Oblique versus Orthogonal Rotation in Exploratory Factor Analysis

Kimani Chege Gabriel

Moi University - Kenya

Abstract: - **Exploratory factor analysis is widely applied by psychometricians and other behavioural science researchers in complex studies involving numerous variables and factors. A variable might be related to more than one factor and therefore a psychometrician should consider this possibility when deciding about how many factors will be considered when analysing the data. The rotation of factors is used to get more interpretable and simplified solutions from the factor extraction results by maximising high item loadings and minimising low item loadings. Rotation helps to deal with data sets where there are large numbers of observed variables that are thought to reflect a smaller number of underlying/latent variables. It is one of the most commonly used inter-dependency techniques and is used when the relevant set of variables shows a systematic inter-dependence and the objective is to find out the latent factors that create a commonality. However, practitioners and researchers often make questionable decisions when conducting these analyses, especially in the choice of the rotation method from among the two; orthogonal and oblique. This paper therefore sought to examine exploratory factor analysis and its relevant protocol, discusses the two factor rotation methods, the operational differences and the parsimoneity of outputs, eigenvectors which are usually at the center of rotation as well as a guide for practitioners in deciding between orthogonal and oblique rotation. Finally the paper gives a parting short in the conclusion section. It is hoped that the paper will present useful insights for practitioners' use.**

I. INTRODUCTION

Factor analysis is a psychometric technique that is used to reduce a large number of variables into fewer numbers of factors. Technically speaking, the technique extracts maximum common variance from all variables and puts them into a common score. In simple terms, factor analysis is a way to take a mass of data and shrinking it to a smaller data set that is more manageable and more understandable. In its internal functioning, it seeks to find hidden patterns, show how those patterns overlap and show what characteristics are seen in multiple patterns.

This paper focused on Exploratory Factor Analysis (EFA) which is a type of factor analysis that is used to find the underlying structure of a large set of variables. It reduces data to a much smaller set of summary variables.

According to Yong and Sean (2013) EFA is an analysis of exploratory type that is used to identify the complex interrelationships among the variables, and group these variables as part of unified concepts. This method helps the psychometrician to draw the main dimensions of the area of interest to derive a theory or a model from the reasonably large set of variables (it is not based on any prior theory).The groups formed from interrelated variables are called factors.

Exploratory Factor Analysis is performed without any prior idea of, which factors indeed subsist and which variables loads to each group formed. In essence, the psychometrician uses conventional procedure and rules to arrange and load the variables on factors and to fix the number of factors. Therefore in its operational mechanics, the EFA explores the data and provides the psychometrician with information about how many factors are needed to best represent it. The correlation between the variables and factors known as factor loading gives the nature of a particular factor.

In general, an Exploratory Factor Analysis prepares the variables to be used for cleaner structural equation modeling. An EFA should always be conducted for new datasets. The functional advantage of an EFA over a CFA (confirmatory Factor Analysis) is that no *a priori* theory about which items belong to which constructs is applied. This makes it possible for EFA to spot problematic variables much more easily than the CFA.

The only catch is the specificity of the nature of data on which EFA can be performed, that is all data must be non-nominal which theoretically belong to *reflective latent* factors**.** The exclusive implication is that formative measures, nominal or categorical data such as gender, marital status etc should not be included in the EFA designs. Again, objective (rather than perceptual) variables should be excluded because they rarely belong to reflective latent factors.

II. EXPLORATORY FACTOR ANALYSIS (EFA) PROTOCOL

An Exploratory Factor Analysis is executed through a series of steps in building clear decision pathway. Of course the logical start point is the definition of the psychometric problem. The figure below shows the systematic steps followed in EFA by psychometricians.

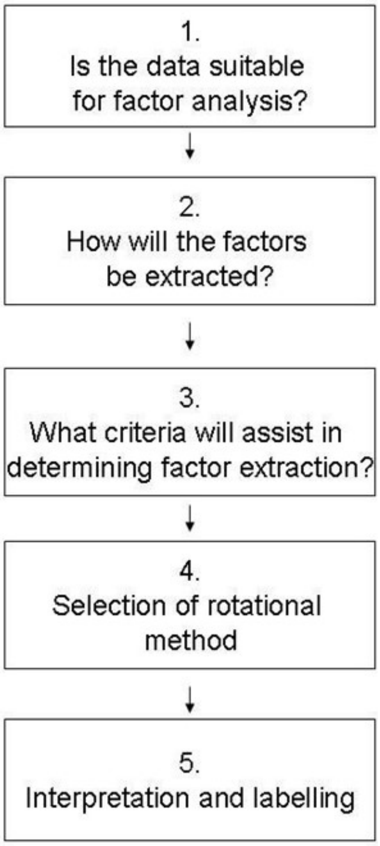

Figure1: Exploratory Factor Analysis (EFA) protocol

III. FACTOR ROTATIONS

During factor analysis, different algorithms or methods are used to achieve the same broad goal- simplification of the factor structure. This involves carrying out rotations. Rotation methods fall into two broad categories: orthogonal and oblique (referring to the angle maintained between the X and Y axes).

