PSYCHOMETRIKA—VOL. 23, NO. 3
SEPTEMBER, 1958

THE VARIMAX CRITERION FOR ANALYTIC ROTATION IN FACTOR ANALYSIS*

HENRY F. KAISER
UNIVERSITY OF ILLINOIS

An analytic criterion for rotation is defined. The scientific advantage of analytic criteria over subjective (graphical) rotational procedures is discussed. Carroll's criterion and the quartimax criterion are briefly reviewed; the varimax criterion is outlined in detail and contrasted both logically and numerically with the quartimax criterion. It is shown that the *normal* varimax solution probably coincides closely to the application of the principle of simple structure. However, it is proposed that the ultimate criterion of a rotational procedure is factorial invariance, not simple structure—although the two notions appear to be highly related. The normal varimax criterion is shown to be a two-dimensional generalization of the classic Spearman case, i.e., it shows perfect factorial invariance for two pure clusters. An example is given of the invariance of a normal varimax solution for more than two factors. The oblique normal varimax criterion is stated. A computational outline for the orthogonal normal varimax is appended.

In factor analysis, an analytic criterion for rotation is defined as one that imposes mathematical conditions beyond the fundamental factor theorem, such that a factor matrix is uniquely determined. Historically, the first such criterion was Thurstone's treatment of the principal axes problem [10]: from any arbitrary factor matrix he suggested rotating under the criterion that each factor successively accounts for the maximum variance. But principal axes have seldom been accepted as psychologically very interesting ([9], p. 139). The rotation problem for psychologically meaningful factors is usually handled judgmentally. Scientifically, however, this procedure is not very satisfactory: the ad hoc quality of subjective rotation makes uniquely determined factors impossible; only factors that are subject to the uncertainties and controversies besetting any a posteriori reasoning can be defined. In contrast, an analytic criterion for rotation would allow factor analysis to become a straightforward methodology stripped of its subjectivity and a proper tool for scientific inquiry.

The Quartimax Criterion

The first analytic criterion for determining psychologically interpretable factors was presented in 1953 by Carroll [1]. In an attempt to provide a

*Part of the material in this paper is from the writer's Ph.D. thesis. I am indebted to my committee, Professors F. T. Tyler, R. C. Tryon, and H. D. Carter, chairman, for many helpful suggestions and criticisms. Dr. John Caffrey suggested the name *varimax*, and wrote the original IBM 602A computer program for this criterion.

I am also indebted to the staff of the University of California Computer Center for help in programming the procedures described in the paper for their IBM 701 electronic computer. Since their installation is partially supported by a grant from the National Science Foundation, the assistance of this agency is acknowledged.

mathematical explication of Thurstone's simple structure, he suggested that for a given factor matrix,

$$f = \sum_{s<t} \sum_{j} a_{js}^2 a_{jt}^2 \tag{1}$$

should be a minimum, where $j = 1, 2, \cdots, n$ are tests, $s, t = 1, 2, \cdots, r$ are factors, and a_{js} is the factor loading of the jth test on the sth factor. It appears that Carroll was motivated in writing (1) primarily by a close inspection of Thurstone's five formal rules for simple structure ([12], p. 335), particularly the requirement that a large loading for one factor be opposite a small loading for another factor.

In his original paper, Carroll provided two numerical examples of the application of his method. Without the restriction of orthogonality, these illustrations gave somewhat equivocal results—while the application of (1) appears to bring one close to the desired simple structure, the criterion has an obvious bias in being too strongly influenced by factorially complex tests.

In the light of later developments, Carroll's criterion should probably be relegated to the limbo of "near misses"; however, this does not detract from the fact that it was the first attempt to break away from an inflexible devotion to Thurstone's ambiguous, arbitrary, and mathematically unmanageable qualitative rules for his intuitively compelling notion of simple structure.

Almost simultaneously with Carroll's development, Neuhaus and Wrigley [7], Saunders [8], and Ferguson [2] proposed what is usually called the quartimax method for orthogonal simple structure. Neuhaus and Wrigley suggest that a most easily interpretable factor matrix, in the simple structure sense, may be found when the variance of all nr squared loadings of the factor matrix is a maximum, i.e.,

$$q_1 = [nr \sum_{j} \sum_{s} (a_{js}^2)^2 - (\sum_{j} \sum_{s} a_{js}^2)^2]/n^2r^2 = \text{maximum}. \tag{2}$$

Saunders' approach requires that the kurtosis (fourth moment over second moment squared) of all loadings and their reflections be a maximum,

$$q_2 = nr \sum_{j} \sum_{s} a_{js}^4/(\sum_{j} \sum_{s} a_{js}^2)^2 = \text{maximum}. \tag{3}$$

