PSYCHOMETRIKA—VOL. 59, NO. 4, 567–579
DECEMBER 1994

SIMPLIMAX: OBLIQUE ROTATION TO AN OPTIMAL TARGET WITH SIMPLE STRUCTURE

HENK A. L. KIERS

UNIVERSITY OF GRONINGEN

Factor analysis and principal component analysis are usually followed by simple structure rotations of the loadings. These rotations optimize a certain criterion (e.g., varimax, oblimin), designed to measure the degree of simple structure of the pattern matrix. Simple structure can be considered optimal if a (usually large) number of pattern elements is exactly zero. In the present paper, a class of oblique rotation procedures is proposed to rotate a pattern matrix such that it optimally resembles a matrix which has an exact simple pattern. It is demonstrated that this method can recover relatively complex simple structures where other well-known simple structure rotation techniques fail.

Key words: factor analysis, component analysis, Promax.

Factor analysis (FA) and principal component analysis (PCA) are popular methods for the exploratory analysis of a data set consisting of observations on (many) variables. The solutions of both methods are partly undetermined. Specifically, the factors are determined up to a (possibly oblique) rotation only. The pattern matrix, with loadings of the variables on the factors, is likewise determined up to a transformation. To eliminate this indeterminacy, the factors are usually rotated such that the corresponding pattern becomes as simple as possible, thus facilitating interpretation of the solution. Here "as simple as possible" is defined as using per variable a minimum number of factors to account for it.

Many proposals for finding transformations that optimize the degree of simplicity (parsimony) of the solution have been made. They are usually called simple structure rotations. Some of these, like Kaiser's (1958) well-known varimax rotation, are limited to orthogonal rotations. However, there are many situations where requiring factors to be orthogonal is unnatural (see, Mulaik, 1972, pp. 224–227). Therefore, oblique simple structure methods are called for. The earliest oblique rotation methods, like Carroll's (1957) oblimin family of methods, were directed at finding a simple structure for the so-called reference factors. One reason for choosing to rotate the reference structure rather than the pattern itself to optimal simplicity may have been the mathematical complexity of the problem of the oblique rotation of a pattern: In order to retain the factors normalized to unit length, the pattern transformation matrix T has to be constrained such that $\text{Diag}\,(T^{-1}T^{-1\prime}) = I$ (e.g., see Mulaik, p. 308). However, as pointed out by Jennrich and Sampson (1966), "Since the reference factors seem more like mathematical abstractions than variables of primary interest, an interest in simple reference structure seems a little strange" (p. 313). They proposed to apply the oblimin criteria to the pattern matrix directly, and solved the mathematical problems associated with such a rotation. Other "direct" oblique rotation methods are the Promax rotation (Hendrickson & White, 1964), and the orthoblique rotations (Harris & Kaiser, 1964).

This research has been made possible by a fellowship from the Royal Netherlands Academy of Arts and Sciences. The author is obliged to Jos ten Berge for helpful comments on an earlier version.

Requests for reprints should be sent to Henk A. L. Kiers, Department of Psychology (SPA), Grote Kruisstr.2/1, 9712 TS Groningen, THE NETHERLANDS.

0033-3123/94/1200-92063$00.75/0

The above mentioned rotation methods aim at optimal simplicity of the pattern as measured by certain simple structure criteria. For instance, quartimin (the simplest member of the oblimin family) finds the oblique rotation for which the sum of all cross-products of squared loadings (pattern elements) is minimal. Alternatively, the varimax criterion maximizes the sum of variances of squared loadings on the factors. The orthoblique methods are usually based on a varimax rotation or on other rotations from the orthomax family (see, e.g., Clarkson & Jennrich, 1988), combined with a column weighting procedure. The oblimin and orthomax criteria can be considered indicators of "pattern simplicity" in that it can be proven that the measures reach their optimum for simple patterns where each variable is associated with only one factor (that is, each variable has "complexity" one). However, it is not clear in what sense these criteria measure the "degree of simplicity" or the "deviation from simplicity" in cases that are less ideal. In addition, for patterns that are simple but do not have complexity one for each variable, the oblimin and orthomax measures need not be optimal. A clearer measure for simplicity, or rather, for the deviation from simplicity is the deviation of an obtained rotated pattern matrix from a truly simple pattern matrix. In fact, the Promax method employs such a measure in its second step. That is, after Promax has found an approximation to a truly simple pattern (based on the varimax rotated pattern matrix, the normalized elements of which are taken to some high power, while signs are left unaffected), in the second step Promax performs an oblique Procrustes rotation of the original pattern toward this approximately simple pattern.

