Books/Book Chapters   School of Management

2012

Exploratory Factor Analysis

Daire Hooper
Technological University Dublin, daire.hooper@tudublin.ie

Recommended Citation

Hooper, D. (2012), 'Exploratory Factor Analysis', in Chen, H. (Ed.), Approaches to Quantitative Research – Theory and its Practical Application: A Guide to Dissertation Students, Cork, Ireland: Oak Tree Press.

Exploratory Factor Analysis

Dr. Daire Hooper

INTRODUCTION

Factor analysis examines the inter-correlations that exist between a large number of items (questionnaire responses) and in doing so reduces the items into smaller groups, known as factors. These factors contain correlated variables and are typically quite similar in terms of content or meaning. Unlike other methods discussed in this book, exploratory factor analysis (EFA) does not discriminate between variables on whether they are independent or dependent, but rather it is an interdependence technique that does not specify formal hypotheses. It is in this sense it is 'exploratory' in nature as it allows the researcher to determine the underlying dimensions or factors that exist in a set of data. The technique is particularly useful for managerial or academic research in reducing items into discrete dimensions that can be summed or aggregated and subsequently used as input for further multivariate analysis such as multiple regression. It is also used extensively in scale development research to condense a large item pool into a more succinct, reliable and conceptually sound measurement instrument. Factor analytic techniques can typically be classified as either exploratory or confirmatory and the former of these is addressed within this chapter using a research example to demonstrate its use.

Dr. Daire Hooper, Dublin Institute of Technology, College of Business, Aungier Street, Dublin 2
e. daire.hooper@dit.ie t. +3531 402 3212

WHEN WOULD YOU USE FACTOR ANALYSIS?

There are a number of reasons why a researcher would use factor analysis. The first is when one wants to determine if a series of dimensions or factors exist in the data and whether they are interpretable in a theoretical sense. For instance, if a researcher collected data from respondents to determine how committed they were to maintaining employment in their organisation, the researcher might utilise Allen and Meyer's (1990) 24-item organisational commitment scale. Allen and Meyer (1990) propose that three sub-dimensions exist within the organisational commitment construct, these being: affective, continuance and normative commitment. Factor analysis can then be used to determine whether this three-factor structure is replicable in the dataset, in other words, to ascertain whether employees conceptually classify organisational commitment along these three dimensions.

Exploratory factor analysis would examine the inter-correlations between all variables on Allen and Meyer's (1990) scale and from that reduce the data into a smaller number of dimensions (factors). The dimensions produced by factor analysis can then be used as input for further analysis such as multiple regression. In the case of the organisational commitment example each of the items on a dimension could be summed to create an aggregate item and subsequently regressed on a dependent variable such as turnover.

The second reason to employ factor analysis would be to refine the number of items on a scale for the purposes of scale development (DeVellis, 2003). Factor analysis allows the researcher to determine the nature and number of latent

Dr. Daire Hooper, Dublin Institute of Technology, College of Business, Aungier Street, Dublin 2

e. daire.hooper@dit.ie t. +3531 402 3212

variables (dimensions/factors) underlying a set of items. One of the critical assumptions associated with scale construction is for items measuring a particular construct to be relatively homogenous or unidimensional (i.e. load together on one factor). To meet this end, factor analysis can be used to determine whether one, or multiple dimensions exist in a set of variables. Scale development is not within the scope of this book, however interested readers can refer to DeVellis's (2003) or Spector's (1992) comprehensive texts on the subject.

What is the difference between factor analysis and principal components analysis?

Too often principal components analysis (PCA) is referred to as exploratory factor analysis but this is an inaccurate classification. To a novice researcher both techniques may appear to be the same – particularly with regard to their execution and output in SPSS – however, mathematically and theoretically they differ considerably. The widespread adoption of principal components analysis can be attributed to it being the default extraction method in both SPSS and SAS (Costello & Osborne, 2005). Holding this default position has more than likely led to PCA being used mistakenly when exploratory factor analysis is more suitable (Park, Daley, & Lemus 2002). The goal of PCA is to reduce the measured variables to a smaller set of composite components that capture as much information as possible in as few components as possible. On the other hand, the goal of exploratory factor analysis (EFA) is to find the latent structure of the dataset by uncovering common factors. Therefore, exploratory factor analysis accounts for shared variance. This is an important distinction from PCA as it fundamentally means EFA is more suitable when

exploring underlying theoretical constructs. There has been much debate over which of these techniques is the true method of factor analysis, with some arguing in favour of exploratory factor analysis (Floyd & Widaman, 1995; Gorsuch, 1990; Snook & Gorsuch, 1989) while others argue there is little difference between the two (Velicer & Jackson, 1990). Principal axis factoring, a type of EFA, is superior to principal components analysis as it analyses common variance only which is a key requirement for theory development. In addition to this, it is a useful technique for identifying items that do not measure an intended factor or that simultaneously measure multiple factors (Worthington & Whittaker, 2006). For these reasons exploratory techniques are most important for theory development and will be employed here.

DATA REQUIREMENTS

Factor analysis is typically a large sample size technique, with correlations less reliable when small samples are used. Recommendations on appropriate sample sizes for factor analysis vary considerably (Fabrigar et al., 1999). Some have suggested a minimum of 300 cases is required, however in reality about 150 should be sufficient (Tabachnick and Fidell 2007) and as few as 100 cases can be adequate in situations where there are a small number of variables. The items themselves must be interval in nature (e.g. Likert scales) and although ideally multivariate normality is a requirement, deviations from this are not usually detrimental to the results. It is also important the researcher assesses for outliers as their presence can alter the factor solution (cf. Tabachnick and Fidell, 2007).

