

2013

Practical Considerations for Using Exploratory Factor Analysis in Educational Research

Amy S. Beavers

John W. Lounsbury

Jennifer K. Richards

Schuyler W. Huck

Gary J. Skolits

See next page for additional authors

Follow this and additional works at: <https://scholarworks.umass.edu/pare>

Recommended Citation

Beavers, Amy S.; Lounsbury, John W.; Richards, Jennifer K.; Huck, Schuyler W.; Skolits, Gary J.; and Esquivel, Shelley L. (2013) "Practical Considerations for Using Exploratory Factor Analysis in Educational Research," *Practical Assessment, Research, and Evaluation*: Vol. 18 , Article 6.

DOI: <https://doi.org/10.7275/qv2q-rk76>

Available at: <https://scholarworks.umass.edu/pare/vol18/iss1/6>

This Article is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Practical Assessment, Research, and Evaluation by an authorized editor of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

Practical Considerations for Using Exploratory Factor Analysis in Educational Research

Authors

Amy S. Beavers, John W. Lounsbury, Jennifer K. Richards, Schuyler W. Huck, Gary J. Skolits, and Shelley L. Esquivel

Beavers et al.: Practical Considerations for Using Exploratory Factor Analysis in

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Copyright is retained by the first or sole author, who grants right of first publication to the *Practical Assessment, Research & Evaluation*. Permission is granted to distribute this article for nonprofit, educational purposes if it is copied in its entirety and the journal is credited. PARE has the right to authorize third party reproduction of this article in print, electronic and database forms.

Volume 18, Number 6, March 2013

ISSN 1531-7714

Practical Considerations for Using Exploratory Factor Analysis in Educational Research

Amy S. Beavers, John W. Lounsbury, Jennifer K. Richards,
Schuyler W. Huck, Gary J. Skolits, and Shelley L. Esquivel
The University of Tennessee

The uses and methodology of factor analysis are widely debated and discussed, especially the issues of rotational use, methods of confirmatory factor analysis, and adequate sample size. The variety of perspectives and often conflicting opinions can lead to confusion among researchers about best practices for using factor analysis. The focus of the present review is to clarify terminology, identify key issues, and clarify areas of debate regarding best practices and functions of factor analytic procedures. The conclusions and implications drawn should be useful to researchers in education, psychology, and cognate social fields who employ factor analytic procedures or evaluate research using factor analytic methods.

Factor analytic procedures are statistical methods used for examining the relationships within a group of observed variables, as measured through questions or items. It is important to note that factor analysis is not a singular statistical method, but rather a group of statistical analyses that share similar methodology and functionality. The theoretical and mathematical variations among the processes allow the analyses to accommodate breadth of purpose and theory in research and result in the widespread use of the tool across disciplines and applications; however, it is the flexibility of the statistical methods that fuel ongoing debate about the appropriate applications of these methods. For over sixty years, researchers from varied social science disciplines have saturated the factor analysis literature with definitions, discussions, and debates concerning factor analysis approaches, applications, and recommendations for most appropriate usage (cf., Garson, 2010; Loo, 1979; Pett, Lackey & Sullivan, 2003; Tabachnick & Fidell, 2001; Velicer & Fava, 1990).

Examination of factors, or dimensions, is most often applied in the development and validation of measures such as personality scales (Schonrock-Adema, Heijne-Penninga, Van Hell & Cohen-Schotanus, 2009); however, it can be used in a variety of measurement

applications (Furr & Bacharach, 2008). Regardless of the setting, within each analysis there are a range of choices and decisions the researcher must make to improve the accuracy of the factor analysis they use and to enhance the quality of the resulting solution (Fabrigar, Wegener, MacCallum & Strahan, 1999). A commonly cited limitation of exploratory factor analysis (EFA) is its level of subjectivity stemming from the many methodological decisions a researcher must make to complete a single analysis, with the accuracy of the results largely dependent upon the quality of these decisions (Henson & Roberts, 2006; Tabachnick & Fidell, 2001).

To further compound the complexity of the analysis, factor analysis is a cyclical process of continually refining and comparing solutions until the most meaningful solution is reached (Tabachnick & Fidell, 2001). This commonsense approach to interpreting the analysis in light of theory and conceptual foundation should be accompanied by a strong theoretical and mathematical justification for the methodological choices and decisions, yet because of the abundance of sources and opinions within the factor analysis literature, it is often difficult for a researcher to determine the most accurate use of this tool within a given research context. The purpose of

the present overview is to provide a step-by-step approach to factor analytic procedures and to offer an evaluation of the theoretical and practical merits associated with four common areas of debate. Recommendations are offered for best factor analytic practice for educational researchers as framed by the following three questions:

1. How large should the sample be to know if the it is “large enough” to produce a reliable factor analytic solution?
2. What is the difference between Component Analysis and Common Factor Analysis?
3. Is it necessary to rotate the initial factor pattern matrix in order to achieve an interpretable and meaningful solution?

In the journey to understand factor analysis so that responsible, methodologically-sound decisions could be made, approximately 45 manuscripts were reviewed. Initially, books and book chapters were consulted to provide a basic overview of the technique and its various components. From these sources, a number of questions and controversies remained. Academic Search Premier was used to identify peer-reviewed journal articles that included overviews, examinations, and/or applications of factor analytic methods and procedures. This process allowed the authors to identify seminal articles as well as trace evidence-based decision making in factor analysis over time. While this does not constitute a meta-analysis, it does provide a broad foundation to ground the recommendations made by the authors for best practices use in EFA. Review of these sources provided additional depth of understanding and allowed the authors to draw conclusions as to general rules-of-thumb for many decision-making controversies often encountered in data analysis of research problems. These guidelines formed the basis for the present article. Where apparent contradictions still existed in the literature, the authors based decisions regarding recommendations on the methodological soundness of the studies consulted and the mathematical foundations of the techniques employed.