Oblique methods allow the factors to correlate (that is, they allow the X and Y axes to assume a different angle than 90^o with smaller angles being manifested more). The diagram below illustrates oblique rotation and the angles of rotation (Rummel, 1970).

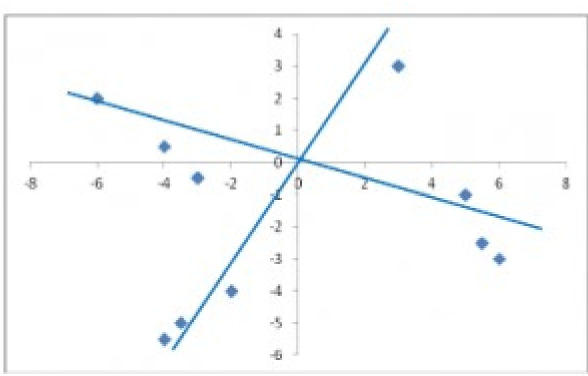

Figure 2: An illustration of oblique rotation

Clearly, the angle between the two factors is now smaller than 90 degrees, meaning the factors are now correlated. In this example, an oblique rotation accommodates the data better than an orthogonal rotation. Consequently, the two axes of the two factors are probably closer together than an orthogonal rotation can make them.

According to Gorsuch (1983), there are over 15 different oblique rotation methods which all assume that the factors are correlated. Of the 15 oblique methods, the most common ones are direct oblimin and promax.

In the case of orthogonal rotations produce factors that are uncorrelated (that is, maintain a 90^o angle between axes). The diagram below illustrates orthogonal rotation and the angles of rotation.

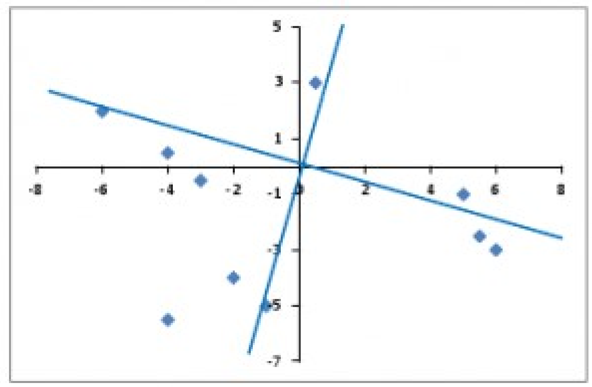

Figure 3: An illustration of orthogonal rotation

Here is a display of the oblique rotation of the axes for our new example, in which the factors are correlated with each other: hence the strong recommendation for adoption of oblique rotation in behavioral sciences.Orthogonal rotation methods assume that the factors in the analysis are uncorrelated. Gorsuch (1983) listed four different orthogonal methods: equamax, orthomax, quartimax, and varimax orthogonal rotation.

Many psychometricians and researchers have (as a matter of tradition) been guided to orthogonal rotation because (the argument went) uncorrelated factors are more easily interpretable.

IV. OPERATIONAL DIFFERENCES AND THE PARSIMONEITY OF OUTPUTS

Assuming there are 5 factors for 100 factored entities (e.g., variables)which are extracted and orthogonally rotated, only 500 factor pattern/structure coefficients are estimated (the5 x 5 factor correlation matrix is not estimated, since it is constrained to have l's on the diagonal and 0's everywhere else).If the same Exploratory Factor Analysis (EFA) factors are rotated obliquely, we shall have 1,010 coefficients(500 factor pattern coefficients, plus 500 factor structure coefficients, plus 10 factor correlation coefficients (the 10 non-redundant off-diagonal entries in the 5 x 5 factor correlation matrix) are estimated. We can argue, however, that only 510 coefficients are estimated in this case, since with either the 10 unique factor correlation coefficients, and either the 500pattern or the 500 structure coefficients, the remaining 500pattern or structure coefficients are fully determined. So, in essence, an oblique factor solution inherently tends to be less parsimonious.