Although the Promax approach is appealing, it has at least two drawbacks. Firstly, the matching procedure by Hendrickson and White (1964) is not optimal subject to the constraint Diag $(T^{-1}T^{-1\prime}) = I$. This constraint is enforced only after the target rotation is computed. Secondly, the procedure for obtaining the target depends on finding a good simple structure of the pattern in advance. Using just an orthomax rotation, or even iterative promax procedures like the one by Digman (1966; see Mulaik, 1972, pp. 302–304) may produce a useful target pattern, but these procedures do not guarantee to yield, either the most simple target, or, among a set of equally simple targets, that target that can be approximated best by a rotated pattern.

In the present paper, we propose a modification of the Promax procedure in which a simple target matrix and the target rotation are found simultaneously. As pointed out by an anonymous reviewer, a similar approach has been developed earlier by Shiba (1972), although that study is restricted to orthogonal rotations. Unfortunately, no sources describing this method seem to have appeared in English. In the present paper, it is proposed to find among all simple target matrices that have a specified number of zero elements, that target matrix that can be approximated best by a rotated pattern (henceforth called the "best" simple target matrix), thus offering a solution to the second problem of the original Promax procedure. It should be emphasized that this procedure does not rely on a subjective choice of the target, but that the target is the (objectively) best simple target with a specified number of zeros. To avoid the first problem of the original Promax procedure, it will be ensured that the target rotation procedure is optimal subject to the constraint Diag $(T^{-1}T^{-1\prime}) = I$. It will be shown that the new rotation method finds that oblique rotation that gives the minimal sum of squares for the p smallest (in the absolute sense) elements of the pattern, where p is to be specified by the user. In this way, the method can be seen to MAXimize the SIMPLIcity of the rotated pattern (for given p). Therefore, the method is called SIMPLIMAX rotation here. By comparing results for different values of p, we can find the solution(s) that have most elements close to zero (e.g., smaller than .05 in the absolute sense). Thus, several SIMPLIMAX analyses can help find the simplest available solution.

The present paper starts by describing an algorithm for SIMPLIMAX rotation. Then, the performance of the SIMPLIMAX rotation on artificial data sets (with known truly simple patterns) is studied. Next, some guidelines are given as to how the number p can be specified, and these are illustrated by an analysis of the well-known Box Data (Thurstone, 1947). The latter analysis also serves to demonstrate that SIMPLIMAX finds the (in this case) known simple structure, whereas five well-known other simple structure rotation techniques fail to do so. The remainder of the paper is devoted to constrained versions of the SIMPLIMAX method. These methods are designed for situations where it is useful to specify the desired complexity for each variable in advance. It is shown that the SIMPLIMAX rotation method can easily be modified to allow for such constraints. The method for rotation to an optimal target with complexity one for each variable is also illustrated by an exemplary analysis.

An Algorithm for SIMPLIMAX Rotation

The problem to be solved by SIMPLIMAX rotation is that of finding the truly simple target matrix with p zero elements that can be approximated best by a rotated pattern matrix (the "best" simple target). If the position of the zero elements were known, we would deal with the situation handled by Lawley and Maxwell (1964), Jöreskog (1965), Gruvaeus (1970), Browne (1972b) and Kashiwagi (1989). However, none of these procedures determine *which* elements should be taken zero. In SIMPLIMAX, we do not only need to perform a target rotation, but also to determine the best positions of the zero elements (i.e., those leading to the best simple target). Explicitly, let $P(m \times r)$ be a given (orthogonal) pattern matrix, $T(r \times r)$ the pattern transformation matrix, and G a target pattern matrix which has at least p zero elements. Then, SIMPLIMAX minimizes

$$\sigma(T, G) = \|PT - G\|^2, \tag{1}$$

over T subject to Diag $(T^{-1}T^{-1\prime}) = I$, *and* over all matrices G that have p zero elements and $(mr - p)$ arbitrary elements. Note that the target employed here differs from the one used in Promax in that it has many elements that are exactly 0 rather than close to zero, and in that it is determined through the minimization process itself rather than in advance.