SAMPLE SIZE

The first question comes during the planning stages: *How large should the sample be to know if it is “large*

enough” to produce a reliable factor analytic solution? The literature contains copious amounts of information in response to this question; however, criteria provided for determining the sufficiency of a sample for factor analytic procedures vary greatly and include a plethora of differing criteria. Sample size requirements may generally be categorized in two ways; a minimum number of cases or a subjects-to-variables ratio (STV) required to achieve an adequate sample. For example, selected criterion suggests the sample size should have:

- 51 more cases than the number of variables (Lawley & Maxwell, 1971).
- At least 10 cases for each item, and the subjects-to-variables [STV] ratio should be no lower than 5 (Bryant & Yarnold, 1995).
- At least 100 cases and a STV ratio of no less than 5 (Suhr, 2006).
- At least 150 - 300 cases (Hutcheson & Sofroniou, 1999).
- At least 200 cases, regardless of STV (Gorsuch, 1983).
- At least 300 cases (Norusis, 2005).

Similar guidelines are provided throughout the literature without clear consensus (Tabachnick & Fidell, 2001; Zhao, 2009). There is, however, general agreement that an inadequate sample size can be detrimental to the factor analytic process and produce unreliable, and therefore, non-valid results (Osborne & Costello, 2004; Pett et al., 2003). How, then, can anyone use factor analysis methods with confidence, assuming that a sample of insufficient size will undermine any meaning produced, without a realistic guide to what sample criteria is “best” or the “right” one? Initial review of the literature suggests that a base number of cases is required and that a ratio of cases to variables should be considered once the base number is met.

Upon closer inspection of the literature, a general opinion has emerged, suggesting that ratio criteria do not provide an accurate guide (Guadagnoli & Velicer, 1988; Hogarty, Hines, Kromrey, Ferron & Mumford, 2005; MacCallum, Widaman, Preacher & Hong, 2001; Osborne & Costello, 2004; Zhao, 2009). Guadagnoli and Velicer (1988) suggest, what has been largely

confirmed in the literature, that the needed sample size is conditional upon the strength of the factors and the items. They provide a new criterion operationalizing these relationships. If the factors have four or more items with loadings of .60 or higher, then the size of the sample is not relevant. If the factors have 10 to 12 items that load moderately (.40 or higher), then a sample size of 150 or more is needed to be confident in the results. Finally, if factors are defined with few variables and have moderate to low loadings, a sample size of at least 300 is needed. Fabrigar et al. (1999) and MacCallum et al. (2001), further support that stable solutions can be reached with samples as low as 100 when three to four strong items (loadings of .70 or greater) comprise a factor, suggesting that weaker relationships need a larger sample size.

A strong solution, made up of stable factors, reduces the influence of the sample size; however, a larger sample size decreases sampling error resulting in more stable solutions (Hogarty et al., 2005). Determination of sample size sufficiency is dependent upon the stability of the solution; therefore, the adequacy of a sample cannot be fully determined until the analysis has been conducted. While the final factor solution can provide enough evidence to suggest that a sample is sufficient, one is still left with the question of how much is enough when collecting the original sample? Because the family of factor analysis procedures are multivariate tools, and multivariate methods require larger sample sizes than do univariate methods, one should plan for a sample of at least 150 cases for initial structure exploration.

CHECK ASSUMPTIONS OF DATA

After the sample data has been obtained, the data used must satisfy the assumptions required of multivariate statistical techniques, including: large sample size, linearity, absence of outliers, continuous data, lack of extreme multicollinearity, and low percentage of missing data (Comrey, 1985; Pett et al., 2003). Factor analysis differs from other multivariate procedures in that there is no separate identification of dependent or independent variables. The relationships between variables are examined without specification of one variables' influence upon another. As a result, multivariate normality is not required within all

methods of extraction in factor analysis (Tabachnick & Fidell, 2001).

EVALUATE FACTORABILITY OF MATRICES

Correlational Values

In addition to meeting assumptions before the factorization of a set of variables, the strength of the relationships and linear relationships are evaluated by reviewing the correlation matrix produced from the data. Generally, correlations exceeding .30 provide enough evidence to indicate that there is enough commonality to justify comprising factors (Tabachnick & Fidell, 2001). If intercorrelations are unexpectedly low, it may be a result of low variance. Samples that are too homogenous can exhibit low variance; consequently, the correlation will be low potentially failing to reveal a factor, or common relationship, that does exist (Fabrigar et al., 1999).

Table 1 includes the correlation matrix for a sample of 5 items in a Teacher Satisfaction and Sense of Belonging Scale (TSSBS).

Table 1. *Example of a Correlation Matrix of Five Items in the TSSBS*

Item	1	2	3	4	5
1	1				
2	.642	1			
3	.542	.801	1		
4	.355	.539	.682	1	
5	.244	.355	.475	.510	1

Note. Pearson's r Correlational values are reported.

With one exception, the intercorrelations exceed .30. The correlation between Item 1 and Item 5 is .244. While this value is below .30, the relationships with other items exceed .30. Therefore, this one correlation is not enough evidence to suggest that factoring would not be beneficial.

