Journal of Modern Applied Statistical Methods

Volume 4 | Issue 1 Article 25

5-1-2005

Exploratory Factor Analysis in Two Measurement Journals: Hegemony by Default

J. Thomas Kellow
University of South Florida-Saint Petersburg

Journal of Modern Applied Statistical Methods
May, 2005, Vol. 4, No.1, 283-287

1538 – 9472/05/$95.00

Brief Report

Exploratory Factor Analysis in Two Measurement Journals: Hegemony by Default

J. Thomas Kellow
College of Education
University of South Florida-Saint Petersburg

Exploratory factor analysis studies in two prominent measurement journals were explored. Issues addressed were: (a) factor extraction methods, (b) factor retention rules, (c) factor rotation strategies, and (d) saliency criteria for including variables. Many authors continue to use principal components extraction, orthogonal (varimax) rotation, and retain factors with eigenvalues greater than 1.0.

Key words: Factor analysis, principal components, current practice

Introduction

Factor analysis has often been described as both an art and a science. This is particularly true of exploratory factor analysis (EFA), where researchers follow a series of analytic steps involving judgments more reminiscent of qualitative inquiry, an irony given the mathematical sophistication underlying EFA models.

A number of issues must be considered before invoking EFA, such as sample size and the relationships between measured variables (see Tabachnick & Fidell, 2001, for an overview). Once EFA is determined to be appropriate, researchers must consider carefully decisions related to: (a) factor extraction methods, (b) rules for retaining factors, (c) factor rotation strategies, and (d) saliency criteria for including variables. There is considerable latitude regarding which methods may be appropriate or desirable in a particular analytic scenario (Fabrigar, Wegener, MacCallum, & Strahan, 1999).

J. Thomas Kellow is an Assistant Professor of Measurement and Research in the College of Education at the University of South Florida-Saint Petersburg. Email him at: kellow@stpt.usf.edu

Factor Extraction Methods

There are numerous methods for initially deriving factors, or components in the case of principal component (PC) extraction. Although some authors (Snook & Gorsuch, 1989) have demonstrated that certain conditions involving the number of variables factored and initial communalities lead to essentially the same conclusions, the unthinking use of PC as an extraction mode may lead to a distortion of results. Stevens (1992) summarizes the views of prominent researchers, stating that:

> When the number of variables is moderately large (say > 30), and the analysis contains virtually no variables expected to have low communalities (e.g., .4), then practically any of the factor procedures will lead to the same interpretations. Differences can occur when the number of variables is fairly small (< 20), and some communalities are low. (p. 400)

Factor Retention Rules

Several methods have been proposed to evaluate the number of factors to retain in EFA. Although the dominant method seems to be to retain factors with eigenvalues greater than 1.0, this approach has been questioned by numerous authors (Zwick & Velicer, 1986; Thompson & Daniel, 1996). Empirical evidence suggests that, while under-factoring is probably the greater danger, sole reliance on the eigenvalues greater than 1.0 criterion may result in retaining factors of trivial importance (Stevens, 1992). Other methods for retaining factors may be more defensible and perhaps meaningful in interpreting the data. Indeed, after reviewing empirical findings on its utility, Preacher and McCallum (2003) reported that "the general conclusion is that there is little justification for using the Kaiser criterion to decide how many factors to retain" (p. 23).

Factor Rotation Strategies

Once a decision has been made to retain a certain number of factors, these are often rotated in a geometric space to increase interpretability. Two broad options are available, one (orthogonal) assuming the factors are uncorrelated, and the second (oblique) allowing for correlations between the factors. Although the principal of parsimony may tempt the researcher to assume, for the sake of ease of interpretability, uncorrelated factors, Pedhazur and Shmelkin (1991) argued that both solutions should be considered. Indeed, it might be argued that it rarely is tenable to assume that multidimensional constructs, such as self-concept, are comprised of dimensions that are completely independent of one another. Although interpretation of factor structure is somewhat more complicated when using oblique rotations, these methods may better honor the reality of the phenomenon being investigated.

Saliency Criteria for Including Variables

Many researchers regard a factor loading (more aptly described as a pattern or structure coefficient) of $|.3|$ or above as worthy of inclusion in interpreting factors (Nunnally, 1978). This rationale is predicated on a rather arbitrary decision rule that 9% of variance accounted for makes a variable noteworthy. In a similar vein, Stevens (1992) offered $|.4|$ as a minimum for variable inclusion as this means the variable shares at least 15% of its variance with a factor. Others (Cliff & Hamburger, 1967) argue for the statistical significance of a variable as an appropriate criterion for inclusion. As Hogarty, Kromrey, Ferron, and Hines (in press) noted, "although a variety of rules of thumb of this nature are venerable, they are often ad hoc and ill advised."

