DOCUMENT RESUME

ED 087 808 TM 003 431

AUTHOR Ambrosino, Robert J.
TITLE A Comparison of Four Methods for Estimating Common Factor Scores.
PUB DATE Nov 73
NOTE 21p.; Paper presented at Northeastern Educational Research Association (Ellenville, N.Y., November 1, 1973)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Comparative Analysis; *Factor Analysis; *Factor Structure; Oblique Rotation; Orthogonal Rotation; *Statistical Analysis; Validity

ABSTRACT

A comparison of four procedures for estimating common factor measurements was made using artificially synthesized "data" matrices. Score estimates were compared with respect to how well they approximated associated true factor scores and the extent of shrinkage in double cross-validation based on random samples. The Horst (1965), Bartlett (1937), and Anderson and Rubin (1956) methods gave what was judged as satisfactory estimates for the (artificial) populations of data. The cross-validational procedures showed the Horn (1965) method to yield highly unstable estimates. It was concluded that the method of using columns of the factor loading matrix as weights to be used in estimating factor measurements cannot be recommended for general applications since this procedure consistently provided highly unstable estimates. (Author)

ED 087808

TM 003 431

A COMPARISON OF FOUR METHODS FOR ESTIMATING COMMON FACTOR SCORES

Robert J. Ambrosino

Albany Medical College

Paper presented at the fourth annual convocation of the Northeastern Educational Research Association - Ellenville, New York - November 1, 1973.

The problem of factor score estimation may be introduced by discussing briefly the classical factor analysis model, which may be expressed in matrix terms as:

$$Z' = FS'_c + US'_u \tag{1}$$

where Z' = (pxN) raw data matrix scaled to have column means of zero and column standard deviations of unity, where p = the number of observed variates and N = the number of individuals or entities.

F = (pxm) f or-loading matrix for m derived common factors.

S'_c = (mxN) matrix o dividual scores on the derived common factors, usually referred to as the common factor score matrix.

U = (pxp) diagonal matrix whose non-zero entries identify standard deviations of the derived uniqueness variables.

S'_u = (pxN) matrix of individual scores on the associated unique factors, usually referred to as the uniqueness factor score matrix.

The model may be classified as an additive, linear model (Thurstone, 1947); equation (1) indicates that data can be represented as a sum of common portions (FS'_c) and unique portions (US'_u). In general, the derived common factors may be correlated or uncorrelated. The derived uniqueness variables are assumed to be uncorrelated among themselves and with the common variables. These assumptions may be represented algebraically in matrix terms as: $S'_uS_u = I$, and $S'_cS_u = \underline{0}$. For present purposes it will also be reasonable to treat common factors as mutually uncorrelated; that is, $S'_cS_c = I$.

Common portions of data, i.e., FS'_c in equation (1) are in general unobservable. Hence, at best it is possible to estimate common factor scores. Geometrically the problem of estimation becomes apparent. The usual geometric mod l for factor analysis represents the p tests as a bundle of unit-length vectors embedded in an N-dimensional Euclidean space. The derived common factors are represented as a set of m linearly independent vectors embedded in the same N-space. But these common factorvectors are outside the space determined by the origin and the end points of the test vectors. Estimates of common factor scores which are based on observed variates are thus often poor because they are based on information within the space of test variables (''eerman, 1964;Thomson,1951).

Theoretical speculation has led to the construction of several methods for estimating factor and component scores. These procedures have been largely developed within a least-squares regression framework. That is, several investigators have approached the problem of estimating common factor scores by first estimating $\underline{F}$ and $\underline{U}^2$ for some selection of $\underline{m}$ less than $\underline{p}$, and then deriving the factor measurements using some form of least-squares analysis treating the estimates of $\underline{F}$ and $\underline{U}^2$ as though they were equal to corresponding population values (see Horst, 1965). Since $\underline{S}_c$ in equation (1) equals the (Nxm) matrix of common factor score measurements, let $\hat{\underline{S}}_c$ represent the (Nxm) matrix of corresponding estimated factor scores, and $\hat{\underline{B}}$ represent the (pxm) matrix of estimation or regression weights. The general problem of estimation for these different least-squares methods may be expressed in matrix terms as: $\hat{S}_c = Z\hat{B}$, where $\hat{\underline{B}}$ depends on $\underline{F}$ and $\underline{U}$ alone, and is chosen in the case of each factor score estimation procedure to minimize certain errors of estimation.

McDonald and Burr (1967) presented the formal properties of four least-squares estimation methods and compared these properties with respect to four generally desirable properties of estimated factor scores. In essence they developed the rationale for estimation and described the differences among estimates given by these procedures with respect to theoretical criteria such as orthogonality, univocality, and conditional unbiasedness, which are discussed below. Harris (1967) discussed these same procedures and added a fifth (relatively crude) procedure that is often used or recommended, but is not generally considered as a standard method of estimation.

