

## Exploratory Factor Analysis (EFA) in Quantitative Researches and Practical Considerations

Nicel Araştırmalarda Açıklayıcı Faktör Analizi (AFA) ve Pratik Hususlar

Lütfi SÜRÜCÜ<sup>1</sup>, Ahmet MASLAKÇI<sup>2</sup>, İbrahim YIKILMAZ<sup>3</sup>

### ABSTRACT

Explanatory factor analysis (EFA) is a multivariate statistical method frequently used in quantitative research and has begun to be used in many fields such as social sciences, health sciences and economics. With EFA, researchers focus on fewer items that explain the structure, instead of considering too many items that may be unimportant and carry out their studies by placing these items into meaningful categories (factors). However, for over sixty years, many researchers have made different recommendations about when and how to use EFA. Differences in these recommendations confuse the use of EFA. The main topics of discussion are sample size, number of items, item extraction methods, factor retention criteria, rotation methods and general applicability of the applied procedures. The abundance of these discussions and opinions in the literature makes it difficult for researchers to decide which procedures to follow in EFA. For this reason, it would be beneficial for researchers to gather different information about the general procedures (sample number, rotation methods, etc.) in the use of EFA. This paper aims to provide readers with an overview of what procedures to follow when implementing EFA and share practical information about the latest developments in methodological decisions in the EFA process. It is considered that the study will be an important guide for the researchers in the development of clear decision paths in the use of EFA, with the aspect of presenting the most up-to-date information collectively.

**Anahtar Kelimeler:** Explanatory factor analysis (EFA), Rotation methods, Sample size, SPSS, Quantitative Research

### ÖZ

Açıklayıcı faktör analizi (EFA), nicel araştırmalarda sıklıkla kullanılan çok değişkenli istatistiksel bir yöntem olup, sosyal bilimler, sağlık bilimleri ve ekonomi gibi birçok alanda kullanılmaya başlamıştır. EFA yardımıyla önemsiz olabilecek çok fazla öğeyi dikkate almak yerine, yapıyı açıklayan daha az sayıdaki öğeye odaklanır ve bu öğeleri anlamlı kategorilere (faktör) yerleştirerek çalışmalarını yürütürler. Bununla birlikte altmış yılı aşkın bir süredir birçok araştırmacı EFA'nın ne zaman ve nasıl kullanılacağına dair birbirinden farklı tavsiyelerde bulunmaktadır. Yapılan tartışma konularının başında örneklem büyüklüğü, madde sayısı, madde çıkarma yöntemleri, faktör tutma kriterleri, döndürme yöntemleri ve uygulanan prosedürlerin genel uygulanabilirliği konuları yer almaktadır. Literatürde yaşanan bu tartışmalar ve görüşlerin bolluğu, araştırmacıların EFA'da hangi prosedürleri izleyeceğine karar vermesini zorlaştırmaktadır. Bu nedenle EFA'nın kullanımındaki genel prosedürlere (örneklem sayısı, döndürme yöntemleri vb.) ait farklı bilgilerin bir araya getirilmesi araştırmacılar için faydalı olacaktır. Bu çalışmanın amacı; EFA uygulanırken hangi prosedürlerinin izleneceği konusunda okuyuculara genel bir bakış açısı sunmak ve EFA sürecindeki metodolojik kararlarla ilgili en son gelişmeler hakkında okuyuculara pratik bilgiler paylaşmaktır. Çalışmanın en güncel bilgileri toplu bir şekilde sunması yönü ile EFA kullanımında net karar yollarının geliştirilmesinde okuyuculara önemli bir kılavuz olacağı değerlendirilmektedir.

**Keywords:** Açıklayıcı faktör analizi (AFA), Rotasyon (döndürme) yöntemleri, Örneklem büyüklüğü, SPSS Nicel Araştırma

<sup>1</sup> Assoc.Prof. Lütfi SÜRÜCÜ, Faculty of Business and Economics, World Peace University Nicosia, lutfi.surucu@wpu.edu.tr, ORCID: 0000-0002-6286-4184

<sup>2</sup> Assoc.Prof. Ahmet MASLAKÇI, Department of Business Administration, Bahçeşehir Cyprus University, ahmet.maslakci@baucyprus.edu.tr, ORCID: 0000-0001-6820-4673

<sup>3</sup> Asst.Prof.Dr. İbrahim YIKILMAZ, Business Management Department, Kocaeli University, ibrahimyklmz@gmail.com, ORCID: 0000-0002-1051-0886

## INTRODUCTION

Explanatory factor analysis (EFA) is a multivariate statistical method frequently used in quantitative research. Its origins date back to the development of the Two Factor Theory by Charles Spearman in the early 1900s. With his work on Spearman's Theory of Personality, he provided the conceptual and theoretical foundations for explanatory factor analysis. In the early years, EFA was based on simple mathematical principles for which the answers were already known and used for simulated data.<sup>1</sup> With the development of computers and more comprehensive mathematical operations, the usage area of factor analysis has now expanded. It has started to be used in many fields such as social sciences, medicine, and economics.

The primary purpose of EFA is to summarize data to easily interpret and understand relationships and patterns of the observed variables in the measurement tool. In other words, the observed variables are regrouped in a limited cluster with fewer latent variables that cannot be observed based on shared variance.<sup>2</sup> By regrouping the observed variables in a limited set, researchers can focus on fewer items that explain the structure, instead of considering too many items that may be unimportant in their studies, and placing these items into meaningful categories (factors) will allow them to easily conduct their studies. For example, let's assume that researchers apply the DASS-21 scale to a group of patients and try to determine their depression, anxiety, and stress levels. Not every item in the survey is likely to measure something unique. Therefore, one group of items will determine depression, another group will assess anxiety, and another group will determine the stress levels of patients. However, although every effort was made to avoid the repetition of questions in the questionnaire design, some items may be closely related due to cultural differences to detect slightly different aspects of the same phenomenon. The most logical way to analyze and interpret the data obtained from the questionnaire is to determine the basic structures that assess depression, anxiety, and

stress. Thus, some excess items that identify different aspects of the same phenomenon will be removed from the study and analyzes could be made easily on the main structure.

EFA consists of a group of statistical analyses that incorporate many mathematical operations and methodologies rather than a single statistical analysis. All factors such as the design of the research, the characteristics of the sample and the data in determining the statistical analyses affect the researcher's decision about which procedures to apply in EFA. However, statistical methods in EFA are flexible and it is up to the personal preference of the researcher which methods to apply. The accuracy of the results is also directly proportional to the quality of the decisions made by the researcher.<sup>3</sup> Therefore, the researcher needs to make careful and informed decisions regarding the use of appropriate procedures in EFA. Several options and methods in the literature are available to help researchers make the right decisions, improve the accuracy of factor analysis and increase the quality of the resulting solution.<sup>4,5</sup> However, for over sixty years, many researchers have made different recommendations about when and how EFA should be used. Differences in these recommendations have led to confusion regarding the use of EFA. In fact, there is not another statistical method in the literature that is designed as powerfully as EFA and causes many debates on its correct use.<sup>6</sup> The primary issues discussed are the sample size, number of items, item extraction methods, factor retention criteria, rotation methods, and general applicability of the applied procedures.

The abundance of these discussions and opinions in the literature makes it difficult for researchers to decide which procedures to follow in EFA. For this reason, it will be useful for readers to review the information on EFA and to gather different information about the general procedures (number of samples, rotation methods, etc.) in the use of EFA. Therefore, the purpose of this paper is to provide readers with an overview of which

procedures to follow when implementing EFA and to give practical information on the latest developments regarding methodological decisions in the EFA process. Thus, our study will be able to guide readers in the use of EFA as a reference in the development of clear decision paths.

The paper is structured as follows. Firstly, readers are informed about the purposes and requirements of factor analysis and testing the suitability of the data for factor analysis. Afterward, the literature about making the data suitable for factor analysis and factor

analysis components are presented. Finally, the study ends with the presentation of the conclusion.

Throughout the article, examples from SPSS outputs are included to provide better information to the readers. Since the same data are not used in every example, it is important to emphasize that different factor solutions are presented. It should also be noted that there may be some differences in the format of reporting in more recent versions of SPSS.

## THEORETICAL FRAMEWORK

### PURPOSE OF FACTOR ANALYSIS

Factor analysis can be used to determine which theoretical constructs lie under a given data set and the extent to which these constructs represent the original variables. In addition, factor analysis can be used to investigate the correlations between observed variables and to model these relationships with one or more latent variables.