While Ferguson, basing his rationale on certain parallels with information theory, calls simply for

$$q_3 = \sum_{j} \sum_{s} a_{js}^4 = \text{maximum}. \tag{4}$$

All these investigators are concerned with attaining a factor matrix with a maximum tendency to have both small and large loadings. While less obviously related to Thurstone's rules than Carroll's criterion, the emphasis on small loadings coincides with Thurstone's requirements of zero loadings. For orthogonal factors, criteria (2), (3), and (4) are equivalent because of the invariance of the sum of the communalities, $\sum_j \sum_s a_{js}^2$. (This term, as well as other constants, disappear when differentiated in the calculus problem involved in finding the required critical point.)

Indeed, it turns out that they are also equivalent to Carroll's criterion in the orthogonal case. Minimizing (1) is equivalent to maximizing (4) since the squared communality of a test is

$$\text{constant} = (\sum_s a_{js}^2)^2 = \sum_s a_{js}^4 + 2 \sum_{s<t} a_{js}^2 a_{jt}^2,$$

and the sum of squared communalities over all tests is

$$\begin{aligned} \text{constant} &= \sum_j \sum_s a_{js}^4 + 2 \sum_j \sum_{s<t} a_{js}^2 a_{jt}^2 \\ &= q_3 + 2f. \end{aligned}$$

Thus, since the quartimax criterion plus twice Carroll's criterion is a constant, maximizing q_3 is equivalent to minimizing f.

Neuhaus and Wrigley realized that none of these criteria can be realistically applied without the aid of an electronic computer—the calculations involved are too extensive for a desk calculator or punched card mechanical computers. Consequently, they programmed the quartimax method for the Illiac* and provided a rather extensive numerical investigation of the empirical properties of the quartimax method.

Their results were perhaps more encouraging than Carroll's. Under the restriction of orthogonality Carroll's criterion (or the equivalent quartimax method) does not show nearly so obvious a bias as does Carroll's criterion when the restriction of orthogonality is removed. However, as an explication of orthogonal simple structure, the quartimax method does have a systematic bias which will be more fully examined in the next section.

The Varimax Criterion

From the outset, the above methods consider all nr loadings simultaneously. In every case, however, these criteria may be applied separately to each *row* of the factor matrix and summed over rows for the final criterion because of the invariance of the communalities. For example, Neuhaus and Wrigley could have defined the *simplicity*, say, of the factorial composition of the jth test as the variance of the squared loadings for this test,

$$q_j^* = [r \sum_s (a_{js}^2)^2 - (\sum_s a_{js}^2)^2]/r^2. \tag{5}$$

*The Illiac is the University of Illinois electronic computer. Subsequently, the quartimax criterion has been programmed for the CRC–102A (Neuhaus), and the IBM 701 (Kaiser). The varimax criterion described in the next two sections has been programmed for SWAC at UCLA (Comrey), the IBM 701 (Kaiser), Illiac (Dickman), and the IBM 650 (Vandenberg).

To obtain the total criterion for the entire factor matrix, (5) could then be summed over all tests to give

$$q^* = \sum_j \{[r \sum_s (a_{js}^2)^2 - (\sum_s a_{js}^2)^2]/r^2\}. \tag{6}$$

Maximizing q^* is equivalent to maximizing q_3, again because constant terms vanish when differentiated.

Equation (6) perhaps provides some insight into the quartimax criterion—its aim is to simplify the description of each row, or test, of the factor matrix. It is unconcerned with simplifying the columns, or factors, of the factor matrix (probably the most fundamental of all requirements for simple structure). The implication of this is that the quartimax criterion will often give a general factor. Under requirement (5) there is no reason why a large loading for each test may not occur on the same factor. In practice, this tendency for the quartimax criterion to yield a general factor is most pronounced when the unrotated factor matrix has a pronounced general factor.

From the simple structure viewpoint, an immediate modification of the quartimax criterion is apparent. Let us define the simplicity of a factor as the variance of its squared loadings,

$$v_s^* = [n \sum_j (a_{js}^2)^2 - (\sum_j a_{js}^2)^2]/n^2. \tag{7}$$

And for the criterion for all factors, define the maximum simplicity of a factor matrix as the maximization of

$$v^* = \sum_s v_s^* = \sum_s \{[n \sum_j (a_{js}^2)^2 - (\sum_j a_{js}^2)^2]/n^2\}, \tag{8}$$

the variance of squared loadings by columns rather than by rows.