Dr. Daire Hooper, Dublin Institute of Technology, College of Business, Aungier Street, Dublin 2

e. daire.hooper@dit.ie t. +3531 402 3212

WORKED EXAMPLE

Examination of the service quality literature finds most authors describe service quality as an overall appraisal of a product or service is dependent on consumers' prior expectations (Grönroos, 1984; Bitner & Hubbert, 1994) and it is this disconfirmation-based definition that prevails most commonly in the literature. Within the services quality literature, two complementary streams of research have evolved and can be broadly categorised as being of either the Scandinavian or American tradition. As previously mentioned, both of these schools of thought agree that consumers arrive at an evaluation of service quality that is based upon disconfirmation theory. This being, prior to consuming a service, consumers hold preconceived ideas of how the service will perform. Once the consumer has experienced the service, they compare performance to their *a priori* expectations in a subtractive manner to determine their perceptions of service quality.

Parasuraman et al.'s (1985; 1988; 1994) model falls into the American tradition and is the most widely cited service quality model in the literature. Building on the premise that quality perceptions are a function of expectations and performance, they developed the Gaps Model and its associated 5-dimension SERVQUAL measurement instrument. Within the Nordic stream of research, Grönroos (1984) proposed that service quality can be described as a two factor structure comprising of both functional and technical elements. The functional element relates to the *way* in which the service is delivered, while the technical element refers to *what* the consumer receives from the service (Brogowicz et al., 1990). This functional aspect of service delivery has been referred to as peripheral

Dr. Daire Hooper, Dublin Institute of Technology, College of Business, Aungier Street, Dublin 2

e. daire.hooper@dit.ie t. +3531 402 3212

to the process while the technical element conceptually constitutes the core or outcome components of the service delivery process (Tripp & Drea, 2002). Writings on this model have been mostly theoretical (Ekinci et al., 1998), however, in more recent years a number of authors have sought to link technical and functional quality dimensions to a variety of constructs such as trust, commitment, satisfaction and loyalty (Lassar et al, 2000; Caceres & Paparoidamis, 2005) and general support has been found for the two dimensional conceptualisation of service quality.

This chapter continues this line of research by examining whether a two-dimensional model of service quality is replicated in the service stations dataset. Service quality was measured using items developed by Grace and O'Cass (2004) as well as a number of self-developed items. These can be found in Table 4.1 below. All items were measured using 7-point scales anchored with 'strongly disagree' (1) and 'strongly agree' (7).

| The service was delivered promptly |
| The service here was reliable |
| The service was efficient |
| The staff were helpful |
| The staff were polite |
| The staff were friendly* |
| The staff were trustworthy |
| The service station provided quality service* |
| The service station provided good service* |
| The service here suited my needs* |

*denotes self-developed items

Table 4.1: Service Quality Items

Factor Analysis Procedure in SPSS

Having decided these service quality items are to be used, the next stage is actually running the factor analysis. A fictitious dataset containing 355 cases was created to demonstrate the technique and can be found on the website that accompanies the book. Once you have opened the file in SPSS select **Analyze/Dimension Reduction/Factor**. At this point, a window will open and you will see all your variables on the left-hand side (see Figure 4.1 below). Select the variables you wish to include in the analysis and move them over to the **Variables** section on the right hand side. For this example we are moving across all ten items with names beginning with SQ (service quality).
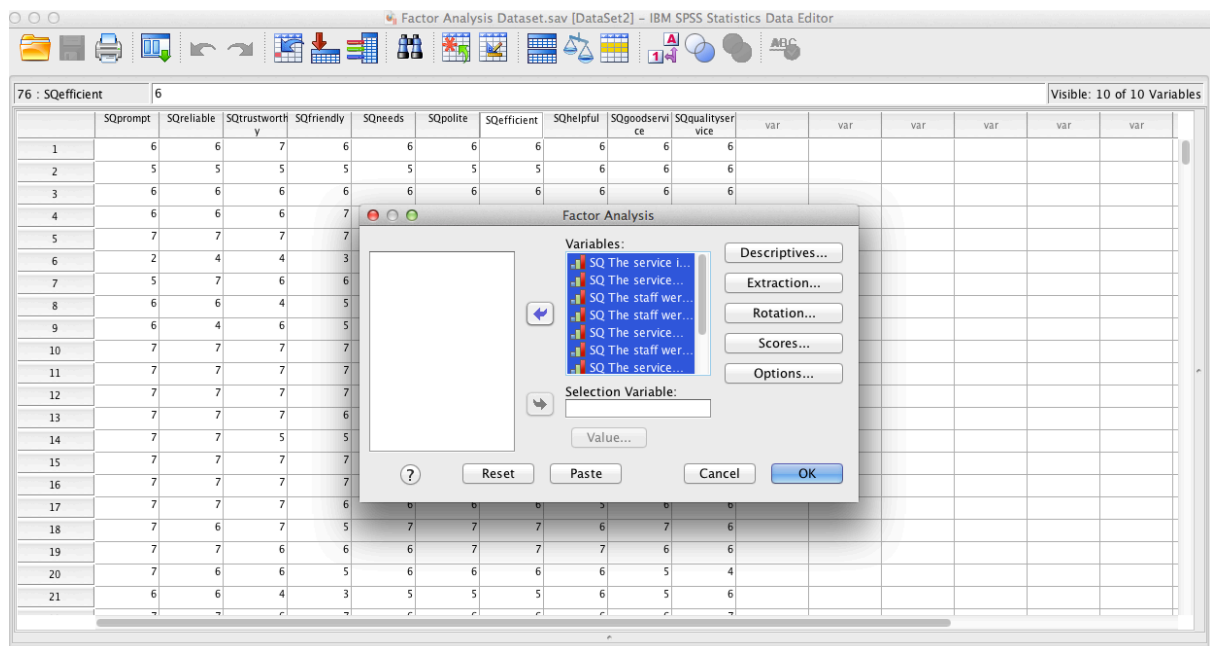