Purpose of the Study

This article does not attempt to provide an introduction to the statistical and conceptual intricacies of EFA techniques, as numerous excellent resources are available that address these topics (e.g., Gorsuch, 1983; Stevens, 1992; Tabachnick & Fidell, 2001; Thompson, 2004). Rather, the focus is on the practices of EFA authors with respect to the above issues. Three of the four important EFA analytic decisions described above are treated by default in SPSS and SAS. These programs are the most widely used analytic platforms in psychology. When conducting EFA in either program, one is guided to (a) use PC as the *extraction method* of choice, (b) use eigenvalues greater than 1.0 to *retain factors*, and (c) use orthogonal (varimax) procedures for *rotation of factors*. Only the fourth decision, *variable retention*, is left solely to the preference of the investigator.

EFA practices in two prominent psychological measurement journals were examined: *Educational and Psychological Measurement* (*EPM*) and *Personality and Individual Differences (PID)* over a six-year period. These journals were chosen because of their prominence in the field of measurement and the prolific presence of EFA articles within their pages. In addition, *EPM* is known for publishing factor analytic studies across a diverse array of specialization areas in education and psychology. While *PID* is concerned primarily with the study of personality, it publishes a great deal of international studies from diverse institutions. These features strengthen the external validity of the present findings.

Methodology

An electronic search was conducted using the PsycInfo database for *EPM* and *PID* studies published from January of 1998 to October 2003 that contained the key word 'factor analysis.' After screening out studies that employed only confirmatory factor analysis or examined the statistical properties of EFA or CFA approaches using simulated data sets, a total of 184 articles were identified. In some instances the authors conducted two or more EFA analyses on split samples. For the present purposes these were coded as separate studies. This resulted in 212 studies that invoked EFA models. Variables extracted from the EFA articles were:

a) factor extraction methods;
b) factor retention rules;
c) factor rotation strategies; and
d) saliency criteria for including variables.

Results

Factor Extraction Methods

The most common extraction method employed (64%) was principal components (PC). The next most popular choice was principal axis (PA) factoring (27%). Techniques such as maximum likelihood were infrequently invoked (6%). A modest percentage of authors (8%) conducted both PC and PA methods on their data and compared the results for similar structure.

Factor Extraction Rules

The most popular method used for deciding the number of factors to retain was the Kaiser criterion of eigenvalues greater than 1.0. Over 45% of authors used this method. Close behind in frequency of usage was the scree test (42%). Use of other methods, such as percent of variance explained logics and parallel analysis, was comparatively infrequent (about 8% each). Many authors (41%) explored multiple criteria for factor retention. Among these authors, the most popular choice was a combination of the eigenvalues greater than 1.0 and scree methods (67%).

Factor Rotation Strategies

Virtually all of the EFA studies identified (96%) invoked some form of factor rotation solution. Varimax rotation was most often employed (47%), with Oblimin being the next most common (38%). Promax rotation also was used with a modest degree of frequency (11%). A number of authors (18%) employed both Varimax and Oblimin solutions to examine the influence of correlated factors on the resulting factor pattern/structure matrices.

Saliency Criteria for Including Variables

Thirty-one percent of EFA authors did not articulate a specific criterion for interpreting salient pattern/structure coefficients, preferring instead to examine the matrix in a logical fashion, considering not only the size of the pattern/structure coefficient, but also the discrepancy between coefficients for the same variable across different factors (components) and the logical "fit" of the variable with a particular factor.

Of the 69% of authors who identified an *a priori* criterion as an absolute cutoff, 27% opted to interpret coefficients with a value of $|.3|$ or higher, while 24% chose the $|.4|$ value. Other criteria chosen with modest frequency (both about 6%) included $|.35|$ and $|.5|$ as absolute cutoff values. For the remaining authors who invoked an absolute criterion, values ranged from $|.25|$ to $|.8|$. A few (3%) of these values were determined based on the statistical significance of the pattern/structure coefficient.

Conclusion

Not surprisingly, the hegemony of default settings in major statistical packages continues to dominate the pages of *EPM* and *PID*. The Little Jiffy model espoused by Kaiser (1970), wherein principal components are rotated to the varimax criterion and all components with eigenvalues greater than 1.0 is alive and well. It should be noted that this situation is almost certainly not unique to *EPM* or *PID* authors. An informal perusal of a wide variety of educational and psychological journals that occasionally publish EFA results easily confirms the status of current practice.

The rampant use of PC as an extraction method is not surprising given its status as the default in major statistical packages. Gorsuch (1983) has pointed out that, with respect to extraction methods, PC and factor models such as PA often yield comparable results when the number of variables is large and communalities (h^2) also are large. Although comforting, authors are well advised to consider alternative extraction methods with their data even when these assumptions are met. When these assumptions are not met, such as "when the rank of the factored matrix is small, there is considerable measurement error, measurement error is not homogeneous across variables, and sampling error is small due to larger sample size, *other extraction methods have more appeal*" (Thompson & Daniel, 1996, p. 202, italics added).

