Multivariate Experimental Clinical Research, Volume 10, No. 2, 1993, 129-135.

SPECIAL REVIEW: EVALUATION OF THE EXPLORATORY FACTOR ANALYSIS PROGRAMS PROVIDED IN SPSSX AND SPSS/PC+

Gregory J. Boyle
University of Queensland

All correspondence should be addressed to Dr. G. J. Boyle, Department of Psychology, University of Queensland, St. Lucia 4067, Australia.

Given the frequent use of SPSSX and SPSS/PC+ exploratory factor analysis in analyzing multivariate psychometric data, it is germane to examine the limitations of the Factor program as it currently exists. Over recent years, the routines in these packages generally have been developed and expanded considerably. In particular, the exploratory factor analysis procedures have been greatly extended and enhanced with inclusion of additional estimation methods such as maximum likelihood, unweighted least squares, generalized least squares, and so on. Likewise, availability of a Scree plot of the latent roots against factor number has facilitated determination of the appropriate factor-extraction number. Provision of a χ^2 goodness-of-fit test, as well as options for sorting and retaining factor loadings greater than a minimum level (usually $\geq \pm .30$), have added to the utility of the SPSSX and SPSS/PC+ statistical routines. These routines have undoubtedly facilitated and stimulated psychometric reports involving the use of factor analysis in the psychological literature.

Despite these evident improvements, a number of useful and practical features still remain absent from the currently available program. For example, the Factor program provides no estimation of simple structure, as indexed by (± .10) hyperplane counts for rotated solutions. Despite provision of *several* types of orthogonal rotation, the *only* oblique method included is direct Oblimin. There is no provision for Promax. There is no provision for Maxplane or Rotoplot or topological rotation. The transformation matrix is not output, for oblique solutions. There is no automatic calculation of the standardized alpha (coefficient of determination) for the loadings of each factor. Nor is there an objective Scree test; inclusion of pertinent algorithms as standard procedures would greatly enhance the utility of the SPSSX and SPSS/PC+ program in multivariate psychometric research.

Such deficiencies are highlighted here in the hope that statistical software manufacturers will incorporate these positive suggestions into subsequent versions of their factor-analytic programs, so that the user at least has the option of choosing whether or not to use these procedures in conducting an exploratory factor analysis. Provision of these additional features in the currently available factor analysis programs would surely enable greater precision in obtaining more refined factor solutions than is currently possible.

HYPERPLANE COUNT

For many years researchers have been routinely calculating the ± .10 hyperplane count as a quantitative index of approximation of the final rotated factor solution to simple structure criteria. This involves determining the percentage of factor loadings less than or equal to, let us say, ± .10 in the rotated factor matrix (factor-pattern matrix for obliquely rotated solutions). This principle was originally enunciated by Thurstone when he recommended that there should be only a few high factor loadings, with the remainder as close to zero as possible. Nonsignificant, trivial factor loadings were regarded by Thurstone as being in what has become known as the "hyperplane." Variables associated with these small loadings are not loaded significantly by the particular factor/latent trait (Cattell, 1978, pp. 102-111; Gorsuch, 1983, pp. 177-182; 340-143). Hyperplane counts for different factor solutions enable comparison of the adequacy of each solution in terms of simple structure criteria. However, the hyperplane count cannot be used in isolation from other considerations, given that it is influenced by variables such as the inter-factor correlations (Gorsuch, p. 199).

Generally, hyperplane counts increase as the number of factors extracted becomes greater (cf. Boyle, 1988). However, there are occasions when the hyperplane count does not increase with additional factors taken out or, more commonly, when the increase is only marginal or trivial (let us say, 2% to 3%, rather than 10% to 12%, for each additional factor extracted). Under these circumstances, it would seem appropriate to accept the more parsimonious solution. In the past, researchers have had to compute the hyperplane count by adding up the number of nonsignificant factor loadings in each column of the rotated factor matrix, and then calculate the percentage hyperplane count as a function of the total number of loadings in the factor matrix. It is a nuisance to have to continually calculate this index of simple structure manually, when the SPSSX or SPSS/PC+ program could very easily produce this information as part of the standard output.