According to Osborne (2015) the fact that more parameters are estimated in an oblique rotation means that oblique solutions almost always better fit sample data than do orthogonal solutions. However, some of this fit involves "over-fitting" sampling error variance. This means that orthogonal solutions, though they may tend to somewhat fit sample data less well, are generally more replicable in future samples, since orthogonal solutions capitalize on less sampling error. Usually, at least in Exploratory Factor Analysis(EFA), somewhat poorer fit is deemed an acceptable tradeoff for better solution replicability (i.e., factor invariance). As the degree of correlation between the factors decreases, both orthogonal and oblique solutions will tend to provide increasingly similar results. Given that oblique solutions are less parsimonious and therefore less replicable, an oblique rotation would therefore only be employed when the benefits of simpler, more interpretable structure outweigh the costs of less replicability (i.e., when the orthogonal factors are not readily interpretable, and the oblique factors are fairly highly correlated but more interpretable)

V. EIGENVECTORS AT THE CENTER OF ROTATION

Eigenvectors are a special set of vectors associated with a linear system of equations (i.e., a matrix equation) that are sometimes also known as characteristic vectors, proper vectors, or latent vectors. . In factor analysis, eigenvalues are used to condense the variance in a correlation matrix. "The factor with the largest eigenvalue has the most variance and so on, down to factors with small or negative eigenvalues that are usually omitted from solutions" (Tabachnick and Fidell, 2007). From the analyst's perspective, only variables with eigenvalues of 1.00 or higher are traditionally considered worth analyzing. However, the other three approaches explained below can provide overriding reasons for selecting other numbers of factors (Gorsuch, 1983).

In matrix algebra, under certain conditions, matrices can be diagonalized. Matrices are often diagonalized in multivariate analyses. In that process, eigenvalues are used to consolidate the variance. The determination of the eigenvectors and eigenvalues of a system is extremely important in psychometrics, physics and engineering, where it is equivalent to matrix diagonalization and arises in such common applications as stability analysis, the physics of rotating bodies, and small oscillations of vibrating systems, to name only a few. Each eigenvector is paired with a corresponding so-called eigenvalue. Mathematically, two different kinds of eigenvectors need to be distinguished: left eigenvectors and right eigenvectors. However, for many problems in psychometrics, it is sufficient to consider only right eigenvectors. The term 'eigenvector' used without qualification in such applications can therefore be understood to refer to a right eigenvector.

The decomposition of a square matrix A into eigenvalues and eigenvectors is known in this work as Eigen decomposition, and the fact that this decomposition is always possible as long as the matrix consisting of the eigenvectors of A is square is known as the Eigen decomposition theorem.

Define a right eigenvector as a column vector $\mathbf{X}_R$ satisfying

$$\mathsf{A}\,\mathbf{X}_R = \lambda_R\,\mathbf{X}_R,$$

where A is a matrix, so

$$(\mathsf{A} - \lambda_R\,\mathsf{I})\,\mathbf{X}_R = \mathbf{0},$$

which means the right eigenvalues must have zero determinant, i.e.,

$$\det(\mathsf{A} - \lambda_R\,\mathsf{I}) = 0.$$

Similarly, define a left eigenvector as a row vector $\mathbf{X}_L$ satisfying

$$\mathbf{X}_L\,\mathsf{A} = \lambda_L\,\mathbf{X}_L.$$

Taking the transpose of each side gives

$$(\mathbf{X}_L\,\mathsf{A})^{\mathrm{T}} = \lambda_L\,\mathbf{X}_L^{\mathrm{T}},$$

which can be rewritten as

$$\mathsf{A}^{\mathrm{T}}\,\mathbf{X}_L^{\mathrm{T}} = \lambda_L\,\mathbf{X}_L^{\mathrm{T}}.$$

Rearrange again to obtain

$$\left(\mathsf{A}^{\mathrm{T}} - \lambda_L\,\mathsf{I}\right)\mathbf{X}_L^{\mathrm{T}} = \mathbf{0},$$

which means

$$\det\left(\mathsf{A}^{\mathrm{T}} - \lambda_L\,\mathsf{I}\right) = 0.$$

Rewriting gives

$$\begin{aligned} 0 = \det\left(\mathsf{A}^{\mathrm{T}} - \lambda_L\,\mathsf{I}\right) &= \det\left(\mathsf{A}^{\mathrm{T}} - \lambda_L\,\mathsf{I}^{\mathrm{T}}\right) \\ &= \det(\mathsf{A} - \lambda_L\,\mathsf{I})^{\mathrm{T}} \\ &= \det(\mathsf{A} - \lambda_L\,\mathsf{I}), \end{aligned}$$

where the last step follows from the identity

$$\det(\mathsf{A}) = \det\left(\mathsf{A}^{\mathrm{T}}\right).$$

Since R=A'A = VD^2V', then RV = D^2V'. So (simplifying the notation) an eigenvector **v** of a matrix **R** is any vector that satisfies this equation: R**v** = λ**v**. R is a square (normally symmetric) matrix, **v** is the eigenvector, λ is the eigenvalue associated with that eigenvector. The eigenvector is a vector which, if pre-multiplied by a matrix, gets you the vector back again (a property called idempotency).

