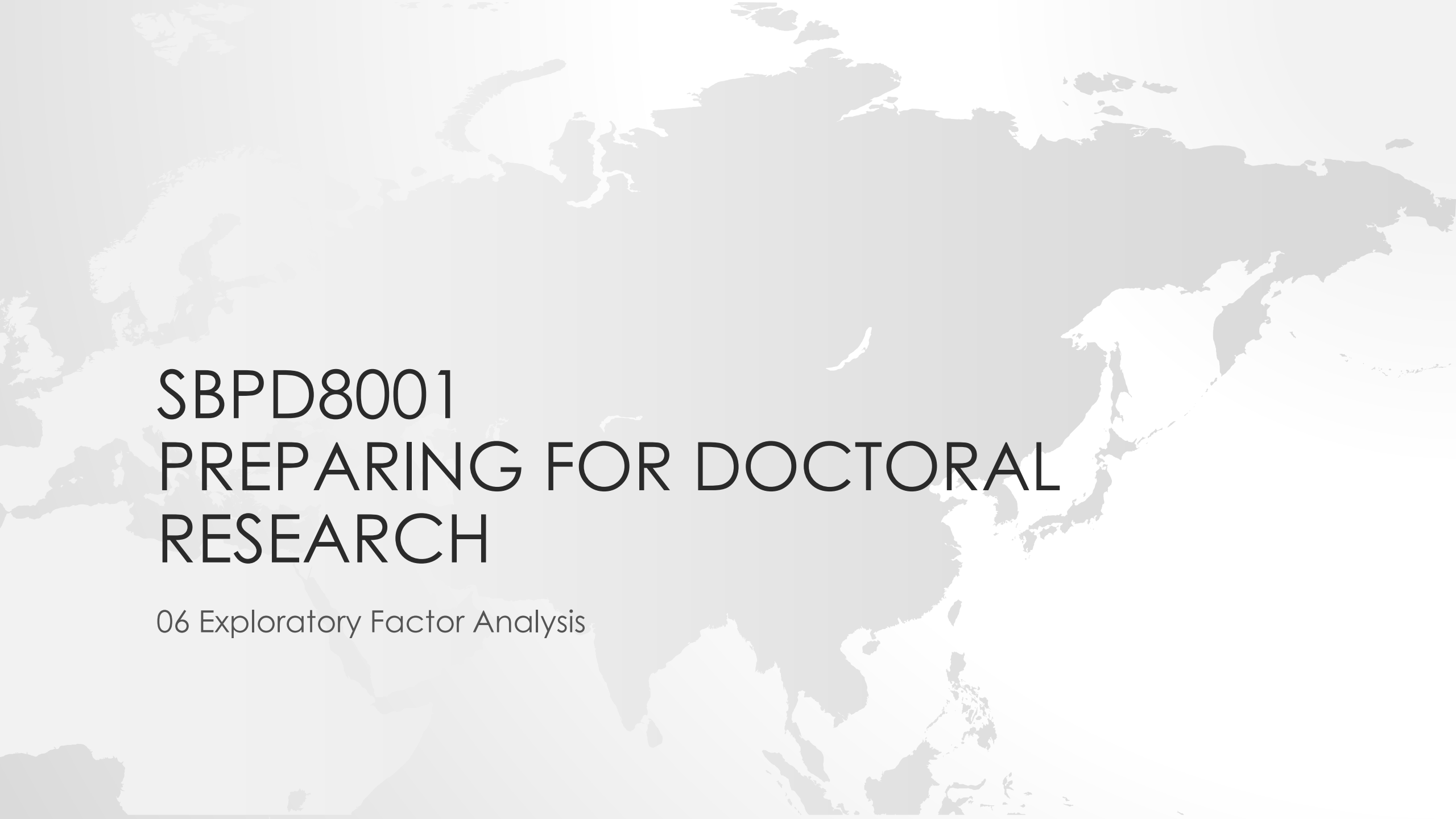


A faint, light gray world map is visible in the background, centered on the Atlantic Ocean. The map shows the outlines of continents and major landmasses.