Researchers want to use simple data to understand any unobservable variable with the observed variables. Factor analysis applies a series of mathematical operations to datasets consisting of these observed variables, combining the common variables in the dataset into descriptive categories and reducing the number of factors. Thus, instead of considering too many variables that may be insignificant, researchers use factor analysis to focus on some basic factors that allows easy interpretation.<sup>7</sup> In line with the existing literature, it could be said that the primary purpose of EFA is to take a relatively large set of variables and reduce them to a smaller and more manageable number while preserving the original variance as much as possible. Finally, it should be noted that factor analysis has many additional uses, such as mapping and scaling, which are beyond the scope of this paper.<sup>8</sup>

### REQUIREMENTS OF FACTOR ANALYSIS

Before conducting EFA analysis, it should be verified whether the data are suitable. To

determine the suitability of the data for factor analysis, researchers can evaluate the suitability of their data for EFA by examining the number of variables, correlation value, missing values, sample size, multicollinearity, and singularity, as well as some other general issues.

### General Considerations

To apply factor analysis to a dataset, it is essential to ensure that univariate and multivariate normality are ensured and that there are no outliers.<sup>1,9</sup> A lack of normality in the data affects the Pearson product-moment correlation coefficients ( $r$ ) between the variables used in the calculation of the EFA results (by reducing the size of the correlation coefficient). In addition, the presence of outliers in the data could also cause artificial factors to be produced. In such cases, the EFA findings are misleading. Therefore, the distribution of the data should be checked as it may affect Pearson correlations. If the multivariate normality assumption is violated when the data distributions are checked, it would be appropriate to choose one of three alternative methods such as transforming the data, correcting the fit indices, or principal axis factoring.<sup>5</sup>

### Number of Variables

One of the EFA requirements is the number of variables in the factor. It is generally accepted that at least three variables are needed to create a factor from the variables that can be observed in EFA. However,

researchers have stated that at least four variables with acceptable reliability ( $>.70$ ) would be sufficient for each expected factor or that a factor with two variables could be formed depending on the research design.<sup>5-7</sup> Examining a total of 1,329 studies published between 2007 and 2017 supporting this view, Goretzko et al. (2021) found that 10.5% of researchers using EFA in their study created a factor with two or fewer variables.<sup>6</sup> As can be seen, although factors with two or fewer variables can be created, they should be interpreted carefully. The formation of two or fewer variable factors should only be considered reliable when the correlation between variables is highly uncorrelated with other variables and the correlation value is higher than .7 ( $r > .70$ ).<sup>7</sup>

Other researchers in the literature have stated that least five or more variables are needed to create a factor.<sup>10</sup> Artificial multiplication of variables to meet this requirement (minimum number of variables) for factor analysis could lead to a violation of local independence. Therefore, artificially increasing the number of variables is not recommended.

### Correlation Value

The basic statistical approach used in factor analysis is the correlation coefficient, which determines the relationship between two variables. While calculating the factor analysis correlation, it is assumed that there is a linear relationship between the factors and variables. Therefore, the correlation value is an important issue to be considered in factor analysis. As mentioned above, a high correlation allows a factor with two or fewer variables to be created, whereas a low correlation value indicates a weak relationship between the variables and prevents factor formation. Therefore, the literature suggests that the correlation coefficient should be at least .32 and that variables with lower correlations should be excluded from the factor analysis.<sup>11</sup> According to statistical scientists, a factor loading of .32 explains about 10% of the overlapping variance and this value constitutes the lower threshold value for factor analysis. Homogeneous

samples cause a decrease in variance and factor loads. For this reason, researchers who want a high factor load are recommended to collect data from a heterogeneous instead of a homogeneous sample group.<sup>12</sup>

### Missing Values

Missing data can affect EFA results. For this reason, it is necessary to check whether the missing data in the data set occurs in a non-random order.<sup>7</sup> It is recommended that data with missing values should be excluded from the factor analysis to avoid incorrect estimations.<sup>3</sup> However, deleting cases with missing data in a list or one-by-one is inefficient and generally not recommended.<sup>13</sup> In such situations, some methods found in many software programs can be used to assign missing data (e.g., regression, mean, multiple, and maximum likelihood). Studies and simulations show that the mean method is acceptable in cases where  $<10\%$  of the data is missing, whereas the regression method is acceptable in cases where  $<15\%$  of the data is missing.<sup>14</sup> It would be a suitable approach to use multiple and maximum likelihood methods in assigning missing data to a greater extent.<sup>13</sup>

### Multicollinearity and Singularity

The literature recommends that it should be checked whether multicollinearity and singularity are present in the dataset. A Squared Multiple Correlation value close to 0 indicates singularity, while a value close to 1 indicates a multicollinearity problem. Therefore, problematic variables should be excluded from the factor analysis.<sup>7</sup>

### Sample Size

The determination of the sample size is at the forefront of the recent discussions on EFA. There are different recommendations and opinions regarding this issue in the literature. The recommendations made for determining the number of samples can generally be categorized in three ways. These are the minimum number of cases, sample to variable ratio ( $N/p$ ), and factor loads required to obtain an adequate sample.

**Number of Sub-threshold Samples:** Many researchers hold the opinion that a sample size of 300 would be sufficient for factor analysis.<sup>3,7</sup> However, some researchers think that smaller sample groups are also sufficient. For example, Sürücü, Sesen, and Maslakçı (2023) considered 200 samples and Beavers (2013) 150 samples to be sufficient.<sup>4,15</sup> Also, Mundfrom, Shaw, and Ke (2005) claimed that even samples smaller than 100 are sufficient for factor analysis when the population is high enough and the factors are represented by many items.<sup>16</sup> Additionally, Winter, Dodou, & Wieringa (2009), expanded on these views, arguing that a sample size of 50 or less may be sufficient in behavioral studies.<sup>17</sup>

Other researchers in the literature have also suggested higher sample numbers. For example, Comrey (1973) claimed that 500 samples are sufficient, whereas Goretzko, Pham, and Bühner (2021) believed that 400 is enough.<sup>6,18</sup> As can be seen, there is still no consensus in the literature regarding the sample size. However, there is general agreement that an insufficient sample size can damage the factor analysis process and produce unreliable and therefore invalid results.<sup>4,19</sup>

In line with the existing literature, the authors think that a sample size of greater than 200 should be used for EFA and therefore 200 should be determined as the lower threshold. For sample sizes less than 200, consideration should be given to factors such as factor loading and several variables. Interpreting factor load and the number of variables requires the researcher to have experience in statistical analysis.. Researchers who do not have sufficient experience in such analysis should prefer sample sizes higher than 200 to obtain more reliable results in their studies. In general, we describe sample sizes of 100 as poor, 150 as moderate, 200 as adequate, 250 as good, 300 as very good, and above 400 as excellent.

**Sample to Variable Ratio (N/p):** Another issue that is evaluated in determining the sample size is the number of variables in the measurement tool. Some researchers think that sampling as much as a certain percentage

(N/p) of the number of variables in the measurement tool is sufficient in determining the sample size. However, there is no consensus in the literature on this issue. For example, it is mentioned that the sufficient sample size for EFA should be 51, 20, 10 or even 6 times the number of variables in the measurement tool.<sup>7,20-22</sup>

Some researchers have expanded on these discussions and proposed different methodologies for determining the number of samples. For example, Suhr (2006) suggested a sample size of at least 100 and a participant-variable ratio (N/p) of at least 5, while Bryant and Yarnold (1995) suggested that the sample size should have at least 10 samples for each variable and the participant-variable (N/p) ratio should not be less than 5.<sup>23,24</sup> Researchers such as Fabrigar, Wegener, MacCallum and Strahan (1999) and Rouquette and Falissard (2011) do not support the idea that the ratio per variable should be used as a guide value for the sample size.<sup>5,25</sup>

Hogarty, Hines, Kromrey, Ferron, and Mumford (2005) as well as MacCallum, Widaman, Zhang, and Hong (1999), who conducted a series of studies based on these discussions in the literature, empirically documented that it is not appropriate to specify a minimum ratio for factor analysis.<sup>10-26</sup> Later studies also showed that these propositions were inconsistent and recommendations about absolute N and N/p ratio were gradually abandoned because they were misunderstood.<sup>26</sup>

**Factor Load:** The responses of the participants to the observed variables could dramatically change the sample size recommended for factor analysis. For this reason, basic rules such as determining a lower threshold or participant-variable ratio in the number of samples can sometimes be misleading.<sup>5,27</sup> Guadagnoli and Velicer (1988) argued that in the literature, the required sample size largely depends on the strength of the factors and items.<sup>28</sup> Thus, the researchers proposed a new criterion that functionalizes the relationships between sample-factor load, which is accepted in the literature.

