DOCUMENT RESUME

ED 109 147 TM 004 578

AUTHOR Hofman, Richard J.
TITLE The Data Generalizability of the Orthotran Solution in Factor Analytic Transformation.

PUB DATE [Apr 75]
NOTE 31p.; Paper presented at the Annual Meeting of American Educational Research Association (Washington, D.C., March 30-April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS *Data Analysis; *Factor Analysis; Factor Structure; Matrices; Measurement Techniques; *Oblique Rotation; Orthogonal Rotation; Statistical Analysis; Transformations (Mathematics); Validity
IDENTIFIERS Orthotran Solution

ABSTRACT

In this paper 12 blind transformation procedures are applied to 18 data sets. The results of the analyses indicate that the orthotran transformation solution is not restricted to particular types of data as are so many other transformation solutions. The evidence presented in this paper strongly suggests that the orthotran solution must be considered as the best blind analytic oblique transformation procedure presently available for general use. That is, the orthotran solution can be used by the average educational researcher without the bothersome consideration of the general nature of the factor matrix to be transformed. (Author)

The Data Generalizability of the Orthotran Solution in Factor Analytic Transformation

Richard J. Hofmann
Miami University

Abstract

In this paper 12 blind transformation procedures are applied to 18 data sets. The results of the analyses indicate that the orthotran transformation solution is not restricted to particular types of data as are so many other transformation solutions.

The evidence presented in this paper strongly suggests that the orthotran solution must be considered as the best blind analytic oblique transformation procedure presently available for general use. That is, the orthotran solution can be used by the average educational researcher without the bothersome consideration of the general nature of the factor matrix to be transformed.

Paper presented at the annual meeting of the American Educational Research Association in Washington, D. C.--April, 1975

The major objective of this study was one of investigating the data generalizability of a new blind transformation procedure, the orthotran solution (Hofmann, 1975). The orthotran solution is intended to be a "workhorse" transformation solution. In order to function in this capacity the orthotran must be operational and available for immediate use, with satisfactory results, without the bothersome prior consideration of the general nature of the factor matrix to be transformed.

The actual procedures for realizing this objective are much like those for validating a test, at least conceptually. One validity question is that of determining how well an orthotran solution will approximate a subjective solution. A second validity question is that of determining whether the quality of an orthotran solution is situation specific. A third validity question is one of determining whether the orthotran is really "better" than existing transformation procedures in terms of approximating subjective solutions.

Clearly the most important consideration is that of data criteria. Eighteen different factor matrices are used in this study. For each factor matrix an ideal subjective solution has been determined either by the author reporting the factor matrix or by a person recognized by the data author as being an excellent "rotator." Solutions were determined by Swineford, Lawley, Thurstone, Holzinger and Horn as well as by others.

Including the orthotran, 12 blind analytic transformation procedures are applied to each of 18 data sets. The transformation solutions are summarily described by four indices. One of the indices is new and is developed in this paper in order to provide a relative index of the adequacy with which a blind solution approximates a subjective solution.

The first part of this paper provides brief descriptions of the 12 blind transformation procedures. The second section describes the four descriptive indices. In the third section the 18 data samples are generally described with regard to a conceptual universe of factor solutions. Also included in this section is a brief description of each data sample. The fourth section provides a discussion and summary of the analyses.

Blind Analytic Procedures

The blind analytic procedures to be discussed in this paper can be subdivided into four general categories:

(a) the class of oblimin solution;

(b) the class of orthomax solutions;

(c) the Harris and Kaiser solutions;

(d) the orthotran solution.

Whereas the oblimin and orthomax solutions are computed directly as a function of either maximizing or minimizing some mathematical criterion, the Harris and Kaiser solutions, gaining in popularity, are two emiprically based solutions, while the orthotran solution is determined primarily through the utilization of principles of artificial intelligence.

The class of orthomax solutions are blind orthogonal solutions; the class of oblimin solutions tend to be blind oblique solutions as do the Harris and Kaiser solutions. The orthotran tends to determine blind oblique solutions but will, by necessity in certain situations, determine solutions identical or very similar to those defined by the class of orthomax solutions. All solutions are determined as "normalized" solutions, variable vectors assumed to be of unit length during the computation process, as opposed to "raw" solutions inasmuch as raw solutions just do not appear to determine as compelling simple structure solutions as do the normalized solutions.

Orthomax Solutions. The class of orthomax solutions all seek to determine an orthogonal solution that will maximize some specific form of the general orthomax criterion (Mulaik, 1972) which is in a very special sense a "weighted" form of the sum of the r column variances of the squared factor coefficients.

It is the weight value, w, selected that determines the specific type of orthogonal solution computed. The quartimax criterion (Carroll, 1953) is defined by (w=0), the varimax criterion (Kaiser, 1958) is defined by (w=1), the equamax (Saunders, 1962) is defined by (w=r/2) and is identical to the varimax for a two factor solution. Although w is specifically defined here it can, technically, be set at any value. This study investigated only the values of (0, 1, r/2) as they represent the prominent choices selected for use in blind orthogonal solutions.

Oblimin Solutions: The class of oblimin solutions are subdivided into two categories: the direct oblimin and the indirect oblimin. The distinction between the two being that the manipulated matrix, the interpretative matrix of interest in the direct oblimin solution, (Jennrich and Sampson, 1966) is the primary pattern matrix, the set of regression coefficients for estimating variable scores from factor scores. The manipulated matrix in the indirect oblimin solution (Carroll, 1957; Harman, 1967) is the reference structure matrix. Once a reference structure matrix has been determined it is usually converted to a primary pattern matrix for interpretation.