Figure 4.1: Selecting variables in Factor Analysis

Then select the **Descriptives** button and in the section marked **Correlation Matrix**, select **Coefficients** and **KMO and Bartlett's test of sphericity** and hit **Continue**.
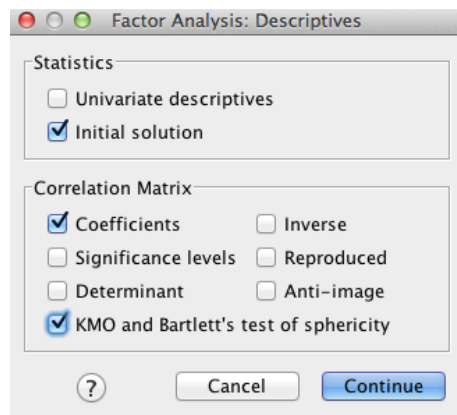
Figure 4.2: Descriptives in Factor Analysis

These are selected to test a number of assumptions associated with Factor Analysis and will be discussed later. Click on the **Extraction** button and in the **Method** section make sure **Principal axis factoring** is selected from the dropdown box (Please note: If you were using Principal Components Analysis you would choose Principal components here). In the **Analyse** section make sure **Correlation matrix** is selected. Under **Display** select **Unrotated factor solution** and tick the check box beside **Scree plot**. In the **Extract** section you will see two options with radio buttons beside each, the first is Eigenvalues greater than 1 and this is the default. For now, leave this as it is and we will return to it later. Click **Continue**.
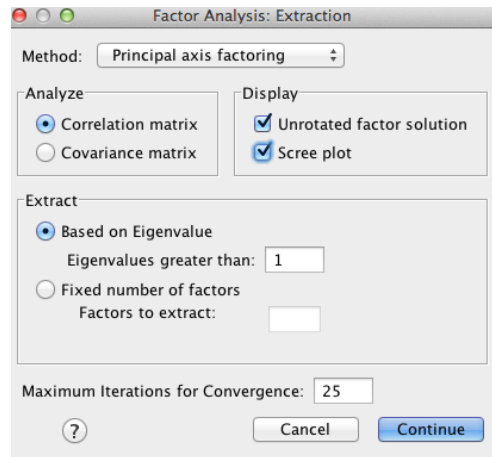
Figure 4.3: Dialog box for factor extraction

Next click on the **Rotation** button and select **Promax**. The default is for Kappa 4 here and we will leave it as it is. We have chosen to use Promax rotation as this is a type of oblique rotation which allows for correlations between factors. There are other oblique rotation methods (e.g. Direct Oblimin), however Promax is generally chosen as it is quicker and simpler. We believe service quality dimensions will be correlated with one another and this is our rationale for choosing this type of rotation. If we were using Principal Components, we would choose Varimax rotation as this is an orthogonal rotation technique which maximises the variances of loadings on the new axes.
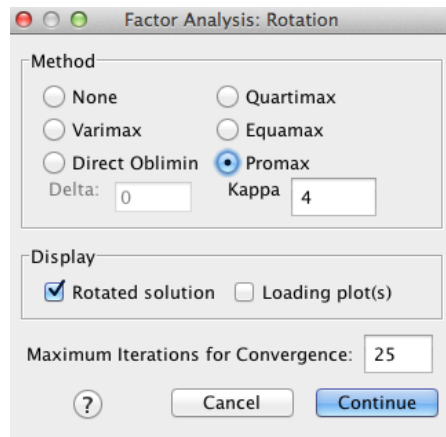
Figure 4.4: Dialog box for factor rotation

Next click on **Options** and make sure the radio button is selected beside **Exclude cases pairwise**. Following this, in the **Coefficient Display Format** section select **Sorted by size**. Sorting by size means factor coefficients will be listed from the largest down to the smallest and this will help when interpreting your results. Then select **Suppress small coefficients** and enter the value .4 in the box beside **Absolute value below**. By choosing this option SPSS will hide coefficients less than .4. This is a useful tool for a number of reasons. Firstly, it helps interpretation as we can see more clearly where particular items load. Secondly, it highlights items with loadings less than .4 on all dimensions. When an item does not load on a dimension (i.e. has loadings less than .4 on all dimensions) it may indicate the item is unreliable and as a result may be a candidate for deletion. Finally, this also shows whether any items cross-load. This means an item is loading on more than one dimension which would lead us to question the reliability of this item.
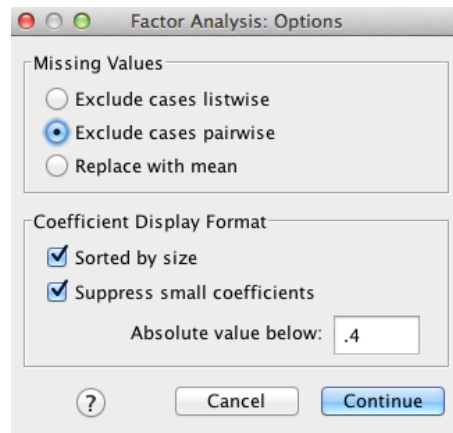
Figure 4.5: Option dialog box in factor analysis

Once all of the above have been selected the next stage is to run the analysis. To do this, click **Continue** and **OK**. The Factor Analysis output will then open in a second window known as your Output file.

Interpretation of Output

Factor analysis produces a considerable amount of output but should not deter students in its interpretation. In this next section the important pieces of information will be explained.

Stage 1 – Testing the Assumptions

The first thing you need to do is to look over the Correlation Matrix to ensure you have correlation coefficients greater than .3 in magnitude. If you do not have any correlations over .3 it might indicate factor analysis is not appropriate. In our example there are quite a number of correlations greater than .3 which tentatively suggests factor analysis is appropriate here (see Table 4.2).