There are different opinions in the literature on determining the number of samples according to the factor load. While some researchers state that factor analysis can be performed with much smaller samples if the factor load is .80 or above ( $n > 150$ ), some researchers think that in a similar situation, if the factor load is greater than .80, a sample size of 50 may be sufficient for factor analysis.<sup>27, 28</sup> As can be seen, the factor load of the item answered by the participants can vary significantly according to the sample size required to complete the factor analysis. This is acceptable because the factor loading of a variable is a measure of how much that variable contributes to that factor. Therefore, high factor loadings indicate that the relevant factors are better explained by the variables.<sup>7</sup> From this perspective, it may be reasonable to have a smaller sample size in data with a high factor loading. Indeed, the literature states that the sample size depends on the strength of factors and items (factor load).<sup>28</sup> According to researchers who support this view, if there are 10 to 12 items with a moderate loading (.40 or higher) in a factor, a sample size of 150 is sufficient to obtain reliable results. Again, according to these researchers, if there are four or more items with a factor loading of .60 or higher in a factor, then a lower threshold value is not needed for the size of the sample. Researchers such as Fabrigar, Wegener, MacCallum, and Strahan (1999) and MacCallum, Widaman, Zhang, and Hong (1999) stated that if there are three or four items with a high factor loading (loadings of .70 or greater) in a factor, stable solutions can be reached with sample sizes as small as 100.<sup>5, 26</sup>

A simulation study by Preacher and MacCallum (2002) on the application of EFA in behavioral genetics clearly showed that EFA can provide reliable solutions even for sample sizes as small as 10 when there are strong items (between .8 and .9) for two factors.<sup>29</sup> These studies showed that for items with high factor loadings, the number of samples required for EFA decreases, and when the items have low factor loadings, a larger sample size is needed.

Authors urge researchers to be cautious about sample size, because EFA generally functions better with larger sample sizes, and larger samples lead to more stable solutions by reducing the margin of error.<sup>4</sup> As can be seen, authors stress the fact that EFA is generally a "large sample" procedure. If the sample size is too small, it is not possible to obtain generalizable or reproducible results.<sup>30</sup> For this reason, we recommend that factor analysis not be performed with sample sizes less than 200, even if the factor load is high.

### TESTING THE SUITABILITY OF THE DATA FOR FACTOR ANALYSIS

After meeting the requirements of factor analysis, it is necessary to test whether the dataset is suitable for EFA. To achieve this, the results of the Kaiser-Meyer-Olkin Measure (KMO) and Bartlett's Sphericity Tests should be firstly checked. Bartlett's Test of Sphericity and the Kaiser-Meyer-Olkin Sampling Adequacy Test (KMO) are widely used in the literature to determine the strength of relationships and evaluate the factorability of variables. While KMO provides information on sample adequacy, Bartlett's Test of Sphericity also provides information on whether the dataset has pattern relationships. A significant KMO value of .06 and above and Bartlett's Test of Sphericity being  $p < .05$  indicate that the data are suitable for factor analysis. An example SPSS output is presented in Table 1.

**Table 1. KMO and Bartlett's Test of Sphericity (SPSS Output)**

KMO and Bartlett's Test	
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.833
Bartlett's Test of Approx. Chi-Square	3903,131
Sphericity	Df 55
	Sig. .000

### Bartlett's Test of Sphericity

Bartlett's Test of Sphericity provides evidence that the observed correlation matrix is statistically different from a single matrix and confirms the existence of linear combinations.<sup>4</sup> If this requirement is not met,

it means that specific and reliable factors cannot be produced.

The null hypothesis of Bartlett's Test of Sphericity states that the observed correlation matrix is equal to the unit matrix, which indicates that the observed matrix is not factorable.<sup>4</sup> In the EFA, the Bartlett's Sphericity Test result should be less than .05. If the result is different, it is recommended to increase the number of samples or remove the items that cause scattered correlation models from the analysis and perform factor analysis again.

### Kaiser-Meyer-Olkin Test of Sampling Adequacy

The Kaiser-Meyer-Olkin Sampling Adequacy Test is a measure of shared variance in items. Researchers give an idea about whether the KMO value sample size is sufficient for EFA. If the KMO value, which ranges between 1 and 0, is 0.6 and above, this indicates a sufficient value for factor analysis. Information on other values is presented in Table 2.

**Table 2. Interpretation guidelines for the Kaiser-Meyer-Olkin Test**

KMO Value	Degree of Common Variance
0.90 to 1.00	Marvelous
0.80 to 0.89	Meritorious
0.70 to 0.79	Middling
0.60 to 0.69	Mediocre
0.50 to 0.59	Miserable
0.00 to 0.49	Unacceptable (No factor)

### Anti-image Correlation Matrix

In addition to KMO and Bartlett's Test of Sphericity, researchers can also verify whether the dataset is suitable for rotation in EFA by checking the diagonal element of the Anti-image correlation matrix. The cutoff point for the value for diagonal elements (elements with superscript 'a') is .5. It is sufficient for these values to be .5 and above. A sample SPSS output is shared in Table 3.

Focusing on the "Anti-image Correlation" line in the notation presented in Table 3, it can be seen that the diagonal elements (items with the 'a' superscript) are .637, .682, and .592, respectively.

**Table 3. Anti-image matrices (SPSS output)**

		Loyalty	Sufficiency	Trust
Anti-image Covariance	Loyalty	.739	-.081	.299
	Sufficiency	.081	.795	.254
	Trust	.299	-.254	.655
Anti-image Correlation	Loyalty	.637 <sup>a</sup>	-.106	.430
	Sufficiency	.106	.682 <sup>a</sup>	.352
	Trust	.430	-.352	.592 <sup>a</sup>

a. Measures of Sampling Adequacy (MSA)

These values indicate that the dataset is suitable for rotation in EFA. In other words, if the values of the diagonal elements are .5 or below, it means that reliable factors cannot be produced from this data set. In such situations, it would be a good practice to increase the sample size or remove the elements that cause the scattered correlation patterns.

### MAKING THE DATA SUITABLE FOR FACTOR ANALYSIS

If it is determined that the data are not suitable for EFA, there are some methods that researchers can follow to ensure compliance. The most simple method is to increase the sample size, which will be beneficial for EFA to give healthy results. Other methods are presented below.

### Correlational Values

One of the main methods of making the data suitable for factor analysis is to remove the variables that cause scattered correlation models from the data set. For this, it is checked whether there are pattern relations between the variables by applying the correlation matrix. Variables with a low correlation coefficient ( $r < +/- .30$ ) and variables with a high correlation coefficient ( $r > +/- .90$ ) in the controls are excluded from the study. A low correlation coefficient of the variable in the correlation matrix shows that this variable does not contribute to the explanation of the related factor. In general, correlations below .30 will fail to reveal the existence of a potential factor or co-relation. Correlations above .30 will provide evidence for researchers to show that there is not sufficient commonality to confirm the contributing factors.<sup>3</sup> However, very high correlation coefficients indicate that there may be a multicollinearity problem ( $r > .9$ ). Therefore, it is recommended that variables with a correlation of  $r > +/- .90$  should be

excluded from the study. Expected values in correlations are in the range of +/- .30 to +/- .90.

### Determinant of the Matrix

In determining the existence of a multicollinearity problem, the determinant score can also be checked. If the determinant score is significantly higher than .00001, this indicates that there is no multicollinearity problem. The determinant of a matrix is the value calculated using values in a square matrix, which reveals the presence or absence of possible linear combinations in the matrix. It has a value between (-1) and (1).<sup>4</sup> Haitovsky's (1969) test can also be used to test the determinant value.<sup>31</sup> However, it is important to emphasize that the calculation of the determinant value does not always give precise results.

## COMPONENTS OF FACTOR ANALYSIS

### Extraction

In the literature, there are two basic extraction methods: component analysis and common factor analysis. Although each orientation has its own characteristics, researchers mostly prefer Principal Components from component analyses and Principal Axis Factoring and Maximum Likelihood methods from common factor analyses. Goretzko, Pham, and Bühner (2021), who reviewed a total of 1,329 articles published between 2007 and 2017, found that 51.3% of researchers used EFA based on Principal Axis Factoring and 16.4% based on Maximum Likelihood.<sup>6</sup> The authors did not provide a statistical value for performing EFA based on the Principal Components estimate.