Since a factor is a vector of correlation coefficients, the most interpretable factor is one based upon correlation coefficients which are maximally interpretable. Those correlations which satisfy this condition are patently obvious: correlations of $\pm$ 1, which indicate a functional relationship, and correlations of zero, which indicate no linear relationship. On the other hand, middle-sized correlations are the most difficult to understand. Thus, it is seen why v_s^* in (7) could be maximized for the maximum interpretability or simplicity of a factor, and more generally, why the interpretability of an entire factor matrix could be considered best when (8) is a maximum.

Criterion (8) is the original *raw* varimax criterion [4]. In the original proposal of this criterion, it was shown to be mathematically equivalent, in the orthogonal case, to minimizing

$$c^* = \sum_{s<t} \{[n \sum_j a_{js}^2 a_{jt}^2 - (\sum_j a_{js}^2)(\sum_j a_{jt}^2)]/n^2\}, \tag{9}$$

i.e., minimizing the covariance of pairs of columns of squared loadings and summing over all possible pairs of columns for the criterion. Criterion (9) then bears the analogous relationship to Carroll's criterion (1) that the varimax criterion (8) does to the quartimax criterion (6).

Some distinctions between quartimax and varimax orthogonal solutions can be illustrated numerically. In Table 1 solutions for Thurstone's eleven-variable box problem ([12], pp. 373–375) are given. It will be noted that the quartimax solution [7] could hardly be called a simple structure. There is a large general factor, and the second factor seems only vaguely concerned

Table 1

Thurstone's 11-Variable Box Problem[a]

	Subjective (oblique)			Quartimax			Raw Varimax		
Test	X	Y	Z	X	Y	Z	X	Y	Z
x	90	05	00	68	65	05	91	19	16
y	04	88	01	83	-47	00	05	93	25
z	03	05	79	42	-08	79	11	17	88
xy	62	63	-06	99	11	-04	64	74	20
yz	-05	54	57	71	-40	56	02	65	75
x^2y	82	37	-01	92	41	03	84	51	22
xy^2	35	76	02	96	-18	03	37	86	28
2x + 2y	53	71	-09	100	00	-07	54	82	18
$(x^2 + y^2)^{\frac{1}{2}}$	52	71	-08	99	-01	-07	53	81	18
$(x^2 + z^2)^{\frac{1}{2}}$	52	-07	65	59	38	68	60	09	77
xyz	42	43	43	88	04	45	48	58	65

[a]Decimal points omitted.

with dimensions of boxes. On the other hand, the raw varimax solution closely parallels Thurstone's original subjective solution, given the restriction of orthogonality.

In Table 2 are solutions for Holzinger and Harman's 24 psychological tests ([3], pp. 229–233). Both the quartimax [7] and the raw varimax methods seem to duplicate the subjectively rotated simple structure patterns. But the respective variance contributions of the factors are perhaps more interesting. It is seen that the dispersion of the $\sum_j a_{js}^2$ for the subjective solution is less than the corresponding figures for the two analytic methods. In other words, Holzinger and Harman have made the factors a little more level or

Table 2

Holzinger and Harman's Twenty-four Psychological Tests[a]

Test	Subjective				Quartimax				Raw Varimax				Normal Varimax			
	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
1	10	32	62	20	37	19	60	07	24	20	65	13	14	19	67	17
2	07	15	41	13	24	07	38	04	17	07	42	08	10	07	43	10
3	10	12	53	13	31	01	48	01	22	02	52	06	15	02	54	08
4	15	18	53	12	36	07	46	-01	27	08	52	04	20	09	54	07
5	75	15	26	15	81	14	-02	-04	78	21	12	06	75	21	22	13
6	72	05	28	25	81	03	00	06	78	10	13	14	75	10	23	21
7	81	08	27	11	85	07	-04	-10	84	15	10	00	82	16	21	08
8	54	26	38	14	66	20	20	-04	60	25	31	05	54	26	38	12
9	76	-04	29	30	86	-06	-02	10	84	01	12	19	80	01	22	25
10	28	66	-19	14	23	70	-12	11	17	71	-08	17	15	70	-06	24
11	27	61	-04	29	31	62	01	23	22	63	06	29	17	60	08	36
12	13	72	09	03	16	69	19	-01	06	70	23	04	02	69	23	11
13	24	63	31	02	35	57	32	-08	24	59	39	-01	18	59	41	06
14	23	19	-02	48	32	19	-03	42	26	20	01	46	22	16	04	50
15	11	14	08	50	25	11	10	45	17	11	13	48	12	07	14	50
16	05	22	34	45	29	13	37	37	17	13	41	41	08	10	41	43
17	15	24	-03	62	28	24	02	57	20	23	05	61	14	18	06	64
18	01	39	20	52	22	32	30	47	08	31	32	51	00	26	32	54
19	12	22	18	39	28	18	19	32	19	18	22	36	13	15	24	39
20	31	18	46	29	52	09	35	14	42	12	43	21	35	11	47	25
21	17	46	33	24	35	38	35	14	23	40	40	20	15	38	42	26
22	31	12	40	40	53	04	30	26	44	06	37	32	36	04	41	36
23	31	29	54	25	55	19	44	09	43	21	52	16	35	21	57	22
24	39	46	14	31	49	43	10	20	40	46	18	27	34	44	22	34
$\sum_j a_{js}^2$	343	292	268	236	559	242	196	142	431	260	264	186	350	244	308	236