To handle the above minimization problem, we introduce an auxiliary set of (indicator) parameters as follows. According to the constraint, G has (at least) p zero elements. To indicate which elements are constrained to be zero and which are not, we use the $m \times r$ binary indicator matrix W which has p zero elements at the positions of the p zero constrained pattern elements, and unit elements elsewhere. The positions of the zero and unit elements in W have to be determined by the minimization procedure itself. Therefore, the matrix W can be viewed as an auxiliary set of parameters in the loss function. Then, writing a_{ij} for element (i, j) of PT, the function to be minimized can be written as

$$\begin{aligned}\tilde{\sigma}(T, G, W) &= \sum_{i=1}^{m}\sum_{j=1}^{r} (a_{ij} - w_{ij}g_{ij})^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{r} (1 - w_{ij})a_{ij}^2 + \sum_{i=1}^{m}\sum_{j=1}^{r} w_{ij}(a_{ij} - g_{ij})^2, \end{aligned} \tag{2}$$

where the first term on the right-hand side pertains to the elements of G that are constrained to zero, and the second term pertains to the unconstrained elements of G. We can now eliminate G from the problem by minimizing $\tilde{\sigma}$ over G, as is done by taking $g_{ij} = a_{ij}$ if $w_{ij} = 1$, and $g_{ij} = 0$ if $w_{ij} = 0$, for $i = 1, \ldots, m, j = 1, \ldots, r$. Then it remains to minimize

$$f(T, W) = \sum_{i=1}^{m} \sum_{j=1}^{r} (1 - w_{ij})a_{ij}^2 \tag{3}$$

over T subject to $T'T = I$ and over W subject to the constraint that it has p zero and $mr - p$ nonzero elements. It follows that the T and W that, together with G chosen such that $g_{ij} = w_{ij}a_{ij}$, $i = 1, \ldots, m, j = 1, \ldots, r$, minimize $\tilde{\sigma}(T, G, W)$ also minimize f(T, W). Hence, the T found by minimizing (3) is the same as that that would be found by minimizing (2). Therefore, from now on we will only consider the problem of minimizing (3) over T and W.

To minimize (3), it is proposed here to alternately minimize f over W for given T, and over T for given W. First, we consider the minimization of f(T, W) over W, that is, the procedure to determine the best positions of the zero elements. Because w_{ij} is constrained to be either one or zero, the minimum of f(T, W) is found by determining the p smallest values of a_{ij}^2 and setting the associated elements w_{ij} of W equal to zero, and the others equal to one. It can be seen now that, according to formula (3), SIMPLIMAX rotation optimizes a straightforward measure for simplicity of a pattern: It minimizes the sum of p smallest squared elements of PT over T and W.

The problem of minimizing (3) over T, for given W, subject to the constraint Diag $(T^{-1}T^{-1\prime}) = I$, has been solved by Browne (1972b). Browne's method minimizes the more general function $\sum_{i=1}^{m} \sum_{j=1}^{r} (1 - w_{ij})(a_{ij} - b_{ij})^2$, where b_{ij} denotes fixed values of a given target matrix. His function reduces to (3) in case all these fixed values equal zero. Browne's procedure for minimizing this function consists of a series of rotations of single columns of loadings in a plane. Each such rotation decreases the function value (or at least does not increase it). Cycling through all possible planes, and repeating this procedure until the function value does not decrease considerably anymore, we have an algorithm for minimizing (3) over T for fixed W. Alternately minimizing f over W considering T fixed, and over T considering W fixed, we can decrease the function value of (3) until convergence of the function value. Rather than actually minimizing f over T (considering W fixed), we can use just one complete cycle of planar updatings of T (which also decreases the function value), then turn to updating W, then perform one cycle for updating T again, etcetera, until convergence.

It should be noted that the algorithm requires an initialization for T. A simple choice is to choose $T = I$, or to choose a random T satisfying the constraint on T. Other useful choices can be obtained from the solutions for T from other simple structure rotation techniques. Such starts are usually referred to as "rational starts". Because the algorithm is rather sensitive to hitting local optima (as will be seen below), it is advised to use many different (among which several random) starts for T in one SIMPLIMAX analysis; the one giving the lowest function value is then considered as the actual SIMPLIMAX solution, and it is hoped that this indeed is the global minimum of the loss function. A computer program using the above algorithm and a procedure for using random restarts can be obtained from the author.

Analyses of Artificial Data

To test the SIMPLIMAX rotation method, 40 truly simple pattern matrices were constructed. The first 20 pertained to $m = 20$ variables and $r = 4$ factors, the last 20 to $m = 30$ variables and $r = 5$ factors. In all cases each variable was associated with at least one factor, and with each factor at least m/r variables were associated. The remaining pattern elements were determined in a random fashion. All nonzero pattern elements exceeded 0.25, so that they can be distinguished clearly from the zero values. For the first ten 20×4 pattern matrices the variables had complexity 1.5 on the average ($p = 50$); in the second series of 20×4 patterns the variables had complexity 1.75 on the average ($p = 45$). For the first ten 30×5 pattern matrices the variables had complexity 1.5 on the average ($p = 105$), for the last ten they had complexity 2 on the average ($p = 90$).