The objective of the oblimin-type solutions is to minimize a criterion which is in a very special sense a weighted form of the r(r-1)/2 column covariances of the squared entries of the solution matrix, either a primary pattern matrix in the case of a direct oblimin approach or a reference structure matrix in the case of an indirect oblimin approach.

As with the orthomax criterion the choice for w determines the specific solution computed. For the indirect solution (w=0) defines the quartimin criterion; (w=.5) defines the biquartimin criterion; (w=1) the covarimin criterion (Harman, 1967). For the direct oblimin, the choice of weights seems to be in need of clarification, but the most prominent weights (Harman, 1967) would seem to be (0, -.05, -1.0) with the zero weight being associated with the direct quartimin (Jennrich & Sampson, 1966). As noted by Harman (1967) there seems to be little relationship between the weights of the direct and indirect oblimin with regard to the "quality" of the solution computed.

Although the range of w is somewhat more restricted (Mulaik, 1972) when used with the class of oblimin solution, than with the class of orthomax solutions there is still a considerable range of choices. The two sets of weight previously noted, (o, -.5, -1; 0, .5, 1), were selected for use in defining the two types of oblimin solutions. Harman (1967) specifies the conseqences, sometimes quite grave, of selecting values outside of the range of those used in this study.

Harris and Kaiser Solutions. The Harris and Kaiser (1964) solutions are two sets of empirically based equations, derived from the general Harris and Kaiser (1964) orthoblique equations, for the quick and sometimes highly efficient computation of oblique factor solutions.

The two solutions are specified as the independent cluster solution and the "A'A proportional to L" solution, henceforth (A'A) solution. The correlations are reproduced from the initial factor matrix and then resolved into r unit length eigenvectors and associated eigenvalues. The independent cluster solution, primary pattern matrix, is a column rescaling of a raw quartimax

rotation applied to the eigenvectors. The A'A primary pattern solution is defined by (a) rescaling the columns of the eigenvectors by the fourth root of the eigenvalues, (b) applying a raw quartimax rotation to the rescaled eigenvectors, (c) rescaling the columns of the rotated matrix (Harris and Kaiser, 1964).

Orthotran Solution. The orthotran solution (Hofmann, 1975) is a type of *ad hoc* solution based upon the general orthoblique equations of Harris and Kaiser (1964). Utilizing principles of artificial intelligence, specifically heuristic search procedures, the orthotran approach determines a solution, primary pattern, from an infinite solution space.

Generally, the orthotran solution functions in conjunction with some orthogonal transformation solution. For the purposes of this study it was used in conjunction with a normal varimax solution as this is the orthogonal solution suggested for use as a blind procedure by Hofmann (1975).

The orthotran has been programmed and is presently being considered for inclusion in all of the major published computer packages. Earlier versions of the orthotran, the obliquimax transformation in particular (Hofmann, 1970), have been used since 1970 with empirical applications already in the published literature.

Descriptive Indices

Four descriptive indices are utilized in this study. Two of the indices are not really new, a third index has just recently been published while the fourth index is being described for the first time in this paper. The new index is descriptive of the degree of congruence one solution has with another. The other three indices represent various ways of summarily describing a factor transformation solution.

Hyperplane Count. The hyperplane count is an attempt to quantify a factor solution by describing it in terms of the number of zero coefficients in the solution matrix. It is just the number of coefficients less that .11 in magnitude for any given factor solution. It is based upon Cattell's (1966) notion that for any given sample of variables, "we should expect any true natural influence to affect only a few of them." Cattell states that there should be a transformation position for each factor such that a majority of the variables have zero loadings.

Unlike the other three indices used in this study the magnitude of the hyperplane count is a function of both the number of variables and number of factors. It may be used for comparisons of different solutions to the same factor matrix.

Average Absolute Factor Intercorrelation. This index unlike the other indices included in this study is defined as a function of the primary factor intercorrelation matrix, as opposed to the primary pattern matrix. It serves as an index of the average obliquity of a factor solution.

To compute this index the *absolute value* of each non-diagonal correlation in the primary intercorrelation matrix is converted to a Fisher z. These values are then averaged. The average Fisher z is then converted back to a correlation, henceforth average factor intercorrelation.

This index is a stable index, being independent of the number of variables and the number of factors, and provides a descriptive index for use in comparing certain aspects of primary factor solutions within and across different data sets.

Average Variable Complexity. This index is defined as a function of the rows of the primary pattern matrix. In the linear description of a given variable any number, r, of common factors may be involved as long as ($r < n$). This

number is referred to as the complexity of the variable. Specifically, variable complexity refers to the number of factors involved in the "row-wise" description of a variable in a factor matrix (Hofmann, 1976).

When there is one substantial entry in a row of a factor matrix, all other row entries being zero, the complexity of the variable associated with the row is 1. The average variable complexity for a factor solution is just the average number of common factors needed to describe the n variables.

This index is a stable index, being independent of the number of variables and relatively independent of the number of factors, usually the average variable complexity will range from an absolute low of 1, a true independent cluster solution, to a high of 2.5 to 3.0. Thus, this index is a descriptor appropriate for use in comparing factor solutions both within and between different data sets.

