PSYCHOMETRIKA—VOL. 31, NO. 3
SEPTEMBER, 1966

ROTATION FOR SIMPLE LOADINGS

R. I. JENNRICH
AND
P. F. SAMPSON

HEALTH SCIENCES COMPUTING FACILITY
UNIVERSITY OF CALIFORNIA, LOS ANGELES

Existing analytic oblique rotation schemes proceed by optimizing a simplicity function applied to the reference structure. This article suggests optimizing a simplicity function applied to primary loadings directly. The feasibility of the suggestion is demonstrated using the quartimin criterion. An algorithm to implement the optimization is derived and the existence of an admissible solution proved. Practical comparisons with the biquartimin method are made using Thurstone's Box Problem and Holzinger and Swineford's Twenty-Four Psychological Tests Problem.

Analytic rotation schemes in factor analysis are based on a number of simplicity criteria. In the case of orthogonal rotations the primary factor loading matrix is identical to the primary factor structure and rotation for simple primary loadings is equivalent to rotation for simple primary structure. In the case of oblique "rotations," however, the matrices are not identical and it appears a choice would have to be made between rotation for simple primary loadings or simple primary structure. In fact, neither of these alternatives are used. Following Thurstone [5] rotations are made instead for simple reference factor structure. The reference factors are defined as the set of variables of unit variance which lie in the common factor space and are biorthogonal to the primary factors. Since the reference factors seem more like mathematical abstractions than variables of primary interest, an interest in simple reference structure seems a little strange. The reference structure, however, is quite similar to the primary loading matrix (the columns of one are scalar multiples of the columns of the other) and it is apparently this similarity which motivates the interest in simple reference structure.

The purpose of this paper is to show that such an indirect method of obtaining simple loadings is not required.

There are a number of popular oblique rotation schemes—oblimax, quartimin, covarimin, biquartimin, and Kaiser-Dickman to name a few. Most of the simplicity criteria used in these schemes are sufficiently complex that it is difficult, if not impossible, to tell in terms of the reference structure or the primary loading matrix what "simplicity" really means. It will be shown here that by applying the fairly simple quartimin criterion to the loadings directly, very satisfactory results are obtained while avoiding the problem of rotating to singularity, which appears to plague some oblique rotation methods.

I. *Basic Concepts and Statement of the Problem*

Each of the n basic variables z_i in a factor analysis model is a sum

$$z_i = c_i + u_i$$

of two orthogonal variables c_i and u_i called the common and unique parts of z_i. Each common part is a linear combination

$$c_i = l_{i1}f_1 + \cdots + l_{im}f_m$$

of m variables $f_1, \cdots, f_m$ called the primary factors. These factors are assumed to be a basis of the space spanned by the common parts (called the common factor space) and to have unit variance. The model is called orthogonal if it is assumed in addition that the primary factors are uncorrelated. In matrix notation, the model may be written

$$z = Lf + u,$$

where $z = \{z_i\}$ is an $n \times 1$ vector of basic variables, $L = \{l_{ij}\}$ is an $n \times m$ matrix of primary factor loadings, $f = \{f_i\}$ is an $m \times 1$ vector of primary factors and $u = \{u_i\}$ is an $n \times 1$ vector of unique parts.

The selection of a primary factor basis for the common-factor space is called the rotation problem. The selection is made in such a way that the corresponding loading matrix L will be as simple as possible. Formal conditions for simplicity have been stated by Thurstone ([5], p. 335). Roughly speaking, L is called simple if it has many nearly zero elements and a few relatively large elements. The goal is to represent each variable with relatively large loadings on one or at most a few factors and nearly zero loadings on the remaining.