[a]Decimal points omitted.

even in their contribution to variance than the analytic criteria. Of the two analytic criteria, the raw varimax solution has given a solution which is closer in this respect to Holzinger and Harman's. It is also noteworthy that as a result of these differences the large loadings of the factors with the larger variance contributions for the analytic methods are larger than the large loadings for the smaller factors, and similarly, the small loadings for the larger factors are larger than the small loadings for the smaller factors. Holzinger and Harman's subjective solution does not show this systematic bias; their solution gives a more equitable patterning of factor loadings.

How this bias may be removed is indicated in the next section. This leads to a revision of the varimax criterion, which appears to have more important characteristics than merely satisfying the rules of simple structure.

Factorial Invariance: Normal Varimax

It seems reasonable to attribute the systematic bias seen in both the quartimax and varimax solutions of the Holzinger-Harman data and other examples [4] to the divergent weights which implicitly are attached to the tests by their communalities. When one deals with fourth-power functions of factor loadings, a test with communality 0.6, for example, would tend to influence the rotations four times as much as a test whose communality was 0.3. Thus, while the most obvious weights have been applied to the tests, namely the square roots of their communalities, after the fact it seems that there is probably a better set of weights—weights which would tend to equalize to a greater extent the relative influence of each test during rotation.

There seems no rational basis for choosing among different weighting schemes. Let us then make the agnostic confession of ignorance which pervades any form of correlational analysis. For the purposes of rotation, weight the tests equally, in the sense that the lengths of the common parts of the test vectors have equal length. (The author is indebted to Dr. D. R. Saunders for this suggestion.) The varimax criterion could then be rewritten as

$$v = \sum_{s} \{[n \sum_{j} (a_{js}^2/h_j^2)^2 - [\sum_{j} (a_{js}^2/h_j^2)]^2]/n^2\}, \tag{10}$$

where h_j^2 is the communality of the jth test. In contrast to (7) and (8), where the variance of the squared correlations of the tests with a factor is maximized, the variance of the squared correlations of the common parts of the tests (the reflections of the tests onto the common-factor space) with a factor is now being maximized. [Note from (10) that we are not advocating a permanent weighting of the tests by a weight inversely as the square root of their communalities. During rotation this weighting extends the common part of each test vector to unit length, but after rotation each of these vectors is shortened to its proper length by reweighting directly as the square root of the test's communality.]

As will be seen in Table 2, under this modification the varimax criterion (the *normal* varimax, since rotation is with respect to normalized common parts of tests) has effectively removed the small but disturbing bias in the raw varimax solution of Holzinger and Harman's example. It also has been shown in a number of other examples [6] that the normal varimax does not seem to deviate systematically from what may be considered the best orthogonal simple structure.*

Thus far, however, merely a numerical-intuitive basis for a weighting procedure which leads to "prettier" results has been provided. Such a basis is quite unsatisfactory theoretically. Indeed, this sort of ad hoc thinking could conceivably lead to a different set of judgmentally determined weights for any particular example—a situation as scientifically reprehensible as the subjective graphical methods.

There is a more fundamental rationale for attempting to establish the normal varimax criterion (10) as a mathematical definition for the rotation

*Professor Andrew Comrey has apparently reached the same conclusion in an extensive application of the normal varimax criterion to interitem correlation matrices of the MMPI (personal communication). A further example, available from the writer, is the normal varimax solution of Thurstone's classic PMA study[11] (dittoed).

problem. Consider the situation illustrated in Fig. 1. There are two clusters of tests, each of which is pure in the sense that the reflections of the test vectors of the cluster onto the two-dimensional, common-factor space are collinear. (While these clusters are drawn less than 90° apart, the following argument is perfectly general.)