Each of the above described pattern matrices was scaled rowwise such that $R = P\Phi P'$ was a correlation matrix, where Φ denotes a factor correlation matrix determined by drawing a random sample of 100 observation units with scores on r (four or five) factors. Next, a principal components analysis was used to obtain, for each correlation matrix, an orthogonal pattern (for the right number of factors). The above procedure ensured that the resulting 40 unrotated pattern matrices could be transformed into a truly simple pattern matrix (with 50, 45, 105, and 90 zero elements, respectively). A first purpose of the present simulation study is to verify whether the new rotation method indeed recovers these truly simple patterns. If it does not, this must be attributed to the local minimum problem of which the method may suffer.

In each analysis we used one rational start for T (based on normalized varimax) and 20 random orthogonal starts. For each analysis we considered the iterative process to have converged if either the difference between function values of two consecutive cycles differed by less than .001%, or if the function value dropped below 10^{-7}. To avoid very long computation times each analysis was allowed a maximum of 100 iterations, but this maximum was reached only occasionally.

For all twenty 20×4 patterns, and for the ten 30×5 patterns with $p = 105$, the truly simple pattern was recovered by at least one of the 21 analyses. For the ten 30×5 patterns with $p = 90$, the truly simple pattern was not recovered in three cases. Apparently, more (randomly started) analyses are needed to obtain the global minimum for data with this fairly high complexity and size. More detailed insight in the local minimum problem can be obtained from Table 1, which lists the average numbers of local minima in percentages, both for the analyses where a random start was employed, and for those where a rational start was used. It should be noted, however, that the SIMPLIMAX method itself (incorporating *all* runs from different starting positions) hit a local optimum in only 3 out of 40 cases.

A second purpose of the present study is to compare the SIMPLIMAX method to some well-known direct rotation methods: Normalized Varimax, Promax (using the power four), Harris and Kaiser's Independent Cluster Rotation (Harris & Kaiser, 1964, pp. 356–360; denoted as HKIC), Harris and Kaiser's orthoblique rotation with M taken to the power $\frac{1}{2}$ (Harris & Kaiser, p. 361; denoted as HKM$^{1/2}$), and Direct Quartimin. Specifically, the main question of interest was to what extent the truly simple pattern was recovered. To compare the resulting rotated matrices with the associated truly simple pattern, the columns of the resulting matrices were permuted and reflected to optimal congruence with the truly simple pattern. Next, the positions of the p smallest (in the absolute sense) elements of the rotated patterns were determined, and an associated indicator matrix was computed. This indicator matrix was compared to the true indicator matrix, and the number of differences was considered as a crude but simple

TABLE 2

Average Misfit Values For Five Simple Structure Rotation Techniques (Number of Cases Without Errors Parenthesized)

m	r	p	Normalized Varimax		Promax		HKIC		HKM$^{1/2}$		Direct Quartimin	
20	4	50	1.2	(7)	1.0	(8)	1.3	(7)	1.1	(7)	1.2	(8)
20	4	45	8.5	(0)	6.1	(2)	2.6	(4)	3.6	(4)	7.8	(2)
30	5	105	5.2	(2)	1.2	(8)	0	(10)	.2	(9)	0.4	(9)
30	5	90	25.5	(0)	19.3	(0)	8.4	(1)	12.4	(1)	13.0	(1)

measure of "misfit" of the rotated pattern solution. It should be noted that a misfit value higher than one implies that the obtained pattern differs substantially from the truly simple one in that a variable is associated with a different factor, so all deviations pertain to differences that are more than just small size differences. Table 2 lists the average numbers of misfits in the four conditions for the five methods. Clearly, there are many instances where all methods fail to find or even approximate the available truly simple pattern, especially in cases with average complexity higher than 1.5. Considering that SIMPLIMAX (as the method incorporating 20 runs from random restarts) obtained the truly simple pattern in all but 3 cases (as was found above), it can be

TABLE 1

Percentages of Local Minima for SIMPLIMAX Rotation of Four Types of Contrived Pattern Matrices

m	r	p	Random Start	Rational Start
20	4	50	61.0 %	20 %
20	4	45	86.5 %	60 %
30	5	105	54.0 %	20 %
30	5	90	96.0 %	100 %

concluded that the new method does a much better job in recovering truly simple patterns, especially in cases where the complexity is considerably larger than one.

Choosing the Number of Zero Pattern Elements in the Target

In practice, the number p of (near) zero elements of the pattern is not given. Several approaches to determining this value can be conceived of. For instance, one may rotate a pattern to simple structure by any simple structure rotation procedure, count the number of "small" elements, and choose this value for p. However, such an approach will fail as soon as the simple structure procedure yields a solution that by no means approximates a truly simple pattern. Therefore, it seems necessary to determine p without recourse to existing simple structure methods.

