DOCUMENT RESUME

ED 346 135 TM 018 438

AUTHOR Daniel, Larry G.
TITLE Bootstrap Methods in the Principal Components Case.
PUB DATE Apr 92
NOTE 34p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Adults; *Estimation (Mathematics); *Factor Analysis; Factor Structure; *Multivariate Analysis; *Reliability; Sampling; Social Science Research
IDENTIFIERS Bem Sex Role Inventory; *Bootstrap Methods; Confidence Intervals (Statistics); FACSTRAP Computer Program; Principal Components Analysis; *Research Replication

ABSTRACT

Some years ago, B. Efron and his colleagues developed bootstrap resampling methods as a way of estimating the degree to which statistical results will replicate across variations in sample. A basic problem in the multivariate use of bootstrap procedures involves the requirement that the results across resamplings must be rotated to best fit in a common factor space before any estimators are averaged. The use of factor analysis for this problem is demonstrated using the responses of 298 persons to items from the Bem Sex-Role Inventory from a study by B. Thompson (1988). The statistical computer program FACSTRAP is used to calculate bootstrap confidence intervals in factor analysis. Bootstrap methods are valuable because they: (1) lend credibility to the factor analysts' choice of the number of factors to extract and interpret; (2) provide evidence for increased confidence in the interpretation of factor meaning; and (3) demonstrate the importance of replication in the social sciences. There are 9 tables of analysis results and a 35-item list of references. (SLD)

ED346135

facstrap.lgd
04/16/92

BOOTSTRAP METHODS IN THE PRINCIPAL COMPONENTS CASE

Larry G. Daniel

University of Southern Mississippi

Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 20-24.

ABSTRACT

Some years ago, Efron and his colleagues developed bootstrap resampling methods as a way of estimating the degree to which statistical results will replicate across variations in sample. A basic problem in the multivariate use of bootstrap procedures involves the requirement that the results across resamplings must be rotated to best fit in a common factor space before any estimators are averaged. The author demonstrates how factor analysis is useful for illustrating this problem and its resolution. The statistical program FACSTRAP is used to calculate boostrap confidence intervals in factor analysis.

BOOTSTRAP METHODS IN THE PRINCIPAL COMPONENTS CASE

Social science research is ultimately a search for truth about the nature of and the relationships among various phenomena. Hence, when using correlational statistical techniques for obtaining scientific results, researchers should be appropriately concerned with the external validity (i.e., the generalizability) of obtained results with respect to a broader population of interest. Since there is always the possibility that results of such analyses will capitalize upon chance, it is desirable that researchers attempt to replicate research findings across various samples selected from a given population. Thus, replication has been regarded as "the cornerstone of science" (Carver, 1978, p. 392). As Holmes (1990, p. 14) has noted:

> Replication, the repeating of studies, is the cornerstone of scientific investigation. Before a study can be considered useful, it must be replicated. . . . In any study, there is the possibility that the results could have been obtained by chance alone. . . . With replication studies, either the original findings will be supported, or disagreements between the original study and the replication will point to problem areas. As these problem areas are investigated, researchers come closer and closer to the truth.

Although replication is crucial, it can involve time-intensive procedures. This is problematic considering that the timeliness of reporting of noteworthy findings is crucial if such findings are to make an impact on future practice. Thus, as Thompson (1992, p. 15) has noted, "it is not always convenient to conduct a replication prior to interpreting results [of a given study]." To combat this problem, researchers have developed several methods which allow the researcher to creatively use a single sample to both generate correlational results and to estimate the degree to which results will generalize. These methods include (a) the "cross-validation" ("holdout," or "invariance") method, (b) the "jackknife" method (Tukey, 1958), and (c) the "bootstrap" method (Diaconis & Efron, 1983; Efron, 1979). A number of comparative summaries of these and other related methods have been provided in the extant literature (e.g., Afifi & Clark, 1984; Cooil, Winer, & Rados, 1987; Daniel, 1989a; Efron & Gong, 1983; Thompson, 1988b, 1992). A brief explanation of these three procedures follows.

The Cross-Validation Method

When employing the cross-validation method, the researcher utilizes a given sample to both derive and validate a given statistical estimator. For example, in the discriminant analysis case, this is done "by randomly splitting the original sample . . . into two [approximately equivalent] subgroups: one for deriving the discriminant function and one for crossvalidating it" (Afifi & Clark, 1984, p. 266). Understandable descriptions of the cross-validation method replete with statistical examples are provided by Thompson (1989) and Ferrell (1992). This method is appealing as it can easily be done within the domain of a single study; yet, it is problematic as it is subject to sample bias, especially when the size of the sample is small.