# SBPD8001 PREPARING FOR DOCTORAL RESEARCH

06 Exploratory Factor Analysis

# TOPICS

- Types of factor analysis
- Steps in Factor Analysis

# TYPES OF FACTOR ANALYSIS

- Exploratory Factor Analysis (EFA)
  - Used to discover underlying structure
  - Principal components analysis (PCA) (Thurstone)
    - Treats individual items or measures as though they have no unique error
  - Factor analysis (principal factor analysis (PFA) or principal axis factoring (PAF)) (Spearman)
    - Treats individual items or measures as having unique error
  - Both PCA and PAF give similar answers most of the time
- Confirmatory Factor Analysis (CFA)
  - Used to test whether data fit a priori expectations for data structure
  - Structural equations modelling

# SINGLE-ITEM VERSUS MULTIPLE-ITEM SCALES

- An abstract or multi-attribute construct cannot be adequately represented by a single-item measure.
- Example: Job satisfaction: How satisfied are you with your job?
  - You are satisfied with the job content but not the salary
- Example: Service quality: How good is McDonald's service?
  - McDonald's employees are responsive (quick) but they lack empathy (not caring enough).
- We need several items to capture an abstract/multi-attribute construct. Yet, when we are dealing with multiple-item scales, we need to find ways to separate items and group them together.
  - Items that do not correlate well with others should be separated into different subscales.
  - Items that correlate strongly with each other should be grouped together.
- Factor analysis can serve this purpose.

# FACTOR ANALYSIS

What is the relationship among these relational benefit items?  
Do thematic categories emerge?

- I believe there is less risk that something will go wrong
- I feel I can trust the service provider
- I have more confidence the service will be performed correctly
- I have less anxiety when I stay in the hotel
- I know what to expect when I check in
- I get provider's highest level of service
- I am recognized by some employees
- I am familiar with the employee(s) who perform(s) the service
- I have developed a friendship with the service providers
- They know my name.
- I enjoy the social aspects

# FACTOR ANALYSIS

- Relational benefits include confidence benefits, social benefits and special treatment benefits (Gwinner, Gremler & Bitner, 1998).
  - Confidence benefits refer to perceptions of reduced anxiety and comfort in knowing what to expect in the service encounter
  - Social benefits include familiarity with the service providers, personal recognition by the firm's people, friendship, rapport and social support.

# PURPOSES

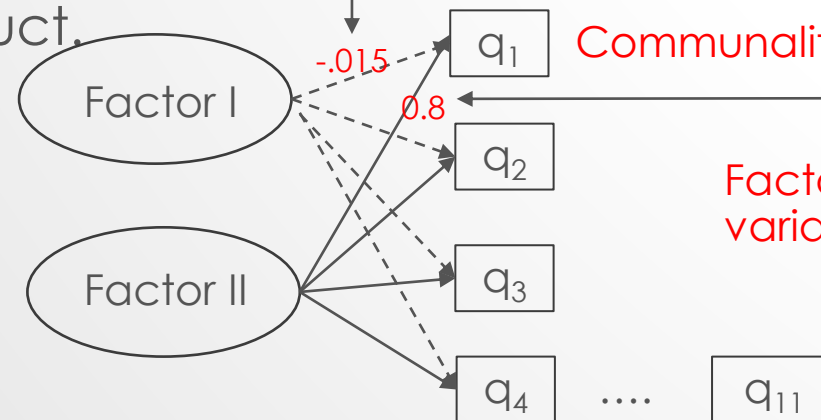
- To assess the degree to which items are tapping the same concept.
- To reduce the number of variables to few factors.
- To find general factors that may affect the things in the variable level

# KEY TERMS

- Factor loadings: A factor loading is the correlation between a variable and a factor that has been extracted from the data.

| Variable                                                                 | Factor I | Factor II |
|--------------------------------------------------------------------------|----------|-----------|
| X <sub>1</sub> I believe there is less risk that something will go wrong | -.015    | .8        |

- Interpretation: Variable q<sub>1</sub> negligibly correlates with Factor I, but highly correlates with Factor II.
- Communalities: A communality is the amount of variance in a variable that is accounted for by the factors together. It is computed by summing squares of factor loadings.
- Interpretation: A large communality value indicates a strong influence by an underlying construct.

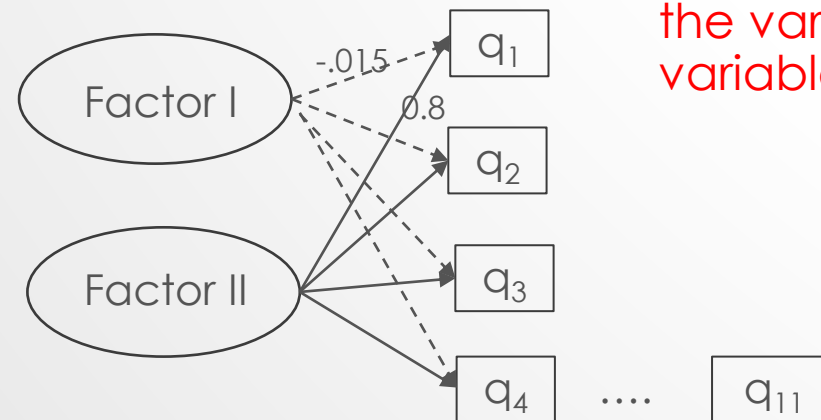


Communality of X<sub>1</sub> =  $(-.015^2 + .8^2) = 0.640225$

Factors I and II account for **64%** of the variance in q<sub>1</sub>

# KEY TERMS

- Eigenvalues: An eigenvalue indicates how much of the total variance in the dataset that is explained by a specific factor.
- Interpretation
  - A large eigenvalue indicates a meaningful factor.
  - An eigenvalue less than 1 implies that the factor accounts for less variance than a single item.