The four desirable properties of estimated factor scores which were discussed by McDonald and Burr are that: (a) estimated factor scores should approximate the associated true factor scores as closely as possible, i.e., the diagonal elements of the (mxm) matrix of cross-correlations between true and estimated scores should approximate unity; (b) the set of $\underline{m}$ estimated score vectors should be mutually orthogonal; (c) each vector of estimated factor scores should correlate zero with each vector of non-corresponding true factor scores (this condition identifies univocal factor scores; see Guilford and Michael, 1943); and (d) the estimated factor scores should be conditionally unbiased estimators of corresponding true factor scores.

Apparently, the choice of initial factoring method makes a difference in the estimation of factor score measurements. Harris (1967) and McDonald and Burr (1967) indicated that the choice of canonical factor analysis (Rao, 1955) as the initial factoring method produces estimated factor scores with particularly desirable properties. Canonical factor analysis is a special case of maximum likelihood factor analysis (Jöreskog, 1967; Lawley, 1940; Rao, 1955). Browne (1968) discussed the properties of several factor analytic techniques and made an empirical comparison of results given by these techniques. He indicated that in general, estimates of factor-loadings given by the maximum likelihood method are theoretically preferable because they are asymptotically efficient and there is a corresponding likelihood ratio test for assessing the fit of the factor model. His results, together with those of Harris and McDonald and Burr, suggest that maximum likelihood factor analysis provides a desirable basis for estimation of factor measurements.

Trites and Sells (1955) compared the unit weighted method and the fractionally weighted method for estimating factor scores, using correlation coefficients. The two methods gave practically identical results. From the standpoint of computation, the unit weighted method was the simpler and it was concluded that this was the more desirable of the two methods for practical applications. Baggaley and Cattell (1956) described certain exact and linear function estimates of oblique factor scores and discussed the conditions under which they were appropriate. They showed the extent of approximation in using factor-loadings in place of the exact regression weights in a 15 factor, 70 variable, 295 person problem. The correlations between the true and estimated scores ranged from .67 to .94 for the approximate procedure. Moseley and Klett (1964) empirically compared the results given by three methods of factor score estimation. Their results indicated that each of the methods were roughly equivalent insofar as the intercorrelations among score estimates and reliabilities were concerned. Horn (1965) described and empirically compared exact and approximate procedures for estimating factor scores, using coefficients of congruence (Tucker, 1951). His exact procedures, i.e., those which use some form of least-squares analysis, were correlated above .90 with one another. His approximate procedures, i.e., those which do not use a least-squares analysis, also correlated above .90 with one another. Wackwitz and Horn (1971) compared estimates given by exact and approximate procedures to true (population) factor scores, using a variety of criteria for comparison. Their results indicated that the weighted salients method (Horn, 1965) produced score estimates more closely matching associated true factor scores than any of the other methods of the study. As can be seen, there appears to have been little empirical research involving the estimation of individual factor measurements.

There appears to be practical value in comparing the results given by different estimation procedures with respect to indices such as predictive validities and the amount of shrinkage to be expected when these validities are examined with cross-validation procedures. The aim of this study was to compare four procedures for estimating common factor measurements using artificially synthesized "data" matrices. In essence, two major questions were raised: (a) how well will the factor score estimates derived using the four procedures approximate associated true factor scores which are available from the data simulation procedures? and (b) what will be the extent of shrinkage in the canonical correlations when the validities of estimates derived using the four estimation procedures are examined by cross-validation procedures?

Methodology

Data Simulation

Data for this study were computer simulated using the classical factor analysis model as represented by equation (1). Eight simple structure factor-loading matrices F were used to develop the common and unique portions of the population data for Ns of 200 and 300. Each of these Fs was chosen from the factor analysis literature with respect to a variety of criteria such as simplicity of loadings, sizes and variabilities of communalities, and ratio of factors to variables (see Table 1). Common and unique factor scores were generated for the population to fit the standard assumptions of mutual and joint orthogonality.

INSERT TABLE 1 ABOUT HERE

Generating the Cross-Validation Samples

For each combination of F and N, the cross-validation sample pairs served as the data base for the estimation of factor scores. These were determined by randomly splitting the (Nxp) population data matrix Z into two (nxp) halves, where n = N/2. For an appropriate permutation of columns, $Z' = [Z'_a \mid Z'_b]$. The synthetic data Z_a and Z_b were represented as $Z'_k = FS'_{ck} + US'_{uk}$, (k = a,b). While S_{ck} do not in general possess the (exact) properties of true (common and unique) parts described above, they were, nevertheless, taken as one representation of common and unique factor scores for the respective half samples. The use of half-sample data of this form is not inconsistent with the approach taken in many applications of factor analysis where it is assumed that a population data matrix fits the common factor model but that samples of observation vectors taken from this population may fit the model only approximately.