The main difference between component and common factor analysis is their purpose. Component analysis aims to preserve the variance of the original measured variables as much as possible and to reduce the number of variables by creating linear combinations. In component analysis, no comments are made about the structures. The purpose of co-factor analysis is to understand the latent (unobserved) variables that explain the relationships between the measured variables.

Because of their different purposes, component analysis and co-factor analysis also differ in their conceptualization of the sources of variance in the measured variables. Co-factor analysis models assume that factors are imperfectly reflected by the measured variables and distinguish between variance in measures due to common factors (factors affecting more than one measure) and variance due to unique factors (factors affecting only one measurement). Component analysis models do not have such separation, and therefore, components contain a mixture of shared and unique variance.<sup>32</sup>

The difference between the objectives shows that there are some differences between component analysis and common factor analysis, both mathematically and theoretically. Mathematically, measurement error, shared variance, and unique variance are included in the component analysis, whereas in common factor analysis, only common variance is included in the analysis to extract the factor solution. Theoretically, component analysis is used as a tool to accurately report and evaluate a large number of variables using fewer components. During all these processes, the dimensions of the data are preserved. On the other hand, common factor analysis allows the investigation of basic structures that cannot be measured directly through variables that are thought to be reflective measures of the structure.<sup>4, 33</sup>

Although there are mathematical and theoretical differences between the two methods, the practical sequence of steps and processes is the same.<sup>19</sup> It is widely known that the results of EFA using both methods are similar.<sup>3, 34</sup> Both component analysis and common factor analysis can mathematically reduce variables to fewer components or factors. However, the precise interpretability and understanding of these values depend on the sequential methods used to extract linear combinations.

Although the produced results are similar, it is important to apply the method that most accurately represents the research purpose and needs. Even if it is purely intuitive, if a researcher's goal is to understand the latent



nature of a set of variables, a direction in which many researchers work, then the use of common factor analysis model is recommended. If the researcher's goal is to reduce the number of variables without interpreting the emerging variables in terms of latent structures, then the use of the component analysis model is an effective decision. Given that most researchers attribute meaning beyond observed variables, common factor analysis will often be the better choice. However, it is recommended that the researcher carefully consider which method to use and decide accordingly.

**Principal Components:** Principal component analysis, which became popular decades ago when computers were slow and expensive, is perceived to be a faster and cheaper alternative to factor analysis.<sup>35</sup> Principal component analysis theoretically assumes that the component is a combination of observed variables or that individual item scores cause the component.<sup>4</sup> Principal component analysis tries to extract the maximum variance from the dataset and reduces many variables to fewer components.<sup>3-7</sup> Therefore, principal component analysis can be considered a data reduction technique.

Researchers prefer to use principal component analysis as a first step to reduce expressions on the scale. Depending on the design of the research, this method can be used to facilitate the interpretation of the data.<sup>7</sup> However, there are debates regarding whether principal component analysis is a factor analysis technique or not, and it has been repeatedly stated that it is not an equivalent alternative to factor analysis.<sup>5, 30, 35</sup> The basic logic underlying this is that the error variance is separated from the shared unique variance of a variable in factor analysis and only the shared variance is considered in the solution. However, principal component analysis does not make any distinction between shared and unique variance. Therefore, component analysis includes shared variance, unique variance, and measurement errors. In addition, it does not partially remove any variance in the variable

while examining the relationships. As a result, since total variance is included in component analysis, some researchers argue that the estimates provided reflect inflated values.<sup>30</sup> In fact, Costello and Osborne (2005) analyzed a data set consisting of 24,599 subjects using principal component analysis and common factor analysis (maximum likelihood) methods, and while factor analysis produced an average variance of 59.8%, principal component analysis generated a variance of 69.6%. The 16.4% excess variance in these analyses performed on the same dataset is an indication that principal component analysis produces inflated item loadings in many cases.<sup>30</sup>

**The Principal Axis Factoring:** The basic axis factor method is based on the concept that all variables belong to the first group and a matrix is calculated when the factor is removed.<sup>7</sup> In the principal axis factor method, factors are successively subtracted until a sufficiently large variance is calculated in the correlation matrix. As can be seen, components are produced in the principal component method, while factors are produced in the principal axis factor method. It is recommended that this method is used when the data violates the assumption of multivariate normality, because the principal axis factor method does not require distribution assumptions and can be used even if the data are not normally distributed.<sup>5, 30</sup>

**Maximum Likelihood:** Maximum likelihood requires multivariate normality and tries to analyze the maximum sampling probability of the observed correlation matrix. This method is beneficial for estimating factor loadings for a population and is more useful for confirmatory factor analysis.<sup>7</sup> In addition, maximum likelihood estimation is an approach preferred by researchers because it offers a large number of fit indices.<sup>5</sup> Another advantage of the maximum likelihood approach is that confirmatory factor analysis includes the maximum likelihood method. Therefore, it allows cross-validation of the EFA and confirmatory factor analysis results. Therefore, we recommend that researchers

who also report confirmatory factor analysis results in their research use this method.

**End Notes and General Comments:** As can be seen, each method presented above has a significant effect on the EFA solution. Although it is noted to have a serious impact, the literature lacks recommendations on which extraction method should be used under what conditions.<sup>5, 6, 30</sup> Many researchers prefer maximum likelihood estimation if the dataset has multivariate normality.<sup>5, 6</sup> The reason why maximum likelihood is preferred is that there are fit indices that can be used for model comparison and evaluation.<sup>6</sup> The fact that maximum likelihood estimation can be applied in many statistical programs (SPSS, R, MPLUS, etc.) is another reason for its preference.

The multivariate normality of Likert-type items is questioned in the literature.<sup>6</sup> Costello and Osborne (2005) stated that principal axis factoring should be preferred if the dataset violates multivariate normality (using Likert-type scales).<sup>30</sup> On the other hand, Yong and Pearce (2013) recommended that principal component analysis be firstly performed to reduce the dimensionality of the data followed by maximum likelihood.<sup>7</sup> According to Yong and Pearce (2013), principal axis factoring should generally be used as an alternative extraction method.<sup>7</sup> Comparing maximum likelihood and principal axis factoring through simulations, DeWinter and Dodou (2012) concluded that maximum likelihood outperforms principal axis factoring when the loadings are unequal and an incomplete inference is given, whereas principal axis factoring is more effective when the factor structure is steep and over-extraction.<sup>36</sup>

Failure to clarify the distinctions between extraction methods leads to difficulties in interpreting the context and reduces the researcher's ability to make theoretically sound decisions. Therefore, as a general understanding, maximum likelihood estimation should be relied on for normally distributed data, whereas principal axis factoring estimation should be preferred for non-normal and ordinal data (especially when Likert-type items with less than five

categories are used). Extraction via maximum likelihood should be limited to situations where unsuitable solutions suffer. Depending on the particular data, more than one method could be attempted and the results for matching patterns can be examined as Widaman (2012) suggests.<sup>37</sup>

### Number of Factors to Retain

In some cases, many of the factors in the dataset do not contribute significantly to the overall solution or are very difficult to interpret. Factors that are difficult to interpret and do not contribute to the solution often cause confusion or error. Therefore, it is not correct to keep the relevant factors in the analysis. Since this decision made by the researcher will directly affect the results, it is important to keep the correct number of factors.

EFA allows for the determination of the number of factors the variables in the dataset are gathered under and which factors are important. The researcher should first evaluate the results and determine the factors that best represent the existing relationships (number of factors to retain). The first factor in the list of factors formed in the analysis results is the one that explains the most variance. In each of the next factors, the amount of variance explained decreases continuously. The amount of variance is used to determine sufficient factors to adequately represent the data.<sup>5</sup> Researchers must decide which factors to keep and which factors to exclude. It should be noted that removing too many factors leads to undesirable error variance, while removing too few factors will cause valuable common variance to be omitted.<sup>7</sup> Therefore, it is important to determine the most appropriate criteria for the study design when deciding on the number of factors to be removed.