The eigenvalues greater than 1.0 criterion was the most popular option for EFA analysts. A number of researchers, however, combined both the eigenvalues greater than 1.0 criterion and the scree test in combination, which is interesting inasmuch as both methods consult eigenvalues, only in different ways. A likely explanation is that both can be readily obtained in common statistical packages.

Other approaches to ascertaining the appropriate number of factors (components) such as parallel analysis (Horn, 1965) and the bootstrap (Thompson, 1988) are available, as are methods based on standard error scree (Zoski & Jurs, 1996). Each of these methods, however, requires additional effort on the part of the researcher. However, EFA authors should consider alternatives for factor retention in much the same way that CFA authors consult the myriad fit indices available in model assessment. As Thompson and Daniel noted, "The simultaneous use of multiple decision rules is appropriate and often desirable" (p. 200).

For authors invoking an absolute criterion for retaining variables, the $|.3|$ level and the $|.4|$ were by far the most popular. Researchers who feel compelled to set such arbitrary criteria often look to textbook authors to guide their choice. The latter criterion can be traced to Stevens (1992), who stated that "It would seem that one would want in general a variable to share *at least* 15% of its variance with the construct (factor) it is going to be used to help name. This means only using loadings (sic) which are about .4 or greater for interpretation purposes" (p. 384). The former rule appears to be attributable to Nunnally (1982), who claimed that "It is doubtful that loadings (sic) of any smaller size should be taken seriously, because they represent less than 10 percent of the variance" (p. 423).

One-third of EFA authors chose not to adhere to a strict, and ultimately arbitrary, criterion for variable inclusion. Rather, these researchers considered the pattern/structure coefficients within the context of the entire matrix, applying various logics such as simple structure and *a priori* inclusion of variables. A (very) few authors considered the statistical significance of the coefficients in their interpretation of salient variables.

Two problems with this approach are that (a) with very large samples even trivial coefficients will be statistically significant, and (b) variables that are meaningfully influenced by a factor may be disregarded because of a small sample size. The issue of determining the salience of variables based on their contribution to a model mirrors that of the debate over statistical significance and effect size. If standards are invoked based solely on the statistical significance of a coefficient, or alternatively, are set based on a strict criterion related to the absolute size of a coefficient related to its variance contribution, it would seem that we would "merely be being stupid in another metric" (Thompson, 2002, p. 30).

Despite criticisms that the technique is often employed in a senseless fashion (e.g., Preacher & MacCallum, 2003), EFA provides researchers with a valuable inductive tool for exploring the dimensionality of data provided it is used thoughtfully. The old adage that factor analysis is as much an art as a science is no doubt true. But few artists rely on unbending rules to create their work, and authors who employ EFA should be mindful of this fact.

References

Cliff, N., & Hamburger, C. D. (1967). The study of sampling errors in factor analysis by means of artificial experiments. *Psychological Bulletin, 68*, 430-445.

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.

Gorsuch, R. L. (1983). *Factor analysis*. Hillsdale, NJ: Earlbaum.

Hogarty, K. Y., Kromrey, J. D., Ferron, J. M., & Hines, C. V. (in press). Selection of variables in exploratory factor analysis: An empirical comparison of a stepwise and traditional approach. *Psychometrika*.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185.

Kaiser, H. F. (1970). A second generation Little Jiffy. *Psychometrika*, *35,* 401–415.

Nunnally, J. C. (1978). *Psychomteric theory* (2nd ed.). New York: McGraw Hill.

Pedhazur, E., & Schmelkin, L. (1991). *Measurement, design, and analysis*. Hillsdale, NJ: Erlabaum.

Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swift's electronic factoring machine. *Understanding Statistics, 2*, 13-43.

Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in *Personality and Social Psychology Bulletin*. *Personality and Social Psychology Bulletin*, *28*, 1629-1646.

Snook, S. C., & Gorsuch, R. L. (1989). Component analysis versus common factor analysis: A Monte Carlo study. *Psychological Bulletin*, *106*, 148-154.

Stevens, J. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Earlbaum.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th Ed.). Needham Heights, MA: Allyn & Bacon.

Thompson, B. (1988). Program FACSTRAP: A program that computes bootstrap estimates of factor structure. *Educational and Psychological Measurement*, *48*, 681-686.

Thompson, B. (2002). What future quantitative social science research could look like: Confidence intervals for effect sizes. *Educational Researcher, 31*(3), 24-31.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.

Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement*, *56*, 197-208.

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.

Zwick, W. R., & Velicer, W. F. (1986). Factors influencing five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432-442.