Determinant of the Matrix

An additional assessment of factorability of the data comes from the *determinant* of the correlation matrix. The determinant of a matrix is a single value

calculated using the values within a square matrix, revealing the presence or absence of possible linear combinations within the matrix. The determinant of a 2 x 2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $ad - cb$. The determinant of the 2 x 2 matrix $\begin{bmatrix} 4 & 1 \\ 3 & 5 \end{bmatrix}$ is equal to $4(5) - 3(1)$, which equals 17. When the determinant does not equal zero, the matrix can be explained by linear combinations; however, if the determinant of a matrix equals zero it is described as a singular matrix. A singular matrix has either an infinite number of linear combinations or there are no possible linear combinations within the matrix. For example, in the matrix $\begin{bmatrix} 2 & 5 \\ 6 & 15 \end{bmatrix}$, the determinant equals $2(15) - 6(5) = 0$. Row one, [2, 5], does not contribute any unique value to the solution and can be expressed entirely as a linear combination of other rows: $3(2x + 5y) = 6x + 15y$.

With the exception of cases where the determinant is zero, the values can be arranged into linear combinations. In factor analysis, these linear combinations are considered factors. A non-zero determinant indicates that a factor or component is mathematically possible; however, it does not offer any indication of the practical meaning or significance of the factors. Because the values of a correlation matrix are restricted to values between -1 and 1, the values for the determinant of a correlation matrix range from 0 to 1. The values seen most often are very small, suggesting that a few linear combinations exist (Pett et al., 2003). For example, Table 2 reports the measures of factorability for the TSSBS correlation matrix used previously (see Table 1).

Table 2. *Example of Measures for Assessing the Correlation Matrix*

Measure	Value
Determinant	1.14 E -11
Bartlett's Test of Sphericity	$p < .0001$
Kaiser-Meyer-Olkin Test of Sampling Adequacy	.889

In addition to using the determinant of a matrix, Bartlett's Test of Sphericity and the Kaiser-Meyer-Olkin Test of Sampling Adequacy (KMO) are commonly used to provide more complex measures for

assessing the strength of the relationships and suggesting factorability of the variables.

Bartlett's Test of Sphericity

The determinant value in the TSSB example is very close to zero. To evaluate if this determinant value is statistically different from zero, Bartlett's Test of Sphericity is used. The null hypothesis of Bartlett's test states that the observed correlation matrix is equal to the identity matrix, suggesting that the observed matrix is not factorable (Pett et al., 2003). In the example used, Bartlett's test produced a significant test result, rejecting the null hypothesis. Bartlett's Test provides evidence that the observed correlation matrix is statistically different from a singular matrix, confirming that linear combinations exist.

Kaiser-Meyer-Olkin Test of Sampling Adequacy

The Kaiser-Meyer-Olkin Test of Sampling Adequacy (KMO) is a measure of the shared variance in the items. Kaiser, Meyer, and Olkin suggest the following guideline for assessing the measure (Friel, n.d.):

Table 3. *Interpretation Guidelines for the Kaiser-Meyer-Olin 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	Don't Factor

INITIAL EXTRACTION

Factoring the matrix begins with the initial extraction of linear combinations. Matrix algebra is used to create linear combinations of items that explain the greatest amount of variance¹ among the items. The initial extraction assumes that each combination is orthogonal (independent or uncorrelated) to the others.

¹ The type of variance included varies by extraction method.

These linear combinations are called factors or components. The first factor extracted accounts for the greatest percentage of variance in the items. The next linear combination attempts to account for the maximum amount of remaining variance that is not included in the first factor. This process continues until all of the variance in the sample is explained (Suhr, 2006). The two most common terms used to discuss the initial extraction are *component analysis* and *common factor analysis*.

What is the difference between Component Analysis and Common Factor Analysis? It is commonly accepted practice to use factor analysis as a broad heading for two distinct statistical techniques: component analysis and common factor analysis (Tabachnick & Fidell, 2001). These terms, as well as other related terms such as component and factor, are often used interchangeably within the literature causing confusion for the reader (Garson, 2010; Furr & Bacharach, 2008; Tabachnick & Fidell, 2001). Both theoretical and mathematical differences exist between component analysis and common factor analysis. The failure to make these distinctions clear leads to difficulty interpreting the context and diminishes the researcher's ability to make theoretically sound decisions.

Component analysis serves as a means to accurately report and evaluate a large number of variables using fewer components, while still preserving the dimensions of the data. It is widely described as a data reduction method used to summarize a large set of variables (Costello & Osborne, 2005; Velicer & Jackson, 1990b). Theoretically, component analysis assumes that the component is a composite of the observed variables, or that the individual item scores *cause* or define the component (DeCoster, 1998). See Figure 1. A student's score on a math test is an illustration of this causal relationship in component analysis. The student's performance on each item comprises the overall test score.

By contrast, common factor analysis allows the exploration of underlying constructs, which cannot be measured directly, through items thought to be reflective measures of the construct (Byrne, 2001). Common factor analysis assumes that individual item

scores are a *result* of an underlying *factor* or construct (DeCoster, 1998). A measure to assess the student's attitudes about math would be an illustration of this relationship. The student's responses to the items are thought to reflect their underlying attitudes about math.

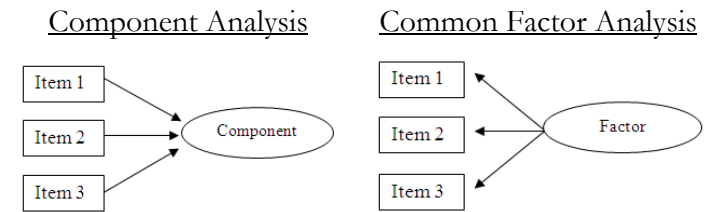


Figure 1. Directional Relationships in Component Analysis and Common Factor Analysis. This figure illustrates the differences of the causal relationships in component analysis and common factor analysis

Mathematically, component analysis and common factor analysis differ in the amount of variance included in the solution. There exist multiple types of variance: *common (shared) variance*, *specific (unique) variance*, and *error variance (measurement error)*. Common variance is the variability present in an item that is shared with other items and factors. Specific variance is the variance resulting from the unique attributes of an item that cannot be explained by other variables or factors. Error variance is associated with the measurement process and is an indication of unreliability.