Factor II accounts for 31.718% of the variance in the dataset (11 variables)

# STEPS IN FACTOR ANALYSIS

1. Choosing the variables
2. Determining sample size
3. Factor extraction
4. Determining number of factors
  - Eigenvalues greater than 1
  - Scree plot
5. Factor rotation
  - Orthogonal (factors uncorrelated)
  - Oblique (factors correlated)
6. Factor naming
7. Conducting reliability analysis
8. Computing summated scores

# CHOOSING VARIABLES AND SAMPLE SIZE

## 1. Choosing the variables

- Variables must be continuous
- To find factors from variables, they (or some of them) must have reasonable correlations (Rule of thumb  $> .3$ )

## 2. Determining sample size

- Min 50, prefer 100
- Min 5 observations/item, prefer 10 observations/item

# 3. FACTOR EXTRACTION

## Principal components analysis

- A composite of the observed variables as a summary of those variables
- Assumes no error in items
- No assumption of underlying construct
- Often used in physical science
- Precise mathematical solutions possible
- Unity inserted on diagonal of matrix

## Principal axis factoring

- Explain relationship between observed variables in terms of latent variables or factors
- Factor is a hypothesized construct
- Assumes error in items
- Precise math not possible, solved by iteration
- Communalities (shared variance) on diagonal

## 4. DETERMINE NUMBER OF FACTORS

- At the first step there are as many factors as there are variables.
- First method: include those factors that have eigenvalue greater than 1.0.
- Second method: Look for bend in plot
- Or a priori conceptual beliefs about the number of factors from past research or theory.

# 5. FACTOR ROTATION

- All original solutions are orthogonal.
- Once you obtain minimal number of factors, you have to interpret them
- Interpreting original solutions is difficult. Rotation aids interpretation.
- You are looking for simple structure
  - Component loadings should be very high for a few variables and near 0 for remaining variables
  - Each variable should load highly on only 1 factor

# 5. FACTOR ROTATION

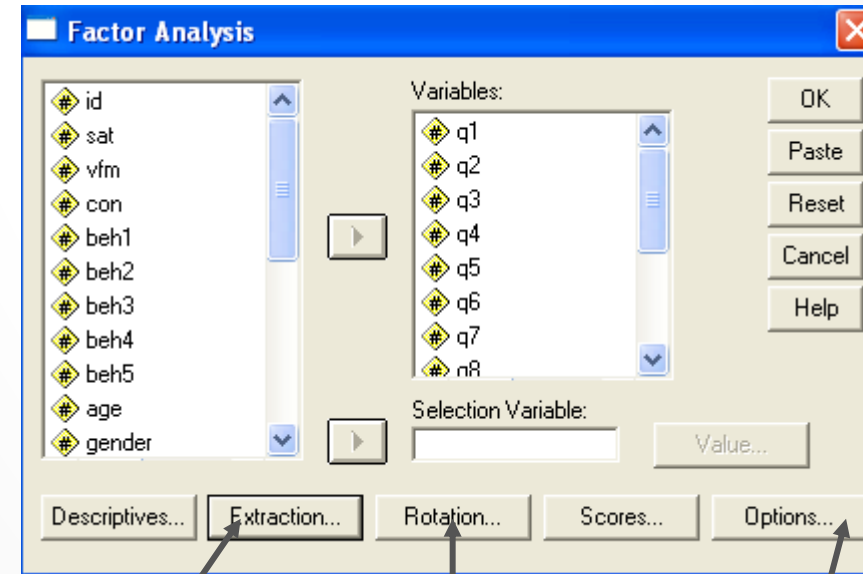
- Two types of rotations
  - Orthogonal (factors uncorrelated)
  - Oblique (factors correlated)
- Nunnally<sup>1</sup> (1978) suggests using orthogonal as opposed to oblique rotations
  - Orthogonal is simpler
  - Leads to same conclusions
  - Oblique can be misleading
- Ford et al.<sup>2</sup> (1986) suggest using oblique unless orthogonality assumption is tenable

1. Nunnally (1978). Psychometric theory.

2. Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel psychology*, 39(2), 291-314.