Table 4.2: Correlation Matrix

	SQ The service in store was delivered promptly	SQ The service here was reliable	SQ The staff were trustworthy	SQ The staff were friendly	SQ The service here suited my needs	SQ The staff were polite	SQ The service was efficient	SQ The staff were helpful	SQ The service station provided good service	SQ The service station provided quality service
SQ The service in store was delivered promptly	1.000	.656	.445	.515	.339	.582	.674	.465	.451	.345
SQ The service here was reliable	.656	1.000	.502	.671	.419	.573	.573	.473	.520	.405
SQ The staff were trustworthy	.445	.502	1.000	.490	.306	.578	.460	.484	.494	.441
SQ The staff were friendly	.515	.671	.490	1.000	.382	.527	.510	.455	.496	.490
SQ The service here suited my needs	.339	.419	.306	.382	1.000	.454	.401	.333	.528	.400
SQ The staff were polite	.582	.573	.578	.527	.454	1.000	.696	.545	.567	.443
SQ The service was efficient	.674	.573	.460	.510	.401	.696	1.000	.580	.567	.466
SQ The staff were helpful	.465	.473	.484	.455	.333	.545	.580	1.000	.419	.375
SQ The service station provided good service	.451	.520	.494	.496	.528	.567	.567	.419	1.000	.667
SQ The service station provided quality service	.345	.405	.441	.490	.400	.443	.466	.375	.667	1.000

Next, check the value of the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) this should be either .6 or above. For our example KMO is .904 which is well within acceptable limits (see Table 4.3 below). The Bartlett's Test of Sphericity should be significant (less than .05) and in this example we have met this criterion as the test is significant (p=.000).

Table 4.3: KMO and Bartlett's Test of Sphericity

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.904
Bartlett's Test of Sphericity	Approx. Chi-Square	1788.071
	df	45
	Sig.	.000

Stage 2 – Deciding on the Number of Factors to Extract

The next decision relates to the number of factors to extract. The number of dimensions selected can be based on a range of criteria and it is widely recommended a variety of approaches are used when making this decision (Fabrigar et al., 1999). According to Tabachnick and Fidell (2007) this stage should take an exploratory approach by experimenting with the different numbers of factors until a satisfactory solution is found. However, in order for you to do this, you will need to familiarise yourself with the different criteria that can be used to determine the number of factors.

The first and most popular method for deciding on the retention of factors is Kaiser's eigenvalue greater than 1 criterion (Fabrigar et al., 1999). This rule specifies all factors greater than one are retained for interpretation. This method offers the advantage of being easy to understand and is also the default method on most programs. Some argue this method oversimplifies the situation and also has a tendency to overestimate the number of factors to retain (Zwick & Velicer, 1986). In fact, this method may lead to arbitrary decisions, for example it does not make sense

Dr. Daire Hooper, Dublin Institute of Technology, College of Business, Aungier Street, Dublin 2
e. daire.hooper@dit.ie t. +3531 402 3212

to retain a factor with an eigenvalue of 1.01 and then to regard a factor with an eigenvalue of .99 as irrelevant (Ledesma and Pedro, 2007). A technique which overcomes some of the deficiencies inherent in Kaiser's approach is Cattell's Scree Test (Cattell and Vogelmann, 1977). The Scree Test graphically presents the eigenvalues in descending order linked with a line. This graph is then scrutinised to determine where there is a noticeable change in its shape and this is known as 'the elbow' or point of inflexion. Once you have identified the point at which the last significant break takes place, only factors *above and excluding* this point should be retained. A priori theory can also drive the process, so if a break was found further along the Scree plot and made theoretical sense, then factor analysis could be re-run specifying the appropriate number of factors.

An alternative criterion is to set a predetermined level of cumulative variance and to continue the factoring process until this minimal value is reached. While no absolute threshold has been adopted, for the social sciences a minimum of 60% cumulative variance is quite commonly accepted (Hair et al, 2006). A final method is Horn's (1965) parallel analysis. Unfortunately this method is not built into the SPSS user-interface, however interested readers can use O'Connor's (2000) syntax if they wish to apply it to their data. Finally, when deciding upon the number of factors, it is strongly advised against underfactoring (choosing too few factors). This is considered a much more serious error than specifying too many (Cattell, 1978) as it can lead to distortions whereby two common factors are combined into a single common factor thus obfuscating the true factor structure.

Dr. Daire Hooper, Dublin Institute of Technology, College of Business, Aungier Street, Dublin 2
e. daire.hooper@dit.ie t. +3531 402 3212

Reverting to our example, if we are to apply Kaiser's eigenvalue greater than 1 criterion we would extract only one factor from the dataset. This is determined by examining the Total Variance Explained table (shown below in Table 4.4) wherein the total eigenvalues for the first dimension is 5.469 which accounts for 54.69% of the variance extracted. If we look to the line below this, we see the second factor has not met the eigenvalue greater than 1 criterion as it has an eigenvalue of .932. As you will recall, Kaiser's eigenvalue greater than 1 criterion has been criticised for its relatively arbitrary selection of factors and here we have a situation where the second factor possesses an eigenvalue of .932 which is reasonably close to the eigenvalue of 1 cut-off point. Given the closeness of these eigenvalues to 1 we may decide to re-run the analysis specifying a two-dimensional solution. However, for now we will proceed by applying each of the other factor extraction criteria to our results as it is recommended to use a combination of criteria to arrive at a final decision.

Table 4.4: Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.469	54.695	54.695	4.988	49.876	49.876
2	.932	9.318	64.013			
3	.693	6.928	70.941			
4	.657	6.565	77.507			
5	.564	5.638	83.145			
6	.507	5.065	88.210			
7	.369	3.686	91.896			
8	.311	3.110	95.007			
9	.263	2.625	97.632			
10	.237	2.368	100.000			

Extraction Method: Principal Axis Factoring.