Component analysis includes all three types of variance and does not partial out any variance from the items when examining the relationships. Because the total variance is included in components analysis, some argue that the estimates provided reflect inflated values (Costello & Osborne, 2005). By contrast, common factor analysis removes specific variance and error variance from the calculations, including only common variance to extract the factor solution.

The literature contains a multitude of studies whose purpose is to determine if the mathematical difference between component analysis and common factor analysis result in a practically different solution (Osborne & Costello, 2004; Costello & Osborne, 2005; Fabrigar et al., 1999). In fact, a special issue of *Multivariate Behavioral Research* (1990) focuses entirely on this topic and suggests that the alignment of the solutions are dependent upon the strength and stability

of the items. Removing sources of unreliability suggests that common factor analysis would produce more accurate solutions; however, it is commonly reported that the results of both processes are similar (Fava & Velicer, 1992; Tabachnick & Fidell, 2001; Velicer & Jackson, 1990a). This proves to be true when the measures used are reliable. The mathematical proofs and research on both sides of this well-established debate are beyond the scope of this review.

In a practical research context, the researcher should be aware that component analysis and common factor analysis function similarly and can produce comparable results. The mathematical and theoretical foundations of the two methods vary; however, the practical sequencing of steps and processes are the same (Pett et al., 2003). Component analysis includes the total variance in the items and has no underlying structural assumptions. Common factor analysis only includes the common variance and hypothesizes that the item responses are a product of an underlying construct. Both component analysis and common factor analysis are mathematically able to reduce variables to a smaller number of components or factors; however, the precise interpretability and understanding of these values vary by the method used to extract these linear combinations.

Methods of Initial Extraction

Component analysis includes the total variance in the initial extraction. *Principal Component Analysis* (PCA) is the most widely used extraction method of component analysis and is most appropriate when the purpose is to reduce the number of items to a smaller number of representative components (Costello & Osborne, 2005; DeCoster, 1998), whereas common factor analysis only includes the common (shared) variance in the extraction. The two most commonly used extraction methods of common factor analysis are *Principal Axis Factoring* (PAF) and *Maximum Likelihood Estimation* (ML)². PAF is appealing because it requires no distributional assumptions and may be used if data are not normally distributed (Fabrigar et al., 1999). ML requires multivariate normality (Pett et al., 2003);

² Additional component and factor extraction methods include image, alpha, unweighted least squares and generalized least squares.

however, the benefit of using ML is that in addition to the correlational estimates, it produces significance tests for each item as well as fit statistics for the structure.

Example. To extract the initial factor solution of the example TSSB scale, Principal Axis Factoring (PAF) is used because the items are believed to reflect the underlying satisfaction and sense of belonging experienced by teachers. In Table 4, the PAF solution is compared to the Principal Component Analysis solution as a means to evaluate the difference between common factor analysis and component analysis in this data. The amount of variance in an item that can be explained by the factor is displayed in a *factor pattern matrix*. The columns of the factor loading matrix represent the factor (component) and the rows display each item or variable.

Table 4. *Example of Initial Extraction for Five Items in the TSSB*

ITEM	Principal Axis Factoring					Principal Component Analysis				
	Factors					Components				
1	.498				.402	.510		.338	-.351	.404
2	.650		.346			.659		.383		-.374
3	.726	-.400				.729	-.407			
4	.758				-.360	.760				-.404
5	.585			.365		.595		-.357	.437	

Note. Loadings of less than $|.32|$ were suppressed.

Even though the results produced are similar, it is important to consider the *best* method that most accurately depicts the purpose and needs of the research hypotheses. Because more variance is included, the PCA solution on the right has more items with cross loadings, meaning an item's variance can be explained by multiple factors. In the PAF solution, Item 1 only loads on factors one and five; whereas, in the PCA solution, Item 1 loads on components one, three and four. Pett et al. (2003) suggest that it is best to compare the PCA solution to the PAF solution, and then use the one that makes the most intuitive sense. This further reinforces the responsibility of the researcher to be thoughtful when making processual choices, having both theoretical and conceptual justification for those decisions. Regardless of choice,

attention must be given to the method in which the variance is accounted for and considered at each step in decision making and interpretation.

Determine the Number of Factors to Retain

Using the results from the initial extraction, the researcher must then determine how many factors should be retained in order to best represent the data and the existing relationships. The first factor accounts for the most variance. The amount of variance explained by each subsequent factor continually decreases (Tabachnick & Fidell, 2001). The objective is to choose enough factors to adequately represent the data, while eliminating factors that are not statistically or theoretically relevant (Fabrigar et al., 1999). The body of literature suggests that choosing to retain more factors than are needed is less detrimental to the analysis than eliminating factors that are needed; however, retaining too many factors can deplete the solution erroneously resulting in weak factor loadings (Pett et al., 2003). Additionally, the literature cautions against using a solution with only one or two factors, as it may not provide an accurate representation of the structure (Fava & Velicer, 1992; Pett et al., 2003).