# FACTOR ANALYSIS

- Select Analyze > Data Reduction > Factor from the menu bar to open the Factor Analysis window.
- Highlight all of the relational benefit items, q1 through q11 from the list of variables on the left, and click on the upper arrow button to move it into the Variables box.



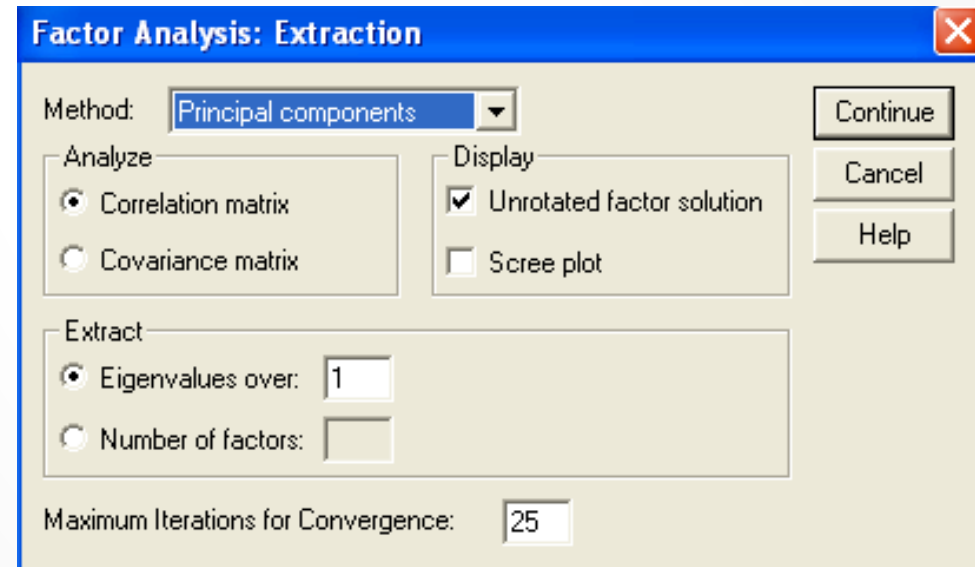
Select an extraction method, obtain a scree plot and control how many factors are extracted.

Specify a rotation method

Sort the display of factor matrices

# FACTOR ANALYSIS

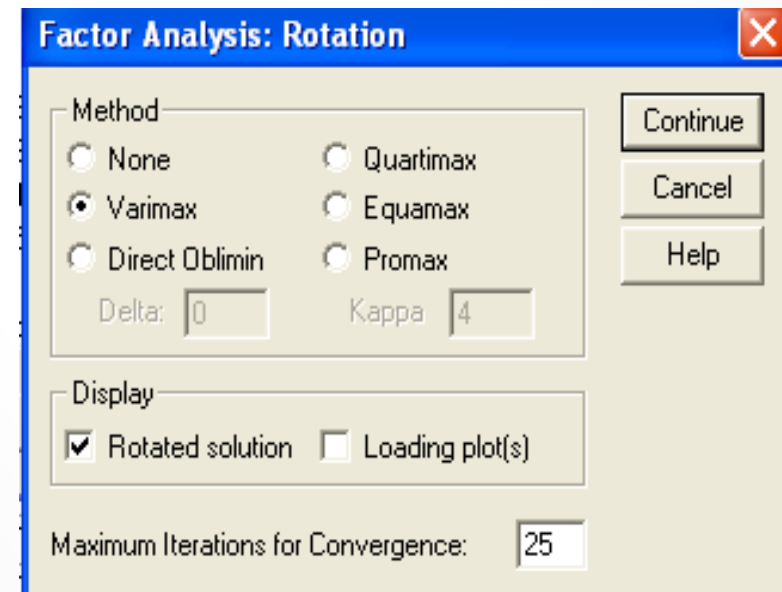
- Click on the Extraction button. A Factor Analysis: Extraction window will appear.
- “Principle components” is the default extraction method.
- By default, factors with eigenvalues greater than 1 will be extracted.
- If you click on the circle next to the word number of factors, you can specify the number of factors you want.
- Click on the Scree plot box. This will prompt SPSS to generate a scree plot which helps you determine the number of factors to be extracted.



- Click on Continue to return to the Factor Analysis window.