We will now to examine the Scree plot (Figure 4.6) to find the point of inflexion (elbow). In our example the most obvious break (point of inflexion) is at Factor 2, suggesting a one-dimensional solution is appropriate. However, a second (albeit much smaller) drop in eigenvalues seems to occur between Factor 2 and 3 which may indicate a two-factor solution is appropriate.
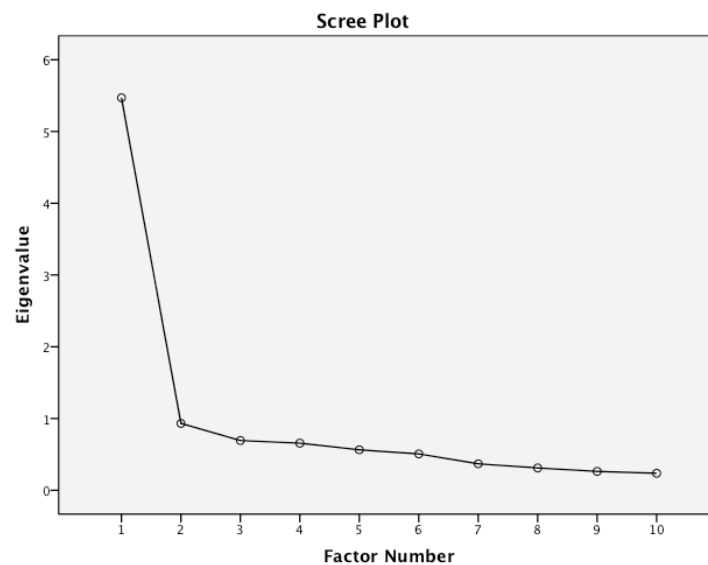

Figure 4.6: Scree plot for exploratory factor analysis

Furthermore, if we apply the cumulative variance criterion, our one factor solution captures 54.695% of the variance which unfortunately does not meet the 60% threshold. This result, combined with our eigenvalue analysis and scree plot inspection would lead us to consider a two factor solution. Coupled with these results we must bear in mind our *a priori* theoretical framework which proposed a two factor solution. Therefore, we will re-run the analysis, this time specifying a two-factor solution. To do this, proceed through all steps described above (i.e.

Dr. Daire Hooper, Dublin Institute of Technology, College of Business, Aungier Street, Dublin 2

e. daire.hooper@dit.ie t. +3531 402 3212

Analyse/Dimension Reduction/Factor etc.), however, when you click on Extraction, rather than leave the default as Based on Eigenvalues, click on **Fixed number of factors** and enter the value 2 in the box beside **Factors to extract**. Select **Continue** and **OK** to re-run factor analysis.
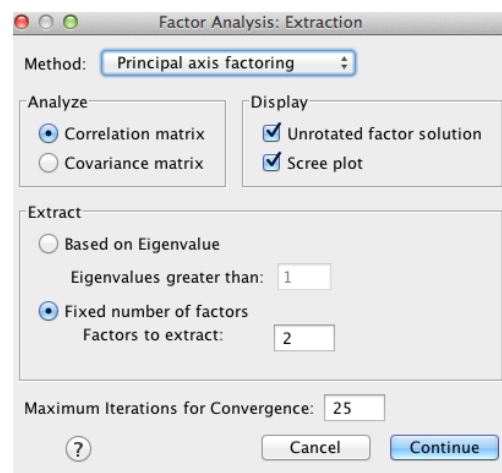

Figure 4.7: Dialog box for factor extraction

New output will be generated and you will notice the correlation matrix, KMO and Bartlett's Test of Sphericity are all the same as our original specification. The only difference is that SPSS has produced a two dimensional solution, rather than a one-dimensional solution. This can be seen by examining the Total Variance Explained table and as shown below.

Table 4.5: Total Variance Explained (Re-specified Solution)

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings[a]
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	5.469	54.695	54.695	5.048	50.476	50.476	4.745
2	.932	9.318	64.013	.541	5.415	55.891	3.986
3	.693	6.928	70.941				
4	.657	6.565	77.507				
5	.564	5.638	83.145				
6	.507	5.065	88.210				
7	.369	3.686	91.896				
8	.311	3.110	95.007				
9	.263	2.625	97.632				
10	.237	2.368	100.000				

Extraction Method: Principal Axis Factoring.

You will note in the section *Extraction Sums of Squared Loadings* that are there two lines of data rather than just one and this reflects the fact that we have constrained the solution to two dimensions. We have now accounted for 64% of variance, or 55% of shared variance in the data. This is a preferable situation to the one-factor solution as when too few factors are included in a model, substantial error is likely (Fabrigar et al 1999).

Stage 3 – Factor Rotation and Interpretation

The next stage is to interpret the factors. Principal axis factoring produces slightly different tables to other forms of factor analysis, however the table you are most interested in is the Pattern Matrix which displays the rotated factor loadings and is used to interpret the dimensions. However, before beginning interpretation, the first thing you need to check is for cross-loadings. A cross-loading is an item with coefficients greater than .4 on more than one dimension. To help with this we requested all loadings less than .4 be suppressed in the output to aid interpretation.

As we can see, our example is free from cross-loadings as all items load on only one dimension. The second thing you need to check is whether there are items that do not load on any of the factors, i.e. have loadings *less* than .4 on all dimensions. Again we can see all items load on either the first or the second dimension providing us with a nice clean solution to interpret. If we found items cross-loading or not loading at all, this would suggest they are poor/unreliable items and may need to be deleted from the analysis. If this were to happen, you would need to re-run your analysis without the offending item.

Having reached a suitable solution, the next stage is to interpret the factors themselves. This has been referred to by some as a 'black art' as there are no hard or fast rules in naming each dimension. However there are a number of guidelines that can aid in the process. Firstly, we can see there are two factors and variables load highly on only one factor. You will also note they are arranged in descending order to help us identify items with substantive loadings.   These variables with higher loadings are used to identify the nature of the underlying latent variable represented by each factor.

Dr. Daire Hooper, Dublin Institute of Technology, College of Business, Aungier Street, Dublin 2

e. daire.hooper@dit.ie t. +3531 402 3212

Table 4.5: Pattern Matrix

	Factor	
	1	2
The service in store was delivered promptly	.902	
The service here was reliable	.760	
The service was efficient	.746	
The staff were polite	.677	
The staff were helpful	.601	
The staff were friendly	.548	
The staff were trustworthy	.464	
The service station provided good service		.870
The service station provided quality service		.791
The service here suited my needs		.443

Extraction Method: Principal Axis Factoring.
Rotation Method: Promax with Kaiser Normalization.
a. Rotation converged in 3 iterations.