As mentioned earlier, the usual method for doing this is somewhat indirect. Let $g_1, \cdots, g_m$ be a set of variables of unit variance in the common-factor space which are bi-orthogonal to the primary factors. These variables, which are called reference factors, uniquely determine the primary factors and are uniquely determined by them. The covariance matrix

$$S = [\text{cov}(z_i, g_j)]$$

of the basic variables with the reference factors is called the reference factor structure. It is easy to show ([3], p. 279) that the reference structure is related to the primary loading matrix by the equation

$$S = LD,$$

where D is a diagonal matrix with diagonal elements $d_{ii} = \text{cov}(f_i, g_i)$. In other words, the columns of the reference structure S are simply scalar multiples of the columns of the primary loading matrix L. Roughly speaking, S will look simple if L does and conversely. The common analytic rotation schemes proceed by defining a function $F(S)$ which measures the simplicity or complexity of the reference structure S. The rotation problem is then solved by selecting the reference factors which maximize or minimize $F(S)$. In perhaps the simplest case (the quartimin method [1]) the function

$$F(S) = \sum_{p<q} \sum_i s_{ip}^2 s_{iq}^2 \tag{1}$$

is minimized.

The authors believe that the reference structure method is unnecessarily complicated and indirect. They suggest that a simplicity function $F(L)$ be applied directly to the loadings and that the rotation problem be solved by maximizing or minimizing $F(L)$. It is always possible (and customary) to begin with an initial set of primary factors which are orthogonal. If A denotes an initial primary loading matrix then L is an admissible primary loading matrix if and only if it can be written in the form $L = AT^{-1}$, where T is a nonsingular normalized matrix (i.e., the rows of T have length one). Similarly S is an admissible reference structure if and only if it can be written in the form $S = AT'$, where again T is a nonsingular normalized matrix. The simple loadings solution amounts to finding a nonsingular matrix T to

$$\text{minimize } F(AT^{-1}) \text{ under the condition } \operatorname{diag}(TT') = I, \tag{2}$$

while the simple reference structure solution amounts to finding a nonsingular matrix T to

$$\text{minimize } F(AT') \text{ under the condition } \operatorname{diag}(TT') = I. \tag{3}$$

Formally the only difference between the methods is that in the simple loadings method T^{-1} replaces T' in the argument of F.

The T's defined by (2) and (3) are in general different and give rise to different loading and primary factor correlation matrices. In the case of the simple loadings solution, the rotated loadings L and the covariance matrix C of the rotated primary factors are given by

$$L = AT^{-1} \quad \text{and} \quad C = TT'.$$

In the case of the simple structure solution, these formulas are a little more complicated

$$L = AT'D^{-1} \quad \text{and} \quad C = D(TT')^{-1}D$$

where

$$D = [\operatorname{diag}(TT')^{-1}]^{-1/2}.$$

It is difficult to understand why the simple loadings method has not been used. One explanation may be Thurstone's leadership in simple structure techniques. Other explanations may include a fear of the complexity of the mathematical details, a fear that the solution methods would behave poorly numerically or that the final solutions would be unsatisfactory. The next three sections will attempt to demonstrate that these fears are unfounded.

II. *Mathematical Details of Method*

The quartimin simplicity function F defined by (1) will be used. The reason for this choice is that both algebraically and conceptually it is about the simplest criterion available and it seems to give very satisfactory results.

The method proceeds by means of a sequence of elementary rotations. Let $f_1, \cdots, f_m$ denote the factors (i.e., primary factors) at an intermediate step, L the corresponding loading matrix, and $C = [c_{ij}]$ the corresponding factor correlation matrix. Choosing two factors, say f_1 and f_2, a simple rotation consists of rotating f_1 in the plane of f_1 and f_2 in such a way that the resulting loading matrix $\tilde{L}$ minimizes $F(\tilde{L})$. The rotated factor

$$\tilde{f}_1 = t_1 f_1 + t_2 f_2 , \tag{4}$$

where t_1 and t_2 are chosen so that $\tilde{f}_1$ has unit length. This amounts to requiring that

$$t_1^2 + 2t_1t_2c_{12} + t_2^2 = 1. \tag{5}$$

Let $l_1, \cdots, l_m$ and $\tilde{l}_1, \cdots, \tilde{l}_m$ denote the columns of L and $\tilde{L}$. By equating common parts it is easy to see that

$$\tilde{l}_1 = \frac{1}{t_1} l_1 , \qquad \tilde{l}_2 = -\frac{t_2}{t_1} l_1 + l_2 \quad \text{and} \quad \tilde{l}_i = l_i \quad \text{for} \quad i \neq 1, 2. \tag{6}$$