Research and experience show that the selection of criteria for the number of factors to be retained is extremely important. Studies clearly show that different techniques often lead to the preservation of a different number of factors.<sup>5, 32, 38</sup> If different criteria lead to a different number of factors, then what techniques should researchers use? The

literature offers many methods to guide researchers to decide which factors to retain. These methods include the eigenvalue  $> 1$  rule, the scree test, Bartlett's chi-square test, partial correlation, bootstrapping, and variance extracted.<sup>39, 43</sup> Among the methods mentioned, the eigenvalue  $> 1$  rule, scree test, checking variance extracted, and Bartlett's chi-square test are the most commonly used as they are the default option in most statistical software packages. However, it would make more sense to use a combination of techniques rather than making decisions based on just one of these techniques. Indeed, the literature warns against using a two or three technique solution as it may not provide an accurate representation of the structure.<sup>4, 19, 34</sup>

**The Eigenvalue  $> 1$  Rule:** Each component and factor has an eigenvalue, which is value that defines the amount of variance in items associated with each factor and that can be explained by that factor.<sup>4</sup> There are different opinions in the literature about determining a lower threshold for eigenvalue. For example, Kaiser (1960) recommended that all factors with eigenvalues above 1 should be kept.<sup>39</sup> On the other hand, Jolliffe (1972), recommended that all factors with eigenvalues above .70 should be retained.<sup>44</sup> Other researchers claim that both criteria could lead to an overestimation of the number of factors to be removed.<sup>9, 30</sup>

In the literature, the eigenvalue  $>1$  rule suggested by Kaiser (1960) is generally accepted and continues to be a method preferred by many researchers.<sup>39</sup> However, it is worth noting that many criticisms have been directed towards the Kaiser Criterion method. Some researchers think that this arbitrary figure, which is used beyond its capabilities, will lead to erroneous results in determining the factors to be retained.<sup>30, 42</sup> For example, Schonrock-Adema, Heijne-Penninga, Van Hell, and Cohen-Schotanus (2009) claimed that the Kaiser criterion tends to overestimate factors.<sup>45</sup> On the other hand, Fabrigar, Wegener, MacCallum, and Strahan (1999) stated that defining a value of one (1) as the lower threshold value is rather arbitrary and criticized the Kaiser criterion in this regard.<sup>5</sup>

Based on the understanding that a single linear combination (factor or component) represents the maximum variance it can explain statistically; It may be conceptually correct to use eigenvalues as an indicator of value to maintain the existing factor structure. However, it is recommended that the scree test be used with eigenvalues so that researchers can determine how many factors to retain.<sup>7</sup>

**Scree Test:** Cattell (1966) developed a graphical method called visual screen, which shows the magnitude of the component eigenvalues against the ordinal numbers of the components.<sup>40</sup> Cattell's Scree Plot test is a graph of factors and their corresponding eigenvalues. The "x" axis in the graph represents the factors and the "y" axis represents the eigenvalues corresponding to these factors (Figure 1). Since the first component constitutes the largest amount of variance, it has the highest eigenvalue and is located on the far left. Afterwards, the eigenvalues decrease continuously and bend in a certain place, forming the breaking point. The break point in the graph, known as the inflection point, determines the number of factors to be kept. Evaluation of the breakpoint in the scree test is highly subjective and requires judgment from the researcher. Therefore, the issue of which factors should be retained is often open to debate. Also, in some cases, data points are clustered at the breakpoint, making it difficult to identify the breakpoint that would allow the researcher to determine the number of factors to keep (Figure 2). In this case, it is recommended that a straight line be drawn along the eigenvalues from the last factor and to determine the place where the deviation occurs as the break point to facilitate the determination of the breakpoint. It may also be helpful to rerun the analysis several times and manually adjust the number of factors to subtract each time in cases where it is difficult to detect the breakpoint.<sup>7, 30</sup> However, the difficulty of precisely determining the cutoff point in such clustered data (Figure 2) often leads to the over-extraction of factors.<sup>4</sup>

If the number of factors obtained from the scree test is different from the estimated

number of factors and the factor structure of the dataset is known, researchers can fix the number of factors to be removed in the factor analysis. Factors with less than three variables or item loadings less than .32 are generally seen as undesirable. The scree test only gives reliable results when the sample size is 200 or more.

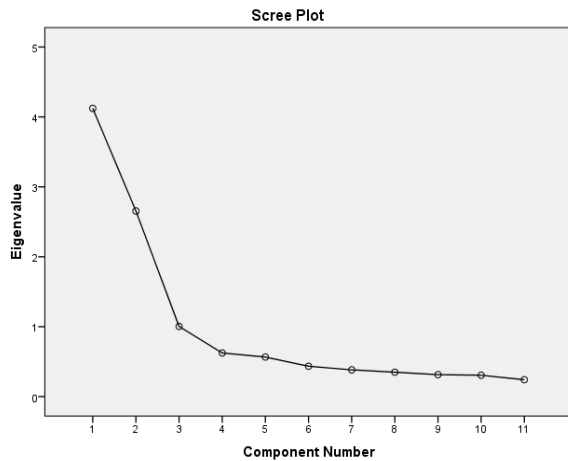


Figure 1. Scree plot

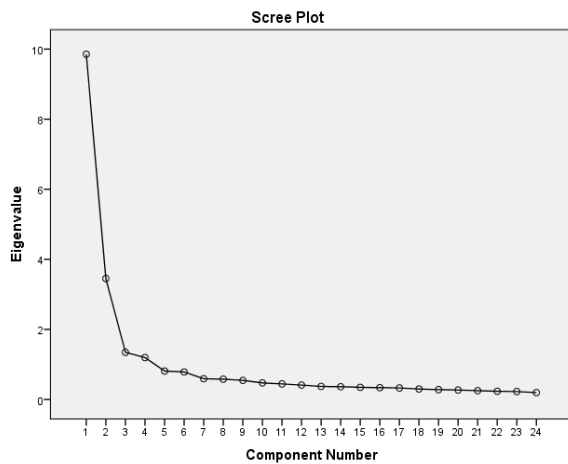


Figure 2. Scree plot

**Variance Extracted:** A third method for determining the number of factors is to keep the factors that explain a certain percentage of the variance extracted. Some statisticians believe that to create a factor from the dataset, the variance of the related factor should be taken into account.<sup>4, 15</sup> Unfortunately, as with other methods, there is no consensus on this method. While some researchers suggest that 75-90% of the variance should be taken into

account, others consider that 50% of the variance explained is sufficient.<sup>4, 15, 19</sup> Accordingly, there are differences in the branch of science in which the research is conducted and in the preference of extraction applied in EFA. For example, component analysis has more variance to explain and therefore means higher percentages of variance explained. In this case, the variance is expected to be high. If the researcher prefers common factor analysis as the extraction method, it is expected that the explained variance percentages will be less. Therefore, when interpreting the percent value of the extracted variance, the preferred extraction method should also be taken into account. In addition, the branch of science in which the research is conducted is also important in determining the variance value. For example, while at least 95% of the variance is explained in natural sciences, this value can be reduced by 50%-60% in the humanities.<sup>21</sup>

These differences in determining the number of factors require researchers to be careful when determining the number of factors to be kept. Many researchers state that it is more appropriate to make decisions with multiple criteria rather than using a single method.<sup>6, 30, 45</sup> In the case of a sufficiently large sample, it is also recommended that researchers split the data set and compare the results of factor retention criteria across the different subsets.<sup>5</sup>

**End Notes and General Comments:** If all data and graphs on factor retention are scattered or uninterpretable, the problem can be resolved by manually determining the number of factors retained. Of course, this preference depends on whether it is known beforehand how many factors the relevant structure has. Sometimes, this problem can be solved by repeating the factor analysis several times. If the factor structure is still not clear after multiple test runs, it can be considered that there is a problem with the item structure, scale design, or the hypothesis itself. Another possibility is that the sample size is too small. In this case, it may be necessary to increase the number of samples.

**Table 4. Sample of total variance explained (SPSS Output)**

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	.122	37.475	37.475	.122	7.475	37.475
2	.654	24.130	61.606	.654	4.130	61.606
3	.004	9.127	70.732	.004	.127	70.732
4	.624	5.674	76.406			
5	.566	5.147	81.553			
6	.435	3.952	85.505			
7	.383	3.480	88.985			
8	.350	3.179	92.164			
9	.314	2.857	95.021			
10	.306	2.783	97.804			
11	.242	2.196	100.000			

### Rotation Methods

Once the number of factors to keep is determined, all other factors are discarded. Items are refactored and forced into a certain number of factors. This result is rotated later. Unrotated factors are ambiguous and therefore need to be rotated to remove uncertainty and give a better interpretation. This process is called factor rotation. The literature often suggests that rotating the factor result is a critical aspect of the interpretation of factors and their variables. Indeed, Tabachnick and Fidell (2001) stated that “none of the subtraction techniques provide an interpretable solution without rotation” (p. 601).<sup>3</sup>

The main purpose of the rotation is to obtain an optimally simple structure that tries to load each variable on as few factors as possible, but while doing this, maximizes the number of high loads on each variable.<sup>7</sup> The simple structure means that each factor has highly loaded variables and the rest are low-loaded. Obtaining an optimal simple structure indirectly facilitates interpretation and allows each factor to define a separate set of interrelated variables.<sup>22</sup> For example, while the items related to organizational commitment will be highly loaded on the factor related to organizational commitment, the job satisfaction factor will be loaded with a low factor load. This will make the data easier to interpret.