The Jackknife Method

The jackknife method for substantiating external validity of research results involves repeated computation of a given statistic, with one case or a given subset of cases omitted from the data set each time the statistic is computed. The ›al of the jackknife procedure is to "average out" the effects of outlying or atypical cases. In each step of the jackknife analysis, "pseudovalues" (Quenouille, 1956) of the statistic are computed. Once a series of analyses has been run with all possible data subsets, in turn, omitted from the sample, the resulting pseudovalues of the selected statistic are averaged to provide a "jackknifed" estimate of the statistic. Stability of the original statistic (run with the entire data set) is assessed by determining whether it falls within confidence intervals for the jackknifed values. Since jackknife methods minimize sample splitting (i.e., the size of any jackknife sample is nearly as large as the original sample), they are particularly useful when sample size is small. For illustrative examples of how to compute the jackknifed value of a statistic, see Crask and Perreault (1977), Daniel (1989a), and Tucker and Daniel (1992).

The Bootstrap Method

The bootstrap method involves "creating" a mock population from a representative sample, from which multiple samples can then be drawn. The method extends the usefulness of the jackknife procedure as it allows for computation of a given statistic across a maximal number of fluctuations in the sample (Lunneborg, 1983). Borrello and Thompson (1989, p. 320) offer the following description of the bootstrap method:

> The bootstrap approach requires the researcher to create a large artificial population of data by copying the data set over and over again into a data file. Dozens or hundreds of random samples are drawn from the population, and statistics are calculated over and over again. The average results are then interpreted for each statistic. The standard deviations of these estimates are especially valuable because they provide an indication of how stable the estimates are across variations in samples. . . . [B]ootstrap analysis involves sampling with replacement and a given subject may be used to generate more than one case of data in a given analysis.

In short, bootstrap methods allow the researcher to simulate an actual population from which realistic samples are drawn. These methods lend the researcher more evidence than either the cross-validation or jackknife procedures regarding the potential for fluctuations in results across samples. Hence, as Efron (1979, p. 1) notes, "bootstrap methods are more widely applicable than the jackknife, and also more dependable." Nevertheless, bootstrap techniques, like other sample-splitting procedures, are neither flawless nor panacean. Thompson (1992, p. 15) succinctly sums up the usefulness of bootstrap methods and the other two aforementioned procedures while also acknowledging their limitations:

> Since all three strategies are typically based on a single sample of subjects in which the subjects usually have much in common (e.g., point in time of measurement, geographic origin) relative to what they would have in common with a separate sample, the three methods all yield somewhat inflated estimates of replicability. Because inflated estimates of replicability provide a better estimate of replicability than no estimate at all (i.e., statistical significance testing), these procedures can still be useful. . .

Purpose of the Present Study

Considering the usefulness of bootstrap techniques, the purpose of the present paper was to apply bootstrap methodology in the interpretation of principal components factor analytic results. Bootstrap logic is especially applicable to factor analytic results as it may provide to factor analysts a wealth of evidence related to significant interpretation issues, including questions about number of factors to extract and interpret, magnitude of structure coefficients, and stability of results across sample fluctuations.

As a rationale for applying these procedures to the interpretation of factor analytic results, the following assumptions were made:

(a) A single representative sample can serve as the source of both initial statistical estimates and replication data.

(b) A given sample representative of the population can be interpreted to actually be many samples.

(c) Bootstrapped estimates of statistics can provide meaningful estimates of the values and relative stability of these statistics in the actual population.

In accomplishing the purposes of the present study, a computer program designed to calculate bootstrap factor analytic statistics was utilized. This program, called FACSTRAP (Thompson, 1988b), computes three sequences of bootstrap analyses each involving different specified sample sizes, using bootstrap samples from a data file based on an original representative sample of data provided by the user. For each sample in a given sequence, FACSTRAP computes the eigenvalues of the first 15 factors, followed by descriptive statistics for the average eigenvalues across the sequence. It also produces means and standard deviations for the average factor structure coefficients across a number of factors to be specified by the user. Other useful statistics are available as well.

The following discussion will provide (a) a brief overview of factor analysis, (b) a review of problems with replicability of factor analytic results, and (c) a logic for understanding how bootstrap methodology can be used to address replicability issues. An example of how the program FACSTRAP can be used to address the replicability issue in the factor analytic case is then presented.

A Brief Overview of Factor Analysis

Factor analysis has enjoyed a place of prominence among the tools of the social scientist. Factor analytic methods have proven useful in social science research in theory development and in test validation (Humphreys, Ilgen, McGrath, & Montanelli, 1969). Considering its usefulness, factor analysis has been described as "one of the most powerful tools yet devised for the study of complex areas of behavioral scientific concern" (Kerlinger, 1986, p. 689), and as "the furthest logical development and reigning queen of the correlational methods" (Cattell, 1978, p. 4).

