A Quick Primer on Exploratory Factor Analysis

Exploratory vs. Confirmatory Factor Analysis

Similarities

- Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) are two statistical approaches used to examine the internal reliability of a measure.
- Both are used to investigate the theoretical constructs, or factors, that might be represented by a set of items (i.e., the *factor structure*).
- Either can assume multiple factors are uncorrelated (*orthogonal*) or correlated.
- Both are used to assess the quality of individual items.
- Both can be used for exploratory or confirmatory purposes.

Differences

- With EFA, researchers usually decide on the number of factors by examining output from a principal components analysis (i.e., eigenvalues are used). With CFA, researchers must specify the number of factors *a priori*.
- CFA requires that a particular factor structure be specified, in which the researcher indicates which items load on which factor. EFA allows all items to load on all factors.
- CFA provides a fit of the hypothesized factor structure to the observed data.
- Researchers typically use maximum likelihood to estimate factor loadings with CFA, whereas maximum likelihood is only one of a variety of estimators used with EFA.
- CFA allows researchers to specify correlated measurement residuals, constrain loadings or factor correlations to be equal to one another, perform statistical comparisons of alternative models, test second-order factor models, and statistically compare the factor structure of two or more groups.

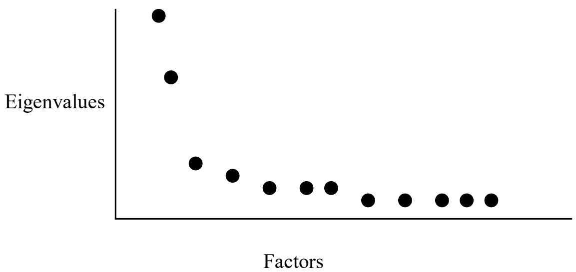
Exploratory Factor Analysis: Purpose

Exploratory factor analysis (EFA) is generally used to discover the factor structure of a measure, examine its internal reliability, and eliminate items. EFA is often recommended when researchers have no hypotheses about the nature of the underlying factor structure of their measure. Exploratory factor analysis has three basic decision points: (1) decide the number of factors, (2) choosing an extraction method, (3) choosing a rotation method.

Exploratory Factor Analysis: Deciding the number of factors

The most common approach to deciding the number of factors is to generate a scree plot, which plots values that are analogous to the proportion of variance accounted for in all of the items by each of the factors estimated in a *principal components analysis* (PCA). PCA is a method of forming a set of weighted composites out of a larger number of variables (often items from a scale). It is a method of data reduction, but not considered a "true" factor analysis, because the item variances are assumed to be fully accounted for by the factors (i.e., no measurement error). PCA is nearly universally used as the first step in an exploratory factor analysis to decide upon the number of factors to extract. The scree plot is a two-dimensional graph with factors on the x-axis and *eigenvalues* on the y-axis. Eigenvalues are produced by the analysis and represent the variance accounted for by each underlying component (often "factor" informally). Eigenvalues are not represented by percentages; they are raw scores that total to the number of items when you start with a correlation matrix (but the sum of item variances if the covariance matrix is used). A 12-item scale will theoretically have 12 possible underlying factors (with a sum of eigenvalues equal to 12 in the correlation matrix case), and each factor will have an eigenvalue that indicates the amount of variation in the items accounted for by each factor. If the first factor has an eigenvalue of 3.0, it accounts for 25% of the variance (3/12=.25). The total of all the eigenvalues will be 12 if there are 12 items, so some factors will have smaller eigenvalues. Eigenvalues are typically arranged in a scree plot in descending order like the following:

Example of a scree plot:



From the scree plot you can see that the first couple of factors account for most of the variance, then the remaining factors all have small eigenvalues. The term “scree” is taken from the word for the rubble at the bottom of a mountain (Cattell, 1966). A researcher might examine this plot and decide there are two underlying factors and the remainder of the factors are just “scree” or error variation. So, this approach to selecting the number of factors involves a certain amount of subjective judgment. A number of studies suggest the scree plot method does pretty well in identifying the correct number of factors (e.g., Hakstian, Rogers, & Cattell, 1982; Tucker, Koopman, & Linn, 1969; but see review by Goretzko et al., 2021), however.

A widely recognized criterion is called the Kaiser-Guttman rule (Kaiser, 1960) and simply states that the number of factors is equal to the number of factors with eigenvalues greater than 1.0. Importantly, this criterion is not consistent across all applications and can lead to over or under extraction of factors (e.g., Cattell & Vogelmann, 1977; Gorsuch, 1983; Zwick & Velicer, 1982; see Preacher & MacCallum, 2003, for additional references). Kaiser-Guttman approach is the most commonly used (Goretzko et al., 2019; Henson & Roberts, 2006), yet the simple scree test is likely better (Velicer, 1976; Velicer et al., 2000). Because of its subjectivity, other more empirical approaches to scree plots have been proposed that use more quantifiable values to find the bend in the plot, including Cattell-Nelson-Gorsuch modified scree test (Gorsuch, 1966; Gorsuch, 1983) and the Zoski-Jurs regression index (1996).

Another widely discussed approach to identifying the number of actors is Horn's (1965) parallel analysis which compares the eigenvalues from the EFA to the eigenvalues obtained from analyzing multiple data sets consisting of random variables, allowing estimation of percentiles for the eigenvalues relative to what is expected by chance (see Dinno, 2009 for a clear explanation of the parallel analysis rationale and method). Although the parallel analysis approach appears to do well at identifying the correct number of factors (e.g., Auerswald & Moshagen, 2019; Fabrigar & Wegener, 2002), it is not as widely available in software packages.