Factoring Method

In the case of each population of data, Z_a and Z_b were used to generate correlation matrices which served as starting points for maximum likelihood factor analysis. Maximum likelihood factor analysis (Lawley, 1940) was chosen as the factoring method of this study because the theoretical properties of factor scores derived using this method are relatively well understood (Harris, 1967; McDonald and Burr, 1967). Computer program UMLFA (unrestricted maximum likelihood factor analysis, Joreskog, 1966) was used to obtain the maximum likelihood estimates of factor-loadings and uniqueness variances used in the factor score estimation. UMLFA provided initial (untransformed) estimates of factor-loadings, $\hat{F}^*_k$ (k = a,b), and orthogonally (using varimax, Kaiser, 1958) transformed versions, $\hat{F}_k$, for each selection of a number of factors, m. This study used derived orthogonal solutions (Fs of the form F_oT, for

F_o an untransformed solution) because this made it possible to make direct comparisons of $\hat{F}_a$ and $\hat{F}_b$ to one another and to the corresponding F which was used in data simulation. In this sense it seems reasonable to use the sets of zero-order correlations computed between resulting matched vectors of true and estimated factor scores as one criterion for examining predictive validities.

Factor-Score Estimation

The reader will recall that for each combination of F and N, the cross-validation sample pairs, Z_a and Z_b served as the data base for factor score estimation. The methods chosen for estimating factor measurements may be identified with respect to the following formulas:

$$\hat{S}_{1k} = Z_k\hat{F}_k(\hat{F}_k'\hat{F}_k)^{-1} \quad \text{Horst (1965)} \tag{2}$$

$$\hat{S}_{2k} = Z_k\hat{U}_k^{-2}\hat{F}_k(\hat{F}_k'\hat{U}_k^{-2}\hat{F}_k)^{-1} \quad \text{Bartlett (1937)} \tag{3}$$

$$\hat{S}_{3k} = Z_k\hat{U}_k^{-2}\hat{F}_k(\hat{F}_k'\hat{U}_k^{-2}R\hat{U}_k^{-2}\hat{F}_k)^{-1/2} \quad \text{Anderson and Rubin (1956)} \tag{4}$$

$$\hat{S}_{4k} = Z_k\hat{F}_k \quad \text{Horn (1965)} \tag{5}$$

$\hat{F}_k$ and $\hat{U}_k$ $(k = a,b)$ represent the maximum likelihood estimates of the corresponding population F and U in the case of each sample of data. R is the variance-covariance (or correlation) matrix associated with Z and $\hat{F}$.

The $\hat{S}_{jk}$ $(j = 1,2,3,4;\ k = a,b)$ are matrices of order (nxm) and may be expressed alternately as $\hat{S}_{jk} = Z_kE_{jk}$, where the E_{jk} are (pxm) matrices of estimation weights corresponding to that portion to the right of the Z_k matrix in (2) - (5).

The Horst, Bartlett, and Anderson and Rubin methods were selected with reference to the theoretical factor score properties discussed by McDonald and Burr. In general, regardless of initial factoring method used, the Horst and Bartlett estimates are univocal and conditionally unbiased estimators of corresponding true scores. Anderson and Rubin estimates are in general orthogonal. What has been termed the Horn method was included in this study because it represents a quick and convenient means of estimation which is sometimes used or recommended for general applications (Horn, 1965; Wackwitz and Horn, 1971).

Examination of Validities

A. Two criteria were used to examine the quality of estimates for each method of estimation, in the case of each sample of data: (1) the set of m zero-order correlations ($r_{\hat{t}e}$) computed between the corresponding vectors of estimated scores identified as $\hat{S}_{jk}$ (j = 1,2,3,4; k = a,b) and true scores identified as S_{ck}; and (2) the set of m canonical correlations (R_{ci}, i = 1,2,...,m) computed between the (optimally) linearly weighted composite of estimated scores, $\hat{S}_{jk}$ and the (optimally) linearly weighted composite of true scores, S_{ck}. For the latter, the (undeviated) root mean square, written RMS = $\left[m^{-1}\sum_{1}^{m} R^2_{ci}\right]^{1/2}$, was computed as a summary index of predictive validity.

B. A (double) cross-validation paradigm was employed to examine the stability of estimates given by the four procedures. The following notation is designed to facilitate an understanding of these cross-validation procedures. In the case of each population of data, consider Z_k as the hypothesis-generating sample and $Z_{k'}$ as the corresponding validation sample (k' = b if k = a, k' = a if k = b). Factor score estimates, $\hat{S}_{jk}$ (j = 1,2,3,4) and estimation weights, E_{jk} were calculated for the initial or hypothesis-generating sample. RMSs were calculated for canonical correlations computed between initial factor score estimates, $\hat{S}_{jk}$ and true scores, S_{ck}.

