DOCUMENT RESUME

ED 072 045 TM 001 520

AUTHOR	Keown, Lauriston L.; Hakstian, A. Ralph
TITLE	Some Notes on the Choice of Measure of Association for the Component Analysis of Likert Scale Data.
INSTITUTION	Alberta Univ., Edmonton. Div. of Educational Research.
REPORT NO	RIR-72-3
PUB DATE	Mar 72
NOTE	22p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Attitude Tests; *Componential Analysis; Correlation; Data Analysis; *Educational Research; *Mathematical Models; Psychometrics; *Statistical Studies; Tables (Data); Technical Reports
IDENTIFIERS	*Likert Type Attitude Scale

ABSTRACT

This study involved an investigation of the use of Pearson r, tetrachoric r, phi, phi/phi maximum, and Kendall Tau-B coefficients as measures of association for the incomplete principal components analysis of simulated Likert scale attitudinal data, based on a known factor pattern and possessing different types of severe departures from normality. The results suggested that in addition to being based on few assumptions, Tau-B was most robust, with respect to distributional distortion, with large or small samples, and this coefficient was followed by the Pearson r. The measures based on 2 x 2 tables--tetrachoric r, phi, and phi/phi maximum--tended to be less robust and were seen to be adversely affected by uneven marginal splits, a condition generally present with Likert scale data. (Author)

research and information report

division of educational research
faculty of education
the university of alberta
edmonton, canada

Division of Educational Research
University of Alberta

RIR-72-3

SOME NOTES ON THE CHOICE OF MEASURE OF ASSOCIATION FOR THE COMPONENT ANALYSIS OF LIKERT SCALE DATA

LAURISTON L. KEOWN and A. RALPH HAKSTIAN

University of Alberta

March, 1972

The contents of this report are not to be cited without permission of the authors.

ABSTRACT

This study involved an investigation of the use of Pearson r, tetrachoric r, ϕ, ϕ/ϕ_{max}, and Kendall Tau-B coefficients as measures of association for the incomplete principal components analysis of simulated Likert scale attitudinal data, based on a known factor pattern and possessing different types of severe departures from normality. The results suggested that in addition to being based on few assumptions, Tau-B was most robust, with respect to distributional distortion, with large or small samples, and this coefficient was followed by the Pearson r. The measures based on 2 x 2 tables--tetrachoric r, ϕ, and ϕ/ϕ_{max}--tended to be less robust and were seen to be adversely affected by uneven marginal splits, a condition generally present with Likert scale data.

SOME NOTES ON THE CHOICE OF MEASURE OF ASSOCIATION FOR THE COMPONENT ANALYSIS OF LIKERT SCALE DATA[1]

LAURISTON L. KEOWN and A. RALPH HAKSTIAN

University of Alberta

One of the most common summative scales for the measurement of attitudes is that known as the Likert scale, in which the respondent indicates his degree of agreement with a particular statement by checking his position on a five- or seven-point continuum (scales with varying numbers of points have been used, but the issue of the optimal number of scale values has been treated elsewhere and is not of interest in the present paper). Often those engaged in research on attitudes and, in particular, in development of attitude measures use correlational and factor analytic procedures to identify homogeneous and theoretically interesting groupings of such scaled items, and the question arises as to the most defensible procedures for such analyses. To begin with, what assumptions can we make regarding the level of measurement scale with such data? Also, what underlying distributional assumptions are warranted? As Carroll (1961) clearly noted, there is nothing preventing our calculating a Pearson product-moment correlation coefficient, for example, on *any* set of data, but the interpretation that attaches to the resulting number has much to do with how closely the data conform to an appropriate statistical model.