Perhaps the most accurate approach will be one that uses a combination of strategies. Auerswald and Moshagen (2019) recommend sequential chi-square tests (Lawley & Maxwell, 1962) together with either the Hull method (Lorenzo-Seva et al., 2011), the Empirical Kaiser Criterion (Braeken & van Assen, 2017), or Horn’s parallel analysis.

Exploratory Factor Analysis: Factor Extraction

Once the number of factors is decided, the researcher runs another factor analysis to get the loadings for each of the factors. To do this, one has to decide which mathematical solution to use to find the loadings. There are about five basic extraction methods (1) PCA, which is the default in most packages. PCA assumes there is no measurement error and is considered not to be a true exploratory factor analysis; (2) maximum likelihood (a.k.a. canonical factoring); (3) alpha factoring, (4) image factoring, (5) principal axis factoring with iterated communalities (a.k.a. least squares), sometimes referred to as "principal factors".

Without getting into the details of each of these, I think the best evidence supports the use of principal axis factoring and maximum likelihood approaches. I typically use the former. Gorsuch (1989) recommends maximum likelihood if only a few iterations are performed (not usually possible in most packages). Snook and Gorsuch (1989) show that PCA can give poor estimates of the population loadings in small samples. With larger samples, most approaches will have similar results.

The extraction method will produce factor loadings for every item on every extracted factor. Researchers hope their results will show what is called *simple structure*, with most items having a large loading on one factor but small loadings on other factors. The measure is usually modified (i.e., items are eliminated) to achieve simple structure.

Exploratory Factor Analysis: Rotation

Once an initial solution is obtained, the loadings are *rotated*. Rotation is a way of maximizing high loadings and minimizing low loadings so that the simplest possible structure is achieved. There are two basic types of rotation: *orthogonal* and *oblique*. Orthogonal rotation implies that the factors are assumed to be uncorrelated with one another. This is the default setting in all statistical packages but is rarely a logical assumption about factors in the social sciences. Not all researchers using EFA seem to be aware of which rotations methods are orthogonal and oblique or the assumptions that orthogonal rotations imply an assumption of uncorrelated factors (completely unrelated constructs). Oblique rotation derives factor loadings based on the assumption that the factors are correlated, and it is probably reasonable to assume factors are correlated for most measures. Even if factors are uncorrelated empirically, there should be no harm in choosing an oblique rotation—the factors will just be shown to be uncorrelated in the output. Some common orthogonal rotations are: varimax, quartamax, equamax. Some common oblique *rotation*s are: oblimin, promax, direct quartimin

I am not an expert on the advantages and disadvantages of each of these rotation algorithms, and they reportedly produce fairly similar results under most circumstances (although orthogonal and oblique rotations will be rather different from one another). I tend to use promax rotation because it is known to be relatively efficient at achieving simple oblique structure.

References and Further Readings

Auerswald, M., & Moshagen, M. (2019). How to determine the number of factors to retain in exploratory factor analysis: A comparison of extraction methods under realistic conditions. *Psychological Methods, 24*(4), 468-491.
Braeken, J., & van Assen, M. A. (2017). An empirical Kaiser criterion. *Psychological Methods, 22*, 450 – 466.
Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.
Cattell, R. B., & Vogelmann, S. (1977). A comprehensive trial of the scree and KG criteria for determining the number of factors. *Multivariate Behavioral Research, 12*(3), 289-325.
Dinno, A. (2009). Exploring the sensitivity of Horn's parallel analysis to the distributional form of random data. *Multivariate Behavioral Research, 44*(3), 362-388.
Fabrigar, L.R ., & Wegener, D.T. (2012). *Factor analysis*. New York, NY: Oxford University Press.
Goretzko, D., Pham, T. T. H., & Bühner, M. (2021). Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Current Psychology, 40*, 3510-3521.
Gorsuch, R. L. (1983). *Factor analysis (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Gorsuch, R.L. (1990). Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research, 25*, 33-39.
Hakstian, A.R ., Rogers, W.D., & Cattell, R. B. (1982). The behavior of numbers factors ruleswith simulated data. *Multivariate Behavioral Research, 17*, 193–219.
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.
Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141–151.
Kim, J.-O., & Mueller, C.W. (1978). *Introduction to factor analysis: What it is and how to do it.* Newbury Park: Sage.
Kim, J.-O., & Mueller, C.W. (1978). *Factor analysis: Statistical methods and practical issues.* Newbury Park: Sage.
Lawley, D. N., & Maxwell, A. E. (1962). Factor analysis as a statistical method. *Journal of the Royal Statistical Society. Series D (The Statistician), 12*(3), 209-229.
Lorenzo-Seva, U., Timmerman, M. E., & Kiers, H. A. (2011). The Hull method for selecting the number of common factors. *Multivariate Behavioral Research, 46*, 340–364.
Pituch, K. A., & Stevens, J. P. (2016). "Chapter 9 Exploratory Factor Analysis." *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS*. New York: Routledge
Preacher, K.J., & MacCallum, R.C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics, 2*, 13-43.
Snook, S.C., & Gorsuch, R.L. (1989). Principal component analysis versus common factor analysis: A Monte Carlo study. *Psychological Bulletin*, *106*, 148-154.
Tabachnick, B.G., & Fidell, L.S. (2013). *Using multivariate statistics* (6th Ed.). Boston: Pearson.
Tucker, L.R., Koopman, R. E.,& Linn, R.L . (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika, 34*, 421–459.
Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. Psychometrika, 41, 321–327.
Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment*: *Honoring Douglas N. Jackson at seventy* (pp. 41–71). Kluwer: Boston.
Zoski, K. W., & Jurs, S. (1996). An objective counterpart to the visual scree test for factor analysis: The standard error scree. *Educational and Psychological Measurement, 56*, 443–451.