In explaining the logic of factor analysis, social scientists (e.g., Cattell, 1988) often conceive of a "data box," i.e., a three-dimensional model for measuring and describing any given psychological or ideological phenomenon. The three dimensions (or <u>modes</u>) which constitute the model are usually considered to be persons, items, and occasions of measurement. Factor analysis generally involves any two of these modes simultaneously, one of which is factored across the other.

In its most common form (known as "R-technique" factor analysis), factor analysis is used to factor items across people:

> In conducting an "R-technique" factor analysis, . . . the researcher first selects a finite set of *p* variables from a universe of possible variables designed to measure a specific construct. The choice of variables may be based on theoretical considerations or on the researcher's own hypothetical notions regarding the nature of the construct. Data are collected on these variables from a sample of persons thought to be representative of the researcher's population of interest. A square (*p* x *p*) matrix of association (correlation matrix) is constructed to determine the intercorrelations among the *p* variables. On the basis of these correlations, a new rectangular correlation matrix is constructed with *p* variables serving as the rows and *m* common factors serving as the columns. (Daniel, 1990, p. 3)

Although R-technique is the most commonly used form of factor analysis, other variations are used with some frequency.

For example, Q-technique factor analysis (Stephenson, 1953) uses the same two modes as R-technique, although the modes are reversed, i.e., the persons are factored across the variables, yielding discrete factors of persons who behave or think differently from persons in other factors in regard to the construct of interest. Daniel (1989c) provides an example of Q-technique factor analysis using data from an actual educational research study.

Factor analytic procedures may be considered as either "exploratory" or "confirmatory" depending on the factor extraction methods employed, and on the degree to which the researcher dictates an a priori desired factor structure (Daniel, 1989b). Furthermore, it is also possible for the researcher to utilize exploratory factor extraction methods followed by confirmatory factor rotation. As Thompson (1988b, p. 682) notes, "Factors can also be extracted using exploratory methods and then rotated in a confirmatory manner to 'best fit' position with a theoretically derived structure matrix." For example, factor matrices derived from two separate studies could be projected into the same factor space to determine the degree to which the factor structures are similar.

Problems with the Replicability of Factor Analytic Results

As previously noted, replication is the hallmark of scientific inquiry. It is important in factor analysis as in all methods to obtain results which are generalizable to the larger population of interest. In discussing threats to the replicability of factor analytic results, Kerlinger (1979, p. 198) notes that "a substantial factor loading can occur by chance. Thus the analyst may try to interpret what is an uninterpretable result." As Humphreys et al. (1969) note, the investigator can capitalize on chance not only during the extraction phase of the given analysis, but also during the rotation phase.

By substantiating a given factor structure across samples, researchers tend to diminish the possibility that results are due merely to artifacts of a given sample. Interestingly, however, it is normal to expect a certain level of fluctuation in factor analytic statistics (e.g., factor structure coefficients, eigenvalues, variable communalities, factor scores) when there are differences in the number of items, the number of observations (e.g., persons), or the number of factors extracted, even when the factors are relatively stable (Humphreys et al., 1969).

As Harman (1967) and Lambert, Wildt, and Durand (1990) have observed, the "number of factors" decision tends to impact the estimation of various factor parameters. In fact, it has been noted that "rotated loadings and factor score coefficients may be substantially distorted if too few or too many factors are retained" (Lambert et al., 1990, p. 34). This observation is particularly problematic in exploratory factor analytic procedures as the number of factors retained and interpreted is often based on one of several mechanical (and to some extent arbitrary) "cutoff" criteria, e.g. Guttman's (1954) "eigenvalue greater than unity criterion" or Cattell's (1966) "scree" test.

Although these criteria may prove useful when eigenvalues are relatively sample invariant, in the absence of information about the variance of the sample eigenvalues, conclusions about factor meanings based on the parameters obtained with the given sample are tenuous at best (Lambert et al., 1990). Eigenvalues just below or just above these cutoff criteria are most subject to variability across samples. For example, using the eigenvalue greater than unity criterion, a factor with an eigenvalue of 1.12 would be retained while one with an eigenvalue of 0.97 would be ignored, even though the relative amount of variance accounted for by each of the two factors would be about the same (Lambert et al., 1990). However, even with relatively small fluctuations in the eigenvalues of either of these factors across a different sample, the decision on whether to interpret the factors could be different.