Letting xy denote the element-wise product of two arbitrary vectors x and y and letting

$$(x, y) = \sum_{i=1}^{n} x_i y_i ,$$

the function $F(\tilde{L})$ can be written in the form

$$F(\tilde{L}) = \sum_{p<q} (\tilde{l}_p^2 , \tilde{l}_q^2) = (\tilde{l}_1^2 , \tilde{l}_2^2) + \left(\tilde{l}_1^2 + \tilde{l}_2^2 , \sum_{q=3}^{m} l_q^2\right) + K,$$

where K is constant with respect to t_1 and t_2. Letting $w = \sum_{q=3}^{m} l_q^2$, the problem reduces to minimizing

$$f = \left(\left(\frac{1}{t_1} l_1\right)^2, \left(\frac{t_2}{t_1} l_1 - l_2\right)^2\right) + \left(\left(\frac{1}{t_1} l_1\right)^2 + \left(\frac{t_2}{t_1} l_1 - l_2\right)^2, w\right) \tag{7}$$

over all t_1 and t_2 satisfying restriction (5). Using the change of variable

$$\gamma = \frac{1}{t_1} , \qquad \delta = \frac{t_2}{t_1} , \tag{8}$$

(5) and (7) become

(9) $$\gamma^2 = 1 + 2c_{12}\delta + \delta^2,$$

and

(10) $$f = a + b\delta + c\delta^2 + d\delta^3 + e\delta^4,$$

where

$$a = (l_1^2, w) + (l_2^2, w) + (l_1^2, l_2^2),$$
$$b = 2c_{12}(l_1^2, w) - 2(l_1^2, l_1l_2) - 2(l_1l_2, w) + 2c_{12}(l_1^2, l_2^2),$$
$$c = (l_1^2, l_1^2) - 4c_{12}(l_1^2, l_1l_2) + (l_1^2, l_2^2) + 2(l_1^2, w),$$
$$d = 2c_{12}(l_1^2, l_1^2) - 2(l_1^2, l_1l_2),$$
$$e = (l_1^2, l_1^2).$$

Hence the problem reduces to minimizing, without restrictions, the fourth degree polynomial given in (10). The minimum may be found by choosing the root δ of the cubic equation

(11) $$b + 2c\delta + 3d\delta^2 + 4e\delta^3 = 0$$

which minimizes f. Then, γ and δ are found from (9) and (11) (the sign of γ is arbitrary) and t_1 and t_2 are obtained from (8). The new loadings matrix $\tilde{L}$ is obtained from (6) and the new correlation matrix $\tilde{C} = \{\tilde{c}_{ij}\}$ of the rotated factors from the equations

(12) $$\tilde{c}_{1j} = t_1c_{1j} + t_2c_{2j}, \qquad j \neq 1;$$
$$\tilde{c}_{ij} = c_{ij}, \qquad i, j \neq 1.$$

This completes a simple rotation. Rotations are performed stepping uniformly through all possible pairs of factors until $F(L)$ converges. The final values of L and C are the loading matrix and correlation matrix of the rotated factor solution.

Harry Harman has suggested that the method described in this section be called the direct quartimin method and that simple loading methods in general be called direct rotation methods. For example, methods based on minimizing the oblimin criterion applied directly to the loadings rather than to the reference structure would be called direct oblimin methods. A FORTRAN IV subroutine which implements the direct quartimin method described here and the direct oblimin methods in general has been written by the authors and may be obtained by writing to the program librarian of the Health Sciences Computing Facility, UCLA, Los Angeles, California 90024. The subroutine also exists as part of a factor analysis program BMDX72 which is available at the same address.