In addition to general recommendations, there are multiple criteria methods to further inform the factor selection using eigenvalues and extracted variance. The eigenvalue is a value associated with each factor describing the amount of variance in the items that can be explained by that factor (Pett et al., 2003). Every factor or component has an eigenvalue. This principle can be observed in mathematics when simplifying multinomial expressions. For example, the binomial $6x + 15y$ can be factored or simplified to $3(2x + 5y)$. A value of “3” is the maximum common amount that can be extracted, or explained. Although the process for determining the eigenvalue based on common variance is significantly more complex, the conceptual principle is the same.

Kaiser Criterion

How much variance does a factor have to explain in order to warrant the retention of a factor (component)? The most commonly used eigenvalue criteria is the *Kaiser Criterion*, which states that factors should be retained if their eigenvalues are greater than

or equal to one (Costello & Osborne, 2005). In a component analysis extraction, such as PCA, where the total variance is accounted for, every item has one unit of variance. If a single component could explain 100% of the variance for all of the items, the eigenvalue for that component would be equal to the total number of items. The reasoning behind the Kaiser Criterion is that a component having an eigenvalue greater than one accounts for more variance than would a single item, thus suggesting merit for combining those items into a factor or component; however, this is only true if each item contributes one unit of variance. Pett et al. (2003) indicate that the Kaiser Criterion should only be used in PCA when the total variance is accounted for in the extraction.

Eigenvalues can be useful if interpreted with an understanding of their conceptual meaning regardless of how much variance was extracted; however, the “cut value” should also reflect this consideration. In common factor analysis extractions, where only common (shared) variance is used, the variance included for each item is less than one. In this case, if a single factor could explain all of the variance in the items, the eigenvalue would still not equal the total number of items. If the Kaiser criterion was used, a factor could account for significant variance but not be retained because the eigenvalue was less than one, resulting in the underextraction of factors.

Example. Using the all items of the TSSBS as an example, the factors were extracted using Principal Axis Factoring, a method that does not include all of the variance in the extraction. Table 5 represents a sample of the variance explained.

Notice that the total initial eigenvalue estimates are different than the total extraction sums. The initial eigenvalue estimates include all of the variance. The PAF solution removes the shared and error variances in extraction, reducing the eigenvalues and percent of variance explained. In a PCA extraction where the total variance is included, no change occurs between the estimated values and the extracted values. For this example, using the Kaiser Criterion, five factors should be retained in order to sufficiently represent the TSSB scale; however, considering that not all the variance is

included, factors six and seven could also be viable linear combinations of the items.

Table 5. Example of Total Variance Explained for a Principal Axis Factoring of the TSSBS

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
	1	13.107	46.810	46.810	12.769	45.603
2	2.057	7.347	54.157	1.737	6.202	51.805
3	1.808	6.459	60.616	1.491	5.326	57.131
4	1.335	4.767	65.383	1.020	3.644	60.774
5	1.230	4.394	69.777	.903	3.226	64.000
6	.963	3.440	73.217			
7	.859	3.069	76.286			

Note. The factors and initial eigenvalues continue until 100% of the variance is accounted for. The full results were not needed for the purposes of this discussion.

The Kaiser Criterion method is often criticized and stated to be used beyond its capabilities resulting in inaccurate determination of factor retention (Costello & Osborne, 2005; Velicer & Jackson, 1990a). Many believe that the Kaiser Criterion tends to overextract factors and cite that it is also capable of underextracting factors as well (Fabrigar et al., 1999; Henson & Roberts, 2006; Schonrock-Adema et al., 2009). Fabrigar et al. (1999) criticizes that the criterion of identifying one as the cut is quite arbitrary. Based on an understanding that the eigenvalue represents the maximum variance that a single linear combination (factor or component) can statistically explain, using the eigenvalues as an indication of value for retaining the factor is conceptually sound; however, the Kaiser Criterion should only be used in component analysis.

Scree Plot

Cattell's Scree Plot is a graphical representation of the factors and their corresponding eigenvalues. The x -axis represents the factors (components) and the eigenvalues are along the y -axis. Because the first component accounts for the greatest amount of variance, it has the highest eigenvalue. The eigenvalues continually decrease resulting in a picture that is often called the "elbow" shape. The scree plot cutoff is quite subjective, requiring that the number of factors be

limited to those occurring before the bend in the elbow (Fabrigar et al., 1999).

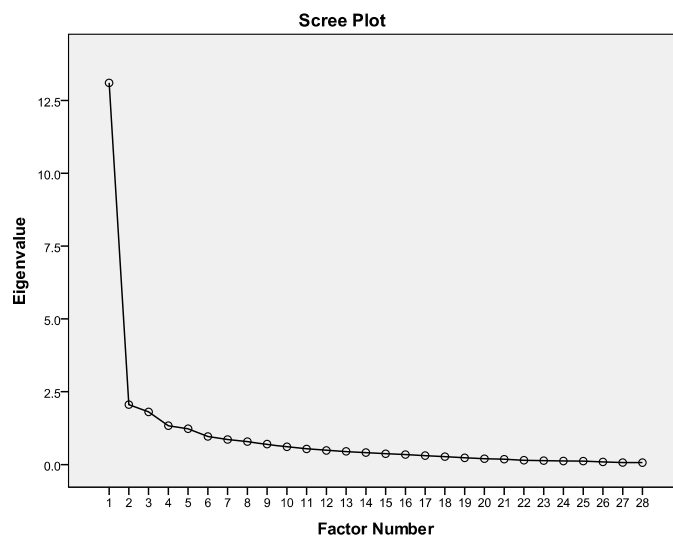


Figure 2. SPSS Scree Plot. This figure demonstrates the scree plot of the eigenvalues and factors from the TSSB extraction seen in Table 5.