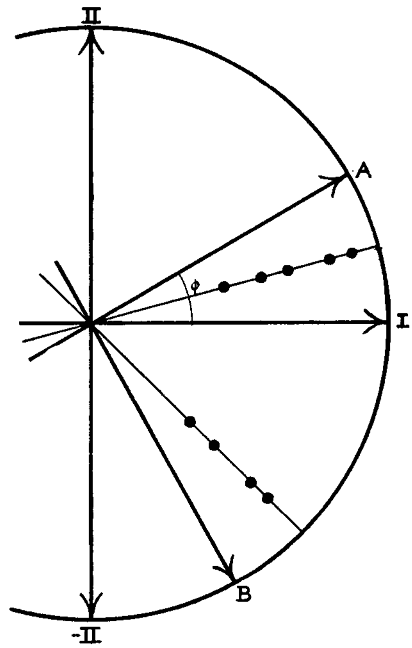

FIGURE 1
Case for which a normal varimax solution is invariant under changes in the composition of the test battery.

It is shown below that the angle of rotation in a plane which maximizes (10) is

$$\phi = \tfrac{1}{4} \arctan \frac{2[n \sum_i u_i v_i - \sum_i u_i \sum_i v_i]}{n \sum_i (u_i^2 - v_i^2) - [(\sum_i u_i)^2 - (\sum_i v_i)^2]}, \tag{11}$$

where

$$u_i = (a_{i1}/h_i)^2 - (a_{i2}/h_i)^2,$$

and

$$v_i = 2(a_{i1}/h_i)(a_{i2}/h_i).$$

Let n_A ($n_A \geq 1$) be the number of tests in the first cluster and n_B ($n_B \geq 1$) be the number of tests in the second cluster ($n = n_A + n_B$). It is readily apparent that all tests of the first cluster have the same values for u_i and v_i .

Let these values be u_A and v_A. Similarly let the values for the second cluster be u_B and v_B. In this case (11) reduces to

$$(12) \qquad \phi = \tfrac{1}{4} \arctan \frac{2n_A n_B(u_A v_A + u_B v_B - u_A v_B - u_B v_A)}{n_A n_B(u_A^2 + u_B^2 - v_A^2 - v_B^2 - 2u_A u_B + 2v_A v_B)} .$$

A most important result is shown in (12). The $n_A n_B$ term may be cancelled, indicating that the angle of rotation does not depend on the number of tests in each cluster, i.e., *for the case illustrated in Fig. 1, the normal varimax solution is invariant under changes in the composition of the test battery.*

This invariance property would seem to be of greater significance than the numerical tendencies of the normal varimax solution to define mathematically the doctrine of simple structure. Although factor analysis seems to have many purposes, fundamentally it is addressed to the following problem. Given an (infinite) domain of psychological content, infer the internal structure of this domain on the basis of a sample of n tests drawn from the domain. The possibility of success in such inferences is obviously dependent upon the extent which a factor derived from a particular battery or sample of tests approximates the corresponding unobservable factor in the infinite domain. If a factor is invariant under changing samples of tests, i.e., shows factorial invariance ([12], pp. 360–361), there is evidence that inferences regarding domain factors are correct.

The normal varimax solution, according to the above result, allows such inferences; regardless of the sampling of tests, for the problem shown in Fig. 1 it is possible to infer precisely the domain normal varimax factors. This is not true for either the quartimax or raw varimax solutions since the angle of rotation is a function of n_A and n_B.

Note that domain normal varimax factors are not said to be more *meaningful* than domain factors according to some different criterion; it is suggested that observed normal varimax factors will have a greater likelihood of portraying the corresponding domain factors.

Although one often gets the impression that simple structure is the ultimate criterion of a rotational procedure, it is suggested here that the ultimate criterion is factorial invariance. The normal varimax solution was originally devised solely for the purpose of satisfying the simple structure criteria. But the fact that it shows mathematically this sort of invariance suggests that Thurstone's reasoning was basically directed toward factorial invariance. The principle of simple structure may probably be considered incidental to the more fundamental concept of factorial invariance. This viewpoint renders meaningless the arguments concerning "psychological reality" of general factors, bipolar factors, simple structure factors, etc.

Admittedly, the result (12) is for a special case. The correlations among the variables within each of the two pure clusters must form a perfect Spearman matrix, and the reduced correlation matrix as a whole must be of rank two. Normal varimax does however give invariant results for two such Spearman clusters simultaneously, and consequently the normal varimax criterion is a two-dimensional generalization of the classic Spearman case. Obviously, the next step would be a generalization along the same lines to the r-dimensional case; thus far, however, work on this problem suggests virtually insuperable mathematical difficulties.

To investigate numerically the tendency of the criterion to give factorially invariant solutions for r larger than two, again consider Holzinger and Harman's empirical data. Taking their centroid loadings, the first five tests were rotated then the first six, etc., systematically until the analysis of all 24 tests was reached, as in Table 2. The results of this application of the normal varimax criterion are given in Table 3. There, the normal varimax loadings for the four factors as a function of the changing number of tests are given.