Index of Congruence. This index is intended to be used as a quantitative summary of the pattern similarities of the coefficients of two (nxr) solution matrices. It is based upon Burt's coefficient of congruence (Cattell, 1966). The formula for determining the coefficient of congruence between a factor in one matrix and a factor in a second matrix is reported by Cattell (1966, p. 196). The coefficient of congruence serves as an index of the degree of pattern similarity of two factors, having a maximum upper limit of 1.00 and lower limit of -1.00.

The index of congruence is a single number with a maximum upper limit of 1.00 and a lower limit of 0. It is an index of the degree of pattern similarity one matrix has with another. In this study it is an index of the adequacy with which a blind solution approximates a subjective solution. It is computed according to the following steps:

(a) compute all possible congruence coefficients of the subjective solution with itself (factor 1 with itself, with factor 2 and so on);

(b) obtain the sum of the absolute values of these congruence coefficients (they will represent the absolute sum of the congruence coefficients of the ideal blind solution with the subjective solution);

(c) compute all possible congruence coefficients of a blind solution with the subjective solution;

(d) obtain the sum of the absolute values of these congruence coefficients;

(e) determine the absolute difference between the the sum obtained in step (b) and step (d) and call this "error of fit;"

(f) express this absolute difference as a ratio to the total of step (b) and call it relative error;

(g) subtract the value determined in step (f) from unity to determine the relative accuracy and call the resulting number the index of congruence.

This index will be unity when a blind solution is identical to the subjective solution. Unfortunately, this index is subject to the same limitation of the coefficient of congruence. It can give a value of unity for patterns of identical shape even though the coefficients may have different levels of magnitude. Also the ideal sum may be obtained by more than the one ideal set of congruence coefficients. However, to the extent that more than two factors are utilized the latter problem is reduced.

This index is a stable index and may be used for comparisons of different solutions to the same factor matrix. That blind solution having the highest index of congruence for a particular data set will be assumed to be the best blind approximation to the subjective solution of that data set.

Data Selection

For illustrative purposes 18 data sets are utilized in this study. Of

particular concern is the representative nature of these data sets. Although only intuitive it is felt that these 18 data sets may be fairly representative of the universe of possible factor matrices with regard to variable complexity and average intercorrelations. A variety of characteristics are manifested by these data: number of factors, variable complexities, hyperplane counts, average factor intercorrelations, heterogeneity of communalities, error factors, bipolarity of factors and perhaps still other important yet unrecognized characteristics. All data sets selected have associated with them solutions that were determined subjectively. These subjective solutions may be thought of as being ideal or criterion solutions.

With two exceptions the data are grouped into two general types: plasmodal and real. Plasmodal data (Cattell, 1966) are generated by physical models having known and easily identified formal properities similar to those of mathematical models, but they are subject to measurement error and have other physical realities not present in abstract mathematically determined data sets. The nature of such data is both well understood and well behaved.

The real data are based upon observations on real people with less than perfect measurement instruments and as such they are prone to numerous predictable and unpredictable problems. It is of the utmost importance to include such data in any generalizability study as it is the satisfactory "handling" of a variety of real data that suggests the generalizability of a transformation procedure. When a procedure performs poorly with plasmodal data one can usually identify the cause of such poor performance through an intimate knowledge of the nature of the data. Unfortunately poor performance with real data is not so easy to explain. For this reason satisfactory performance of a blind transformation procedure with a variety of real data is perhaps more laudable than satisfactory performance with plasmodal or true artificial data.

Table 1

Identification, Descriptions and Source of Illustrative Data Sets

Data Identification	Data Type	Number Variables	Factors	Sample	Source
(1) Thurstone's Cylinders	Plasmode	7	2	27	Thurstone, 1947 p. 119
(2) 8 Physical Variables	Real	8	2	305	Holzinger and Harman,1941,p.2
(3) Swineford	Real	9	3	504	Swineford,1948 p. 17-18
(4) Lawley	Real	9	3	73	Lawley and Maxwell, 1971,p
(5) 13 Psychological	Real	13	3	145	Holzinger and Harman,1941,p.5
(6) Cups of Coffee A	Plasmode	15	6	80	Cattell and Sul van,1962,p.191
(7) Cups of Coffee B	Plasmode	15	6	80	Cattell, and Su van,1962,p.191
(8) Thurstone's Trapezoids	Plasmode	15	4	32	Thurstone,1947, p.432
(9) Box Problem	Plasmode	20	3	20	Thurstone,1947, p.136
(10) Horn	Real	20	5	172	Horn,1963,p.127
(11) Coan's Eggs	Plasmode	21	6	100	Coan,1959,p.158
(12) 24 Psychological	Real	20	4	145	Holzinger and Harman,1941,p.2
(13) Thurstone's orthogonal	Artificial	25	5	---	Thurstone,1947, p. 254
(14) Pemberton	Real	25	8	154	Pemberton,1952, p. 276

Table 1 (continued)

Identification, Descriptions and Source of Illustrative Data Sets

Data Identification	Data Type	Number Variables	Factors	Sample	Source
(15) Ball Problem	Plasmode	32	4	80	Cattell and Dickman, 1962, p.
(16) Degan Data	Real	32	9	367	Degan, 1952, p. 3
(17) Kelley's Data	Real	40	11	442	Kelley, 1964, p. 3
(18) Hofmann's Oblique	Artificial	8	2	---	--------------