Employing Bootstrap Methodology in Factor Analysis

Although bootstrap procedures have been employed with a variety of statistical estimators, their application in the study of factor eigenvalues has been largely unexplored (Lambert et al., 1990). As previously noted, one advantage of using bootstrap procedures with any statistical procedure is that the techniques allow the researcher an opportunity to build a distribution for the given statistic using the "mega file" created from the data in hand. By building such a distribution for factor eigenvalues, the researcher could then compute confidence intervals around each eigenvalue based on its standard deviation, and thereby estimate "normal" fluctuations in its value across samples (Thompson, 1988b). Similar distributions for factor structure coefficients could be computed as well, and the researcher could use these distributions to support or reject original beliefs about which items most clearly define a given factor.

This information could prove extremely useful in factor retention and interpretation decisions, especially when employing various frequently-used interpretation criteria. As Thompson (1988b, pp. 682-683) notes:

> If variations in samples do not have much effect on the estimates, then these standard deviations will be small and the researcher can reasonably vest more confidence in a belief that results will generalize. Thus, the bootstrap approach may better support inferences about invariance or replicability of results in new samples. . .

Currently, there are several statistical computer programs available for performing factor analysis bootstrapping procedures (e.g., Lambert et al., 1990; Thompson, 1988b). As noted by Thompson (Scott, Thompson, & Sexton, 1989; Thompson, 1988b), one of the difficulties of applying bootstrap procedures in factor analysis is the requirement that factors are presented in a common result space. Similar factors may be extracted across various sample, yet the factors may not occur in the same order as ranked by magnitude of their eigenvalues (e.g., Factor II from one analysis may occur as Factor III in another analysis). Hence, it has been recommended that each alternative factor analytic run should be rotated to best fit to a given "target" matrix derived from the original factor run, or else based on an ideal matrix of ones, negative, ones, and zeroes (Thompson, 1988b).

Procedures

An intact data set (Thompson, 1988a) was utilized for the purpose of the present investigation. The data included the responses of 298 persons to 40 items from the the Bem Sex-Role Inventory (Bem, 1981), including 20 "masculine" items and 20 "feminine" items. An initial principal components factor analysis was performed using these data. The analysis yielded 13 factors with prerotational eigenvalues greater than unity. Based on the scree plot of the eigenvalues, two factors were extracted. The varimax rotated factor matrix for this solution is presented in Table 1.

INSERT TABLE 1 ABOUT HERE

A population target matrix of ones, negative ones, and zeroes was specified based on the results of this original analysis run with the entire sample of 298 persons. This target matrix, which is presented in Table 2, was input into the program FACSTRAP (Thompson, 1988b). The Table 1 matrix was projected into the same factor space as this target matrix, resulting in the "Procrustean-rotated" matrix shown in Table 3.

INSERT TABLES 2 AND 3 ABOUT HERE

Bootstrap samples were then selected from the original set of 298 subjects using n's of 100, 200, and 600. Ten samples were selected for each of these three n sizes, and the FACSTRAP program was utilized to compute estimates for the first 15 eigenvalues as well as Procrustes rotated factor matrices for each of the 30 resultant analyses. The program then was used to compute average factor analytic statistics across the three sets of analyses.

Results

Table 4 presents the descriptive statistics for the eigenvalues of the first five factors across the 10 bootstrap samples of 100 subjects each. Even though the FACSTRAP analysis yielded estimates of the first 15 eigenvalues, only five are reported for this analysis as well as the subsequent analyses as the results focus largely on only the first two factors. The average factor structure coefficients for items across these two factors along with their standard deviations are presented in Table 5.

INSERT TABLES 4 AND 5 ABOUT HERE

In general, these results confirm the original factor structure as shown in the Table 1 matrix. A scree plot of the eigenvalues would confirm the appropriateness of the two-factor solution. In addition, the make-up of the factors would remain relatively constant as suggested by the saliency of the items across the two factors. Interestingly, however, some of the items' structure coefficients are more negligible in magnitude across the averaged bootstrap matrix as compared to the matrix derived from the original result even for the factor with which they are most highly correlated. For the most part the standard deviations of the structure coefficients tend to be small, suggesting that there is a relatively high degree of stability in the results. Nevertheless, there is also a general trend for the standard deviations associated with smaller structure coefficients to be somewhat larger.

Table 6 presents the descriptive eigenvalue statistics for the 10 principal components analyses run with 200 subjects each. The first five eigenvalues are summarized in the table. The average factor structure coefficients for the first two factors along with their standard deviations are presented in Table 7.