Thus, we *could* calculate Pearson *r*'s with Likert scale data. On the other hand, one could disregard the assumption of interval-scale measurement by dichotomizing the obtained distributions, presumably at a point as near as possible to the median, and using measures of association based on 2 x 2 tables--

a procedure that has often been recommended for attitude scale data. An obvious choice for such a measure of association is a point statistic such as the $\underline{\phi}$ coefficient, which has been discussed in this context by Demaree and Smith (1957). This coefficient has been seen to result, in some cases, in spurious "difficulty" factors, however (Carroll, 1945; Ferguson, 1941), that is, one or more factors reflecting the fact that the items differed in difficulty level. This shortcoming of the $\underline{\phi}$ coefficient should not, logically, be present with attitudinal data. Further, the limits for $\underline{\phi}$ are ±1 only if the marginal distributions are of identical shape. This latter defect can be compensated for somewhat by dividing an obtained $\underline{\phi}$ coefficient by the maximum possible value of $\underline{\phi}$ for that particular pair of marginal distributions (Cureton, 1959), but as Carroll (1961) demonstrated, the underlying correlation surface of the resulting $\underline{\phi}/\underline{\phi}_{max}$ statistic is somewhat improbable. Nonetheless, at least one factor analysis text (Guertin and Bailey, 1970) recommends $\underline{\phi}/\underline{\phi}_{max}$ in such situations. Probably the most widely recommended measure of association for data in a 2 x 2 table has been the tetrachoric correlation coefficient, $\underline{r}_t$ (see, for example, Carroll, 1961; Wherry and Gaylord, 1944). Although no assumptions are made, with $\underline{r}_t$, regarding the scale of measurement of the variables involved, the assumption <u>is</u> made of an underlying bivariate normal correlation surface. Strictly speaking, $\underline{r}_t$, as calculated by its most precise formula, does not share the defect of $\underline{\phi}$ of being dependent upon the similarity of shape of the marginal distributions.

One solution to the problem of assumptions regarding underlying bivariate distributions is to use a nonparametric or distribution-free measure of association, such as Kendall's Tau, or the version suited for ties, Tau-B (Kendall, 1962; Kruskal, 1958). Carroll (1961), however, pointed out that Tau does not very effectively adjust for the scaling error resulting from broad grouping or <u>censoring</u>. Several possible measures of association exist, then, for use with data such as Likert scale items, and the question arises as to which are best as starting points for the subsequent factor analytic treatment of the data. The purpose of the present study was to compare the measures of association noted, $\underline{r}$, $\underline{r}_t$, $\underline{\phi}$, $\underline{\phi}/\underline{\phi}_{max}$, and Tau-B, in terms of the stability and robustness of correlation and rotated factor matrices over Likert scale data varying as to the shape of distributions of the variables.

Method

Two factorially simple data sets were generated using computer simulation procedures, according to the common-factor model:

$$Z = XF' + U, \qquad (1)$$

where $\underline{Z}$, of order $\underline{N}$ persons x $\underline{p}$ variables, is the matrix of standardized manifest variables, $\underline{X}$, of order $\underline{N}$ persons x $\underline{k}$ common factors, is the matrix of common-factor scores, again scaled to have zero mean and unit variance, $\underline{F}$, of order $\underline{p}$ x $\underline{k}$, is the primary-factor pattern matrix, and $\underline{U}$, of order $\underline{N}$ x $\underline{p}$, is the matrix of the unique parts of the manifest variables. In the present case, for simplification, the columns of $\underline{X}$ were approximately orthogonal (i.e., the factors were hypothesized to be strictly uncorrelated in the population), and were approximately normal. Two such matrices, $\underline{X}$, were generated, one for $\underline{N}$ = 200, the other, for $\underline{N}$ = 100, each with $\underline{k}$ = 5 columns. A complexity-one orthogonal factor pattern matrix, $\underline{F}$--for 20 variables and five factors, and displayed in Table 1--was constructed and the common parts of the variables were simulated by $\underline{XF'}$. The pattern coefficients were chosen so that the variables so generated varied in communality from .36 to .72, with a mean communality of .52. Next, the matrix $\underline{U}$, either 200 x 20 or 100 x 20, containing the unique parts of the variables was generated by first constructing matrices of approximately normally distributed, approximately orthogonal columns of

Insert Table 1 about here

scores and then scaling these columns to the metric of the uniquenesses, the latter dictated by the earlier fixed communalities.

The columns of the matrix $\underline{Z}$--the latter formed by fully column-standardizing $\underline{XF}' + \underline{U}$--containing near-normal manifest variable scores, were then transformed to seven-point Likert scale variables, according to the five distributions shown in Table 2. The logic employed in constructing the Likert scale distributions was that, given an underlying normal population distribution for each attitude variable, distortions due to (1) sampling and (2) phrasing of the item stems could result in distributions similar to those displayed in Table 2. Additionally, the censoring (see Carroll, 1961) involved in Likert scaling constitutes an additional scaling error that can be seen to contribute to a certain degree of distortion.