Note that factors A and C are essentially invariant from the outset; the loading changes, while somewhat systematic, are negligible—24 appears to be essentially infinite. On the other hand, factors B and D show more movement before they become stable. The reason for this is readily apparent. For both B and D, this movement occurs only while they are underdetermined, i.e., only while they contain no appreciable loadings. However, once they pick up a test or two with high loadings, their ambiguous definition abruptly stops, and they settle down to exhibit a degree of invariance similar to the

Table 3A

Normal Varimax Loading Changes for Holzinger and Harman's Factor A (n = 5, 6, ..., 24)[a]

	n																			
Test	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	19	19	18	19	18	16	16	16	16	16	15	16	15	16	16	15	15	15	14	14
2	12	12	12	13	12	12	12	12	12	11	11	11	11	12	11	11	11	11	10	10
3	15	16	15	16	16	17	17	17	17	16	16	16	16	17	17	16	16	16	15	15
4	21	21	21	22	21	21	21	21	21	21	21	21	21	21	21	20	20	20	20	20
5	79	78	78	79	78	76	75	76	76	76	75	75	75	76	76	75	75	75	75	75
6		78	77	77	78	77	76	77	77	76	75	75	75	75	75	75	75	75	75	75
7			83	84	83	82	82	82	82	82	82	82	82	83	82	82	82	82	82	82
8				59	58	55	55	55	55	55	55	55	55	55	55	55	55	55	54	54
9					82	83	82	83	83	81	80	80	80	80	80	80	80	80	80	80
10						18	15	17	17	16	16	16	15	16	16	16	16	15	16	15
11							19	21	21	19	18	18	18	18	18	18	18	18	18	17
12								03	02	03	03	03	03	03	03	03	03	03	02	02
13									18	19	19	20	20	20	20	19	19	19	18	18
14										23	21	22	21	22	22	22	22	22	22	22
15											11	12	11	12	12	12	12	12	12	12
16												09	09	10	09	09	09	09	09	08
17													13	14	14	14	14	14	14	14
18														01	01	00	00	00	00	00
19															13	13	13	13	13	13
20																35	35	35	35	35
21																	16	16	16	15
22																		36	36	36
23																			35	35
24																				34

Table 3B

Normal Varimax Loading Changes for Holzinger and Harman's Factor B (n = 5, 6, ..., 24)[a]

Test	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	10	10	10	10	07	12	16	19	21	20	20	19	19	19	18	19	19	19	19	19
2	01	01	01	01	02	03	05	07	08	08	07	07	07	07	06	07	07	07	07	07
3	-07	-07	-07	-06	-04	-04	-01	01	02	03	02	02	02	02	02	02	02	02	02	02
4	-02	-03	-02	-01	-04	02	05	07	09	09	09	09	09	03	08	08	09	09	08	09
5	-00	03	05	04	-00	19	22	21	22	22	21	21	20	20	20	20	20	21	20	21
6		-03	-02	-05	03	10	14	12	13	11	10	10	09	08	08	08	09	09	09	10
7			-03	-03	-08	12	16	15	16	16	16	16	15	15	15	15	15	16	16	16
8				10	02	21	24	25	26	26	26	26	25	25	25	25	25	25	25	26
9					03	03	08	05	05	03	02	01	00	-00	-01	-01	-00	00	-00	01
10						73	74	74	74	72	71	71	70	69	69	69	69	70	70	70
11							68	67	67	64	62	62	61	60	59	59	60	60	60	60
12								67	68	69	69	69	69	69	68	69	69	69	69	69
13									57	59	59	59	59	59	59	59	59	59	59	59
14										22	19	19	16	15	15	14	15	15	15	16
15											10	10	07	06	06	06	06	06	06	07
16												12	10	09	08	08	09	09	09	10
17													19	17	17	17	17	17	17	18
18														26	25	25	26	26	26	26
19															14	14	14	15	15	15
20																09	10	10	10	11
21																	37	38	37	38
22																		04	04	04
23																			20	21
24																				44

[a]Decimal points omitted.