When the EFA is used as a tool to identify psychological constructs and develop related questionnaires, there is some uncertainty about which rotation to apply.<sup>6</sup> Although oblique rotation is generally preferred in the literature on EFA, there is no consensus on the need to use this method clearly.<sup>5, 6, 30, 46</sup> Browne (2001) strongly recommended a multi-method approach rather than choosing a single method.<sup>47</sup> Although there are different recommendations in the literature for rotation preference, the choice of the best rotation method is at the researcher’s discretion to some extent. The effect of the selected rotation on the results may vary according to the number of samples on which the research is conducted. For example, Costello and Osborne (2005) applied both orthogonal and oblique rotations in a dataset belonging to a large sample group of 24,599 subjects and reported that similar findings were achieved.<sup>30</sup>

Among the rotation methods, there are two types of rotation methods: orthogonal (varimax, quartimax and equamax) and oblique (direct oblimin, promax). The main difference between rotation methods is related to the direction of rotation of the factors. Fabrigar, Wegener, MacCallum, and Strahan (1999) stated that if the relationship between factors is unknown, an oblique rotation should be used first.<sup>5</sup> In fact, it is seen that the oblique rotation method is mostly preferred (approximately 74.4%) in studies.<sup>6</sup>

Since the oblique rotation also includes the relationship between the factors, it gives better results compared to orthogonal solutions and is much more representative of the theoretical relationships. The fact that most of the literature uses oblique rotations in determining whether the orthogonal or oblique rotation is appropriate in EFA seems somewhat controversial. Orthogonal rotations are still frequently reported in the literature. For example, Tabachnick and Fidell (2007) recommended the use of orthogonal rotations if correlations between factors are low.<sup>3</sup> When the existing literature is evaluated, it is seen that orthogonal rotation is more appropriate in cases where the factors are conceptually independent (low correlation), whereas

oblique rotations are more appropriate in social science studies where the factors are generally related to each other (high correlation).

Detailed information on both rotations is presented below so that researchers can decide which method to use in their studies.

**Orthogonal Rotation:** Orthogonal rotation is a rotation in which the factors are rotated 90 degrees from each other.<sup>8</sup> In this rotation, the factors are assumed to be unrelated. This method is considered less valid as it is thought that the factors are related to some extent in the real environment.<sup>30</sup>

Researchers who prefer orthogonal rotation mostly use the Varimax and Quartimax rotation techniques. Varimax loads as few variables as possible into a factor to extract more factors. Thus, a structure with more factors is created. On the other hand, in Quartimax, a structure with few factors is created by gathering as many variables as possible under a single factor. More than half of researchers (53%) prefer the Varimax rotation technique, which tries to maximize the variance of square loads on a factor, in their studies.<sup>5</sup> Fabrigar, Wegener, MacCallum, and Strahan (1999) documented in a study on the same data that Varimax produced significantly less "cross-loading" than rotation. Conversely,<sup>5</sup> Goretzko, Pham, and Bühner (2021), evaluated that Varimax rotation methods should be used in complex structures where the number of cross-loadings and load coefficients are higher, and Quartimax is a more appropriate technique when less or smaller cross-loadings are expected.<sup>6</sup>

**Oblique Rotation:** Oblique rotation is a rotation in which factors are rotated obliquely. In this rotation, factors are assumed to be related. Fabrigar, Wegener, MacCallum, and Strahan (1999) emphasized that oblique rotations can be used even when the factors are not significantly correlated.<sup>5</sup> Oblique rotations, which are more suitable for social science research, explain the relationships between factors and produce more realistic results than orthogonal rotations. However, oblique rotation has a more complex structure

than orthogonal rotation. The results are presented using the pattern matrix and structure matrix. The pattern matrix shows the value reflecting the relationships between the variable and the factor when the variance of other factors is removed. The structure matrix, on the other hand, shows the relationship between the variable and the factors.

Researchers who prefer the oblique rotation method mostly use direct oblimin and promax rotation techniques. Direct oblimin finds the oblique solution that balances the criteria of: (a) each variable being relatively single-factorial (ideally one high loading and other near-zero loads) and (b) minimizing the covariance between variables on the factors. Promax starts with an oblique rotation and uses it to calculate a target matrix.<sup>32</sup> The final solution for Promax is the oblique solution that most closely matches the target matrix.

In general, direct oblimin is preferred because it simplifies the structure of the output, while promax is preferred because of its speed in large datasets.<sup>7</sup> Researchers who prefer the oblique rotation method should use zero (0) values for "Delta" and four (4) values for "Kappa". Manipulating the delta or Kappa changes the amount that the rotation procedure "allows" for factors to correlate, adding complexity to the interpretation of results.

### Factor Scores

Factor scores can be considered as variables that can be used for further statistical analysis, or they can be used to overcome the multicollinearity problem since unrelated variables can be produced.<sup>7</sup> Although three methods produce factor scores, they work according to two basic logics. The first of these is the Bartlett method and regression methods, which are associated only with their factors and produce unbiased scores. Another method is the Anderson-Rubin method, which produces unrelated and standardized scores. Although it varies depending on the research design, the Bartlett method is the most easily understood by researchers.<sup>3</sup>

Factor loads express the contribution of the variable to the formation of the relevant

factor. Therefore, factor loads need to be controlled. The signs of factor loadings indicate the direction of the correlation and do not affect the interpretation of the number of factors to be retained.<sup>12</sup> However, the values of the factor load are important in determining the number of factors to be kept. In this regard, while some researchers think that the loading should be .60 or greater, others think that the loading should be .50 or greater.<sup>28-30</sup> Some researchers think that the factor loading should be .70 or higher because this situation can be explained by the factor for approximately 50% of the variance of that item.<sup>4</sup> Tabachnick and Fidell (2007), on the other hand, stated that correlation coefficients above .30 may be sufficient.<sup>3</sup> Hair, Anderson, Tatham, and Black (1995) classified factor loads as:  $\pm 0.30$ =minimal,  $\pm 0.40$ =important, and  $\pm 0.50$ =practically significant using another general rule.<sup>21</sup> As can be seen, different cut-off points for factor loading have been proposed in the literature.

Researchers need to set a cut-off point for the loading value to facilitate interpretation. According to a detailed review of the literature, it is recommended that the lower cut-off value be set as .50 (where there are 4 or fewer variables in the factor) or .40 (where there are more than 4 variables in the factor) depending on the number of variables in the factor. The choice of cut-off point may vary depending on the ease of interpretation, including how complex variables are handled. In addition, there is a view that the larger the sample size, the more it is permissible for a factor to be low factor-loaded.<sup>48</sup>

Although high factor loadings seem to be preferred by researchers, it is also important to examine low loadings to confirm the identification of factors as well as to determine under which factor the high loadings accumulate.<sup>49</sup> For example, it is not preferable to load a variable of job satisfaction under a factor called organizational commitment. The expectation here is that variables with high factor loads of organizational commitment are gathered under the organizational commitment factor, while the variables with low factor loads are

gathered in the job satisfaction factor. In other words, several items must be cross-loaded for each factor to define a separate set of interrelated variables.<sup>7</sup> The presence of cross-loading indicates that the factors are more stable.