The E_{jk} derived using the initial or hypothesis-generating sample were then applied to the data of the corresponding validation sample, creating four new sets of (cross-validity) factor score estimates identified as $\hat{S}^*_{jk} = Z_k E_{jk'}$. RMSs were calculated for canonical correlations computed between cross-validity estimates, $\hat{S}^*_{jk}$ and true scores, S_{ck}. The reader may wish to refer to Figure 1, which is a schematic diagram of the cross-validation procedures described above.

- - - - - - - - - - - - - -

INSERT FIGURE 1 ABOUT HERE

- - - - - - - - - - - - - -

The shrinkage between RMSs computed for each method identified as $\hat{S}_{jk}$ and $\hat{S}^*_{jk}$ corresponds to the stability of the estimates based on the initial hypothesis-generating sample $\underline{Z}_k$. (Note- when $\hat{\underline{F}}_k$ and $\hat{\underline{U}}_k$ are closer approximations to the corresponding population $\underline{F}$ and $\underline{U}$ than are $\hat{\underline{F}}_{k'}$ and $\hat{\underline{U}}_{k'}$, $\underline{\text{RMSs}}$ computed for estimates based on weights given by $\underline{E}_{jk}$ will be higher than those computed for estimates based on weights given by $\hat{\underline{E}}_{jk'}$. That is, the (direction of) shrinkage will be positive. If, however, $\hat{\underline{F}}_{k'}$ and $\hat{\underline{U}}_{k'}$ provide the closer approximations to $\underline{F}$ and $\underline{U}$, then it is possible to have negative shrinkage. That is, $\underline{\text{RMSs}}$ computed for estimates based on the validation sample weights, $\underline{E}_{jk'}$ will be higher than those computed for estimates based on weights given by the original hypothesis-generating sample. Thus, in the sense that negative shrinkage can be observed here, the present study is an unconventional cross-validation study).

Results

Several summary statistics based on results from each of the comparisons noted above, are presented below for each method of estimation and for each selection of $\underline{F}$ and $\underline{N}$, and for all values of m/p, the ratio of the number of factors to the number of variables for the respective solution. These statistics allow the reader readily to examine for himself the quality of estimates given by the different methods with respect to the factor score criteria of this study.

For each estimation method and for each combination of $\underline{F}$ and $\underline{N}$, the following summary statistics are given: (1) the average zero-order correlation ($\bar{\underline{r}}_{te}$) between the true and estimated factor scores; (2) the average canonical root mean square ($\overline{\underline{\text{RMS}}}$), which is included as a summary index of predictive validity; and (3) the average residual between the canonical root mean square ($\overline{\underline{\text{RMS}}}_r$) computed for the hypothesis-generating and the validation estimates. All averages are simply unweighted means computed across results for both halves of data.

Tables 2 - 5 include these summary statistics. Each table contains statistics for a single estimation method. The reader should recognize that a single (small value) selection of $\underline{m}$ was always used when N = 200. The final (row) entries in each table represent (unweighted) averages for each statistic, computed across all the initial $\underline{F}$s included. Although different combinations of row complexities, (sizes and variabilities of) communalities and ratios of factors to variables were represented, it seems reasonable to summarize results using these composite statistics. For all these tables, there are no entries for the Browne (1968) data with N = 200 and m = 3, for the Overall and Porterfield (1963) data with N = 200 and m = 3, and for the entire Conry (1965) data. For no clearly apparent reason, the r_{te}s, canonical Rs and root mean squares computed for these data were extremely small. Thus, these data were censored for this summary in view of the writer's belief that these data sets are not generally comparable with the resr.

The data of tables 2,3, and 4 indicate that the Horst, Bartlett and Anderson and Rubin methods gave what may be judged as satisfactory results for each combination of F and N with respect to the estimates approximating associated true scores and with respect to cross-validational shrinkage. Inspection of tables 2 - 5 indicates that there were only small differences among these summary statistics given for N = 200 and N = 300, when using the same ratio of factors to variables.

Perhaps the most striking observation from these tables is the degree of similarity of results for the first three methods of factor score estimation and the finding of essentially no shrinkage for these methods. Had the ratio of sample to population size, n/N been smaller, the probability of negative shrinkage would no doubt have been lowered; that negative shrinkage was occasionally found for sampling ratios of 1/2 clearly does not imply that opposite sample Fs will produce higher cross-sample validities in general. Nevertheless, it at least seems reasonable to suggest that score estimates based on the first three methods are apt to be relatively stable for many applications.

Another finding which is manifest from a study of the summary tables is that average r_{te}s or RMSs for the individual Fs are quite highly correlated with the average communalities for these Fs. This is of course not surprising. The one point that is interesting, however, is that $\bar{r}_{te}$s are greater than .70 despite the fact that the smallest average communalities are about .50. Highest levels of r_{te}, about .90 were reached for the F matrix from the work of Wiggins and Lovell (1965), the average communality for which was .69.