Suppose X is a case-by-variable matrix (e.g., the columns of X give responses for each case on a series of attitude questions such as 'Should man eat where he *worketh*?' or 'Should citizens be allowed to own guns?') and R is the matrix of correlations among the variables of X. Then the eigenvectors of R (multiplied by their eigenvalues) are known as the factor loadings and are literally the correlations of the each variable in X with an underlying factor or principal component.

VI. DECIDING BETWEEN ORTHOGONAL AND OBLIQUE ROTATION

But how should I choose which one to use? The decision to rotate orthogonally or obliquely is often difficult for psychometricians and researchers and is largely based on the goal of the analysis. If the goal of the analysis is to generate results that best fit the data, then oblique rotation seems to be the logical choice.

Conversely, if the replicability of the factor analytic results is the primary focus of the analysis, then an orthogonal rotation might be preferable since results from orthogonal rotation tend to be more parsimonious.

Tabachnick and Fiddell (2007) argue that "Perhaps the best way to decide between orthogonal and oblique rotation is to request oblique rotation [e.g., direct oblimin or promax from SPSS] with the desired number of factors (Brown, 2009) and look at the correlations among factors...if factor correlations are not driven by the data, the solution remains nearly orthogonal. Look at the factor correlation matrix for correlations around .32 and above. If correlations exceed 0.32, then there is 10% (or more) overlap in variance among factors, enough variance to warrant oblique rotation unless there are compelling reasons for orthogonal rotation.

One way we know if we have selected an adequate rotation method is if the results achieve simple structure . Bryant and Yarnold (1995) define simple structure as:a condition in which variables load at near 1 (in absolute value) or at near 0 on an eigenvector (factor). Variables that load near 1 are clearly important in the interpretation of the factor, and variables that load near 0 are clearly unimportant. Simple structure thus simplifies the task of interpreting the factors.

Using logic like that in the preceding quote, Thurstone (1947) first proposed and argued for five criteria that needed to be met for simple structure to be achieved:

1. Each variable should produce at least one zero loading on some factor.
2. Each factor should have at least as many zero loadings as there are factors.
3. Each pair of factors should have variables with significant loadings on one and zero loadings on the other.
4. Each pair of factors should have a large proportion of zero loadings on both factors (if there are say four or more factors total).
5. Each pair of factors should have only a few complex variables.

In order to understand Thurstone's five criteria, you will need to understand a few more concepts:

1. What's a zero loading? One rule of thumb (after Gorsuch, 1983) is that zero loadings includes any that fall between -.10 and +.10.
2. What's a significant loading? With a sample size of say 100 participants, loadings of .30 or higher can be considered significant, or at least salient (Kline, 2002). With much larger samples, even smaller loadings could be considered salient, but in language research, researchers typically take note of loadings of .30 or higher.
3. And what are complex variables? Simply put, these are variables with loadings of .30 or higher on more than one factor.

Two elements typically whether orthogonal and obliquerotation strategies will generate similar or identical results (i) the degree of correlation between the factors, and (ii) the factor to variable ratio. When the ratio of variables to factors is small, both rotation strategies will produce similar results, as simple structure will tend to be the same regardless of the type of rotation.

Further, if the correlation between the factors is small (that is, factor correlation coefficients closer to zero), then orthogonal and oblique rotation strategies will generally produce similar, if not identical, results. In a nutshell, choosing a rotation strategy to employ in factor analysis is not an arbitrary decision; rather, the appropriate choice of either an orthogonal or oblique rotation largely depends on the goals of the analysis (best fit to data or replicability of the analysis), the factor to variable ratio, and the degree of correlation between the factors*(*Warne and Larsen, 2014).

There is also an argument in favor of orthogonal rotation as the mathematics are simpler, and that made a significant difference during much of the 20th century when EFA was performed by hand calculations or much more limited computing power. Orthogonal rotations are generally the default setting in most statistical computing packages.