III. *Existence and Nonsingularity of Solution*

In this section we will prove that as the common factors approach linear dependence, the simplicity criterion $F(L)$ approaches infinity and also that there exists a loading matrix L which minimizes $F(L)$. While it may be argued that these are merely mathematical niceties, it is not clear that other rotation schemes enjoy them. This will be discussed in the next section.

LEMMA 1. *For any* l_{ij} ,

$$F(L) \geqq \left(\frac{l_{ij}^2 - |l_{ij}|}{m - 1}\right)^2.$$

PROOF.

$$\text{cov}\,(c_i\,, f_j) = l_{i1}\,\text{cov}\,(f_1\,, f_j) + \cdots + l_{im}\,\text{cov}\,(f_m\,, f_j).$$

Thus, $|l_{ij}| \leq 1 + |l_{i1}| + \cdots + |l_{i,j-1}| + |l_{i,j+1}| + \cdots + |l_{im}|$ and thus for some

$$k \neq j, \qquad |l_{ik}| \geqq \frac{|l_{ij}| - 1}{m - 1}.$$

The required result follows from formula (1).

THEOREM 1. *If L is a factor loading matrix and C is the corresponding factor correlation matrix, then $F(L) \rightarrow \infty$ as* det $(C) \rightarrow 0$.

PROOF. Since $L = AT^{-1}$ and $C = TT'$,

$$\det\,(L'L) = \frac{\det\,(A'A)}{\det\,(C)} \rightarrow \infty \quad \text{as} \quad \det\,(C) \rightarrow 0.$$

The required result follows from Lemma 1.

THEOREM 2. *There exists a loading matrix L which minimizes $F(L)$.*

PROOF. Let K be the greatest lower bound of $F(L)$ over all admissible loading matrices L. Since $L = AT^{-1}$ and $C = TT'$ it follows from Theorem 1 that $F(AT^{-1}) \rightarrow \infty$ as det $(T) \rightarrow 0$. Since the set of all (not necessarily non-singular) normalized matrices T is closed and bounded it follows by continuity that there exists a non-singular normalized matrix T such that $F(AT^{-1}) = K$ and hence an admissible loading matrix $L = AT^{-1}$ which minimizes $F(L)$.

In all fairness it should be pointed out that it has not been proved that the algorithm presented in the previous section converges to a loading matrix L which minimizes $F(L)$ or even that it converges at all. No such proofs exist for any of the common analytic rotation schemes. For all methods the fact that they converge to a loading or structure matrix is a matter of experience and that the matrix optimizes the required criterion a matter of faith.

IV. *Some Comparisons with Other Methods*

The oblimin methods minimize the function

$$G(S) = \sum_{p \neq q} \left(\sum_i s_{ip}^2 s_{iq}^2 - \frac{\gamma}{n} (\sum_i s_{ip}^2)(\sum_i s_{iq}^2) \right), \qquad 0 \leqq \gamma \leqq 1$$

of the reference structure S. When $\gamma = 0, 1/2, 1$ these methods are called the quartimin [1], biquartimin [2], and covarimin methods [4]. In this section, a simple example will be used to raise some objections to these methods while in the next section more practical comparisons will be made.

For the example, we will begin with orthogonal factors and the loading matrix

$$\begin{bmatrix} .960 & .140 \\ .480 & .070 \\ -.560 & .300 \end{bmatrix}$$

The loading matrices of the simple loading and several oblimin solutions are

$$\underset{\text{Simple loading}}{\begin{bmatrix} .970 & .000 \\ .485 & .000 \\ .000 & .635 \end{bmatrix}}; \quad \underset{\gamma = 0}{\begin{bmatrix} .970 & .000 \\ .485 & .000 \\ .000 & .635 \end{bmatrix}}; \quad \underset{\gamma = 1/4}{\begin{bmatrix} .951 & -.028 \\ .475 & -.014 \\ -.117 & .547 \end{bmatrix}};$$

$$\underset{\gamma = 1/2}{\begin{bmatrix} .940 & -.051 \\ .470 & -.026 \\ -.217 & .486 \end{bmatrix}}; \quad \underset{\gamma = 3/4}{\begin{bmatrix} .933 & -.084 \\ .467 & -.042 \\ -.294 & .456 \end{bmatrix}}; \quad \underset{\gamma = 1}{\begin{bmatrix} .928 & -.144 \\ .464 & -.072 \\ -.350 & .459 \end{bmatrix}}.$$