INSERT TABLES 6 AND 7 ABOUT HERE

Once again, these results do not yield too many surprises, with the magnitude of the eigenvalues suggesting a scree solution of two factors, and with items relatively salient with the factors they were intended to be identified with. What is interesting, however is the increased stability of both the eigenvalues and the structure coefficients as judged by those statistics' standard deviations when compared to the standard deviations of these statistics across the average of the 100 person sequence of 10 factor runs. Hence, increasing sample size in the bootstrap sequence seems to enhance the stability of the various statistics of interest.

Descriptive statistics for the eigenvalues across the first five factors for 10 bootstrap sample of 600 subjects each are presented in Table 8. Average factor structure coefficients for the 40 items across these 10 analyses are presented in Table 9.

INSERT TABLES 8 AND 9 ABOUT HERE

These results further confirm the original factor interpretations, with larger structure coefficients tending to be more highly stable than smaller ones, and with standard deviations of the structure coeffients and the eigenvalues as a whole tending to be smaller than those computed using either the 100- or 200-subject bootstrap sample sequences.

Discussion

The foregoing results illustrate several important points about the usefulness of bootstrap procedures in interpreting the results of principal components factor analysis. First, bootstrap methods lend credibility to the factor analyst's choice of the number of factors to extract and interpret. In the example presented here, the choice of two factors was confirmed across realistic bootstrap samples ranging in size from 100 to 600.

Second, bootstrap methods can provide evidence for increased confidence in researchers' interpretation of the meaning of factors. That the stability of the estimators increased with the n of subjects is particularly noteworthy, indicating that bootstrapping is particularly valuable in situations involving small samples. The foregoing results certainly suggest that the original sample utilized for the study is indeed representative of a larger population of interest, and further lend credence to the assumption that the defined factors do indeed represent constructs of interest rather than meaningless conglomerations of items due to chance relationships.

Finally, bootstrap methods demonstrate well the importance of replication in the social sciences. Obviously, nothing can ever take the place of true replication of research findings. However, bootstrap methods may be beneficial in helping researchers to think beyond the immediate findings of their research to project with some confidence notions regarding the impact of the findings to a larger population of interest.

REFERENCES

Afifi, A. A., & Clark, V. (1984). Computer-aided multivariate analysis. Belmont, CA: Lifetime Learning Publications.

Bem, S. L. (1981). Bem Sex-Role Inventory. Palo Alto, CA: Consulting Psychologists Press.

Borrello, G. M., & Thompson, B. (1989). A replication bootstrap analysis of the structure underlying perceptions of stereotypic love. Journal of General Psychology, 116, 317-327.

Carver, R. P. (1978). The case against statistical significance testing. Harvard Educational Review, 48, 378-399.

Cattell, R. B. (1966). The scree test for the number of factors. Multivariate Behavioral Research, 1, 245-276.

Cattell, R. B. (1978). The scientific use of factor analysis in behavioral and life sciences. New York: Plenum.

Cattell, R. B. (1988). The data box: Its ordering of total resources in terms of possible relational systems. In J. R. Nesselroade & R. B. Cattell (Eds.), Handbook of multivariate experimental psychology (2nd ed.) (pp. 69-130). New York: Plenum.

Cooil, B., Winer, R. S., & Rados, D. L. (1987). Cross-validation for prediction. Journal of Marketing Research, 24, 271-279.

Crask, M. R., & Perreault, W. D. (1977). Validation of discriminant analysis in marketing research. Journal of Marketing Research, 14, 60-68.

Daniel, L. G. (1989a, January). Use of the jackknife statistic to establish the external validity of discriminant analysis results. Paper presented at the annual meeting of the Southwest Educational Research Association. (ERIC Document Reproduction Service No. ED 305 382)

Daniel, L. G. (1989b, November). Comparisons of exploratory and confirmatory factor analysis. Paper presented at the annual meeting of the Mid-South Educational Research Association, Little Rock, AR. (ERIC Document Reproduction Service No. ED 314 447)

Daniel, L. G. (1989c, November). Stability of Q-factors across two data collection methods. Paper presented at the annual meeting of the Mid-South Educational Research Association, Little Rock, AR. (ERIC Document Reproduction Service No. ED 314 439)

Daniel, L. G. (1990, November). Common factor analysis or components analysis: An update on an old debate. Paper presented at the annual meeting of the Mid-South Educational Research Association, New Orleans. (ERIC Document Reproduction Service No. ED 325 531)

Diaconis, P., & Efron, B. (1983). Computer-intensive methods in statistics. Scientific American, 248(5), 116-130.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. Annals of Statistics, 7, 1-26.

Efron, B., & Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. American Statistician, 37, 36-48.

Ferrell, C. M. (1992, January). Statistical significance, sample splitting, and generalizability of results. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston.

Guttman, L. (1954). Some necessary conditions for common factor analysis. Psychometrika, 19, 149-161.