ROTATION STRATEGIES

SPSSX and SPSS/PC+ provide several options for orthogonal rotation of factor solutions. However, only *one* method of oblique rotation is provided (direct Oblimin). This situation represents a definite deficiency in the SPSSX and SPSS/PC+ packages. Gorsuch (1983, pp. 203-204) has listed no fewer than 15 major methods of oblique rotation. In virtually all instances in psychometric research, factors are actually obliquely related. Only in very special situations are factors truly orthogonal and unrelated to each other (Loo, 1979). It is a fundamental principle in behavioral and life contexts that "things tend to go together" to some degree or other. As Boyle (1988, p. 743) indicated, the popular use of orthogonal rotation is rather disconcerting, particularly as the method often fails to even moderately achieve simple structure. Moreover, in the unlikely event that hyperplanes are actually orthogonal, an oblique solution conveyed to *maximum* simple structure approximation inevitably will terminate at the appropriate special orthogonal position (Cattell, 1978). Only after finding nonsignificant, trivial correlations between obliquely rotated factors is one entitled, *a posteriori*, to conduct an orthogonal rotation of the factor matrix. *A priori* considerations alone can never fully justify the use of orthogonal rotation in *exploratory* factor analysis. Routine application of orthogonal rotation makes a mockery of this reality of the interrelatedness of psychometric dimensions in nature.

Especially disconcerting is that there is no provision for undertaking a Promax oblique rotation in the SPSSX or SPSS/PC+ programs, as they currently are structured. This is unfortunate, given that Promax (Hendrickson & White, 1964) is one of the best analytic rotation strategies, allowing control over both the degree of obliquity and extent of approximation to simple structure via the power (k) function, and particularly since the method has been available for over two decades. Users of the SPSSX and SPSS/PC+ factor programs at least should be given the option of carrying out a Promax rotation as an alternative to direct Oblimin alone. This is a major limitation of SPSS Inc. products as they currently stand. Indeed, MacCallum (1983) pointed to this deficiency several years ago, and still Promax is not included as a rotational option. Moreover, SPSSX and SPSS/PC+ do not make any provision for other oblique rotational strategies such as Maxplane, Functionplane, Optres, Orthotran, and so on. SPSS, Inc. has apparently been slow to catch up with certain of its competitors in this regard.

TOPOLOGICAL ROTATION

As for the issue of topological rotation in addition to analytic rotation alone, SPSSX and SPSS/PC+ have failed to include, for instance, Rotoplot (Cattell, 1978, pp. 136-151), which enables an even greater approximation to *maximum* simple structure of the final rotated solution than is possible with analytic rotation alone. Several studies have suggested the efficacy of improving simple structure by means of computer-assisted graphic oblique transformation (cf. Boyle & Stanley, 1986). On the basis of their review of the relevant literature, Boyle and Stanley reported that Rotoplot typically produced an average 8% increase in the ± .10 hyperplane count, additional to that gained with direct Oblimin alone. Some studies displayed up to 15% gain in the hyperplane count. These findings confirm that the analytic rotation options provided in SPSSX and SPSS/PC+ may enable only a rough approximation to simple structure, in many instances. Accordingly, the substantive psychological interpretation of derived factors is likely to alter appreciably, where the hyperplane count can be improved significantly over that initially obtained by analytic rotation alone.

Boyle and Stanley (1986) demonstrated empirically that rotoplot "polished" solutions were more readily interpretable in psychological terms, being less complex with reduced overlap of significant factor loadings for each factor/latent trait. Findings indicated that Rotoplot reduced many of the factor-pattern loadings, leaving the larger ones more clearly prominent, and producing results somewhat akin to those attained using Promax. However, best approximations to simple structure were reported for Rotoplot-"polished" Maxplane, Promax, and Harris-Kaiser methods respectively, indicating, therefore, that use of Rotoplot can even improve the hyperplane count beyond that obtained through analytic Promax alone. Rotoplot "cleaned up" the resultant factor solution, thereby facilitating its substantive interpretability. Given the greater complexity of the analytic solution, the improvement to simple structure due to Rotoplot was evident. Use of up to 20 or 30 Rotoplot cycles may be required to clarify the factor solution in some instances. Kline (1979, p. 37) has advocated use of Rotoplot, stating that attainment of simple structure requires "demonstration that further rotations produce a drop in the hyperplane count, the application of Bargmann's . . . statistical test for simple structure and the fact that the percentage of variables in the hyperplane lies between 55% and 85% . . . topological programs by these criteria of simple structure do seem more efficient than others."