A perfect cluster solution exists and the simple loading and quartimin methods produce it. The remaining oblimin methods fail to obtain the "perfect" solution. This observation suggests the use of the simple loadings or quartimin method but the latter method, and to an extent all the oblimin methods with small γ are plagued by a problem which is called the problem of rotating to singularity. More precisely stated, the problem is that in some cases, as the iterations employed in the solution proceed, the reference factor correlation matrix becomes singular. To the authors' knowledge it has never been shown for the oblimin methods that there exists a set of linearly independent reference factors whose structure S minimizes the criterion $G(S)$. It is not clear whether the problem of rotating to singularity is due to the non-existence of a solution S or simply to problems in numerical analysis.

Carroll has suggested two solutions to the problem. One is to use a high value of γ which tends to make the reference factor correlations smaller ([3], p. 324). The other is to put an arbitrary upper bound on the correlations. Either of these alternatives can make it impossible to rotate to a perfect cluster solution when one exists. They also introduce added complexity into the algebraic formulation of the criterion G and into its interpretation.

On the other hand, the simple loadings method uses an algebraically simple criterion, can produce perfect cluster solutions when they exist and, thanks to Theorem 1, will not rotate to singularity.

V. *Practical Comparisons with the Biquartimin Method*

It remains to be demonstrated that the simple loadings method given in Section II works for practical problems—that for real data the method in fact gives loadings that look simple, or at least loadings which look as simple as those of other analytic oblique rotation methods. To do this the simple loadings and biquartimin methods will be compared using two well-known problems from the literature—Thurstone's Box Problem ([5], p. 140) and Holzinger and Swineford's Twenty-Four Psychological Tests Problem ([3], p. 135).

For the purpose of making comparisons with the results in the literature, Kaiser communality normalizations ([3], p. 325) have been used for both the simple loadings and oblimin methods. In both examples the initial factors are orthogonal.

Table 1 contains the simple loadings and biquartimin solutions to the

TABLE 1

Thurstone's Box Problem

Initial Loadings			Simple Loadings			Biquartimin Loadings		
.659	-.736	.138	-.035	1.020	-.056	-.002	1.003	-.036
.725	.180	-.656	1.011	-.032	-.022	.980	.020	.029
.665	.537	.500	-.046	-.066	1.018	-.016	-.047	1.001
.869	-.209	-.443	.765	.456	-.045	.756	.488	.003
.834	.182	.508	-.073	.344	.879	-.032	.353	.870
.836	.519	.152	.364	-.071	.834	.377	-.035	.840
.856	-.452	-.269	.525	.727	-.079	.531	.742	-.037
.848	-.426	.320	-.045	.866	.363	-.004	.859	.373
.861	.416	-.299	.787	-.092	.452	.775	-.041	.484
.880	-.341	-.354	.648	.610	-.059	.648	.633	-.013
.889	-.147	.436	-.066	.656	.644	-.022	.656	.645
.875	.485	-.093	.611	-.092	.653	.611	-.046	.673
.667	-.725	.109	.000	1.006	-.066	.031	.990	-.045
.717	.246	-.619	.989	-.087	.044	.959	-.034	.091
.634	.501	.522	-.091	-.039	.997	-.059	-.023	.978
.936	.257	.165	.327	.222	.725	.347	.250	.736
.966	-.239	-.083	.450	.629	.244	.465	.648	.276
.625	-.720	.166	-.073	.998	-.041	-.039	.980	-.024
.702	.112	-.650	.977	.023	-.071	.947	.072	-.020
.664	.536	.488	-.036	-.068	1.009	-.005	-.049	.992