Harman, H. H. (1967). Modern factor analysis (2nd ed.). Chicago: University of Chicago Press.

Hinkle, D. E., & Winstead, W. H. (1990, April). Bootstrap methods: A very leisurely look. Paper presented at the annual meeting of the American Educational Research Association, Boston. (ERIC Document Reproduction Service No. ED 318 751)

Holmes, C. B. (1990). The honest truth about lying with statistics. Springfield, IL: Charles C. Thomas.

Humphreys, L. G., Ilgen, D., McGrath, D., & Montanelli, R. (1969). Capitalization on chance in rotation of factors. Educational and Psychological Measurement, 29, 259-271.

Kerlinger, F. N. (1979). Behavioral research: A conceptual approach. New York: Holt, Rinehart, and Winston.

Kerlinger, F. N. (1986). Foundations of behavioral research (3rd ed). New York: Holt, Reinhart, and Winston.

Lambert, Z. V., Wildt, A. R., & Durand, R. M. (1990). Assessing sampling variation relative to numbers of factors criteria. Educational and Psychological Measurement, 50, 33-48.

Lunneborg, C. E. (1983, August). Efron's bootstrap with some applications in psychology. Paper presented at the annual meeting of the American Psychological Association, Anaheim, CA. (ERIC Document Reproduction Service No. ED 241 554)

Quenouille, M. H. (1956). Notes on bias in estimation. Biometrika, 43, 353-360.

Scott, R. L., Thompson, B., & Sexton, D. (1989). Structure of a short form of the questionnaire on resources and stress: A bootstrap factor analysis. Educational and Psychological Measurement, 49, 409-419.

Stephenson, W. (1953). The study of behavior: Q-technique and its methodology. Chicago: University of Chicago Press.

Thompson, B. (1988a). [Bootstrap data from Bem Sex Role Inventory]. Unpublished raw data.

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

Thompson, B. (1989). Statistical significance, result importance, and result generalizability: Three noteworthy but somewhat different issues. Measurement and Evaluation in Counseling and Development, 22, 2-6.

Thompson, B. (1992, April). The use of statistical significance tests in research: Some criticisms and alternatives. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Tucker, M. L., & Daniel, L. G. (1992, January). Investigating result stability of canonical function equations with the jackknife technique. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston.

Tukey, J. W. (1958). Bias and confidence in not-quite large samples. Annals of Mathematical Statistics, 29, 614.

Table 1
Varimax Rotated Two-Factor Solution Using Original Sample
($\underline{n}$ = 298)

	Factor I	Factor II
Item1	.19493	-.37889
Item2	.17518	-.38570
Item3	.16021	-.35839
Item4	.22808	-.55195
Item5	-.11821	-.52797
Item6	.12677	-.65976
Item7	-.10674	-.32439
Item8	-.06872	-.59485
Item9	.13436	-.61207
Item10	-.09345	-.52323
Item11	.14351	-.46348
Item12	-.08901	-.34501
Item13	.11658	-.35009
Item14	.11986	-.45518
Item15	.12970	-.45320
Item16	.16182	-.35069
Item17	-.28116	-.42322
Item18	.06935	-.16374
Item19	.25822	-.40750
Item20	.11935	-.64954
Item21	.60658	.01195
Item22	.59244	-.01662
Item23	.60768	-.06985
Item24	.57032	-.03721
Item25	.54270	-.14260
Item26	.62027	-.15151
Item27	.60304	-.14913
Item28	.61596	.04132
Item29	.58906	.12570
Item30	.74256	-.00947
Item31	.34769	-.05202
Item32	.51041	-.18136
Item33	.08968	.20884
Item34	.28268	-.06850
Item35	.52269	-.14569
Item36	.28519	.26064
Item37	.02580	.15445
Item38	.06030	.28680
Item39	.22250	.12391
Item40	.34720	.31287

Table 2
Population Target Matrix

	Factor I	Factor II
Item1	0	-1
Item2	0	-1
Item3	0	-1
Item4	0	-1
Item5	0	-1
Item6	0	-1
Item7	0	-1
Item8	0	-1
Item9	0	-1
Item10	0	-1
Item11	0	-1
Item12	0	-1
Item13	0	-1
Item14	0	-1
Item15	0	-1
Item16	0	-1
Item17	0	-1
Item18	0	-1
Item19	0	-1
Item20	0	-1
Item21	1	0
Item22	1	0
Item23	1	0
Item24	1	0
Item25	1	0
Item26	1	0
Item27	1	0
Item28	1	0
Item29	1	0
Item30	1	0
Item31	1	0
Item32	1	0
Item33	0	1
Item34	1	0
Item35	1	0
Item36	1	0
Item37	0	1
Item38	0	1
Item39	1	0
Item40	1	0