Cross-loading is when an item is loaded in two or more factors. Cross-loading is a situation encountered by numerous researchers and is likely to happen in many empirical studies. The important aspect here is the factor load value of the cross-loaded variable. If the cross-loaded item has a value of .32 or higher, it is recommended that the item be removed from the search, because a factor load of .32 shows that it explains about 10% of the overlapping variance and exceeds the lower threshold value. However, depending on the design of the study, an item in a cross-loading situation may not be excluded from the analysis based on the assumption that it is the latent nature of the factor involved. In addition, another situation to be considered in the case of cross-loading is whether the high load is under the correct factor and whether there is a difference of .1 or more between the cross loads. For example, if it is assumed that a variable of organizational commitment has a factor load of .562 on the organizational commitment factor, and a factor load of .401 is cross-loaded on the job satisfaction factor, it is recommended that it not be excluded from the study because the high factor load of this variable is under the right factor (i.e., in the organizational commitment factor) and there is a value of more than .1 between the cross-loads. After the cross-loaded variables are removed from the analysis, the data should be reanalyzed after those variables have been removed for a more refined solution.

The removal of any variable from the research is an important decision. Removing too many variables or keeping variables that do not have enough load cause the relevant factor to not be defined correctly. A factor that is not defined correctly is not considered stable and robust and causes unhealthy research results. For a factor to be considered stable and robust, it must contain at least three

to five items with significant loadings.<sup>30</sup> More importantly, the variables and factors must be conceptually meaningful. Finally, theoretical information can sometimes be more important than a statistical measurement. Thus, while

items are expected to have a significant attribution indicating a statistically valuable contribution, the conceptual significance of that item should be examined before removing a variable from the factor.

## CONCLUSION AND RECOMMENDATIONS

With statistical software becoming increasingly user-friendly, EFA has become accessible and more easily applicable for today's researchers. Although it has become easier for researchers to access EFA, the procedures for EFA still do not seem to be clearly understood.

The main reason underlying the inability to understand all aspects of EFA is that there are different opinions and recommendations in the literature about which procedures should be applied in EFA. Other reasons may include the lack of adequate EFA training for researchers or their unfamiliarity with the complex procedures of EFA. Regardless of the reason, researchers' suboptimal decisions regarding EFA produce skewed and potentially meaningless solutions. In particular, the quality of EFAs reported in psychology research is ordinary and very low, and many researchers continue to use suboptimal methodology.<sup>5, 50</sup> To overcome this problem and improve this situation, there is a need for a systematic, evidence-based guide for researchers with intermediate and lower statistical education to allow EFA studies to be conducted properly.<sup>51</sup> Therefore, this study is a guide for different views and practices in the literature on EFA.

Which procedures will be applied in EFA and which methodological paths will be followed vary according to the specific case. Therefore, researchers who conduct EFA analysis should evaluate each case separately.<sup>6</sup> Based on this approach, we realize that it is difficult to make general recommendations on properly conducting EFA. We also acknowledge that EFA is a very complex analysis and includes many procedures. We also understand that differences in the recommendations made in the literature for EFA also lead to confusion

regarding the use of EFA. However, although EFA is a complex statistical approach, the approaches adopted in the analyses are sequential and linear with many options. In other words, there are several options and methods in the literature that need to be implemented to help researchers make the right decisions, improve the accuracy of factor analysis and increase the quality of the resulting solution.

Firstly, researchers should determine the research objectives precisely and choose the most appropriate procedures in this regard while conducting EFA. Regardless of the procedures chosen, researchers should transparently report what goals they want to achieve, what procedures they have implemented to achieve those goals and their research findings. Such a course of action allows readers to evaluate the research solution and compare it with subsequent research findings. The ability of readers to independently evaluate the results obtained in an EFA study will increase confidence in the results of previously reported research.

Researchers must ensure that the dataset from EFA meets the requirements for EFA. To apply factor analysis to the dataset, it is essential to ensure that univariate and multivariate normality are ensured and that there are no outliers.<sup>1,9</sup> The lack of normality in the data affects the Pearson's Product-Moment Correlation coefficients ( $r$ ) between the variables used in the calculation of the EFA results (reduces the size of the correlation coefficient). In addition, the presence of outliers in the data can also cause artificial factors to be produced. In such cases, the EFA findings are misleading. Therefore, the distribution of data, which may affect Pearson correlations, should be controlled.



Missing data can affect EFA results. For this reason, it is necessary to check whether the missing data in the data set occurs in a non-random order. It is recommended that data with missing values should not be included in the factor analysis to avoid incorrect estimations in factor analysis.<sup>3</sup> However, deleting the missing data cases in a list or one-by-one is a waste of effort and time. In such situations, methods such as regression, "mean, multiple and maximum likelihoods", which are found in many software programs and used to assign missing data, can be used. Studies and simulations show that the mean method is acceptable in cases where <10% of the data is missing, and the regression method is acceptable in cases where <15% of the data is missing (Schumacker, 2015). It would be a correct approach to use multiple and maximum likelihood methods in assigning missing data to a greater extent.<sup>52</sup>

Researchers are expected to be cautious with sample size, because EFA generally works better with larger sample sizes and larger samples reduce the margin of error, leading to more stable solutions. Authors think that a sample size of greater than 200 should be used to create reliable factor patterns and produce healthy results for EFA, and therefore, 200 should be determined as the lower threshold. In sample sizes less than 200, consideration should be given to factors such as factor loading and the number of variables. If the sample size is too small, it may not be possible to obtain generalizable or reproducible results. It should not be forgotten that repetition is a fundamental principle of science. For this reason, researchers who do not have enough experience in the analysis should prefer sample sizes higher than 200 to reach healthy and reproducible findings in their studies. In general, we describe 100 as being poor, 150 as moderate, 200 as adequate, 250 as good, 300 as very good, and above 400 as excellent. Finally, we would like to emphasize that homogeneous samples cause a decrease in variance and factor loads. For this reason, we recommend that researchers who want a high factor load should collect data from a heterogeneous instead of a homogeneous sample group.

When applying EFA to datasets with sample sizes of 200 or more, authors consider that a factor load of .30 is sufficient and it would be appropriate to exclude items with lower factor loads from the study. Because the factor loading of an item is a measure of how much that item contributes to that factor, low factor loadings indicate that related factors are not well explained by items. Therefore, it would be appropriate to exclude that item from the analysis.

Researchers should check the suitability of the dataset for EFA before performing EFA analysis. For this, the results of the Kaiser-Meyer-Olkin Measure (KMO) and Bartlett's Sphericity Test could be checked. A significant KMO value of .06 and above and Bartlett's Test of Sphericity being  $p < .05$  indicate that the data are suitable for factor analysis.

While performing factor analysis, the literature does not provide a clear picture to clarify the distinctions between extraction methods. This complexity causes difficulties for researchers in interpreting analyses and reduces the researcher's ability to make sound theoretical decisions. Although the extraction methods applied can affect the EFA solution, the similarity in the results produced is very high. Despite the high similarity, it is important for researchers to apply the most accurate method that best meets the purpose and needs of the research. Therefore, as a general understanding, maximum likelihood estimation should be relied on for normally distributed data, and principal axis factoring estimation should be preferred for non-normal and ordinal data (especially when Likert-type items with less than five categories are used). Another advantage of the maximum likelihood approach is that confirmatory factor analysis includes the maximum likelihood method. Therefore, it allows cross-validation of EFA and confirmatory factor analysis results. Therefore, we recommend the use of the maximum likelihood method by researchers who will also report confirmatory factor analysis results in their research.

EFA analysis provides researchers with an idea about how many factors are collected

under the variables in the dataset and which factors are important. In some cases, many of the factors in the dataset do not contribute significantly to the overall solution or are very difficult to interpret. Factors that are difficult to interpret and do not contribute to the solution often cause confusion or error. Therefore, it is not correct to keep the relevant factors in the analysis. The researcher needs to keep the correct number of factors, as this decision will directly affect the results. To arrive at the right solutions, researchers must correctly interpret the EFA results and decide which factors should be retained. The literature offers many methods to guide researchers in deciding which factor to retain.