One fact that was initially unsettling is that, depending on the population F, the RMSs sometimes increased and sometimes decreased as the ratio, m/p was increased. For the Browne, Overall and Porterfield and Wiggins and Lovell Fs, it is seen that RMSs increase markedly as m/p is increased. But for the Bechtoldt (1961) and the Emmett (1949) data, the opposite effect occurs; as the half-sample m/p ratio is increased to the population ratio, the RMSs decrease. The reason apparently, is that for the latter two data sets (Bechtoldt and Emmett), population Fs have relatively lower complexity rows for an orthogonal solution than do the three former sets. It is suggested that when the sample ratios m/p were set to be smaller than the population ratio for the complex Fs of Browne, Overall and Porterfield and Wiggins and Lovell data, the maximum likelihood factoring procedure resulted in derived half-sample Fs whose columns may not have clearly matched any of those of the associated population F; that the half-sample Fs were instead "stretched" across the true common space. As the sample ratio m/p was increased to the population ratio, for the more complex data, the individual factors tended better to match the population factors, thus the RMSs reached their highest values for largest m/p ratios for these data.Perhaps the conclusion that one ought to make in this context is that RMSs associated with the largest values of m/p are the ones which ought to be given primary attention for interpretation of all methods. When one compares the four methods using these rows alone, however, the same general conclusions are reached about the relative merits of these four methods of estimation. This can be seen by inspecting the final row entries of each of these tables.

As must be true based on analytical study, the average root mean square statistic ($\overline{\text{RMS}}$) was identical for both the Horst and the Horn methods and for both the Bartlett and Anderson and Rubin methods, across the different specifications of F included in these tables. However, as can be seen from tables 2 and 5, the true-estimated correlations were distinctly lower for the Horn method than for the Horst method within each half sample for each specification of F and N.

Despite the fact that the Horn method appeared to give what may be judged as satisfactory true-estimated correlations and canonical root mean squares, the cross-validational procedures showed this method to yield highly unstable estimates. As can be seen from table 5, the cross-validational shrinkage was substantially higher for this method across all specifications of F and N, than the shrinkage for the other three methods. Recalling information from table 5, the extent of shrinkage was generally low for the first or larger factors, but extremely high for the subsequent factors, when using what has been termed the Horn method.

No noticeable patterns in the results given by the four estimation methods were observed between Ns of 200 and 300, across all specifications of initial population Fs. Apparently, the difference in size of N was not large enough to produce a noticeable effect for these data.

- - - - - - - - - - - - - - - - -
INSERT TABLES 2-5 ABOUT HERE
- - - - - - - - - - - - - - - - -

Conclusions

Based on the results presented above, the following general conclusions may be drawn. The first three methods of this study (Horst, Bartlett and Anderson and Rubin) gave practically identical results with respect to approximating true factor scores and with respect to cross-validational shrinkage. Each of these methods gave what may be judged as satisfactory results for these (artificial) populations of data. It has been shown that for general applications, it is not unreasonable to expect vectors of estimated factor scores based on any of the first three methods of this study to correlate upwards of .70 with underlying true factor scores when average communalities for the initial population Fs are above .50.

Despite the fact that the crude method attributed to Horn appeared to give satisfactory within sample true-estimated correlations and canonical root mean squares, the cross-validational procedures of this study showed this method to yield highly unstable estimates. The conclusion here, is that the method of selecting columns of F to be used as weights for estimating factor measurements cannot be recommended for general applications since this procedure consistently provided highly unstable estimates.

Extensions which might be considered in future studies of the present type include the following: A single cross-validation sample pair (Z_a and Z_b) was analyzed for each combination of F and N. In this study, the objective was to examine a wide variety of Fs and selections of m and N, on the assumption that point estimates of validity coefficients would suffice for the inter-method comparisons of the different methods for estimating factor scores. Studies which included several cross-validation sample pairs in the case of each population of data, would make it possible to generate approximations to the distributions of validities and shrinkage for each of several estimation methods across all selections of F and N.

To provide a greater opportunity to observe the effect of sample size on the quality of estimates given by the four respective methods, future studies might include a wider variety of Ns, for example, $N = 200$, 500, 700, and 1000. A larger set of Ns would also make it possible to generate cross-validation samples whose sample sizes were some fraction of the initial population size other than one half. For each of the Ns in this expanded set, the analyses of future studies probably should include only those selections of m that are equal to the number of factors for the associated population F.

It might be interesting to include other common factor solutions to derive $\hat{F}_a$ and $\hat{F}_b$ for the different specifications of F and N. Normal varimax was used to obtain the orthogonally transformed versions of $\hat{F}_a$ and $\hat{F}_b$ in the case of each sample of data. This transformation algorithm had also been used to derive several of the initial population Fs. Oblique transformations (see Harris and Kaiser, 1964; Hofmann, 1970) perhaps ought to be investigated in future studies of this kind.