# FACTOR ANALYSIS

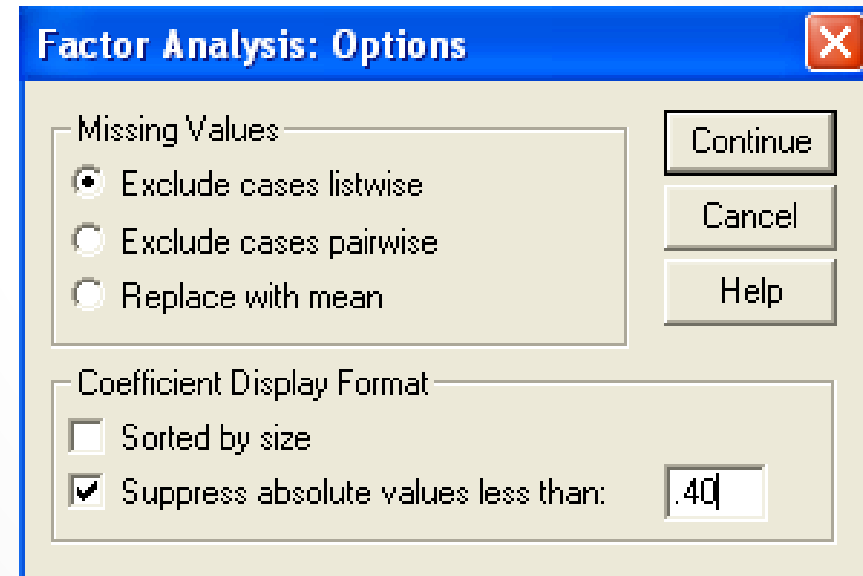
- Click on the Rotation button. A Factor Analysis: Rotation window will appear.
- Click on the circle next to the word Varimax.
- SPSS will rotate the solution to find the best fit for the data while minimizing the number of variables that have high loadings on each factor.
- Click on Continue to return to the Factor Analysis window.



The screenshot shows the 'Factor Analysis: Rotation' dialog box in SPSS. The 'Method' section has five radio buttons: 'None', 'Varimax' (which is selected), 'Direct Oblimin', 'Quartimax', and 'Equamax'. Below these are two input fields: 'Delta' with the value '0' and 'Kappa' with the value '4'. The 'Display' section has two checkboxes: 'Rotated solution' (checked) and 'Loading plot(s)' (unchecked). At the bottom, there is a label 'Maximum Iterations for Convergence:' followed by an input field containing the value '25'. On the right side of the dialog, there are three buttons: 'Continue', 'Cancel', and 'Help'.

# FACTOR ANALYSIS

- Click on the Option button. A Factor Analysis: Option window will appear.
- Click on the box next to “suppress absolute value less than”
- This will prompt SPSS to suppress coefficients with absolute values less than the specified value.
- The default is 0.10. To override the default, enter 0.4.
- Click on Continue to return to the Factor Analysis window.
- Click on OK to run the analysis.



### Communalities

|                                                                | Initial | Extraction |
|----------------------------------------------------------------|---------|------------|
| I believe there is less risk that something will go wrong      | 1.000   | .640       |
| I feel I can trust the service provider                        | 1.000   | .783       |
| I have more confidence the service will be performed correctly | 1.000   | .818       |
| I have less anxiety when I stay in the hotel                   | 1.000   | .599       |
| I know what to expect when I check in                          | 1.000   | .270       |
| I get provider's highest level of service                      | 1.000   | .405       |
| I am recognized by some employees                              | 1.000   | .765       |
| I am familiar with the employee(s) who perform(s) the service  | 1.000   | .736       |
| I have developed a friendship with the service providers       | 1.000   | .855       |
| They know my name.                                             | 1.000   | .869       |
| I enjoy the social aspects of the relationship                 | 1.000   | .814       |

Extraction Method: Principal Component Analysis.

- Communality refers to the amount of variance in an item that is accounted for by the factors together.
- Large communalities are required for a valid and useful factor solution
- It is equal to the sum of squares of all the item factor loadings.
- $0.64 = 0.371^2 + 0.709^2$

### Component Matrix<sup>a</sup>

|                                                           | Component |      |
|-----------------------------------------------------------|-----------|------|
|                                                           | 1         | 2    |
| I believe there is less risk that something will go wrong | .371      | .709 |

- $0.64 = -0.015^2 + 0.8^2$

### Rotated Component Matrix<sup>a</sup>

|                                                           | Component |      |
|-----------------------------------------------------------|-----------|------|
|                                                           | 1         | 2    |
| I believe there is less risk that something will go wrong | -.015     | .800 |