In our example we can see the variables loading on the first factor all relate to the service process, or the human element of the service delivery. The second dimension contains three items and appear to be evaluative items whereby the respondents are providing the service with an overall rating. These two dimensions are in keeping with our proposed theory which stated consumers perceive services along two discrete, yet related dimensions. The first of these is the functional dimension and corresponds to the way in which the service is delivered. By and large this is dependent on the service delivery process and it is the frontline employees that play a key role here. All items on dimension one relate to the role of the employee and are in keeping with our understanding of functional service quality and as such as will be named 'Functional Service Quality'. The items on the second dimension can be regarded as outcome-type items as they refer to 'what' kind of service the customer received and for this reason we will name the dimension 'Technical Service Quality'.

Reporting Factor Analysis Results

When reporting factor analysis there are a number of key pieces of information you need to include so a reader can assess the decisions you made. It is essential you report the extraction technique used, rotation technique (used Promax, Varimax etc.), total variance explained, intial eigenvalues and rotated eigenvalues. You will also need to include a table of loadings showing all values (not just those in excess of .4) in the Pattern Matrix. As an oblique rotation method was used you should also report the Structure Matrix.

For our example the results would be described along the following lines:

'Ten service quality items (Grace and O'Cass, 2004) were subjected to principal axis factoring to assess the dimensionality of the data. The Kaier-Meyer-Olkin was .904 which is well above the recommended threshold of .6 (Kaiser, 1974) and the Bartlett's Test of Sphericity reached statistical significance indicating the correlations were sufficiently large for exploratory factor analysis.

Two factors were extracted explaining 64.01% of the variance. This was decided based on eigenvalues, cumulative variance and inspection of the scree plot. Factors were obliquely rotated using Promax rotation and interpretation of the two factors was in keeping with Grönroos's (1984) two dimensional theory of service quality. Items that load on the first

dimension suggests it represents Functional Service Quality and the second dimension suggests it represents Technical Service Quality'.

Table 4.6: Pattern Matrix for Coefficients

	Factor	
	1	2
	Functional Service Quality	Technical Service Quality
The service in store was delivered promptly	**.902**	-.165
The service here was reliable	**.760**	.037
The service was efficient	**.746**	.081
The staff were polite	**.677**	.159
The staff were helpful	**.601**	.076
The staff were friendly	**.548**	.210
The staff were trustworthy	**.464**	.238
The service station provided good service	.011	**.870**
The service station provided quality service	-.045	**.791**
The service here suited my needs	.170	**.443**
% of variance explained	**54.69%**	**9.31%**

Reliability Analysis

If you are to use scales in your research it is essential that they are reliable. Reliability refers to how free the scale is from random error and is frequently measured using a statistic known as Cronbach's alpha (α). Cronbach's alpha is a measure of internal consistency which means the degree to which items in your scale measure the same underlying attribute or construct. Cronbach's alpha ranges from 0 to 1 with higher values indicating high levels of reliability. Nunnally (1978)

recommends a minimum of .7, however alpha values increase with scale length so checking for unidimensionality via exploratory factor analysis is key here.

For our example we are going to test the reliability of all items on the Functional Service Quality dimension. To do this click on **Analyze/Scale/Reliability Analysis**. Move all seven Functional Service Quality variables to the Items field and click **Statistics** (shown in Figure 4.8 below). In the **Descriptives for** section select **Item** and **Scale if item deleted** click **Continue** and then **OK**.
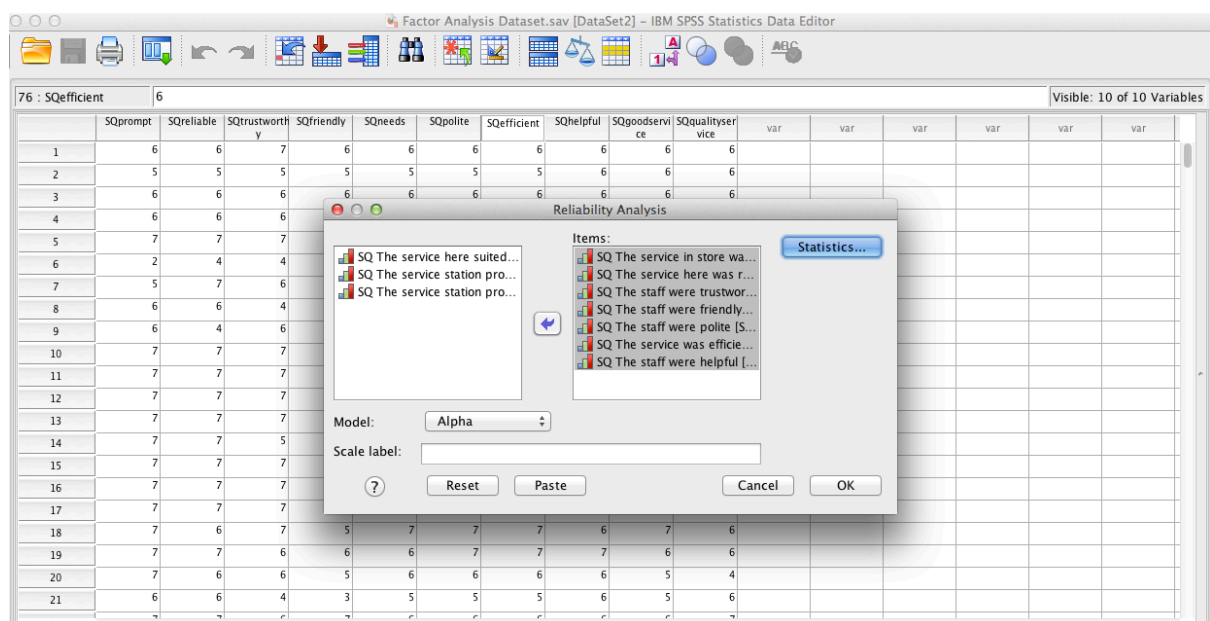