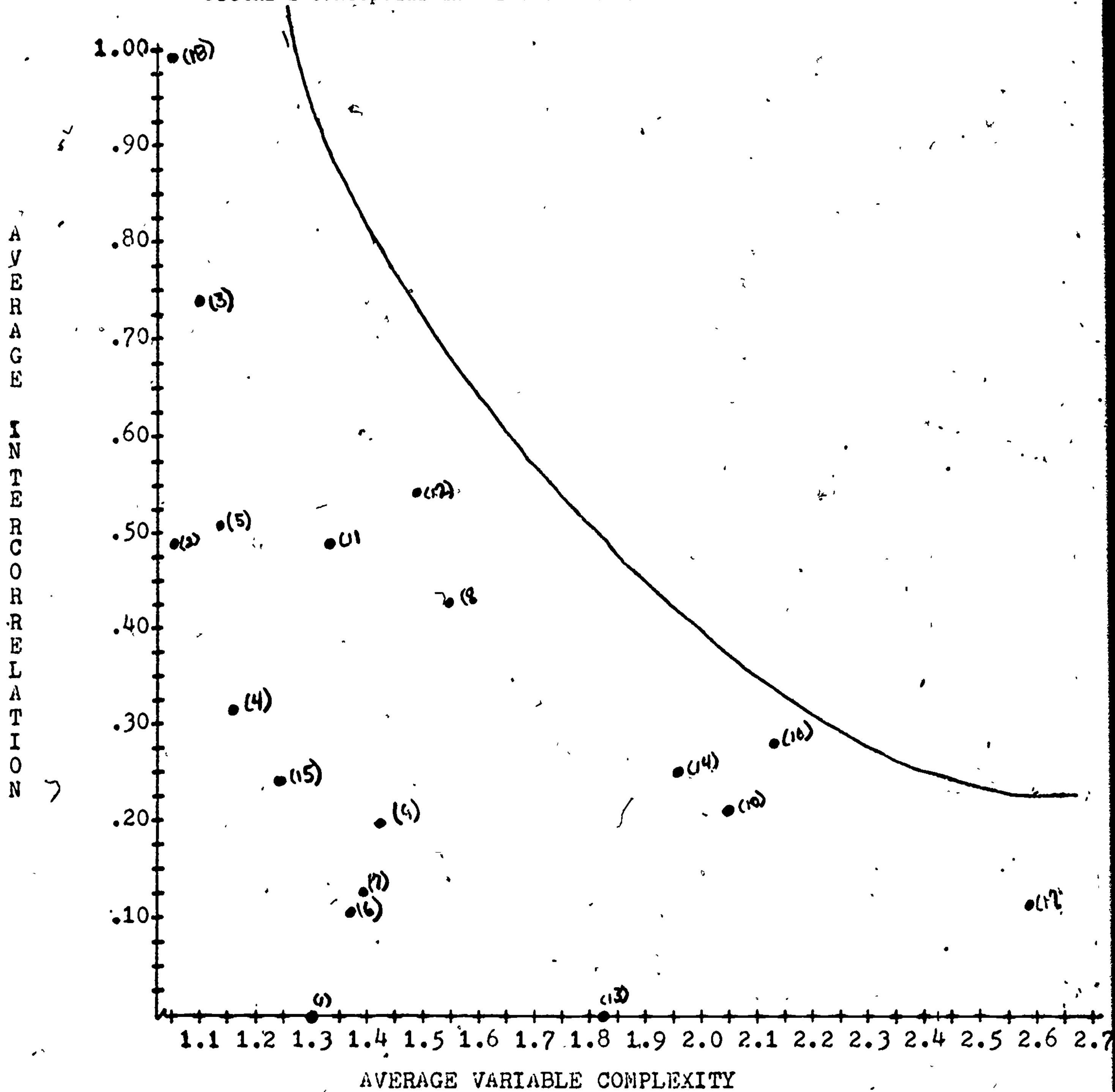
Two of the data sets are artificial, from the point of view that the factor solutions were mathematically determined. Both artificial data sets were generated in order to define extreme average factor intercorrelations. One data set was generated specifically for this study in order to define an average factor intercorrelation of .99.(It can be obtained by request from the author). The second data set was generated by Thurstone (1947) to define an average factor intercorrelation of zero.

The data are identified and partially summarized in Table 1. Their subjective solutions are summarily described with the descriptive indices in Table 2. All descriptions for the subjective solutions are based upon primary pattern matrices.

Conceptual Universe of Factor Solutions. As previously noted two of the descriptive indices used in this study may be used for intersolution comparison: average variable complexity and average factor intercorrelation. In Figure 1 the 18 subjective solutions have been plotted as points in a two-space with average variable complexity and average factor intercorrelation serving as coordinates. The abcissa is defined according to average variable complexity and ranges from 1 to 2.7. The ordinate is defined according to average factor intercorrelation and ranges from 0 to 1.0. These two indices are reported in Table 2 for the subjective solutions.

Note that in Figure 1 the solution points tend to be bounded by a hyperbola that becomes asymptotic to the two axes. The implication of this bound seems clear. Generally speaking the greater the obliquity of the solution the lower is the variable complexity, alternatively, the more complex solutions tend to be associated with relatively uncorrelated factors. A majority of the solutions tend to have an average correlation less than .50, this observation is

FIGURE I Conceptual universe of factor solutions



supported by Rummel (1970) and Cattell (1966). Also it may be noted that a majority of the solutions have an average variable complexity less than 2.0.