Insert Table 2 about here

In the seven-point distribution referred to as normal, only this latter scaling error was operative. In the rectangular distribution, some abnormal polarization was additionally presumed operative, so that the extreme and next-to-extreme categories were, in effect, chosen relatively more frequently than in the underlying distribution. The opposite distortion is seen in the central distribution, in which an abnormally large number of middle or undecided responses were presumably recorded. The positive skew distribution reflects distortion due to a trend toward the left-hand pole of the scale for each variable. In the mixed skew distribution--the only one in which the same distortion was not presumed for each variable--the odd-numbered variables were positively skewed, and the even-numbered, negatively skewed (the latter skew indicating a trend toward the right-hand pole of the scale).

For each distribution, five correlation matrices were constructed. In the first, standard Pearson product-moment correlation coefficients, $\underline{r}$, were calculated among the Likert scaled variables. Secondly, tetrachoric correlations were obtained, r_t, by dichotomizing each distribution at a point yielding as nearly as possible, equal proportions in each category. Although a reasonably precise formulation is available for computing r_t (see Kendall and Stuart, 1958), it involves the evaluation of a complex series expression, and the computing algorithm has been found to be relatively unstable, often yielding no usable value for r_t. As reviewed by Castellan (1966), however, several approximations for r_t exist, and although not the best approximation by Castellan's standards, that obtained by the so-called cosine-pi formula is a generally good approximation and is unquestionably the one in greatest current use. This formula, then, was used in computing r_t in the present study; that is

$$r_t \simeq \cos \pi \left[\frac{1}{1 + \sqrt{bc/ad}} \right],$$

where a, b, c, and d are obtained from the familiar bivariate frequency table

$$\begin{array}{c|c|c|} & - & + \\ \hline + & a & b \\ \hline - & c & d \\ \hline \end{array} \quad . \tag{2}$$

As pointed out by Glass and Stanley (1970), this approximation for r_t will be in error as a function of the degree of departure of the marginal proportions from .50. As found by Castellan (1966), however, most approximations of r_t share this defect.

Next, standard ϕ coefficients were obtained, again by first dichotomizing the distributions at a point as near as possible to the median. Following the computations of the matrices of ϕ coefficients, new matrices in which the ϕ coefficients were divided by the maximum possible (in absolute value) ϕ coefficient for the particular distributions involved, yielding the coefficient ϕ/ϕ_{max}. If we use the fourfold layout in (2), ϕ/ϕ_{max} is given, as shown by Cureton (1959), by

$$\phi/\phi_{max} = (bc - ad)/b'c', \ \phi \text{ positive;} \tag{3}$$

$$\phi/\phi_{max} = (bc - ad)/a'd', \phi \text{ negative}, \qquad (4)$$

where, in (3), b' and c' are obtained by constructing a table with the same marginal frequencies as earlier, but with a or d (whichever is smaller) set to zero, and, in (4), a' and d' are obtained by, again maintaining the marginal frequencies, but setting the smaller of b or c to zero.

Finally, Kendall Tau-B coefficients were obtained (see Kendall, 1962). Tau-B--the version of Kendall's rank-order correlation coefficient suitable for data with tied ranks--equals the standard Tau (or Tau-A) coefficient if no ties exist, but, of course, with attitude scale data many ties will be present.

In general, matrices of r_t, ϕ/ϕ_{max}, or Tau-B are not gramian--a fact that has occasionally been noted as a criticism of these coefficients, insofar as subsequent factor analytic work is concerned. Other things being equal, however, we regard this defect as relatively minor. Additionally, it can be seen from Table 2, that for the coefficients requiring a median dichotomization, r_t, ϕ, and ϕ/ϕ_{max}, some departure from equal proportions in each category was inevitable--a situation certainly true of "real" attitude scale data.

After computation of the R matrices, principal components analyses were performed, and the component pattern matrix obtained in each case was rotated to a varimax orthogonal simple structure. These factor analytic procedures were not intended to represent, necessarily, an optimal analytic strategy, but rather the one that is by far the most commonly used.