Primary Factor Correlations

1.000			1.000			1.000		
.000	1.000		.334	1.000		.242	1.000	
.000	.000	1.000	.337	.249	1.000	.248	.196	1.000

Box Problem. Included are the initial loadings, the rotated loadings, and the primary factor correlations. The rotated loadings look fairly similar. This similarity along with their common similarity to Thurstone's graphical solution ([5], p. 228) can be seen in Fig. 1 which is an extended vector representation on the second and third original factors. The primary factor correlations for the simple loadings method are a little higher than those for the biquartimin method.

The simple loadings and biquartimin solutions to the Twenty-Four Psychological Tests Problem are very similar as can be seen from Table 2. The maximum absolute difference in the loadings is .071 and the average absolute difference is .016. All of the loadings except one nearly zero loading have the same sign. Again, the correlations for the simple loadings method are a little higher than those for the biquartimin method.

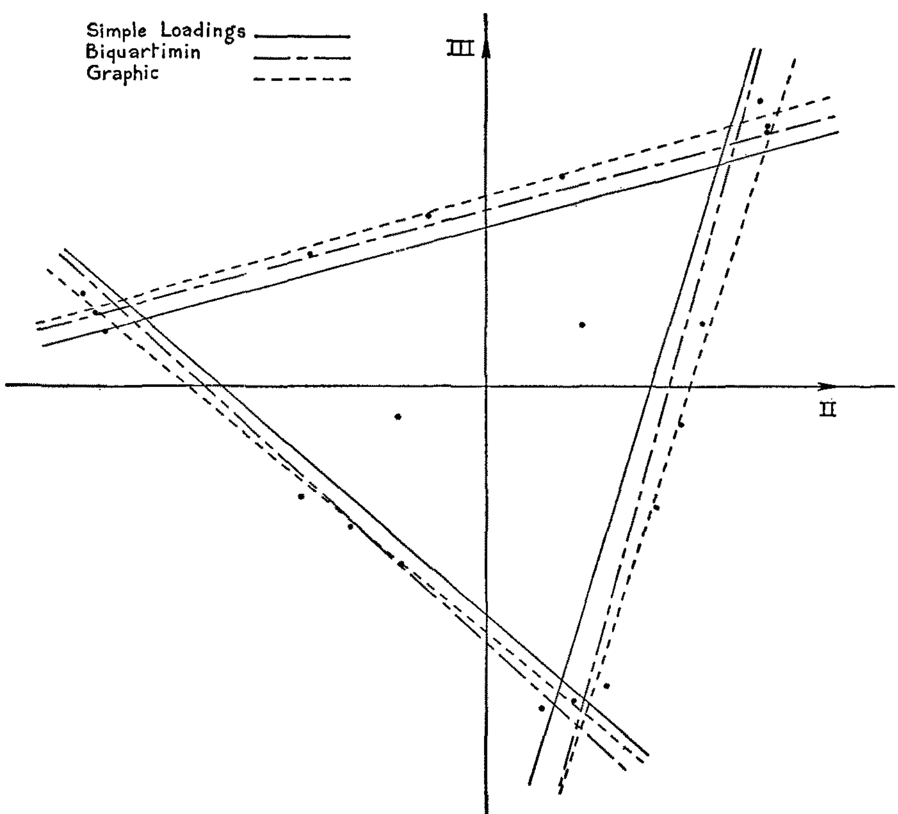

FIGURE 1

Comparison of Oblique Rotations, Extended Vector Representation

TABLE 2

Twenty-Four Psychological Tests Problem

Initial Loadings				Simple Loadings				Biquartimin Loadings			
.608	-.116	.300	-.250	.665	.116	-.004	.054	.666	.103	.015	.058
.372	-.119	.207	-.135	.435	.015	.019	.032	.436	.008	.031	.033
.427	-.220	.262	-.155	.554	-.049	.061	.000	.555	-.057	.074	.000
.477	-.211	.206	-.184	.540	.022	.107	-.030	.545	.012	.117	-.025
.668	-.306	-.344	.108	.024	.113	.779	-.033	.083	.120	.748	-.009
.661	-.337	-.258	.216	.039	-.031	.771	.079	.099	-.018	.742	.095
.652	-.396	-.384	.124	.013	.062	.872	-.096	.076	.069	.836	-.069
.662	-.225	-.153	-.060	.245	.176	.512	-.042	.285	.174	.497	-.021
.664	-.394	-.240	.308	.023	-.140	.831	.134	.087	-.122	.800	.146
.462	.455	-.365	-.136	-.238	.717	.097	.123	-.208	.711	.089	.149
.569	.397	-.208	-.063	-.091	.576	.085	.260	-.060	.574	.082	.276
.484	.360	-.149	-.388	.138	.721	-.087	-.039	.146	.701	-.080	-.012
.608	.130	-.099	-.402	.328	.594	.078	-.121	.341	.573	.082	-.092
.442	.199	-.013	.293	-.110	.059	.152	.501	-.076	.077	.149	.492
.407	.170	.146	.266	.035	-.037	.025	.525	.058	-.021	.032	.508
.523	.077	.300	.076	.357	-.013	-.053	.418	.368	-.008	-.035	.404
.492	.317	.082	.338	-.091	.066	.032	.668	-.059	.086	.037	.652
.547	.307	.248	.072	.229	.171	-.165	.531	.242	.176	-.144	.518