The hyperbola has been included in Figure 1 and that portion of the figure between the axes and the hyperbola may be thought of as a conceptual universe of factor solutions. Most properly determined factor solutions would fall within that bounded region.

Discussion and Summary of Analyses

In conducting the analyses of this study the following sequence of steps was adhered to for each data source:

(a) an orthogonal factor matrix was obtained from the literature source, usually a centroid solution, with the exception of data set 18 which was defined mathematically in a principal axis form determined from an artificial factor solution matrix;

(b) the orthogonal matrix was subjected systematically to all 12 blind transformation procedures;

(c) the subjectively determined solution was obtained from the literature source and if it was an oblique solution reported as a reference structure solution it was rescaled to a primary pattern solution following Thurstone (1947);

(d) for each solution the average primary intercorrelation, the hyperplane count, and the average variable complexity was determined (see Table 2);

(e) each blind transformation solution was compared to the subjective solution to determine the index of congruence. (Final column of Table 2).

In summarizing the results for discussion purposes the single most important consideration was felt to be one of addressing the adequacy with which the blind solutions approximated the subjective solutions. To this end for each data set the highest index of congruence, the smallest overall percentage of relative

Table 2. Summary

Data	Transformation		Average Factor r	Hyper-plane Count	Average Variable Complexity	Index of Congruence
1	Direct	w=0	30	1	1.23	.85
	Oblimin	w=-.5	24	1	1.23	.82
		w=-1	20	3	1.23	.80
	Indirect	w=0	48	0	1.22	.79
	Oblimin	w=.5	27	0	1.24	.88
		w=1	00	3	1.32	.99
	Quartimax		00	0	1.24	.87
	Varimax		00	3	1.32	.99
	A'A		23	1	1.49	.96
	Indep. Cluster		43	0	1.21	.79
	Orthotran		05	3	1.31	.99
	Subjective		00	14	1.30	--
2	Direct	w=0	47	8	1.01	1.00
	Oblimin	w=-.5	37	7	1.02	.94
		w=-1	34	6	1.02	.92
	Indirect	w=0	01	0	1.62	.58
	Oblimin	w=.5	27	4	1.04	.89
		w=1	00	0	1.14	.78
	Quartimax		00	0	1.14	.78
	Varimax		00	0	1.14	.78
	A'A		25	0	1.78	.60
	Indep. Cluster		48	8	1.01	1.00
	Orthotran		46	8	1.01	1.00
	Subjective		48	8	1.01	-
3	Direct	w=0	30	14	1.16	.96
	Oblimin	w=-.5	22	14	1.16	.98
		w=-1	18	13	1.16	1.00
	Indirect	w=0	42	13	1.15	.93
	Oblimin	w=-.5	19	14	1.17	1.00
		w=1	10	5	1.17	.75

Table 2 Summary (continued)

Data	Transformation	Average Factor r	Hyper-plane Count	Average Variable Complexity	Index of Congruence
	Quartimax	00	12	1.19	.94
	Varimax	00	12	1.34	.80
	Equamax	00	7	1.41	.91
	A'A	20	5	1.54	.71
	Indep. Cluster	38	13	1.16	.95
	Orthotran	37	13	1.16	.98
	Subjective	74	15	1.10	-
4	Direct w=0	31	12	1.17	.99
	Oblimin w=-.5	26	11	1.19	.96
	w=-1	24	11	1.19	.94
	Indirect w=0	45	11	1.17	.99
	Oblimin w=.5	20	9	1.21	.91
	w=1	09	2	1.44	.70
	Quartimax	00	3	1.34	.76
	Varimax	00	3	1.34	.76
	Equamax	00	2	1.34	.76
	A'A	20	1	2.34	.54
	Indep. Cluster	40	13	1.16	.94
	Orthotran	39	11	1.16	.94
	Subjective	32	11	1.18	-
5	Direct w=0	48	23	1.20	.81
	Oblimin w=-.5	49	19	1.25	.79
	w=-1	45		1.30	.76
	Indirect w=0	21	17	1.14	.88
	Oblimin w=.5	16	18	1.17	.91
	w=1	Factor Collapse			
	Quartimax	00	16	1.25	.82
	Varimax	00	15	1.28	.82
	Equamax	00	7	1.88	.62

Table 2. Summary (Continued)

Data	Transformation		Average Factor r	Hyper-plane Count	Average Variable Complexity	Index of Congruence
5	A'A		25	12	1.69	.69
	Indep. Cluster		74	15	1.48	.90
	Orthotran		22	19	1.18	.93
	Subjective		50	19	1.14	-
6	Direct	w=0	09	58	1.41	.89
	Oblimin	w=-.5	10	54	1.47	.96
		w=-1	09	55	1.47	.96
	Indirect	w=0	49	24	2.35	.98
	Oblimin	w=.5	07	55	1.39	.99
		w=1	Factor Collapse			
	Quartimax		00	53	1.40	.99
	Varimax		00	52	1.41	.97
	Equamax		00	50	1.44	.98
	A'A		07	22	2.28	.73
	Indep. Cluster		14	51	1.41	.87
	Orthotran		12	54	1.36	.94
	Subjective		10	52	1.38	-
7	Direct	w=0	12	56	1.32	.93
	Oblimin	w=-.5	11	57	1.34	.98
		w=-1	10	58	1.35	1.00
	Indirect	w=0	Factor Collapse			
	Oblimin	w=.5	Factor Collapse			
		w=1	Factor Collapse			
	Quartimax		00	48	1.38	.96
	Varimax		00	49	1.38	.90
	Equamax		00	48	1.41	.84
	A'A		08	29	2.43	.68
	Indep. Cluster		20	49	1.39	.83
	Orthotran		10	57	1.34	1.00
	Subjective		13	50	1.39	-