These methods can be specified as eigenvalue  $> 1$  rule, scree test, Bartlett's chi-square test, partial correlation, bootstrapping, and checking variance extracted. Among the aforementioned methods, the eigenvalue  $> 1$  rule, scree test, checking variance extracted, and Bartlett's chi-square test are the most commonly used as they are a default option in most statistical software packages. However, it would make more sense to use a combination of techniques rather than making a decision based on just one of the techniques presented. Indeed, the literature warns against using a two or three technique solution as it may not provide an accurate representation of the structure.<sup>4, 19, 34</sup>

#### REFERENCES

1. Child, D. (2006). "The essentials of factor analysis. (3rd ed.)". New York, NY: Continuum International Publishing Group.
2. Bartholomew, D, Knotts, M. and Moustaki, I. (2011). "Latent variable models and factor analysis: A unified approach. (3rd ed.)". West Sussex, UK: John Wiley & Sons.
3. Tabachnick, B. G. and Fidell, L. S. (2007). "Using multivariate statistics (5th ed.)". Boston, MA: Allyn & Bacon.
4. Beavers, A. S, Lounsbury, J. W, Richards, J. K, Huck, S. W, Skolits, G. J. and Esquivel, S. L. (2013). "Practical considerations for using exploratory factor analysis in educational research". *Practical Assessment, Research, and Evaluation*, 18 (1), 6–16.
5. Fabrigar, L. R, Wegener, D. T, MacCallum, R. C. and Strahan, E. J. (1999). "Evaluating the use of exploratory factor analysis in psychological research". *Psychological Methods*, 4 (3), 272–299.
6. Goretzko, D, Pham, T. T. H. and Bühner, M. (2021). "Exploratory factor analysis: Current use, methodological developments and recommendations for good practice". *Current Psychology*, 40 (7), 3510–3521.
7. Yong, A. G. and Pearce, S. (2013). "A beginner's guide to factor analysis: Focusing on exploratory factor analysis". *Tutorials in Quantitative Methods for Psychology*, 9 (2), 79–94.
8. Rummel, R. J. (1970). "Applied factor analysis. Evanston", IL: Northwestern University Press.
9. Field, A. (2009). "Discovering statistics using SPSS: Introducing statistical method (3rd ed.)". Thousand Oaks, CA: Sage Publications.
10. Hogarty, K. Y, Hines, C. V, Kromrey, J. D, Ferron, J. M. and Mumford, K. R. (2005). "The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination". *Educational and Psychological Measurement*, 65 (2), 202–226.
11. Tabachnick, B. G, Fidell, L. S. and Ullman, J. B. (2007). "Using multivariate statistics (Vol. 5, pp. 481–498)". Boston, MA: Pearson.
12. Kline, P. (1994). "An easy guide to factor analysis". New York, NY: Routledge.
13. Baraldi, A. N. and Enders, C. K. (2010). "An introduction to modern missing data analyses". *Journal of School Psychology*, 48 (1), 5–37.
14. Schumacker, R. E. (2015). "Learning statistics using R". Thousand Oaks, CA: Sage.
15. Sürücü, L, Şeşen, H. and Maslakçı, A. (2023). "Regression, Mediation/ Moderation, and Structural Equation Modeling with SPSS, AMOS, and PROCESS Macro". France, Livre de Lyon.
16. Mundfrom, D. J, Shaw, D. G. and Ke, T. L. (2005). "Minimum sample size recommendations for conducting factor analyses". *International Journal of Testing*, 5 (1), 159–168.
17. Winter, J. C, Dodou, A. and Wieringa, P. A. (2009). "Exploratory factor analysis with small sample sizes". *Multivariate behavioral research*, 44 (2), 147–181.
18. Comrey, A. L. (1973). "A first course in factor analysis". New York: Academic.
19. Pett, M. A, Lackey, N. R. and Sullivan J.J. (2003). "Making Sense of Factor Analysis: The use of factor analysis for instrument development in health care research". California: Sage Publications Inc.
20. Lawley, D. N. and Maxwell, A. E. (1971). *Factor analysis as a statistical method*. Butterworths: United Kingdom.
21. Hair, J, Anderson, R. E, Tatham, R. L. and Black, W.C. (1995). "Multivariate data analysis". New Jersey: Prentice-Hall Inc.
22. Cattell, R. B. (1973). "Factor analysis". Westport, CT: Greenwood Press.
23. Suhr, D. (2006). "Exploratory or Confirmatory Factor Analysis". SAS Users Group International Conference (pp. 1 - 17). Cary: SAS Institute, Inc.
24. Bryant, F. B. and Yarnold, P. R. (1995). "Principal-components analysis and exploratory and confirmatory factor analysis". In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99–136). American Psychological Association.
25. Rouquette, A. and Falissard, B. (2011). "Sample size requirements for the internal validation of psychiatric scales". *International Journal of Methods in Psychiatric Research*, 20 (4), 235–249.
26. MacCallum, R. C, Widaman, K. F, Zhang, S. and Hong S. (1999). "Sample size in factor analysis". *Psychological Methods*, 4 (1), 84–99.
27. Sappas, K. and Zeller, R. A. (2002). "Minimizing sample size when using exploratory factor analysis for measurement". *Journal of Nursing Measurement*, 10 (2), 135–53.

28. Guadagnoli, E. and Velicer, W. F. (1988). "Relation to sample size to the stability of component patterns". *Psychological Bulletin*, 103 (2), 265–275
29. Preacher, K. J. and MacCallum, R. C. (2002). "Exploratory factor analysis in behavior genetics research: Factor recovery with small sample sizes". *Behavior Genetics*, 32 (1), 153–161.
30. Costello, A. B. and Osborne, J. (2005). "Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis". *Practical Assessment, Research, and Evaluation*, 10 (1), 7–27.
31. Haitovsky, Y. (1969). "Multicollinearity in regression analysis: A comment". *Review of Economics and Statistics*, 51 (4), 486–489.
32. Conway, J. M. and Huffcutt, A. I. (2003). "A review and evaluation of exploratory factor analysis practices in organizational research". *Organizational Research Methods*, 6 (2), 147–168.
33. Byrne, B. M. (2001). "Structural equation modeling with AMOS - Basic concepts, applications, and programming". LEA, ISBN 0-8058-4104-0
34. Fava, J. L. and Velicer, W. F. (1992). "The effects of overextraction on factor and component analysis". *Multivariate Behavioral Research*, 27 (1), 387–415.
35. Gorsuch, R. L. (1997). "Exploratory factor analysis: Its role in item analysis". *Journal of Personality Assessment*, 68 (3), 532–560.
36. DeWinter, J. C. F. and Dodou, D. (2012). "Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size". *Journal of Applied Statistics*, 39 (4), 695–710.
37. Widaman, K. F. (2012). "Exploratory factor analysis and confirmatory factor analysis". In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol 3: Data analysis and research publication* (pp. 361–389). Washington, DC: American Psychological Association.
38. Zwick, W. R. and Velicer, W. F. (1986). "Comparison of five rules for determining the number of components to retain". *Psychological Bulletin*, 99 (3), 432–442.
39. Kaiser, H. F. (1960). "The application of electronic computers to factor analysis". *Educational and Psychological Measurement*, 20 (1), 141–151.
40. Cattell, R. B. (1966). "The scree test for the number of factors". *Multivariate Behavioral Research*, 1 (1), 245–276.
41. Bartlett, M. S. (1951). "A further note on tests of significance in factor analysis". *British Journal of Psychology*, 4 (1), 1–2.
42. Velicer, W. F. (1976). "Determining the number of components from the matrix of partial correlations". *Psychometrika*, 41 (1), 321–327.
43. Thompson, B. (2004). "Exploratory and confirmatory factor analysis". Washington, DC: American Psychological Association.
44. Jolliffe, I. T. (1972). "Discarding variables in a principal component analysis, I: Artificial data". *Applied Statistics*, 21 (1), 160–173.
45. Schonrock-Adema, J., Heijne-Penninga, M., Van Hell, E. A. and Cohen-Schotanus, J. (2009). "Necessary steps in factor analysis: enhancing validation studies of educational instruments". *Medical Teacher*, 31 (1), 226–232.
46. Baglin, J. (2014). "Improving your exploratory factor analysis for ordinal data: A demonstration using FACTOR". *Practical Assessment, Research & Evaluation*, 19 (1), 5–10.
47. Browne, M. W. (2001). "An overview of analytic rotation in exploratory factor analysis". *Multivariate Behavioral Research*, 36 (1), 111–150.
48. Stevens, J. P. (2002). "Applied multivariate statistics for the social sciences (4th ed.)". Hillsdale, NS: Erlbaum.
49. Gorsuch, R. L. (1983). "Factor analysis (2nd ed.)". Hillside, NJ: Lawrence Erlbaum Associates.
50. Norris, M. and Lecavalier, L. (2010). "Evaluating the use of exploratory factor analysis in developmental disability psychological research". *Journal of Autism and Developmental Disorders*, 40 (1), 8–20.
51. Watkins, M. W. (2018). "Exploratory factor analysis: A guide to best practice". *Journal of Black Psychology*, 44 (3), 219–246.
52. Baraldi, A. N. and Enders, C. K. (2010). "An introduction to modern missing data analyses". *Journal of School Psychology*, 48 (1), 5–37.