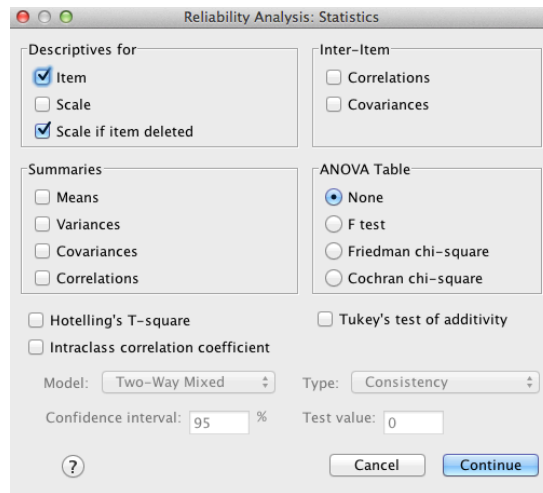
Figure 4.8: Reliability Analysis

Figure 4.9: Calculating Cronbach's alpha

In the table marked Reliability Statistics the first column provides us with the alpha coefficient which is .833 and is well above Nunnally's .7 threshold. In the table marked **Item-Total Statistics** look to the **Cronbach's Alpha if Item Deleted** column to determine if the alpha value would change substantially if we delete particular items. If there are values for some items higher than your Cronbach's alpha you might want to re-run Cronbach's alpha excluding this item. For our example there appear to be no problems here so we can proceed to use this scale in further analysis.

Table 4.7: Reliability Statistics

Reliability Statistics

Cronbach's Alpha	N of Items
.883	7

Table 4.8:Item Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
SQ The service in store was delivered promptly	37.60	16.643	.697	.863
SQ The service here was reliable	37.67	16.700	.738	.859
SQ The staff were trustworthy	37.78	15.972	.615	.876
SQ The staff were friendly	38.01	15.385	.665	.870
SQ The staff were polite	37.46	17.687	.746	.864
SQ The service was efficient	37.58	16.718	.733	.860
SQ The staff were helpful	37.66	16.297	.624	.873

In order to use these items in further analyses we must calculate the total scale scores for each of the dimensions. Summating scales is common practice in research and is done to allow us to perform statistical tests that require continuous variables (correlation, multiple regression, ANOVA, MANOVA all use continuous variables). Before calculating a total score, check that no items on your scale are negatively worded. If items are negatively worded they will need to be reverse coded in SPSS (i.e. if you have a scale ranging from 1- 7, reverse coding means replacing all 1 with 7, 2 with 6, all the way to 7 replaced with 1.). In our example all items on the Functional Service Quality scale are positively worded so we can proceed to add all items.

In SPSS select **Transform/Compute Variable**. In **Target variable** type in a name for the new summated item you are to create. For our example we will enter

FunctionalSQ. From the list on the left-hand side select the first item from the Functional Service Quality scale (SQprompt) and move it to the **Numeric Expression** box and click on the + on the calculator. Proceed in this manner until all Functional Service Quality items are in the box.
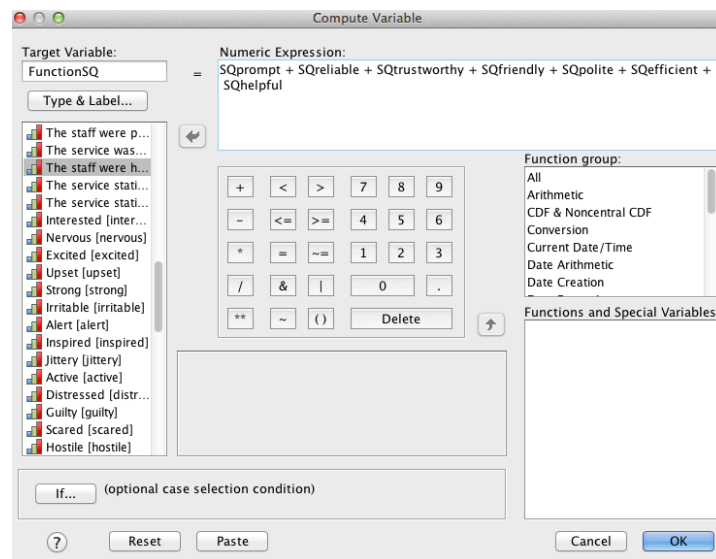

Figure 4.9: Computing variables

Once all items have been entered hit **OK** and SPSS will generate a new item which will be listed after all other variables in the dataset. This new item can now be used in other analyses that require continuous variables.

SUMMARY

This chapter has introduced the reader to exploratory factor analysis and has demonstrated how it can be used to assess the dimensionality of a dataset. In

particular, a two dimensional factor structure of service quality was found. The following summarises the steps undertaken when using factor analysis.

- Ensure your sample size is sufficiently large (minimum of 150 or 10 cases per item). Items should be interval in nature.

- Check your correlation matrix to ensure there are a reasonable number of correlations greater than .3.

- Check KMO and Bartlett's Test of Sphericity, these should be over .6 and under .05 respectively.

- Choose an extraction method. For this example Principal Axis Factoring was chosen. If you were purely interested in data reduction, rather than theory building, Principal Components Analysis would be more suitable.

- Choose a rotation method. Here an oblique method (Promax) was chosen as factors were expected to be correlated with one another. Where factors are not expected to correlate, orthogonal methods such as Varimax can be used.

- Decide on the number of factors to extract. The default in SPSS is Kaiser's eigenvalue greater than 1 criterion. It is recommended that a number of factor extraction methods are used. The example given here relied on a priori theory, the scree plot and the percentage of variance extracted.

- Using the high loading items, interpret dimensions in the Pattern Matrix.

- To use dimensions for further analysis, sum items to form a composite item.