Table 3C

Normal Varimax Loading Changes for Holzinger and Harman's Factor C (n = 5, 6, ..., 24)[a]

Test	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	70	70	70	70	70	70	70	69	68	68	66	67	67	67	66	67	67	67	67	67
2	45	44	45	44	44	45	44	44	44	44	44	43	43	43	43	43	43	43	43	43
3	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54
4	55	55	55	55	55	55	55	55	54	54	54	54	54	54	54	54	54	54	55	54
5	23	23	24	22	23	23	23	21	21	21	21	20	20	20	19	20	20	20	22	22
6		24	25	24	24	24	24	23	23	23	22	21	21	21	20	21	21	21	23	23
7			21	20	21	21	21	20	19	19	19	19	19	18	18	19	19	19	21	21
8				39	40	40	39	38	37	37	37	37	37	36	36	37	37	37	38	38
9					23	23	23	22	22	22	21	20	21	20	19	20	20	20	22	22
10						04	01	-03	-04	-05	-06	-06	-06	-07	-07	-07	-07	-07	-06	-06
11							16	13	11	11	09	09	09	08	07	07	07	07	08	08
12								25	23	23	23	23	23	23	23	23	23	22	23	23
13									41	41	41	41	41	41	41	41	41	41	41	41
14										08	05	04	05	03	02	03	03	03	04	04
15											15	14	15	13	12	13	13	13	14	14
16												42	42	41	40	41	41	41	41	41
17													08	06	04	05	06	06	06	06
18														32	31	31	31	31	32	32
19															22	23	23	23	24	24
20																47	47	47	48	47
21																	42	41	42	42
22																		41	41	41
23																			57	57
24																				22

Table 3D

Normal Varimax Loading Changes for Holzinger and Harman's Factor D (n = 5, 6, ..., 24)[a]

Test	n = 5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	01	-00	00	01	-07	01	01	01	02	08	13	14	15	17	19	18	19	17	17	17
2	01	01	02	01	00	02	02	02	01	05	08	09	09	10	11	11	10	10	10	10
3	-00	01	01	-00	05	-00	01	-00	-01	02	06	07	07	08	09	09	09	08	08	08
4	-03	-03	-03	-04	-01	-04	-03	-04	-04	01	05	06	06	07	09	08	08	08	07	07
5	-00	-06	-03	-02	-09	-06	-06	-05	-05	07	12	12	14	14	15	14	14	14	14	13
6		06	09	09	05	05	06	06	05	17	21	21	22	21	22	22	22	22	22	21
7			-06	-06	-06	-10	-10	-09	-10	03	07	07	09	08	09	09	09	08	08	08
8				-04	-14	-07	-08	-07	-07	04	10	10	12	12	14	13	13	13	12	12
9					15	11	12	12	11	22	26	26	27	25	26	26	26	26	25	25
10						-00	-07	-00	02	15	19	19	23	24	25	25	25	24	24	24
11							07	12	15	27	31	32	35	36	37	37	37	37	37	36
12								-14	-11	-00	05	05	09	11	13	12	12	11	11	11
13									-17	-05	00	01	04	07	08	08	07	07	07	06
14										47	49	49	50	50	50	50	50	50	50	50
15											49	49	50	50	50	50	50	50	50	50
16												42	42	44	45	44	44	44	44	43
17													64	64	64	64	64	64	64	64
18														54	55	55	54	54	54	54
19															40	40	40	39	39	39
20																26	26	26	25	25
21																	27	26	26	26
22																		36	36	36
23																			22	22
24																				34

[a]Decimal points omitted.

other two factors, which had high loadings from the beginning. For $n = 24$, there appear to be good approximations to the domain normal varimax factors.

The Oblique Case

If the restriction of orthogonality is relaxed, it is impossible to apply directly the quartimax criterion (4) or the normal varimax criterion (10). This is because interfactor relationships are not considered when the criteria are in this form, and when applied all factors will collapse into the same factor—that one factor which best meets the criterion. However, Carroll's version of the quartimax criterion explicitly considers interfactor relationships and an oblique solution is attainable. As suggested by (9), if

$$(13) \qquad c = \sum_{s<t} \{[n \sum_j (a_{js}^2/h_j^2)(a_{jt}^2/h_j^2) - (\sum_j a_{js}^2/h_j^2)(\sum_j a_{jt}^2/h_j^2)]/n^2\},$$

it may be shown that in the orthogonal case $v = -2c$. This alternative form of the normal varimax may then be used to obtain oblique factors. The mathematical problem of minimizing (13) is exactly analogous to Kaiser's [5] treatment for Carroll's criterion. Computationally, the (iterative) solution involves finding the latent vector associated with the smallest latent root of a constantly changing symmetric matrix of order r.

Computational Appendix

To compute an orthogonal normal varimax solution, the following procedure is suggested. The first step is to normalize the rows of the arbitrary reference factor matrix (e.g., principal axes or centroids) by dividing each element by h_i . Rotation to the direction of the normal varimax factors may then be carried out with respect to these normalized loadings.

The criterion (10) will be applied to two factors at a time. For this purpose, the following notation for an orthogonal rotation is convenient.