Since the coefficients compared are not all based upon identical underlying theories regarding distributions and operationalizations of relationship, the comparisons were, as much as possible, made within a certain coefficient across the various data distortions. Thus, although r_t and Tau-B, for example, are commonly seen as substitutes for r under certain conditions, they are based upon different assumptions and may be of considerably different magnitude, both from each other, and each from *r*. Thus, r_t computed by the cosine-pi formula will, given a departure from equal marginal splits, be systematically larger than the associated *r* calculated on the underlying bivariate normal correlation surface. Tau-B generally errs on the other side of *r*. The central issue, then, in this study was: given a known factor structure, and correlation and component pattern coefficients for each measure of association for normally distributed Likert scale variables, which measure is most robust (yielding values close to those for the normal data) with regard to distributions that depart substantially from normal? We thus were less concerned with the values of all the correlation and pattern coefficients in comparison with the true Pearson *r*'s calculated on the underlying bivariate normal surface, than we were with the *change* in these values--within a certain measure of association--as the distributions were distorted away from the normal.

To assess robustness with regard to departures from normality, three measures were used. The first two involved computing the root mean square (*RMS*) statistic between either (1) the obtained values of the measure of association or (2) the component pattern coefficients of the *normal* data, on the one hand, and of *each distorted distribution*, on the other. In the present context, the *RMS* is given by

$$\mathrm{RMS}_r = \sqrt{[2/p(p-1)]\sum_{j<k=1}^{p-1}\sum_{k=j+1}^{p}(\hat{\rho}_{d_{jk}} - \hat{\rho}_{n_{jk}})^2}\ , \qquad (5)$$

where RMS_r denotes the *RMS* computed on correlation matrices, $\hat{\rho}_d$ is the value of the measure of association for the particular distorted data, and $\hat{\rho}_n$ is the value of the measure of association for the normal data. In the case of the component pattern matrices, *F*, in (1), we have

$$\mathrm{RMS}_f = \sqrt{(1/pk)\sum_{i=1}^{k}\sum_{j=1}^{p}(f_{d_{ij}} - f_{n_{ij}})^2}\ , \qquad (6)$$

where f_d is the value of the rotated component pattern coefficient for the distorted data, and f_n is the normal counterpart. Finally, hyperplane-counts (number of component pattern coefficients in the range 0 ± .10) were made on each rotated pattern matrix, in order to determine which measure of association yielded a pattern matrix with hyperplane count closest to the true value of 80, as seen in Table 1. It is noted in passing that although the *common-factor* model was assumed to hold for the population, the *component* model was employed in the analyses in this study. We believe this to be a close reflection of common practice--a practice that is not entirely indefensible, if we conceptualize the component model as representing a close and easily obtained approximation to the more theoretically complex common-factor model.

Results and Discussion

The values of the *RMS* coefficients computed on both the R and F matrices appear in Table 3. These values are for the generated data set with $N = 200$.

Insert Table 3 about here

Table 3 is to read as follows: the *RMS*--as defined in (5)--reflecting discrepancies between entries in (1) the R matrix containing Pearson r coefficients for the normally distributed Likert data, and in (2) the R matrix containing Pearson r coefficients for the rectangularly distributed data was found to be .0279. The analogous *RMS*--this time as defined in (6)--for entries of the rotated component pattern matrices, again based on Pearson r's and involving the normal and rectangular distribution, was found to be .0296. The actual *RMS* values can be roughly interpreted as "average distances" between corresponding elements of the two matrices involved in each case.

From Table 3, it is clear that the effects of distribution distortion were least--over all distortions--in both the correlation and component pattern matrices, when Tau-B was used. The next most robust measure of association, at least according to the present criterion, was the Pearson r, followed by ϕ, ϕ/ϕ_{max}, and lastly r_t. It should be reiterated that the particular computational formula for $\underline{r}_t$ employed--the cosine-pi formula--yields a substantially biased result in the face of severe departures from equal marginal splits, a fact that undoubtedly helps account for the large RMS values for $\underline{r}_t$ when the results for the central data (for which the marginal splits were the most uneven) were compared with those for [illegible] normal data. With the rectangular data, on the other hand (for which the marginal splits were the most nearly even), the marginal splits cannot be blamed for the relative lack of robustness for $\underline{r}_t$ in the face of distribution distortion. Castellan (1966), for example, found little bias in this estimate of $\underline{r}_t$ for marginal splits as nearly equal as those in the rectangular data. As might be expected, almost perfect agreement was found between the rank of the measures of association in terms of the RMS coefficients computed on the correlation matrices and those computed on the component pattern matrices.