Of course, further variations on the present theme could take on many forms. Multivariate analysis typically involves the estimation of so many parameters that one cannot in a single study vary all relevant dimensions of parameter investagation. While analytical studies are clearly essential for methodological progress, studies of the present variety appear to have considerable value for refining knowledge and for making judgments about practical uses of quantitative methods.

References

Anderson, T.W. and Rubin, H. "Statistical inference in factor analysis." Procedings of the Third Berkely Symposium on Mathematical Statistics and Probability, 5(1956), 11-150.

Baggaley, A. and Cattell, R.B. "A comparison of exact and approximate linear function estimates of oblique factor scores." British Journal of Statistical Psychology, IX(1956), 83-86.

Bartlett, M.S. "The statistical conception of mental factors." British Journal of Psychology, XXVIII(1937), 97-104.

Bechtoldt, H.P. "An empirical study of the factor analysis stability hypothesis." Psychometrika, 26(1961), 405-432.

Browne, M.W. "A comparison of factor analytic techniques." Psychometrika, 33(1968), 267-334.

Conry, R.F. "Canonical relationships among measures of concept formation, information processing, and other hypothesized cognitive factors." Unpublished Master's thesis, University of Wisconsin, 1965.

Emmett, W.G. "Factor analysis by Lawley's method of maximum likelihood." British Journal of Psychology:Statistical Section, 2(1949), 90-97.

Guilford, J.P. and Michael, W.B. "Approaches to univocal factor scores." Psychometrika, 13(1948), 1-22.

Harris, Chester W. "On factors and factor scores." Psychometrika, 32(1967), 363-379.

Harris, Chester W. and Kaiser, H. "Oblique factor analytic solutions by orthogonal transformations." Psychometrika, 29(1964), 347-362.

Heerman, E.F. "The geometry of factorial indeterminacy." Psychometrika, 29(1964), 371-381.

Hofmann, R.J. "The obliquimax transformation." Unpublished Doctoral thesis, State University of New York at Albany, 1970.

Horn, J.L. "An empirical comparison of methods for estimating factor scores." Educational and Psychological Measurement, 25(1965), 313-322.

Horst, Paul. Factor Analysis of Data Matrices. New York: Holt Rinehart and Winston, 1965.

Jöreskog, K.G. "UMLFA - a computer program for unrestricted maximum likelihood factor analysis." Research Memorandum 66-20. Educational Testing Service, 1966.

Jöreskog, K.G. "Some contributions to maximum likelihood factor analysis." Psychometrika, 32(1967), 443-482.

Kaiser, H. "The varimax criterion for analytic rotation in factor analysis." Psychometrika, 23(1958), 187-200.

Lawley, D.N. "The estimation of factor loadings by the method of maximum likelihood." Procedings of the Royal Society of Edinburg, 60(1940), 64-82.

Maxwell, E.A. "Recent trends in factor analysis." Journal of the Royal Statistical Society, Series A, 124(1961), 49-59.

McDonald, R.P. and Burr, E.J. 'A comparison of four methods of constructing factor scores." Psychometrika, 32(1967), 381-401.

Moseley, E.C. and Klett, C.J. "An empirical comparison of factor scoring methods." Psychological Reports, 14(1964), 179-184.

Overall, J.E. and Porterfield, J.L. "Powered vector method of factor analysis." Psychometrika, 28(1963), p. 420.

Rao, C.R. "Estimation and tests of significance in factor analysis." Psychometrika, 20(1955), 93-112.

Thomson, G.H. The Factorial Analysis of Human Ability. London: University of London Press, 1951.

Thurstone, L.L. Multiple Factor Analysis. Chicago: University of Chicago Press, 1947.

Trites, D.K. and Sells, S.B. "A note on alternative methods for estimating factor scores." Journal of Applied Psychology, 39(1955), 455-456.

Tucker, L.R. "A method for synthesis of factor analysis studies." Department of the Army, Personnel Research Section Report No. 984, 1951.

Wackwitz, J.H. and Horn, J.L. "On obtaining the best estimates of factor scores within an ideal simple structure." Multivariate Behavioral Research, 6(1971), 389-408.

Wiggins, J.S. and Lovell, V.R. "Communality and favorability as sources of method variance in the MMPI." Educational and Psychological Measurement, 25(1965), 399-41[illegible].

*-The following entry, while out of order, was also referenced in this paper

Harman, H.H. Modern Factor Analysis. Chicago: University of Chicago Press, 1967.