Table 2. Summary (Continued)

Data	Transformation		Average Factor r	Hyper-plane Count	Average Variable Complexity	Index of Congruence
8	Direct	w=0	28	24	1.61	.96
	Oblimin	w=-.5	26	25	1.65	.93
		w=-1	25		1.68	.91
	Indirect	w=0	Factor Collapse			
	Oblimin	w=.5	22	23	2.01	.75
		w=1	10		1.59	.95
	Quartimax		00	15	2.01	.99
	Varimax		00	16	1.53	.79
	Equamax		00	13	1.84	.76
	A'A		23	15	2.01	.85
	Indep. Cluster		50	15	1.65	.94
	Orthotran		29	26	1.62	.94
	Subjective		43	30	1.57	-
9	Direct	w=0	31	27	1.40	.92
	Oblimin	w=-.5	27	27	1.41	.94
		w=-1	26	27	1.41	.95
	Indirect	w=0	20	10	1.94	.99
	Oblimin	w=.5	23	27	1.41	.97
		w=1	10	6	1.61	.83
	Quartimax		00	17	1.51	.87
	Varimax		00	15	1.52	.89
	Equamax		00	15	1.52	.89
	A'A		23	9	1.99	.92
	Indep. Cluster		44	6	1.42	.81
	Orthotran		19	27	1.42	1.00
	Subjective		19	27	1.43	-
10	Direct	w=0	11	47	1.76	.97
	Oblimin	w=-.5	11	44	1.80	.90
		w=-1	11		1.81	.92

Table 2. Summary (Continued

Data	Transformation	Average Factor r	Hyper-plane Count	Average Variable Complexity	Index of Congruence
10	Indirect =0	Factor Collapse			
	Oblimin w=.5	12	42	1.72	.81
	w=1	Factor Collapse			
	Quartimax	00	42	1.75	.84
	Varimax	00	45	1.76	.93
	Equamax	00	40	1.87	.98
	A'A	08	33	2.23	.77
	Indep. Cluster	18	48	1.93	.83
	Orthotran	16	44	1.71	.83
	Subjective	20	55	2.06	-
11	Direct w=0	27	72	1.47	.95
	Oblimin w=-.5	23	67	1.62	.99
	w=-1	23	65	1.67	.99
	Indirect w=0	31	41	1.89	.87
	Oblimin w=.5	Factor Collapse			
	w=1	Factor Collapse			
	Quartimax	00	65	1.40	.98
	Varimax	00	69	1.44	.95
	Equamax	00	49	1.40	.82
	A'A	18	32	2.33	.79
	Indep. Cluster	35	66	1.42	.88
	Orthotran	11	64	1.40	.97
	Subjective	48	87	1.34	-
12	Direct w=0	38	44	1.51	.83
	Oblimin w=-.5	32	44	1.53	.78
	w=-1	31	45	1.53	.76
	Indirect w=0	19	12	2.24	.24
	Oblimin w=.5	32	46	1.53	.77
	w-1	14	6	2.18	.32
	Quartimax	00	36	1.72	.58

Table 2. Summary (Continued)

Data	Transformation		Average Factor r	Hyper-plane Count	Average Variable Complexity	Index of Congruence
12	Varimax		00	23	1.85	.43
	Equamax		00	18	1.87	.43
	A'A		26	22	2.10	.34
	Indep. Cluster		56	39	1.50	.94
	Orthotran		51	40	1.49	.93
	Subject		54	42	1.49	-
13	Direct	w=0	13	61	1.76	.83
	Oblimin	w=-.5	13	63	1.76	.84
		w=-1	12	63	1.76	.84
	Indirect	w=0	28	11	2.62	.81
	Oblimin	w=.5	12	60	1.75	.84
		w=1	11	53	1.93	.84
	Quartimax		00	66	1.77	.99
	Varimax		00	67	1.78	.99
	Equamax		00	70	1.78	1.00
	A'A		12	18	2.77	1.00
	Indep. Cluster		26	38	1.82	.70
	Orthotran		05	65	1.77	.98
	Subjective		00	70	1.83	-
14	Direct	w=0	20	116	2.02	.91
	Oblimin	w=-.5	19	112	2.04	.87
		w=-1	19	110	2.06	.85
	Indirect	w=0	14	59	2.89	.71
	Oblimin	w=.5	Factor Collapse			
		w=1	Factor Collapse			
	Quartimax		00	101	2.09	.88
	Varimax		00	89	2.35	.63

Table 2. Summary (Continued)