This subjectivity is apparent when examining the scree plot of the eigenvalues of the TSSB extraction (See Figure 2). Where does the "bend" occur? Should the cutoff be at the third factor, the fourth, or even the sixth? The difficulty identifying the precise cut point most often leads to overextraction of factors (Henson & Roberts, 2006).

Variance Extracted

A third selection method based on similar conceptual structure is to retain the number of factors that account for a certain percent of variance extracted. The literature varies on how much variance should be explained before the number of factors is sufficient. The majority suggest that 75 – 90% of the variance should be accounted for (Garson, 2010; Pett et al., 2003); however, some indicate as little as 50% of the variance explained is acceptable. As with any criteria method solely depending on variance, this seemingly broad standard must be viewed in relation to the foundational differences between extraction methods. The amount of variance that was included for extraction must be considered when interpreting the value of percent of variance extracted. Component analysis includes more variance to be explained,

suggesting that higher percentages of explained variance are expected than would be required when only common variance is included.

Practical Implications

It is best to evaluate the initial extraction with multiple criterion methods and by comparing the factors suggested to retain (Costello & Osborne, 2005; Schonrock-Adema et al., 2009). Using the eigenvalues, the scree plot, and the percent of variance extracted, the TSSB example could require the retention of between two and eight factors (see Table 5). It is acceptable practice to vary the number of factors retained and compare the solutions. Ultimately, the decision of how many factors to retain should be made based on comprehensibility and interpretability in the context of the research (Suhr, 2006). Are the factors represented by multiple variables that share a conceptual meaning? This evaluation is more relevant after the matrix has been rotated; however, it should be considered when retaining factors as well.

FACTOR ROTATION

Factor rotation is readily accepted as a sequential step when conducting a factor analysis. As discussed previously, the factors, or components, and the factor loadings of each variable are linear combinations of these relationships. The mathematical purpose of factor analysis is to summarize the relationships among variables and the factors. These linear combinations do not have a single, unique solution (Fabrigar et al., 1999). There exist an infinite number of rotations (alternative solutions) that all explain the same amount of variance (DeCoster, 1998; Tabachnick & Fidell, 2001). After the number of factors to include has been determined, all other factors are discarded. The items are factored again, forced into a specified number of factors. That solution is then rotated. This is called factor rotation.

Is it necessary to rotate the initial factor pattern matrix in order to achieve an interpretable and meaningful solution? The literature frequently suggests that rotating the initial factor solution is critical for interpretation of the factors and indicator variables. This stance is presented quite consistently, and is widely followed without much explanation:

- Tabachnick and Fidell (2001) state that “none of the extraction techniques routinely provide an interpretable solution without rotation” (p. 601).
- Fabrigar, Wegener, MacCallum, & Strahan (1999) provide that it is “usually necessary for a researcher to select a method for rotating the initial factor analytic solution to a final solution that can be more readily interpreted” (p. 273).
- Child (1990) explains, “Most factor analysts agree that direct solutions are not sufficient. Adjustment to the frames of reference by rotation methods improves the interpretation of factor loadings by reducing some of the ambiguities which accompany the preliminary analysis” (as cited in Suhr, 2006, p. 3)

These comments are representative examples of the information found within the literature. The use of words that allow for exceptions, such as “routinely” in Tabachnick and Fidell (2001) and “usually” in Fabrigar et al. (1999) leave question as to whether is it ever appropriate to use only the initial factor pattern matrix. In his 1947 work, Thurstone (as cited in Fabrigar et al., 1999) suggested that the most easily interpretable solution is the “simple structure” solution. He also indicated that, because there are an infinite number of solutions, the component matrix should be rotated in order to produce a solution with this simple structure.

By understanding how the variance is removed, the initial solution can be interpreted and examined as a series of linear combinations. However, the initial solution may not be the most useful. Rotations create a statistically comparable solution that is usually more meaningful and easy to interpret.

Simple Structure

Simple structure is achieved when each factor is represented by several items that each load strongly on that factor only (Pett et al., 2003; Tabachnick & Fidell, 2001). Practically, “several items” is generally considered to be at least three to five items with strong loadings (Guadagnoli & Velicer, 1988). An item is considered to be a good identifier of the factor if the loading is .70 or higher and does not significantly cross load on another factor greater than .40 (Garson, 2010).

These guidelines vary slightly within the literature. Tabachnick & Fidell (2001) suggest that the secondary loading (or cross-loading) should be no greater than .32. Costello and Osborne (2005) suggest that a loading of .50 is enough to be considered “strong,” while Guadagnoli and Velicer (1988) state that the loading should be .60 or greater. Generally, a communality (loading) of .70 or greater is ideal because that suggests that approximately 50% of the variance of that item is accounted for by the factor.

Rotations are very similar to many other more familiar mathematical concepts used more frequently in basic math and algebra. An infinite number of numerical combinations could represent the ratio 2:3, such as 4:6, 10:15, or even 160:240, without altering the meaning of the relationship. In algebra and trigonometry, this multiplicative property is often used to make the operation easier to solve. For example, the equation $-1/3x - 2/5y = -7/15$ may be solvable, but multiplying the equation by -15 will simplify the equation without altering the values: $-15(-1/3x - 2/5y = -7/15) = 5x + 6y = 7$. Although it is much more intricate, the rotation of the initial factor solution is grounded in the same mathematical properties. Herein, the certain rotational techniques can serve to create a more interpretable solution without altering the structural relationships. Although the initial solution is capable of being interpreted, the factors and communalities are more easily identified through the use of rotational methods to reach a simple structure solution.