TABLE 1

Sources for F Matrices Which Were
Used in Construction of Population Data

Source for F	X_{h^2}	s_{h^2}	m/p
Bechtoldt (1961)	.661	.158	6/17
Browne (1968)	.472	.293	4/12
Conry (1965)	.657	.130	6/17
Emmett (1949)	.639	.140	4/9
Harman (1967)	.500	.140	4/20
Maxwell (1961)	.535	.230	4/10
Overall-Porterfield (1963)	.673	.080	5/15
Wiggins-Lovell (1965)	.694	.110	3/13

TABLE 2

Summary Ststistics for the Horst (1965)
Classical Least-Squares Method of Estimation

$\underline{F}$	m/p	$\bar{r}_{te}$	$\overline{RMS}$	$\overline{RMS}_r$	$\bar{r}_{te}$	$\overline{RMS}$	$\overline{RMS}_r$
		N = 200			N = 300		
Bech-toldt	4/17	.893	.900	.004	.888	.897	-.002
	6/17				.883	.895	-.015
Browne	3/12				.822	.814	.003
	4/12				.798	.851	.000
Emmett	2/9	.847	.856	-.008	.848	.854	-.002
	3/9				.801	.828	-.008
	4/9				.622	.738	-.003
Harman	2/20	.775	.782	-.001	.787	.794	.002
	3/20				.706	.765	.001
	4/20				.818	.843	-.001
Maxwell	3/10	.736	.780	-.006	.762	.798	.007
	4/10				.740	.808	-.010
Overall-Porter.	3/15				.828	.856	.002
	5/15				.905	.923	.001
Wiggins-Lovell	2/13	.874	.884	-.001	.888	.892	-.002
	3/13				.932	.936	.000
Average		.825	.840	-.002	.814	.852	-.002
					(.814)	(.870)	(-.005)

Note - Entries in parentheses are those averages computed for selections of the sample ratio, m/p, that was equal to the ratio, m/p, for the respective population $\underline{F}$.

TABLE 3

Summary Statistics for the Bartlett (1937) Method of Estimation

F	m/p	$\bar{r}_{te}$ (N = 200)	$\overline{RMS}$ (N = 200)	$\overline{RMS}_r$ (N = 200)	$\bar{r}_{te}$ (N = 300)	$\overline{RMS}$ (N = 300)	$\overline{RMS}_r$ (N = 300)
Bech-toldt	4/17	.893	.901	.005	.891	.898	.008
	6/17				.855	.870	-.006
Browne	3/12				.772	.816	.007
	4/12				.812	.825	.007
Emmett	2/9	.853	.860	-.002	.854	.860	-.002
	3/9				.792	.824	.001
	4/9				.586	.723	.008
Harman	2/20	.791	.800	-.003	.762	.786	.005
	3/20				.686	.767	-.004
	4/20				.817	.841	-.004
Maxwell	3/10	.724	.768	-.012	.754	.779	-.001
	4/10				.726	.806	-.016
Overall-Porter.	3/15				.838	.861	-.001
	5/15				.905	.924	.000
Wiggins-Lovell	2/13	.879	.888	-.003	.898	.904	-.004
	3/13				.935	.937	-.002
Average		.828	.843	-.002	.805	.839	.000
					(.806)	(.846)	(-.002)

TABLE 5

Summary Statistics for the Anderson and Rubin (1956) Method of Estimation

$\underline{F}$	m/p	$\bar{r}_{te}$ (N = 200)	$\overline{RMS}$ (N = 200)	$\overline{RMS}_r$ (N = 200)	$\bar{r}_{te}$ (N = 300)	$\overline{RMS}$ (N = 300)	$\overline{RMS}_r$ (N = 300)
Bech-toldt	4/17	.898	.901	.005	.895	.898	.008
	6/17				.864	.870	-.006
Browne	3/12				.807	.816	.007
	4/12				.816	.825	.007
Emmett	2/9	.814	.860	-.002	.856	.860	-.002
	3/9				.805	.824	.000
	4/9				.556	.723	.008
Harman	2/20	.790	.800	-.003	.741	.786	.005
	3/20				.695	.767	-.004
	4/20				.826	.841	-.004
Maxwell	3/10	.734	.768	-.012	.767	.790	.010
	4/10				.751	.806	-.016
Overall-Porter.	3/15				.838	.861	-.001
	5/15				.907	.924	.000
Wiggins-Lovell	2/13	.880	.888	-.003	.898	.904	-.005
	3/13				.935	.937	-.002
Average		.823	.842	-.003	.810 (.808)	.840 (.846)	.000 (-.002)