It is apparent from Table 3 that the various distribution distortions have differing effects upon the actual changes in the values of the measures of association. A non-normal distribution that tends toward rectangularity can be seen to cause the least disturbance, with the skewed distributions having the next least effect. The central distribution appears to result in the greatest change in computed values of the various coefficients, and this distribution can certainly be expected with "real" data whenever the combination of respondents and items favors a middle or relatively noncommittal response. It is not immediately clear to the authors why the $\underline{\phi}$ coefficient appeared as robust as it did with this distribution distortion. The overall implications of Table 3 appear to be that the measures of association computed from 2 x 2 tables--$\underline{r}_t$, $\underline{\phi}$, and $\underline{\phi}/\underline{\phi}_{max}$--are least robust when applied with non-normal data.

A second criterion of robustness with regard to distribution distortion used in the present study was the degree to which the "true" zero orthogonal factor loadings were recovered in the varimax-rotated component pattern matrices. The hyperplane of each rotated component was taken to be the range 0 ± .10, and the various obtained solutions were compared in terms of how closely each came to yielding the "correct" value of 80 entries in this range. The results are given in Table 4.

Insert Table 4 about here

An additional criterion that could be used is the decision regarding the number of components in the data, using the well-known Kaiser-Guttman rule of the number of components with associated latent roots greater than one. It can be seen from Table 4 that for sample size 100, ϕ, and ϕ/ϕ_{max} led to a decision of six, rather than the correct five, components for all distributions including the normal. It may well be that the sixth factor that emerged with the smaller N was a "difficulty" factor, since although Likert items do not strictly vary in difficulty, the broad censoring of the data resulting from using a seven-point scale can be expected to lead to quite different marginal splits between variables, particularly with centrally-biased data. A similarly incorrect decision was reached in the case of r_t for all data sets except the normal and rectangular.

A near-normal distribution of Likert scale values results in some departure from an equal marginal split--as can be seen from Table 2--for those coefficients computed from 2 x 2 tables. Theoretically, therefore, r could be expected to be least adversely affected with a normal Likert distribution. As can be seen from Table 4, the ranking of the various measures of association, in terms of the degree to which the "true" hyperplane-count of 80 was approximated, is identical to that for the RMS values. One reason for the overall superiority of Tau-B on the hyperplane-count criterion is undoubtedly the slight difference in metric between this measure of association and the Pearson r-related coefficients, but this fact is probably less important than the relative robustness of Tau-B. As before, and again likely, to a large degree, a function of uneven marginal splits, r_t ranked last on this criterion.

Implications

It seems clear from the present study that the well-known technique of splitting attitudinal items at the median and computing either ϕ, ϕ/ϕ_{max}, or r_t coefficients, in order to compensate for severe departures from normality in the item scale distributions, may--in addition to sacrificing information--lead to less than optimal correlational and factorial results. In the case of Likert-type attitude scales, equal median splits on the variables are, of course, scarcely possible, and, thus, the measures of association that depend upon even marginal splits, such as ϕ, are clearly inappropriate, as Carroll (1961) pointed out, and as the present empirical findings show. Carroll's criticisms of ϕ/ϕ_{max}, coupled with the relatively poor results, in the present study, when this measure was used suggest that ϕ/ϕ_{max} should not be used. Although the results with r_t were not encouraging in the present study, they do not directly contradict Carroll's (1961) recommendation of this coefficient in the face of scaling error. Carroll correctly noted that the more precise tetrachoric correlation coefficient is independent of the dichotomization points in the 2 x 2 table, and had we utilized such coefficients in the present study, r_t almost certainly would have ranked more highly on the various criteria used. Unfortunately, as noted earlier this coefficient can and frequently does present computing problems, and, in the present study, the cosine-pi approximations used yielded rather biased values when the dichotomization points were associated with uneven marginal frequencies. Additionally, r_t depends, for interpretation, upon the validity of the assumption of an underlying normal bivariate surface, with linear regressions, and this assumption may often, with attitude scale variables, be difficult to make.