There does not seem to be a compelling reason for modern researchers to default to orthogonal rotations. In the social sciences (and many other sciences, such as biomedical sciences) we generally expect some correlation among factors, since behavior is rarely partitioned into neatly packaged units that function independently of one another.

Therefore using orthogonal rotation potentially results in a less useful solution where factors are correlated. Remembering that Exploratory Factor Analysis (EFA) is an exploratory technique (not a confirmatory technique), we should be looking for the clearest solution possible. Besides, there does not appear to be a drawback to using oblique rotation even if the factors are truly uncorrelated.

Oblique rotations do not force factors to be correlated, and so in that instance, the factors would be allowed to assume a correlation of zero, and the solution would be the same as that of an orthogonal rotation.

Oblique rotation output is only slightly more complex than orthogonal rotation output, but should yield either identical or superior results to that of orthogonal rotations. In SPSS output the rotated factor matrix is interpreted after orthogonal rotation; the rotated factor matrix represents both the loadings and the correlations between the variables and factors.

In contrast, when using oblique rotation the pattern matrix is examined for factor/item loadings and the factor correlation matrix reveals any correlation between the factors. The pattern matrix holds the loadings (which are of most interest), and each row of the pattern matrix can be thought of as a regression equation where the standardized observed variable is expressed as a function of the factors, with loadings as the regression coefficients. The structure matrix holds the correlations between the variables and the factors, which are generally of less interest in exploratory applications (Gorusch, 1983)

There are a variety of choices in each category. Varimax rotation is by far the most orthogonal rotation, likely because it is the default in many software packages, but also because it was developed as an incremental improvement upon prior algorithms quartimax, and equamax. There is no widely preferred method of oblique rotation; all tend to produce similar results (Fabrigar et al., 1999), and it seems generally fine to use the default settings in software packages. Common oblique rotations you will see include: direct oblimin, quartimin, and Promax

The mathematical algorithms for each rotation are different, and beyond the scope of this brief technical note. Note that for all rotations, the goal is the same: simplicity and clarity of factor loadings.

VII. CONCLUSION

In conclusion, the study recommends that practitioners will usually find it useful to try one oblique rotation method (for example, direct oblimin or promax, while examining the factor correlation matrix for values over ±0.32, using the criterion explained in Tabachnick & Fiddell, 2007) and one orthogonal rotation method (for example, the ever-popular varimax rotation). Also consider whether there are any theoretical reasons why an orthogonal method might be preferable to an oblique method or vice versa. Above all, the rotated results should be examined for simple structure, at least following Kline's (2002) relatively flexible definition: "...that each factor should have a few high loadings with the rest of the loadings being zero or close to zero..." (that is, less than ±0.10 after Gorsuch, 1983). The view of the study is that in the modern era of high-power computing, orthogonal rotations are probably not a best practice, as oblique rotations can accurately model uncorrelated and correlated factors, whereas orthogonal rotations cannot handle correlated factors as effectively. Thus, there is little cost to using oblique rotations regardless of the underlying relatedness of the factors.

REFERENCES

[1]. Brown, J. D. (2009). 'Statistics Corner - questions and answers about language testing statistics: Principal components analysis and exploratory factor analysis—Definitions, differences, and choices.' *In Shiken: JALT Testing & Evaluation SIG Newsletter*, 13(1), 26-30.

[2]. Bryant, F. B., & Yarnold, P. R. (1995). 'Principal-components analysis and confirmatory factor analysis.' In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99-136). Washington, DC: American Psychological Association.

[3]. Fabrigar, L. R, Duane, T. W., Robert, C. M. & Erin, J. S. (1999). Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *In Psychological Methods*,4(3), 272-299

[4]. Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

[5]. Kline, P. (2002). *An easy guide to factor analysis.* London: Routledge.

[6]. Osborne, J. W. (2015).What is Rotating in Exploratory Factor Analysis*? Practical Assessment, Research & Evaluation*, 20(2).

[7]. Rummel, R.J. (1970). *Applied factor analysis.* Evanston, IL: Northwestern University Press

[8]. Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.

[9]. Thurstone, L. L. (1947*). Multiple factor analysis: A development and expansion of vectors of the mind.* Chicago: University of Chicago.

[10]. Warne, R. T.&Larsen, R. (2014). 'Evaluating a proposed modification of the Guttman rule for determining the number of factors in an exploratory factor analyses. *Psychological Test and Assessment Modeling*. 56: 104–123.

[11]. Yong, A. G. & Sean, P. (2013). 'A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis.' *In Tutorials in Quantitative Methods for Psychology*, 9(2), 79-94.