TABLE 5

Summary Statistics for the Horn (1965)
Method of Estimation

$\underline{F}$	m/p	$\bar{r}_{te}$ (N = 200)	$\overline{RMS}$ (N = 200)	$\overline{RMS}_r$ (N = 200)	$\bar{r}_{te}$ (N = 300)	$\overline{RMS}$ (N = 300)	$\overline{RMS}_r$ (N = 300)
Bech-toldt	4/17	.744	.900	.372	.748	.897	.208
	6/17				.680	.895	.228
Browne	3/12				.831	.814	.272
	4/12				.785	.852	.309
Emmett	2/9	.743	.856	.279	.743	.871	.294
	3/9				.706	.835	.285
Harman	2/20	.607	.782	.091	.648	.794	.196
	3/20				.619	.765	*
	4/20				.668	.843	.328
Maxwell	3/10	.712	.780	.209	.710	.798	.197
	4/10				.666	.808	.284
Overall-Porter.	3/15				.820	.856	.099
	5/15				.876	.923	.181
Wiggins-Lovell	2/13	.861	.884	*	.872	.892	*
	3/13				.894	.936	.206
Average		.733	.840	.237	.764	.852	.223
					(.754)	(.870)	(.231)

Note - Asterisk identifies those cases where the $\underline{RMS}$ was not calculated for the validation estimates and therefore no residual was found.

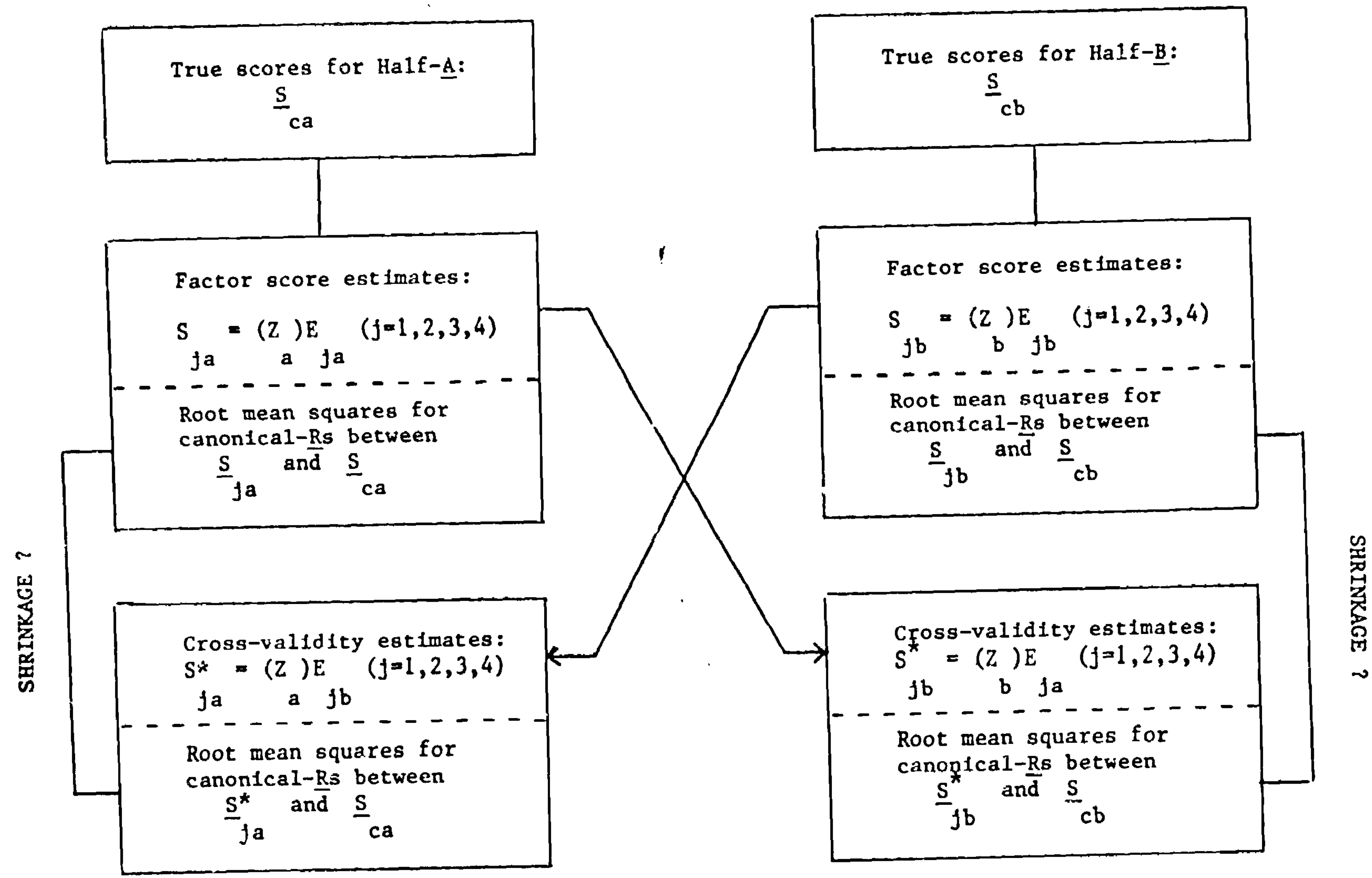

FIGURE 1 - Schematic diagram of factor score cross-validation procedures.