The indication is that it may well be wisest to forego median dichotomization of the Likert scale variables and to compute either standard Pearson r's or, perhaps better, Tau-B coefficients. Obtained Pearson r coefficients for Likert scale data, of course, are not independent of the shapes of the marginal distributions and the effects of scaling error, and r_{max} may well be less than unity. This coefficient does appear, however, from the present study, to be relatively robust in such a situation, and does, of course, imply a gramian R matrix. For R to be gramian is not particularly important in the present context of an incomplete component analysis, that is, an analysis in which considerably fewer than p components are retained, but is, of course, necessary if a common-factor or image analysis is planned, for which rescaling in the metric of the uniquenesses is required.

Tau-B, on the other hand, is not based upon the assumptions about the underlying distributions and regressions that r is, and, consequently, appears more appropriate in the case of attitude scale variables, for which it may also generally be most reasonable to expect only an ordinal scale of measurement. Tau-B does not meet one of Carroll's (1961) requirements for a parametric measure of association based on a 2 x 2 table, namely, that it should be equal to the Pearson r calculated on the underlying correlation surface. It is not, however, susceptible to unequal polychotomizations, and, as has been seen, is relatively stable regardless of the shape of the distributions of the variables.

In this study, a very narrow slice of the overall issue has been studied, and further research is definitely needed before optimal strategies for the factor analytic treatment of attitudinal data are generally available. Such future research could profitably be directed at (1) an investigation of the effects of all distortions occurring simultaneously in the same data set, (2) a more systematic study of the effects of N, and (3) a similar analysis as that in the present study, but involving a more inherently factorially complex factor pattern matrix than the unifactor configuration used in this study.

REFERENCES

Carroll, J. B. The effect of difficulty and chance success on correlations between items or between tests. Psychometrika, 1945, 10, 1-19.

Carroll, J. B. The nature of the data, or how to choose a correlation coefficient. Psychometrika, 1961, 26, 347-372.

Castellan, N. J. On the estimation of the tetrachoric correlation coefficient. Psychometrika, 1966, 31, 67-73.

Cureton, E. E. Note on ϕ/ϕ_{max}. Psychometrika, 1959, 24, 89-91.

Demaree, R. G. and Smith, R.W. Measures of relationship not restricted by differences in proportions in respective categories. Unpublished paper, University of Illinois. Cited in R. B. Cattell, Personality and motivation structure and measurement. New York: World Book, 1957.

Ferguson, G. A. The factorial interpretation of test difficulty. Psychometrika, 1941, 6, 323-329.

Glass, G. V and Stanley, J. C. Statistical methods in education and psychology. Englewood Cliffs, New Jersey: Prentice-Hall, 1970.

Guertin, W. H. and Bailey, J. P. Introduction to modern factor analysis. Published by the authors. Distributed by Edwards Bros. Inc., Ann Arbor, Michigan, 1970.

Kendall, M. G. Rank correlation methods. New York: Hafner, 1962.

Kendall, M. G. and Stuart, A. Advanced theory of statistics (Vol. I, 2nd ed.). London: Griffin, 1958.

Kruskal, W. H. Ordinal measures of association. Journal of the American Statistical Association, 1958, 53, 814-861.

Wherry, R. J. and Gaylord, R. H. Factor pattern of test items and tests as a function of the correlation coefficient: content, difficulty, and constant error factors. Psychometrika, 1944, 9, 237-244.

FOOTNOTE:

[1] The authors thank Mr. David Blackmore for his assistance with some of the computations in the study.