Table 3
Procrustean Rotated Factor Matrix for Original Sample
(n = 298)

	Factor I	Factor II
Item1	.17321	-.38940
Item2	.15311	-.39499
Item3	.13970	-.36687
Item4	.19653	-.56395
Item5	-.14785	-.52044
Item6	.08963	-.65988
Item7	-.12490	-.31784
Item8	-.10222	-.59002
Item9	.09957	-.61868
Item10	-.12286	-.51711
Item11	.11709	-.47085
Item12	-.10836	-.33943
Item13	.09661	-.35612
Item14	.09395	-.46122
Item15	.10389	-.45980
Item16	.14174	-.35927
Item17	-.30463	-.40666
Item18	.03456	-.61667
Item19	.23479	-.42144
Item20	.08246	-.65525
Item21	.60629	-.02234
Item22	.59055	-.05006
Item23	.60276	-.10407
Item24	.56730	-.06938
Item25	.53377	-.17304
Item26	.61072	-.18632
Item27	.59365	-.18297
Item28	.61731	.00645
Item29	.59522	.09222
Item30	.74084	-.05142
Item31	.34420	-.07158
Item32	.49935	-.20991
Item33	.10134	.20344
Item34	.27836	-.08437
Item35	.51363	-.17499
Item36	.29946	.24411
Item37	.03449	.15274
Item38	.07641	.28294
Item39	.22914	.11114
Item40	.36433	.29275

Table 4
Descriptive Statistics for Eigenvalues Across 10 Bootstrap Samples of '00 Subjects Each

	I	II	III	IV	V
Mean	6.61	4.49	2.81	2.44	2.14
SD	0.63	0.68	0.26	0.22	0.16
Minimum	5.63	3.52	2.48	2.08	1.93
Maximum	7.60	5.44	3.22	2.84	2.38
Range	1.97	1.92	0.74	0.75	0.45

Table 5
Average Factor Structure Coefficients Based on 10 Bootstrap Samples of 100 Subjects Each

	Factor I	Factor II
Item1	.154(.166)*	-.342(.133)
Item2	.154(.115)	-.428(.079)
Item3	.144(.138)	-.374(.075)
Item4	.221(.106)	-.519(.080)
Item5	-.144(.120)	-.557(.092)
Item6	.116(.083)	-.618(.076)
Item7	-.114(.133)	-.286(.139)
Item8	-.108(.094)	-.570(.070)
Item9	.107(.111)	-.626(.044)
Item10	-.116(.093)	-.528(.098)
Item11	.102(.103)	-.456(.075)
Item12	-.101(.185)	-.290(.085)
Item13	.146(.150)	-.383(.112)
Item14	.106(.154)	-.454(.099)
Item15	.097(.076)	-.451(.117)
Item16	.140(.116)	-.355(.167)
Item17	-.301(.085)	-.360(.095)
Item18	.039(.155)	-.557(.067)
Item19	.237(.107)	-.447(.086)
Item20	.056(.102)	-.626(.111)
Item21	.636(.078)	-.036(.107)
Item22	.550(.178)	-.119(.212)
Item23	.600(.065)	-.124(.122)
Item24	.560(.140)	-.091(.117)
Item25	.558(.076)	-.153(.181)
Item26	.607(.099)	-.237(.104)
Item27	.588(.078)	-.213(.160)
Item28	.621(.101)	-.019(.146)
Item29	.618(.092)	.047(.104)
Item30	.732(.057)	-.077(.099)
Item31	.326(.105)	-.039(.173)
Item32	.505(.052)	-.203(.079)
Item33	.040(.143)	.185(.090)
Item34	.322(.064)	-.091(.079)
Item35	.542(.060)	-.138(.118)
Item36	.280(.127)	.240(.151)
Item37	.008(.156)	.176(.164)
Item38	.005(.187)	.205(.115)
Item39	.202(.140)	.075(.094)
Item40	.343(.123)	.268(.117)

*Structure coefficients are presented first followed by their standard deviations in parentheses.