While inclusion of Rotoplot would be a welcome addition to the SPSSX package, its utility would seem even more discernible in the SPSS/PC+ program. Interactive computer-assisted graphic oblique rotation is particularly suited for use with the PC, as provided in the Stat I computer package (wherein Rotoplot is included as an easily used addition to analytic rotation — Brennan & Nitz, 1986, p. 174). The Stat I package is clearly in advance of SPSS/PC+ in this regard. Rotoplot, however, does require access to the transformation matrix associated with the analytic rotation. The SPSSX and SPSS/PC+ programs fail to report the transformation matrix for oblique solutions, so that one is presently required to compute this matrix "by hand" in order to use Rotoplot as an "add-on" to the SPSSX or SPSS/PC+ output alone. This is done by inverting the unrotated factor matrix (Vo), producing a new matrix (Vo^1), which is premultiplied by the corresponding factor-pattern matrix (Vr). A shift matrix is generated from visual inspection of the positioning of the reference vectors with respect to the factor loadings. This shift matrix operates on the existing transformation matrix such that a new transformation matrix is produced enabling, hopefully, an improved factor pattern with each additional Rotoplot cycle. The factor rotation equation is $Vo.L = Vr$ ---(1), where L is the transformation matrix. Use of Rotoplot requires that Eqn.(1) is premultiplied by the inverse of Vo such that $Vo^{-1}.Vo.L = Vo^{-1}.Vr$ ---(2). Since the identity matrix $(I_1) = Vo^{-1}.Vo$, $L = Vo^{-1}.Vr$ ---(3). With non-square matrices, it is not possible to obtain a classical inverse, necessitating estimation of a "pseudo-inverse." By postmultiplying Vo by Vo^{-1}, one obtains $Vo.Vo^{-1} = I_2$ ---(4), such that I_2 is not an identity matrix. $Vo.Vo^{-1} = Vo^{-1}.Vo$, as the order of I_1 is different from that of I_2, where Vo^{-1} is not a classical inverse of Vo.

While topological rotation strategies such as Rotoplot can be very helpful in certain instances, this is not true for every factor analysis. Indeed, Boyle and Stanley (1986) reported only marginal improvement to the (± .10) hyperplane count in their own use of Rotoplot, although they carried out only five Rotoplot cycles. As they reported, however, several studies have shown significant and substantial improvement in approximation to simple structure as a direct result of employing Rotoplot. Hence, Rotoplot would seem to be a useful addition to the SPSSX and SPSS/PC+ factor analysis programs. Implementation in the SPSS/PC+ program should overcome the main disadvantage of Rotoplot, namely that it can be time-consuming in some instances. It would be nice if SPSS, Inc. would consider including Rotoplot, especially in future versions of the SPSS/PC+ package.

OBJECTIVE SCREE TEST

On a different note, the ability to output a Scree plot of latent roots against factor number in the SPSSX and SPSS/PC+ packages is undoubtedly a great asset. This allows subjective interpretation of the Scree test, which admittedly encourages the researcher to take account of the Scree plot. However, there is no provision in the current versions of SPSSX and SPSS/PC+ for *objective* assessment of the Scree plot, such as the automated Scree test of Barrett and Kline (1982) or the objective CNG Scree test (Gorsuch, 1983, p. 168), both of which are based on computer algorithms, thereby avoiding this subjectivity in interpretation. It would be relatively easy for SPSS, Inc. to include a computer algorithm for objective interpretation of the Scree test. Use of these algorithms has enabled more accurate predictions of the appropriate factor extraction number than does subjective interpretation of the Scree plot as currently output from the SPSSX and SPSS/PC+ packages. The objective Scree tests (cf. Gorsuch & Nelson, 1981) compare favorably with other indices of the appropriate number of factors, such as Velicer's (1976) MAP test, or Revelle and Rocklin's (1979) VSS test. It would be useful if some form of objective Scree test, along with other objective indices of the appropriate factor-extraction number could be reported in the standard factor-analytic output. The researcher, then, can more conveniently evaluate this information, and so reach an informed decision as to the number of factors. In this way, interpretation of the Scree plot which is currently subjective, undoubtedly resulting in discrepant estimation of the number of factors by different investigators (and even the same investigator at different times), would be standardized for all researchers using the SPSSX and SPSS/PC+ programs. This situation must be regarded as a desirable goal, putting the use of exploratory factor analysis onto a more reliable footing than presently is the case. The problem of the number of factors is often a complex and difficult one. Provision of further information as suggested here should be a beneficial addition to the SPSSX and SPSS/PC+ packages.