TABLE 1

Input Orthogonal Factor Pattern Matrix
(Decimal Points Omitted)

70	00	00	00	00
80	00	00	00	00
60	00	00	00	00
75	00	00	00	00
00	65	00	00	00
00	75	00	00	00
00	70	00	00	00
00	80	00	00	00
00	00	60	00	00
00	00	70	00	00
00	00	75	00	00
00	00	65	00	00
00	00	00	70	00
00	00	00	75	00
00	00	00	85	00
00	00	00	65	00
00	00	00	00	70
00	00	00	00	75
00	00	00	00	65
00	00	00	00	80

TABLE 2

Standard Scores Corresponding to Lower Boundaries of Each Likert Scale Value and Approximate Percentages (in Parentheses) at Each Value for the Five Distributions

	Scale Value						
Distribution	1	2	3	4	5	6	7
1. Normal	-∞	-1.65	-1.00	-.35	.35	1.00	1.65
	(5)	(11)	(20)	(28)	(20)	(11)	(5)
2. Rectangular	-∞	-1.08	-.58	-.20	.20	.58	1.08
	(14)	(14)	(14)	(16)	(14)	(14)	(14)
3. Central	-∞	-2.33	-1.65	-.84	.84	1.65	2.33
	(1)	(4)	(15)	(60)	(15)	(4)	(1)
4. Positive Skew	-∞	-.52	.25	.67	1.04	1.48	2.05
	(30)	(30)	(15)	(10)	(8)	(5)	(2)
5. Mixed Skew							
(a) Even-Numbered	-∞	-2.05	-1.48	-1.04	-.67	-.25	.52
	(2)	(5)	(8)	(10)	(15)	(30)	(30)
(b) Odd-Numbered	-∞	-.52	.25	.67	1.04	1.48	2.05
	(30)	(30)	(15)	(10)	(8)	(5)	(2)

TABLE 3

Root Mean Square Coefficients between both the Correlation Matrices (R) and the Rotated Component Pattern Matrices (CP) for the Distorted Distributions and those for the Corresponding Normal Distributions (N = 200)

	Distribution							
	Rectangular		Central		Positive Skew		Mixed Skew	
Measure of Association	R	CP	R	CP	R	CP	R	CP
r	.0279	.0296	.0585	.0456	.0438	.0410	.0447	.0416
r_t	.0757	.0500	.2200	.1319	.1245	.0839	.1117	.0747
ϕ	.0625	.0541	.0064	.0002	.0839	.0752	.0646	.0561
ϕ/ϕ_{max}	.0754	.0478	.1120	.0680	.1360	.0826	.1282	.0825
Tau-B	.0210	.0173	.0529	.0414	.0238	.0207	.0253	.0215

TABLE 3

Root Mean Square Coefficients between both the Correlation Matrices (R) and the Rotated Component Pattern Matrices (CP) for the Distorted Distributions and those for the Corresponding Normal Distributions ($N = 200$)

Measure of Association	Distribution							
	Rectangular		Central		Positive Skew		Mixed Skew	
	R	CP	R	CP	R	CP	R	CP
r	.0279	.0296	.0585	.0456	.0438	.0410	.0447	.0416
r_t	.0757	.0500	.2200	.1319	.1245	.0839	.1117	.0747
ϕ	.0625	.0541	.0064	.0002	.0839	.0752	.0646	.0561
ϕ/ϕ_{max}	.0754	.0478	.1120	.0680	.1360	.0826	.1282	.0825
Tau-B	.0210	.0173	.0529	.0414	.0238	.0207	.0253	.0215

TABLE 4

Total Number of Entries, of the Rotated Component Pattern Matrices, in the Range 0 ± .10 ($p = 20$; $k = 5$; $N = 100$ and 200)

Measure of Association	Distribution									
	Normal		Rectangular		Central		Positive Skew		Mixed Skew	
	N=100	N=200	N=100	N=200	N=100	N=200	N=100	N=200	N=100	N=200
r	57	74	61	76	58	72	54	75	60	70
r_t	29	58	35	60	--[a]	46	--[a]	64	--[a]	58
ϕ	--[a]	64	--[a]	68	--[a]	64	--[a]	69	--[a]	66
ϕ/ϕ_{max}	--[a]	60	--[a]	58	--[a]	63	--[a]	67	--[a]	62
Tau-B	61	77	67	72	57	75	65	77	65	75

[a] A solution involving six, rather than five, factors--according to the Kaiser-Guttman rule--was obtained, thus rendering the component pattern and hyperplane-count not directly comparable with those for the "correct" five factor solutions.