$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix} \cdot \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} X_1 & Y_1 \\ X_2 & Y_2 \\ \vdots & \vdots \\ X_n & Y_n \end{bmatrix},$$

where x_i and y_i, the present normalized loadings, are constants, and X_i and Y_i , the desired normalized loadings, are functions of ϕ, the angle of rotation.

It is immediately seen that

$$X_i = x_i \cos\phi + y_i \sin\phi, \tag{14}$$

$$Y_i = -x_i \sin\phi + y_i \cos\phi. \tag{15}$$

Thus,

$$dX_i/d\phi = Y_i, \tag{16}$$

$$dY_i/d\phi = -X_i. \tag{17}$$

According to (10), in this plane,

$$n^2 v_{xy} = n \sum (X^2)^2 - (\sum X^2)^2 + n \sum (Y^2)^2 - (\sum Y^2)^2 \tag{18}$$

should be a maximum. Differentiating (18) with respect to ϕ, using (16) and (17), and setting the derivative equal to zero,

$$n \sum XY(X^2 - Y^2) - \sum XY \sum (X^2 - Y^2) = 0. \tag{19}$$

To solve (19) for ϕ in terms of x_i and y_i , substitute the values of X_i and Y_i from (14) and (15), consult a table of trigonometric identities, and, after a good deal of algebraic manipulation,

$$\phi = \tfrac{1}{4} \arctan \frac{2[n \sum (x^2 - y^2)(2xy) - \sum (x^2 - y^2) \sum (2xy)]}{n\{\sum [(x^2 - y^2)^2 - (2xy)^2]\} - \{[\sum (x^2 - y^2)]^2 - [\sum (2xy)]^2\}}. \tag{20}$$

If $u_i = x_i^2 - y_i^2$ and $v_i = 2x_i y_i$, (20) reduces to the form (11) above.

Of course, (11) or (20) is only a necessary condition for a maximum. By taking the second derivative of (18) sufficient conditions for a maximum may be found. These are summarized below.

		sign of numerator +	sign of numerator −
sign of denominator	+	0° to $+22\frac{1}{2}°$	0° to $-22\frac{1}{2}°$
	−	$+22\frac{1}{2}°$ to $+45°$	$-22\frac{1}{2}°$ to $-45°$

The sign of numerator and denominator refer to the right-hand member of (20); the values in the cells refer to ϕ.

These single-plane rotations are made on factors 1 with 2, 1 with 3, $\cdots$, 1 with r, 2 with 3, $\cdots$, 2 with r, $\cdots$, $(r - 1)$ with r, 1 with 2, $\cdots$ iteratively until $r(r - 1)/2$ successive rotations of $\phi = 0$ are obtained, i.e., until the process converges. (It was shown [6] that v in (10) cannot be greater than $(r - 1)/r$, and since each successive application of (20) can result only in a non-decrease of v, this iterative procedure must converge.) After convergence, each normalized test vector is restored to its proper length by multiplying by h_j .

Since this article was accepted for publication, the author has prepared a detailed outline for coding an electronic computer program for the varimax criterion. This (dittoed) paper is available from the writer.

REFERENCES

[1] Carroll, J. B. An analytical solution for approximating simple structure in factor analysis. *Psychometrika*, 1953, **18**, 23–38.
[2] Ferguson, G. A. The concept of parsimony in factor analysis. *Psychometrika*, 1954, **19**, 281–290.
[3] Holzinger, K. J. and Harman, H. H. *Factor analysis.* Chicago: Univ. Chicago Press, 1941.
[4] Kaiser, H. F. An analytic rotational criterion for factor analysis. *Amer. Psychologist*, 1955, **10**, 438. (Abstract)
[5] Kaiser, H. F. Note on Carroll's analytic simple structure. *Psychometrika*, 1956, **21**, 89–92.
[6] Kaiser, H. F. The varimax method of factor analysis. Unpublished doctoral dissertation, Univ. California, 1956.
[7] Neuhaus, J. O. and Wrigley, C. The quartimax method: an analytical approach to orthogonal simple structure. *Brit. J. statist. Psychol.*, 1954, **7**, 81–91.
[8] Saunders, D. R. An analytic method for rotation to orthogonal simple structure. Princeton: Educational Testing Service Research Bulletin 53–10, 1953.
[9] Thomson, G. H. *The factorial analysis of human ability.* (5th ed.) New York: Houghton Mifflin, 1951.
[10] Thurstone, L. L. *Theory of multiple factors.* Ann Arbor: Edwards Bros., 1932.
[11] Thurstone, L. L. Primary mental abilities. *Psychometric Monogr. No. 1*, 1938.
[12] Thurstone, L. L. *Multiple-factor analysis.* Chicago: Univ. Chicago Press, 1947.

Manuscript received 11/15/57
Revised manuscript received 1/9/58