Data	Transformation		Average Factor r	Hyper-plane Count	Average Variable Complexity	Index of Congruence
14	Equamax		00	80	2.62	.46
	A'A		18	63	3.93	.50
	Indep. Cluster		42	108	1.93	.90
	Orthotran		16	111	2.01	.87
	Subjective		25	131	1.97	-
15	Direct	w=0	19	69	1.26	.93
	Oblimin	w=-.5	16	68	1.29	.94
		w=-1	15	63	1.30	.92
	Indirect	w=0	Factor Collapse			
	Oblimin	w=.5	09	64	1.39	.87
		w=1	27	42	1.64	.73
	Quartimax		00	53	1.35	.75
	Varimax		00	53	1.48	.81
	Equamax		00	54	1.49	.80
	A'A		13	23	1.72	.55
	Indep. Cluster		23	70	1.24	.97
	Orthotran		26	68	1.24	1.00
	Subjective		24	74	1.25	-
16	Direct	w=0	14	148	2.07	.92
	Oblimin	w=-.5	13	146	2.15	.96
		w=-1	13	146	2.20	.98
	Indirect	w=0	Factor Collapse			
	Oblimin	w=.5	.09	157	1.95	.88
		w=1	Factor Collapse			
	Quartimax		00	131	2.24	.93
	Varimax		00	118	2.26	.99
	Equamax		00	124	2.43	.90
	A'A		11	110	3.50	.79
	Indep. Cluster		26	149	2.14	.79
	Orthotran		18	162	2.01	.85
	Subjective		28	170	2.05	

Table 2. Summary (Continued)

Data	Transformation		Average Factor r	Hyper-plane Count	Average Variable Complexity	Index of Congruence
17	Direct	w=0	15	294	2.38	.93
	Oblimin	w=-.5	14	294	2.38	.93
		w=-1	14	289	2.48	.96
	Indirect	w=0	Factor Collapse			
	Oblimin	w=.5	Factor Collapse			
		w=1	Factor Collapse			
	Quartimax		00	256	2.61	.81
	Varimax		00	268	2.52	.96
	Equamax		00	248	2.75	.71
	A'A		13	186	4.78	.78
	Indep. Cluster		32	280	2.42	.73
	Orthotran		16	297	2.32	.84
	Subjective		13	289	2.60	-
18	Direct	w=0	04	8	1.04	.81
	Oblimin	w=-.5	03	8	1.04	.81
		w=-1	02	8	1.04	.81
	Indirect	w=0	20	8	1.04	.80
	Oblimin	w=.5	20	8	1.04	.80
		w=1	00	0	1.98	.63
	Quartimax		00	8	1.00	.81
	Varimax		00	0	1.98	.63
	A'A		19	0	1.01	.83
	Indep. Cluster		99	8	1.01	.88
	Orthotran		99	3	1.50	.95
	Subjective		99	8	1.01	-

error in approximating the subjective solution, was identified. It was decided that within each data set all indices of congruence within 5 percentage points of the highest index of congruence for that data set would be considered as being amongst the best congruence values for the set of solutions determined for that given data set, 5 percentage points being an arbitrary but necessary error band. In Table 3 the 12 blind solutions are summarized with regard to their frequency of defining an adequate solution relative to the subjective criterion solutions, as indices of congruence.

Clearly evident in Table 2 is the susceptability of the indirect oblimin toward factor collapse, coefficients greater than 1.75 in the pattern matrix and factor intercorrelations of 1.00. Although not reported in Table 2 it was found that direct oblimin weights that were positive and less than unity, e.g. .5, will also result in frequent factor collapse. It would seem that restricting at least the direct oblimin weights to zero or some value less than zero might serve as a guard against factor collapse. The use of the orthotran and Harris and Kaiser models, both models being based upon the orthoblique equations, will not by definition ever result in factor collapse.

For all data sets there were indices of congruence greater than .90. This would suggest that all subjective solutions were reasonable from the perspective that indeed they could be achieved by a blind transformation solution. However for some data sets, there were blind solutions having a relatively low index of congruence but having a hyperplane count considerably higher than the other blind solutions. This would suggest that there may be more than one adequate solution for a given data set, at least in terms of hyperplane count (see the orthotran solution of data set 16).

Harman (1967) has indicated that as the indirect oblimin weight varies from zero to unity the factors become less correlated. This study presents overwhelm-

ing evidence contradicting this generalization. When there are two factors it appears as though the indirect oblimin weight of unity will always define an orthogonal solution. However, in many of the instances of singularity with an indirect oblimin weight of unity the average factor intercorrelation (not reported) was very high, above .90. This contradictory evidence can also be noted in the Table 2 values associated with data sets 2, 9, 12, and 15.

If one were to plot the blind solutions in the two-space used for the subjective solutions certain "solution regions" would appear. These solution regions would represent the domains into which a subjective solution would have to appear in order for the blind procedure defining the domain to provide an adequate solution. These regions are bounded by very irregular lines and seem only to exist for the poorer blind procedures. This would suggest that there are additional dimensions of description for factor solutions that appear to have an influence on certain blind procedures. Although we note this interesting observation it is not of great consequence because there are blind procedures which functioned quite well with the data samples regardless of the data properties.

No one orthogonal solution showed enough consistency in providing accurate representation of the subjective solutions to really warrant further consideration as a pragmatic blind transformation procedure. Although it is frequently argued that orthogonal solutions provide an interpretative picture of the underlying factor structure with a poor hyperplane count the results of this study do not support this particular argument. The index of congruence does not really take into account the levels of the coefficients only their patterns with regard to those of the subjective solution. Therefore if an orthogonal solution, regardless of its hyperplane count, provides a reasonably good repre-

Table 3. Frequency summary of best blind solution relative to subjective solution.