# Determining number of factors

**Extraction** Sums of Squared Loadings

**Rotation** Sums of Squared Loadings

## Total Variance Explained

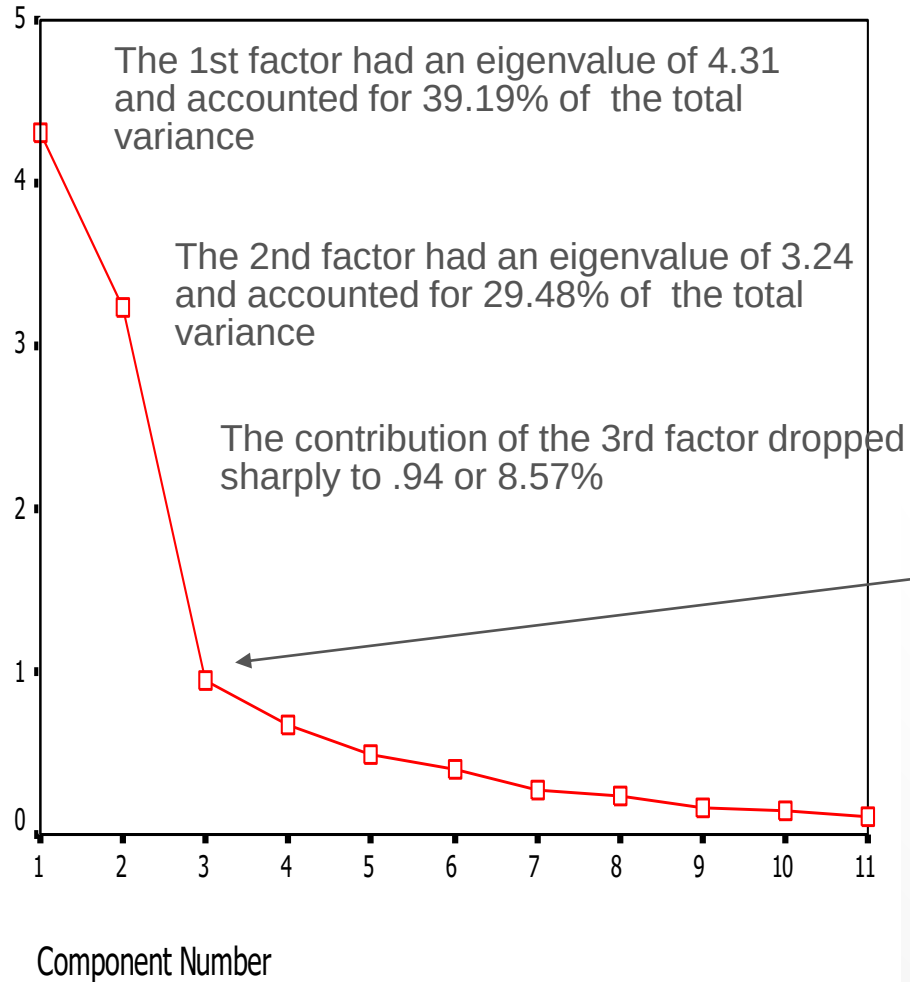
| Component | Initial Eigenvalues |               |              | Extraction Sums of Squared Loadings |               |              | Rotation Sums of Squared Loadings |               |              |
|-----------|---------------------|---------------|--------------|-------------------------------------|---------------|--------------|-----------------------------------|---------------|--------------|
|           | Total               | % of Variance | Cumulative % | Total                               | % of Variance | Cumulative % | Total                             | % of Variance | Cumulative % |
| 1         | 4.311               | 39.193        | 39.193       | 4.311                               | 39.193        | 39.193       | 4.065                             | 36.952        | 36.952       |
| 2         | 3.242               | 29.477        | 68.670       | 3.242                               | 29.477        | 68.670       | 3.489                             | 31.718        | 68.670       |
| 3         | .942                | 8.567         | 77.236       |                                     |               |              |                                   |               |              |
| 4         | .678                | 6.161         | 83.397       |                                     |               |              |                                   |               |              |
| 5         | .487                | 4.427         | 87.824       |                                     |               |              |                                   |               |              |
| 6         | .399                | 3.627         | 91.451       |                                     |               |              |                                   |               |              |
| 7         | .279                | 2.535         | 93.986       |                                     |               |              |                                   |               |              |
| 8         | .235                | 2.134         | 96.120       |                                     |               |              |                                   |               |              |
| 9         | .169                | 1.532         | 97.653       |                                     |               |              |                                   |               |              |
| 10        | .150                | 1.362         | 99.015       |                                     |               |              |                                   |               |              |
| 11        | .108                | .985          | 100.000      |                                     |               |              |                                   |               |              |

Extraction Method: Principal Component Analysis.