CONCLUSIONS

While these recommendations apply specifically to the SPSSX and SPSS/PC+ packages, most, if not all, of them apply equally to other well-known statistical packages such as SAS and BMDP. However, given that SPSS, Inc. products are among the most frequently used in multivariate psychometric investigations, it seems germane to focus attention, initially at least, on SPSSX and SPSS/PC+. Certainly, with the advent of confirmatory factor-analytic methods employed alone or, more fruitfully, as part of an overall structural equation modeling approach using, let us say, LISREL in SPSSX (cf. Cuttance & Ecob, 1987; Jöreskog & Sörbom, 1981, 1985), COSAN (McDonald, 1985), or EQS in BMDP (Bentler, 1984), reliance on exploratory factor-analytic methods has declined considerably. Nevertheless, this does not invalidate the usefulness of exploratory factor-analytic methodology as such, and it is often used very fruitfully in conjunction with confirmatory methods of analysis. Inclusion of the above suggestions into future versions of SPSSX and SPSS/PC+ would undoubtedly enhance the precision with which exploratory factor analysis can be employed using these programs. It is important that as many as possible of the decision points in exploratory factor analysis be put onto a more objective basis than currently provided in the SPSSX and SPSS/PC+ programs.

REFERENCES

Barrett, P. T., & Kline, P. (1982). Factor extraction: An examination of three methods. *Personality Studies and Group Behavior*, *2*, 84-98.

Bentler, P. M. (1984). *Theory and Implementation of EQS: A Structural Equations Program.* Los Angeles: BMDP Statistical Software.

Boyle, G. J. (1988). Exploratory factor analytic principles in motivation research. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of Multivariate Experimental Psychology* (pp. 742-745). New York: Plenum.

Boyle, G. J., & Stanley, G. V. (1986). Application of factor analysis in psychological research: Improvement of simple structure by computer-assisted graphic oblique transformation: A brief note. *Multivariate Experimental Clinical Research*, *8*, 175-182.

Brennan, J. M., & Nitz, L. H. (1986). *Stat I: A Statistical Toolbox. Version 3.0.* Kailua, Hawaii: Sugar Mill Software.

Cattell, R. B. (1978). *The Scientific Use of Factor Analysis in Behavioral and Life Sciences.* New York: Plenum.

Cuttance, P., & Ecob, R. (1987). *Structural Modeling by Example: Applications in Educational, Sociological, and Behavioral Research.* New York: Cambridge.

Gorsuch, R. L. (1983). *Factor Analysis* (2nd ed.). Hillsdale, N.J.: Erlbaum.

Gorsuch, R. L., & Nelson, J. (1981). CNG Scree test: An objective procedure for determining the number of factors. Paper presented at the annual meeting of the Society for Multivariate Experimental Psychology.

Hendrickson, A. E., & White, P. O. (1964). Promax: A quick method for rotation to oblique simple structure. *British Journal of Statistical Psychology*, *17*, 65-70.

Jöreskog, K. G., & Sörbom, D. (1981). *LISREL: Analysis of Linear Structural Relationships by the Method of Maximum Likelihood.* Chicago: National Educational Resources.

Jöreskog, K. G., & Sörbom, D. (1986). *PRELIS: A Program for Multivariate Data Screening and Data Summarization: A Preprocessor for LISREL.* Mooresville, IN: Scientific Software Inc.

Kline, P. (1979). *Psychometrics and Psychology.* London: Academic Press.

Loo, R. (1979). The orthogonal rotation of factors in clinical research: A critical note. *Journal of Clinical Psychology*, *35*, 762-765.

MacCallum, R. (1983). A comparison of factor analysis programs in SPSS, BMDP, and SAS. *Psychometrika*, *48*, 223-231.

McDonald, R. P. (1985). *Factor Analysis and Related Methods*. Hillsdale, N.J.: Erlbaum.

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika, 44*, 443-460.

Poon, W. Y., & Lee, S. Y. (1987). Maximum likelihood estimation of multivariate polyserial and polychoric correlation coefficients. *Psychometrika, 52*, 409-430.

Revelle, W., & Rocklin, T. (1979). Very simple structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research, 14*, 403-414.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41*, 321-327.