Orthogonal and Oblique Rotations

There are two main types of rotational methods: orthogonal and oblique. Orthogonal rotations (varimax, quartimax, and equimax) are appropriate when the purpose for the factor analysis is to generate factor scores (PCA) or when the theoretical hypotheses concern uncorrelated dimensions (Loo, 1979). Of the orthogonal types of rotations, varimax is generally regarded as best and is most widely used (Fabrigar et al., 1999; Loo, 1979). Loo (1979) cautions that orthogonal rotations are not always theoretically appropriate and do “not reflect interrelationships that probably exist in much clinical data, such relationships

might be better described and interpreted in terms of oblique solutions” (p. 763).

Oblique rotations account for the relationships between the factors, which often is more appropriate within social science research. Fabrigar et al. (1999) emphasize that oblique rotations can be used even when the factors are not significantly correlated. If the factors are not correlated, then the rotation will provide estimates of the factor correlations that are close to zero. Oblique rotational methods include direct oblimin, promax, orthoblique and procrustes. There is not a single best method recommended for oblique rotations and the method choice often depends on the options available through the software used (DeCoster, 1998; Fabrigar et al., 1999).

Example. To demonstrate the differences in an orthogonal and oblique rotation, the TSSB solution was rotated using an orthogonal rotation and an oblique rotation (See Table 6). The values in the orthogonal factor matrix on the left represent maximized relationships of each item with the factor. Each relationship is assessed independently. Conceptually interpreted, in the orthogonal solution, factor one can explain .391 of the variance associated with the responses in item 2, and factor two is able to explain .497 of the variance, suggesting that factor two is more representative of the item than factor one.

The oblique rotation accounts for relationship between factors before determining an item’s relationship to the factor. This solution is presented using two matrices: the pattern matrix and the structure matrix. The pattern matrix values reflect the relationships between the item and the factor when the variance of the other factors are removed. In the oblique pattern matrix in Table 6, once the relationship between the factors is removed, factor two can additionally account for .429 of the variance associated with the responses for item 2. The factor structure does not factor out influences of other factors and provides the correlations between item and factor without controlling for shared variance. The structure matrix demonstrates that all of the items are related to both factors.

Table 6. *Example of Rotational Comparison of 10 Items in the TSSBS*

Rotation Matrix	Orthogonal: Varimax		Oblique: Direct Oblimin			
	Rotated Factor Matrix		Pattern Matrix		Structure Matrix	
	Factor		Factor		Factor	
ITEM	1	2	1	2	1	2
1		.571		.657	.367	.587
2	.391	.497		.429	.556	.607
3		.761		.818	.583	.820
4	.350	.688		.699	.593	.768
5		.731		.776	.578	.796
6	.395	.660		.637	.624	.759
7	.417	.404	.336		.543	.530
8	.443	.374	.386		.555	.512
9	.844		.960		.888	.580
10	.740		.838		.782	.515

Note. Maximum Likelihood Estimation used for Initial Extraction

The orthogonal rotation in Table 6 has more cross-loadings than the oblique solution. The oblique rotation controls for the shared variance between the factors. By first acknowledging that most all of the items and factors are related, as displayed in the structure matrix, oblique rotations allows more apparent differences to emerge in the pattern matrix. Because oblique solutions can incorporate the relationships between the factors, they usually fit the data better than the orthogonal solution (Henson & Roberts, 2006). In this case where the factors are strongly correlated ($r = .709$), the oblique solution is much more representative of the theoretical relationships.

Determining whether an orthogonal or oblique rotation is most appropriate seems to be controversial simply for the fact that much of the literature supports use of oblique rotations, yet orthogonal rotations are still the most commonly used and reported in studies using factor analysis (Costello & Osborne, 2005). Darlington (n.d.) suggests that this may be a result of more complex interpretation of the oblique solutions. Tabachnick and Fidell (2001) recommend using orthogonal rotations if the correlations between factors are low; however, unless strong theoretical foundation

exists to suggest the factors are not correlated, allowances should be made enabling the true relationships among factors to be reflected in the solution. Fabrigar et al. (1999) suggest that if the relationship among factors is unknown an oblique rotation should be used first. Then, if the correlations are low, opt to use an orthogonal rotation. Regardless of methodological choice, theoretical and mathematical limitations and meaning must be acknowledged. Just as decisions for extracting initial factors and determining the number of factors to retain, rotational type should be made based on theoretical purpose of the research. If the factors are conceptually independent, then orthogonal rotation is acceptable; however, oblique rotations are generally more appropriate for social science research where the factors are usually related.

INTERPRETING THE FACTOR SOLUTION

Because factor analytic processes are iterative, much of the evaluation has occurred throughout each subsequent step. The rotated factor solution is useful to examine and further refine the factors. Mathematical and conceptual examination is required for accurate interpretation of both the items and the factors. The items should possess a significant loading, indicating a statistically valued contribution; however, an item's conceptual significance should be examined before an item is removed from the set. Theoretical knowledge is more relevant than a statistical measure. If an item is not significantly correlated to any of the factors (generally considered to be less than .30) and does not provide a conceptually vital dimension to the measure, the item should be removed. Additionally, a complex variable, or a variable that loads on more than one factor, should be removed if the cross-loading is greater than .40 (Schonrock-Adema et al., 2009). Once the weak items have been removed, the data should be factored again without the presence of that item for a more refined solution (Pett et al., 2003). Interpretation of the factor also requires that each factor be sufficiently identified. This means that a factor contain at least three to five items with significant loadings in order to be considered a stable and solid factor (Costello & Osborne, 2005). More importantly, the items and the factors should make sense conceptually.