- Examine the Total Variance Explained table to determine how many factors emerged from the data.
- To determine the number of factors look at the second column which represents the Eigenvalues.
- As you can see there can be as many factors as there are items (i.e., there are 11 Eigenvalues and 11 items on the relational benefit scale).
- Since we don't want to retain factors which account for little variance, we can examine only those factors that have an Eigenvalue greater than 1. In this case, there are two viable factors which account for 68.67% of the variance.

# Determining number of factors

Scree Plot



- Alternatively, the number of factors to be selected can be determined by the number of factors before an 'elbow'.
- The scree plot also shows that 2 factors should be selected (number of factors = 3-1)
- Eigenvalue and scree plot may yield different results. As a general principle, factor structure should be determined by the theoretical model.
- In this case the theoretical model consists of 2 factors.

**Rotated Component Matrix<sup>a</sup>**

|                                                                | Component |      |
|----------------------------------------------------------------|-----------|------|
|                                                                | 1         | 2    |
| I believe there is less risk that something will go wrong      | -.015     | .800 |
| I feel I can trust the service provider                        | .017      | .885 |
| I have more confidence the service will be performed correctly | .067      | .902 |
| I have less anxiety when I stay in the hotel                   | .071      | .770 |
| I know what to expect when I check in                          | -.029     | .518 |
| I get provider's highest level of service                      | .183      | .610 |
| I am recognized by some employees                              | .870      | .087 |
| I am familiar with the employee(s) who perform(s) the service  | .858      | .002 |
| I have developed a friendship with the service providers       | .922      | .072 |
| They know my name.                                             | .931      | .039 |
| I enjoy the social aspects of the relationship                 | .900      | .062 |

Extraction Method: Principal Component Analysis.  
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Factor loadings < .4 have been suppressed

**Rotated Component Matrix<sup>a</sup>**

|                                                                | Component |      |
|----------------------------------------------------------------|-----------|------|
|                                                                | 1         | 2    |
| I believe there is less risk that something will go wrong      |           | .800 |
| I feel I can trust the service provider                        |           | .885 |
| I have more confidence the service will be performed correctly |           | .902 |
| I have less anxiety when I stay in the hotel                   |           | .770 |
| I know what to expect when I check in                          |           | .518 |
| I get provider's highest level of service                      |           | .610 |
| I am recognized by some employees                              | .870      |      |
| I am familiar with the employee(s) who perform(s) the service  | .858      |      |
| I have developed a friendship with the service providers       | .922      |      |
| They know my name.                                             | .931      |      |
| I enjoy the social aspects of the relationship                 | .900      |      |

Extraction Method: Principal Component Analysis.  
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

**Rotated Component Matrix<sup>a</sup>**

|                                                                | Component |      |
|----------------------------------------------------------------|-----------|------|
|                                                                | 1         | 2    |
| I believe there is less risk that something will go wrong      |           | .800 |
| I feel I can trust the service provider                        |           | .885 |
| I have more confidence the service will be performed correctly |           | .902 |
| I have less anxiety when I stay in the hotel                   |           | .770 |
| I know what to expect when I check in                          |           | .518 |
| I get provider's highest level of service                      |           | .610 |
| I am recognized by some employees                              | .870      |      |
| I am familiar with the employee(s) who perform(s) the service  | .858      |      |
| I have developed a friendship with the service providers       | .922      |      |
| They know my name.                                             | .931      |      |
| I enjoy the social aspects of the relationship                 | .900      |      |

Extraction Method: Principal Component Analysis.  
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