Table 6
Descriptive Statistics for Eigenvalues Across 10 Bootstrap Samples of 200 Subjects Each

	I	II	III	IV	V
Mean	6.46	4.56	2.33	2.10	1.82
SD	0.50	0.29	0.18	0.12	0.15
Minimum	5.62	4.10	2.13	1.96	1.64
Maximum	7.07	5.03	2.56	2.29	2.12
Range	1.45	0.93	0.44	0.33	0.47

Table 7
Average Factor Structure Coefficients Based on 10 Bootstrap Samples of 200 Subjects Each

	Factor I	Factor II
Item1	.176(.095)*	-.319(.067)
Item2	.162(.084)	-.442(.060)
Item3	.217(.114)	-.352(.048)
Item4	.163(.083)	-.568(.077)
Item5	-.117(.069)	-.464(.061)
Item6	.090(.058)	-.668(.059)
Item7	-.139(.085)	-.352(.061)
Item8	-.094(.111)	-.602(.071)
Item9	.094(.089)	-.618(.029)
Item10	-.129(.069)	-.498(.066)
Item11	.136(.086)	-.498(.047)
Item12	-.123(.093)	-.316(.089)
Item13	.116(.068)	-.339(.108)
Item14	.154(.082)	-.477(.064)
Item15	.092(.056)	-.479(.103)
Item16	.149(.056)	-.382(.088)
Item17	-.325(.081)	-.367(.090)
Item18	.016(.035)	-.595(.034)
Item19	.276(.086)	-.432(.077)
Item20	.073(.088)	-.658(.045)
Item21	.646(.042)	-.018(.073)
Item22	.605(.063)	-.059(.062)
Item23	.580(.058)	-.123(.070)
Item24	.555(.082)	-.107(.065)
Item25	.532(.070)	-.160(.094)
Item26	.612(.069)	-.138(.074)
Item27	.614(.041)	-.217(.069)
Item28	.646(.051)	.020(.082)
Item29	.598(.082)	.104(.056)
Item30	.757(.038)	-.040(.070)
Item31	.341(.108)	-.100(.095)
Item32	.506(.048)	-.220(.078)
Item33	.163(.094)	.213(.095)
Item34	.308(.120)	-.086(.108)
Item35	.503(.069)	-.168(.077)
Item36	.311(.051)	.273(.070)
Item37	.051(.062)	.222(.088)
Item38	.062(.048)	.252(.094)
Item39	.233(.098)	.084(.142)
Item40	.399(.064)	.242(.091)

*Structure coefficients are presented first followed by their standard deviations in parentheses.

Table 8
Descriptive Statistics for Eigenvalues Across 10 Bootstrap Samples of 600 Subjects Each

	I	II	III	IV	V
Mean	6.21	4.37	2.08	1.91	1.74
SD	0.18	0.21	0.06	0.07	0.05
Minimum	5.82	3.93	2.00	1.80	1.65
Maximum	6.38	4.59	2.17	2.01	1.83
Range	0.56	0.66	0.17	0.21	0.18

Table 9
Average Factor Structure Coefficients Based on 10 Bootstrap Samples of 600 Subjects Each

	Factor I	Factor II
Item1	.183(.033)*	-.341(.038)
Item2	.170(.052)	-.385(.054)
Item3	.172(.047)	-.349(.039)
Item4	.187(.039)	-.571(.033)
Item5	-.129(.046)	-.516(.043)
Item6	.100(.038)	-.655(.032)
Item7	-.120(.043)	-.329(.054)
Item8	-.102(.057)	-.606(.034)
Item9	.101(.039)	-.616(.023)
Item10	-.121(.037)	-.520(.040)
Item11	.126(.053)	-.465(.042)
Item12	-.121(.082)	-.328(.042)
Item13	.089(.036)	-.333(.051)
Item14	.125(.062)	-.472(.048)
Item15	.106(.062)	-.467(.047)
Item16	.161(.051)	-.366(.047)
Item17	-.345(.035)	-.391(.048)
Item18	.010(.041)	-.612(.020)
Item19	.259(.042)	-.420(.061)
Item20	.072(.045)	-.642(.042)
Item21	.617(.045)	-.025(.054)
Item22	.607(.032)	-.061(.039)
Item23	.591(.019)	-.107(.022)
Item24	.575(.042)	-.095(.031)
Item25	.533(.039)	-.173(.040)
Item26	.609(.026)	-.188(.040)
Item27	.597(.028)	-.181(.052)
Item28	.611(.044)	.014(.053)
Item29	.588(.038)	.080(.034)
Item30	.748(.023)	-.061(.034)
Item31	.337(.058)	-.091(.025)
Item32	.495(.032)	-.225(.040)
Item33	.122(.065)	.199(.053)
Item34	.292(.071)	-.113(.051)
Item35	.521(.047)	-.169(.053)
Item36	.293(.070)	.263(.048)
Item37	.025(.056)	.174(.054)
Item38	.052(.042)	.271(.052)
Item39	.229(.053)	.108(.059)
Item40	.383(.032)	.299(.076)

*Structure coefficients are presented first followed by their standard deviations in parentheses.