FINAL THOUGHTS

Every step of the process in a factor analysis requires the researcher to be firmly grounded in contextual theory and fundamental understanding of factor analysis methodology. The greatest difference in methods centers on how the technique accounts for variance and relationships between the factors. Regardless of choice, decisions should be supported by strong theoretical and mathematical justification, providing credibility to the final outcome.

While factor analysis contains many variations, the process may be summarized as a series of mathematical iterations designed to create linear combinations in order to explain the data. Each iteration reveals new information, further expanding the researcher's understanding of the relationships. Based on the new perspective, the structure is refined until the solution reached is parsimonious, mathematically sound, and theoretically grounded. The factor structure should continue to be tested and refined to more fully understand the relationships in different contexts; however, just because more testing of the structure is desired, does not mean that the current solution is not useful. When evaluated in light of the strength of methodological purpose and use, the solution can be very meaningful and contribute significantly to relevant research. It is important to remember that factor analysis is a mathematical process. While the matrix relationships are not elementary, they are still simply an indication of how the responses provided for each item relate to others. Factor analysis should always be interpreted in light of theory and common sense.

REFERENCES

- Byrne, B. M. (2001). *Structural equation modeling with AMOS - Basic concepts, applications, and programming*. LEA, ISBN 0-8058-4104-0
- 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, 1-9. Retrieved March 5, 2013 from <http://pareonline.net/getvn.asp?v=10&n=7>
- Darlington, R. (n.d.). *Factor Analysis*. Retrieved March 5, 2013, from <http://psych.cornell.edu/Darlington/factor.htm>
- DeCoster, J. (1998). *Overview of factor analysis*. Retrieved March 5, 2013, from <http://www.stat-help.com/notes.html>
- Fabrigar, L.R., Wegener, D.T., MacCallum, R.C. & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272-299.
- Fava, J.L. & Velicer, W.F. (1992). The effects of overextraction on factor and component analysis. *Multivariate Behavioral Research*, 27, 387-415.
- Garson, D. (2010, February 11). *Factor analysis*. Retrieved March 5, 2013, from <http://faculty.chass.ncsu.edu/garson/PA765/factor.htm>
- Gorsuch, R. (1990). Common factor analysis versus component analysis: some well and little known facts. *Multivariate Behavioral Research*, 25, 33-39.
- Guadagnoli, E. & Velicer, W.F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, 103, 265-275.
- Henson, R.K. & Roberts, J.K. (2006). Use of exploratory factor analysis in published research: common errors and some comment on improved practice. *Educational and Psychological Measurement*, 66, 393-416.
- Hogarty, K.Y., Hines, C.V., Kromrey, J.D., Ferron, J.M & Mumford, K.R. (2005). The quality factor solutions in exploratory factor analysis: the influence of sample size, communality, and overdetermination. *Educational and Psychological Measurement*, 65, 202-226.
- Loo, R. (1979). The orthogonal rotation of factors in clinical research: a critical note. *Journal of Clinical Psychology*, 35, 762-765.
- MacCallum, R.C., Widaman, K.F., Preacher, K.J & Hong, S. (2001). Sample size in factor analysis: the role of model error. *Multivariate Behavioral Research*, 36, 611-637.
- Osborne, J.W. & Costello, A.B. (2004). Sample size and subject to item ratio in principal components analysis. *Practical Assessment, Research & Evaluation*, 9, 1-15. Retrieved March 5, 2013 from <http://pareonline.net/getvn.asp?v=9&n=11>
- Pett, M., Lackey, N. & Sullivan, J. (2003). *Making sense of factor analysis*. Thousand Oaks: Sage Publications, Inc.
- Schonrock-Adema, J., Heijne-Penninga, M., Van Hell, E.A. & Cohen-Schotanus, J. (2009). Necessary steps in factor analysis: enhancing validation studies of educational instruments. *Medical Teacher*, 31, e226-e232.
- Suhr, D. (2006). Exploratory or Confirmatory Factor Analysis. *SAS Users Group International Conference* (pp. 1 - 17). Cary: SAS Institute, Inc.

Tabachnick, B. & Fidell, L. (2001). *Using multivariate statistics*. Needham Heights: Allyn & Bacon.

Velicer, W.F. & Fava, J.L. (1990). An evaluation of the effects of variable sampling on component, image, and factor analysis. *Multivariate Behavioral Research, 25*, 193-209.

Velicer, W.F. & Jackson, D.N. (1990a). Component analysis versus common factor analysis: some further observations. *Multivariate Behavioral Research, 25*, 97-114.

Velicer, W.F. & Jackson, D.N. (1990b). Component analysis versus common factor analysis: some issues in selecting an appropriate procedure. *Multivariate Behavioral Research, 25*, 1-28.

Zhao, N. (2009, March 23). *The minimum sample size in factor analysis*. Retrieved March 5, 2013, from <http://www.encorewiki.org/display/~nzhao/The+Minimum+Sample+Size+in+Factor+Analysis>

Citation:

Beavers, Amy S., Lounsbury, John W., Richards, Jennifer K., Huck, Schuyler W., Skolits, Gary J. & Esquivel, Shelley L. (2013). Practical Considerations for Using Exploratory Factor Analysis in Educational Research. *Practical Assessment, Research & Evaluation, 18*(6). Available online: <http://pareonline.net/getvn.asp?v=18&n=6>

Corresponding Author:

Amy Beavers
The University of Tennessee
2509 River Drive
Knoxville, TN 37996-4539
Email: abeavers [at] utk.edu