## Factor rotation

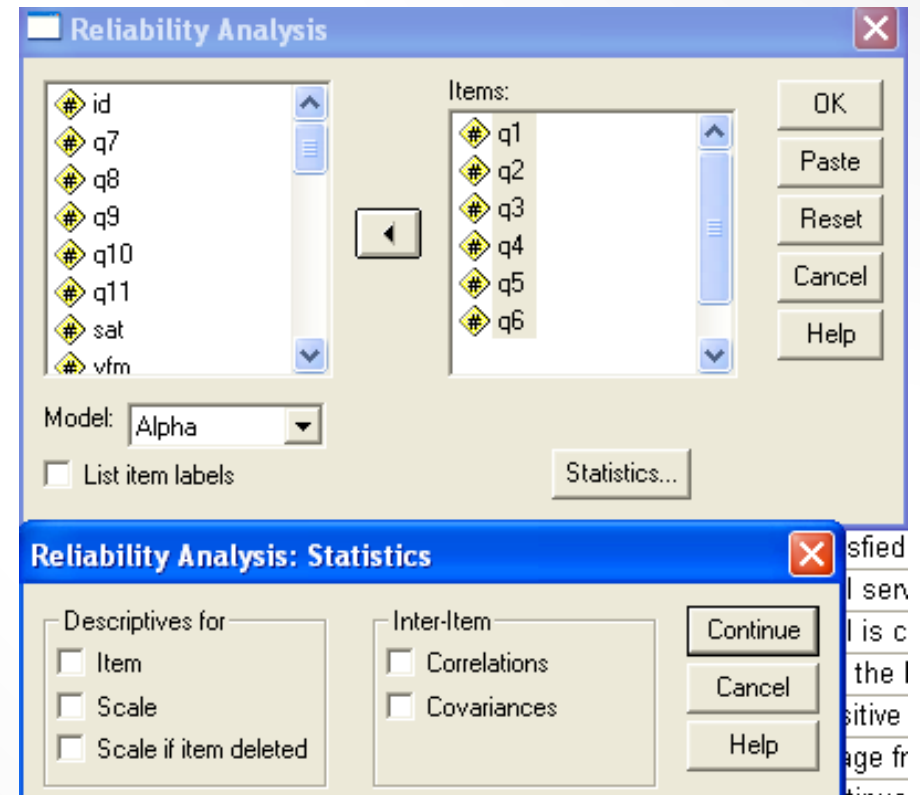
- There are two columns in the Rotated Component Matrix, one to represent each of the two factors with Eigenvalues greater than 1.
- Look at the first item (i.e., I believe there is less risk that something will go wrong). You will see a number in each of the two columns. These numbers are called **factor loadings**.
- If an item has a factor loading greater than .40 in any of the columns it is considered to be representative of that factor. For example, the first item (0.800) is representative of the second factor.
- If an item has a factor loading of .40 or higher in more than one column, it is considered to be cross loaded, meaning it is representative of both factors. You might then need to drop the item and rerun the factor analysis.

## 6. FACTOR NAMING

- Once you have identified the items in a particular factor, you will need to identify the underlying meaning of the factor.
- In this case, you will need to determine what type of relational benefit is being tapped by each of the two factors.
- Factor 1: Social benefits
- Factor 2: Confidence benefits

# 7. CONDUCTING RELIABILITY ANALYSIS

- **Key question:** Are the subscales reliable? Are they measuring the same sub-construct?
- Select Analyze > Scale > Reliability Analysis from the menu bar to open the Reliability Analysis window.
- Highlight q1 to q6 from the list of variables on the left, and click on the arrow button to move it into the items box.
- Click on the Statistics button. A Reliability Analysis: Statistics window will appear.
- Click on the box next to “scale if item deleted”
- Click on Continue to return to the Reliability Analysis window.
- Click on OK to run the analysis.



# 7. CONDUCTING RELIABILITY ANALYSIS

- Cronbach's alpha coefficient measures the extent to which individual items that constitute a multiple-item scale correlate with one another.
- It takes a value of 1.0 when the total score variance is perfectly correlated.
- Alpha should be at least .70.
- In this case alpha is 0.846
- Note that deletion of Q5 would increase alpha to .8591.
- Repeat the same procedure for Q7 to Q11

| Reliability Analysis                 |              |             |              |        |
|--------------------------------------|--------------|-------------|--------------|--------|
|                                      |              | Corrected   |              |        |
| Scale Mean                           | Scale Var if | Item-total  | Alpha if     |        |
| if Item Deleted                      | Item Deleted | Correlation | Item Deleted |        |
| Q1                                   | 24.9126      | 15.8256     | 0.6655       | 0.8136 |
| Q2                                   | 24.699       | 15.9379     | 0.7807       | 0.7901 |
| Q3                                   | 24.5437      | 16.2309     | 0.8209       | 0.7855 |
| Q4                                   | 24.3204      | 16.7101     | 0.6494       | 0.8161 |
| Q5                                   | 24.466       | 19.6434     | 0.3945       | 0.8591 |
| Q6                                   | 25.0194      | 17.8428     | 0.4868       | 0.8484 |
| Reliability Coefficients             |              |             |              |        |
| N of Cases = 103.0    N of Items = 6 |              |             |              |        |
| Alpha = .8460                        |              |             |              |        |

# 8. COMPUTING SUMMATED SCORES

- Select Transform > Compute from the menu bar to open the Compute Variable window.
- Enter “factor1” to the Target Variable window.
- Enter mean (Q1, Q2, Q3, Q4, Q5, Q6) to the Numeric Expression box
- Click on OK to compute factor1.
- A new item will be generated at the rightmost column
- Repeat the same procedure for factor2