References

Allen, N. J. & Meyer, J. P. (1990). The Measurement and Antecedents of Affective, Continuance and Normative Commitment to the Organization. *Journal of Occupational Psychology*, *63*(1), 1-18.

Babakus, E. & Boller, G. W. (1992). An Empirical Assessment of the SERVQUAL Scale. *Journal of Business Research*, *24*(2) 253-268.

Bitner, M. J. & Hubbert, A. R. (1994). Encounter Satisfaction Versus Overall Satisfaction Versus Quality. In R. T. Rust & R. L. Oliver (Eds), *Service Quality: New Directions in Theory and Practice* (pp 72-94). Thousand Oaks, CA: Sage.

Brogowicz, A. A., Delene, L. M., & Lyth, D. M. (1990). A Synthesised Service Quality Model with Managerial Implications. *International Journal of Service Industry Management 1*(1): 27-45.

Buttle, F. (1996). SERVQUAL: review, critique, research agenda. *European Journal of Marketing*, *30*(1), 8-32.

Caceres, R. C. & Paparoidamis, N. G. (2005). Service Quality, Relationship Satisfaction, Trust, Commitment and Business-to-Business Loyalty. *European Journal of Marketing*, *41*(7/8), 836-867.

Dr. Daire Hooper, Dublin Institute of Technology, College of Business, Aungier Street, Dublin 2
e. daire.hooper@dit.ie t. +3531 402 3212

Cattell, R. B. & S. Vogelmann, S. (1977). A Comprehensive Trial of the Scree and KG Criteria for Determining the Number of Factors. *Multivariate Behavioral Research*, *12*(3), 289.

Costello, A. B. & Osborne, J. W. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment, Research & Evaluation*, *10*(7), 1-9.

DeVellis, R. F. (2003). *Scale Development: Theory and Applications* (2nd ed.). Thousand Oaks, CA: Sage.

Fabrigar, L. R., MacCallum, R. C., Wegener, D. T., & Strahan R. (1999). Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *Psychological Methods 4*(3), 272-299.

Floyd, F. J. & Widaman K. F. (1995). Factor Analysis in the Development and Refinement of Clinical Assessment Instruments. *Psychological Assessment*, *7*(3), 286-299.

Gorsuch, R. L. (1990). Common Factor Analysis versus Component Analysis: Some Well and Little Known Facts. *Multivariate Behavioral Research*, *25*(1), 33.

Grace, D. & O'Cass, A. (2004). Examining Service Experiences and Post-

Consumption Evaluations. *Journal of Services Marketing, 18(*6), 450-461.

Grönroos, C. (1984). A Service Quality Model and its Marketing Implications. *European Journal of Marketing, 18*(4), 36-44.

Hair, J. S., Black, W. C., Babin, B. J., Anderson, R. E. & Tatham, R. L. (2006). *Multivariate Data Analysis*. New Jersey: Prentice-Hall.

Horn, J. L. (1965). A Rationale and Test for the Number of Factors in Factor Analysis. *Psychometrika, 30*(2), 179-183.

Lassar, W. M., Manolis, C., & Winsor, R.D. (2000). Service Quality Perspectives and Satisfaction in Private Banking. *Journal of Services Marketing, 14*(3), 244-271.

Ledesma, R. D. & Pedro, V. M. (2007). Determining the Number of Factors to Retain in EFA: an easy-to-use computer program for carrying out Parallel Analysis. *Practical Assessment, Research & Evaluation, 12*(2), 1-11.

Nunnally, J. C. (1978). *Psychometric Theory*. New York: McGraw-Hill.

O'Connor, B. P. (2000). SPSS and SAS Programs for Determining the Number of Components Using Parallel Analysis and Velicer's MAP Test. *Behavior Research Methods, Instruments and Computers*, *32*(3), 396-402.

Dr. Daire Hooper, Dublin Institute of Technology, College of Business, Aungier Street, Dublin 2
e. daire.hooper@dit.ie t. +3531 402 3212

Parasuraman, A., Zeithaml, V. A., & Berry, L. (1985). A Conceptual Model of Service Quality and Its Implications For Future Research. *Journal of Marketing*, *49*(4), 441-450.

Parasuraman, A., Zeithaml, V. A., & Berry, L. (1988). SERVQUAL: A Multiple Item Scale for Measuring Customer Perceptions of Service Quality. *Journal of Retailing*, *64*(1), 12-40.

Parasuraman, A., Zeithaml, V. A., & Berry, L. (1994). Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research. *Journal of Marketing*, *58*(1), 111-124.

Park, H. D., R. Dailey, et al. (2002). "The Use of Exploratory Factor Analysis and Principal Components Analysis in Communication Research." <u>Human Communication Research</u> 28(4): 562-577.

Snook, S. C. & Gorsuch, R. L. (1989). Component Analysis Versus Common Factor Analysis: A Monte Carlo Study. *Psychological Bulletin*, *106*(1), 148-154.

Spector, P. E. (1992). *Summated Rating Scale Construction*. Newbury Park, CA: Sage Publications.

Tripp, C. & Drea, J. T. (2002). Selecting and Promoting Service Encounter Elements in Passenger Rail Transportation. *Journal of Services Marketing*, *16*(5), 432-442.

Dr. Daire Hooper, Dublin Institute of Technology, College of Business, Aungier Street, Dublin 2
e. daire.hooper@dit.ie t. +3531 402 3212

Velicer, W. F. & Jackson, D. N. (1990). Component Analysis versus Common Factor Analysis: Some Issues in Selecting an Appropriate Procedure. *Multivariate Behavioral Research*, *25*(1), 97-114.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using Multivariate Statistics*. New York: Allyn and Bacon.

Worthington, R. L. & Whittaker, T. A. (2006). Scale Development Research: A Content Analysis and Recommendations of Best Practices. *The Counselling Psychologist*, *34*(6), 806-838.

Zwick, W. R. & Velicer, W. F. (1986). Comparison of Five Rules for Determining the Number of Components to Retain. *Psychological Bulletin*, *99*(3), 432-442.