.452	.125	.129	.111	.150	.065	.031	.372	.168	.072	.039	.365
.612	-.174	.128	.004	.397	-.003	.263	.158	.421	-.002	.265	.161
.601	.114	.080	-.171	.339	.324	.032	.149	.353	.316	.042	.159
.608	-.144	.145	.136	.317	-.086	.276	.295	.246	-.070	.277	.291
.691	-.164	.129	-.116	.494	.109	.243	.087	.516	.104	.247	.095
.654	.151	-.150	-.003	.054	.371	.266	.223	.091	.372	.260	.237

Primary Factor Correlations

1.000				1.000				1.000			
.000	1.000			.313	1.000			.295	1.000		
.000	.000	1.000		.434	.313	1.000		.341	.261	1.000	
.000	.000	.000	1.000	.376	.412	.405	1.000	.329	.338	.337	1.000

Conclusion

It has been shown both from a mathematical and practical point of view that it is possible to solve the oblique rotation problem by applying a simplicity criterion directly to the primary factor loadings rather than to the reference factor structure. While it is not the intent of the authors to promote any particular criterion it turns out that the quartimin criterion in addition to being algebraically and conceptually simple, works quite well. The criterion produces perfect cluster loadings when they exist and cannot rotate to singularity, a problem which appears to plague the criterion when applied to the reference structure.

It would be interesting to investigate the behavior of other existing criteria applied to the primary loadings rather than the reference structure.

The authors, for example, have used other oblimin criteria with apparent success. Since some of the elements of the primary factor loading matrix become infinite as the factors become dependent (Lemma 1), while the elements of the reference factor structure matrix are always bounded by one, the problem of rotating to singularity seems less likely to plague a criterion applied to primary loadings than one using reference structure.

REFERENCES

[1] Carroll, J. B. An analytical solution for approximating simple structure in factor analysis. *Psychometrika,* 1953, **18**, 23–28.
[2] Carroll, J. B. Biquartimin criterion for rotation to oblique simple structure in factor analysis. *Science,* 1957, **126**, 1114–1115.
[3] Harman, H. H. *Modern factor analysis.* Chicago: Univ. Chicago Press, 1960.
[4] Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika,* 1958, **23**, 187–200.
[5] Thurstone, L. L. *Multiple-factor analysis.* Chicago: Univ. Chicago Press, 1947.

Manuscript received 10/18/65
Revised manuscript received 1/15/66