Blind Procedure		Frequency	Percent Hits
Direct	w=0	8	44
Oblimin	w=-.5	9	50
	w=-1	8	44
Indirect	w=0	4	22
Oblimin	w=.5	3	17
	w=1	2	11
Quartimax		6	33
Varimax		7	38
Equamax		4	22
A'A		2	11
Independent Cluster		8	44
Orthotran		15	83

sentation of the patterns of the coefficients in the subjective solution, then one would expect a high index of congruence, but this occurred for less than half of the data sets, mainly with those having an average intercorrelation of approximately zero.

For 13 of the 18 data sets one or more of the direct oblimin solutions provided a highly accurate solution. Unfortunately we were unable to determine either *a priori* or *a posteriori* which particular direct oblimin weight would define either the best solution or the poorest solution. According to the results presented in Table 3 if one were to select any particular weight to use consistently, they would obtain highly accurate solutions, based upon our data samples, at best only 50 percent of the time and at worst only 44 percent of the time. Alternatively if they were to vary the weights from solution to solution they might obtain an accurate soltuion only 17 percent of the time.

The Harris and Kaiser (1964) A'A soltuion appeared to be particularly inappropriate for the data samples. For a number of data sets it defined extremely complex solutions. Alternatively the independent cluster model which assumes variables of unit complexity did comparatively well in defining adequate solutions, which is probably why it is so popular.

The orthotran solution unlike the rest of the blind solutions had an overwhelming tendency, 83 percent of the time, to define a solution that was accurate with regard to the subjective solution.

Conclusions

Several important conclusions seem warranted, however caution must be exercised with regard to the absolute generalizability of the findings of this study. Although every attempt was made to obtain data samples from a variety of sources in order to obtain data representative of the universe of possible factor solutions, we have no guarantee that this was accomplished. A second

major limitation of this study is that the index of congruence may be a very rough measure of the fit of a transformed solution to a subjective solution and the 5 percent range may be an underestimate of the error of this index.

Certainly it is in order to recommend the total abandonment of the indirect oblimin solutions. As a group they appear to be highly susceptible to factor collapse. They also had the lowest frequency of defining adequate solutions.

The Harris and Kaiser (1964) A'A solution should also be abandoned. Although it does not transform to singularity it does tend to transform to complexity. It defined the most complex solution for 15 of the 18 data sets.

Assuming the 5 percent range is appropriate it may be concluded that the orthotran solution is the only blind transformation solution, of those studied, that is data generalizable. However until its publication both as a manuscript (Hofmann, 1975) and as a computer program one should use a direct oblimin (Jennrich and Sampson, 1966) solution with one particular weight. The choice for the direct oblimin weight should be either zero or negative one-half.

References

Coan, R. W. A comparison of oblique and orthogonal factor solutions. Journal of Experimental Education, 27, 1959, 151-166.

Carroll, J. B. An analytic solution for approximation simple structure in factor analysis. Psychometrika, 18, 1953, 23-28.

Carroll, J. B. Biquartimin criterion for rotation to oblique simple structure in factor analysis. Science, 126, 1957, 1114-1115.

Cattell, R. B. The meaning and strategic use of factor analysis in Handbook of Multivariate Experimental Psychology (Ed. R. B. Cattel). Chicago Rand-McNally, 1966, 174-243.

Cattell, R. B. and Sullivan, W. The scientific nature of factors: A demonstration by cups of coffee. Behavioral Science, 7, 1962, 184-193.

Degan, J. W. Dimensions of functional psychosis. Psychometric Monograph, 6, 1952.

Harman, H. Modern factor analysis. (Second Edition). Chicago, University of Chicago Press, 1967.

Harris, C. W. and Kaiser, H. F. Oblique factor analytic solutions by orthogonal transformations. Psychometrika, 29, 1964, 347-362.

Hofmann, R. J. The obliquimax transformation, unpublished Doctoral Dissertation, State University of New York, Albany, New York, 1970.

Hofmann, R. J. The orthotran solution, xerox manuscript, Miami University, Oxford, Ohio 45056.

Hofmann, R. J. Indices descriptive of factor complexity. Journal of General Psychology (In Press for 1976).

Holzinger, K. J. and Harman, H. H. Factor analysis: a synthesis of factorial methods. Chicago,University of Chicago Press, 1941.

Horn, J. Second-order factors in questionnaire data. Educational and Psychological Measurement, 23, 1963, 117-134.

Jennrich, R. E. and Sampson, P. F. Rotation for simple lo adings. Psychometrika, 31, 1966, 313-323.

Kelley, H. P. Memory abilities: a factor analysis, Psychometric Monograph, 11, 1964.

Lawley, D. N. and Maxwell, A. E. Factor analysis as a statistical method (second edition), 1971.

Mulaik, S. A. The foundations of factor analysis. New York, McGraw-Hill, 1972.

Pemberton, C. The closure factors related to other cognitive processes. Psychometrika, 17, 1952, 267-288.

Rummel, R. J. Applied factor analysis. Evanston, Northwestern University Press, 1970.

Saunders, D. R. Trans-varimax: some properties of the ratiomax and equamax criteria for blind orthogaonal rotation. Paper presented at the annual meeting of the American Psychological Association, St. Louis, 1962.

Swineford, F. The nature of the general, verbal, and spatial bi-factors. Supplementary Educational Monograph, Chicago, The University of Chicago Press, 1947.

Thurstone, L. L. Multiple factor analysis. Chicago, The University of Chicago Press, 1947.