A reasonable approach to determine p is analogous to the scree-test used in factor analytic techniques for determining the number of factors. To determine the most useful p value, one can analyze the data by using a range of p values and compare associated function values. As p increases, the function value usually increases as well. The function value is the sum of squared smallest elements, so the increase can be seen as an indication for the impact of adding one more zero constraint on the total size of the smallest elements. As a cut-off point for p one can take that value whereafter the function value shows a considerable relative increase.

Obviously, this procedure has certain drawbacks. First of all, the question how to choose the range of p values to check remains unanswered. The safest way is to try all possible values (from 0 to mr), but more efficient procedures are desirable. For instance, we can always find an oblique rotation yielding at least $r(r - 1)$ exact zeros, and, in practice, we will never be able to find more than $(m - 1)r$ zeros (because that would imply that at least one column of the pattern matrix becomes zero, and hence one factor or component would be superfluous). Usually, it will not be difficult to indicate a maximum desirable average complexity for the variables; in addition, one may specify a maximum function value, or, a maximum average size of the smallest elements, above which one is no longer inclined to consider the smallest elements really small. A second problem arises in cases where there is no clear jump in the function values. Then it will be hard to determine a cut-off point at all. In all problematic cases, one may also compare solutions in terms of interpretability, where one may use substantive information in addition to the numerical information yielded by the method itself. It should be noted that, like determining the number of factors in a factor analysis, this procedure has a nonnegligible subjective component.

SIMPLIMAX Applied to the Pattern for the Box Data

To illustrate the SIMPLIMAX method, and to show that determining the value for p can be quite simple indeed, we reanalyzed the pattern matrix for Thurstone's (1947) Box Data, as given by Cureton and Mulaik (1975). The Box Data have been constructed by Thurstone as an example where three factors are used to construct 26 variables by highly different but simple nonlinear combinations of the factors (for which the formulas are mentioned in Table 3). Because of the nonlinearity of the construction formulas, there need not be a truly simple pattern for these data, but it is expected (and found for instance by graphical rotation) that a good approximation to a truly simple pattern can be obtained.

The reported unrotated pattern matrix was first analyzed by the five simple structure methods that were used in the previous section. None of these recovered the pattern specifying from which factors the variables were originally constructed. For

TABLE 3

SIMPLIMAX and HKM$^{1/2}$ Rotated Patterns for Box Data

Variable Formula	SIMPLIMAX			HKM$^{1/2}$		
	1	2	3	1	2	3
x	.99	-.01	-.01	.64	.70	-.29
y	.06	.94	.04	.65	-.24	.63
z	.01	.05	.97	-.21	.57	.74
xy	.64	.65	-.02	.84	.23	.23
xz	.60	.00	.65	.23	.82	.31
yz	-.02	.61	.64	.23	.17	.87
x^2y	.83	.39	.01	.80	.47	.02
xy^2	.38	.82	.03	.78	.02	.44
x^2z	.79	-.02	.42	.40	.82	.07
xz^2	.45	.02	.86	.09	.83	.52
y^2z	-.02	.76	.45	.38	.01	.83
yz^2	-.02	.43	.78	.07	.32	.86
x/y	.75	-.83	.02	-.06	.81	-.73
y/x	-.75	.83	-.02	.06	-.81	.73
x/z	.81	.02	-.82	.75	.07	-.82
z/x	-.81	-.02	.82	-.75	-.07	.82
y/z	-.02	.86	-.81	.76	-.78	-.03
z/y	.02	-.86	.81	-.76	.78	.03
2x+2y	.55	.72	-.02	.83	.14	.29
2x+2z	.56	-.03	.69	.17	.82	.33
2y+2z	.00	.61	.63	.24	.18	.86
$(x^2+y^2)^{1/2}$	.54	.71	-.01	.82	.14	.30
$(x^2+z^2)^{1/2}$	.53	-.01	.68	.16	.79	.34
$(y^2+z^2)^{1/2}$	.02	.61	.60	.27	.17	.83
xyz	.45	.48	.46	.49	.45	.52
$(x^2+y^2+z^2)^{1/2}$	.34	.53	.48	.45	.36	.60

instance, all these methods assigned complexity two or more to the first three variables, whereas these variables should each reflect simply one of the original constituting factors. Next, we applied SIMPLIMAX, choosing the values $p = 25, \ldots, 30$ (where it should be noted that the number of zero values corresponding to Thurstone's construction formulas is $p = 27$). For each analysis we used one rational start (from the normalized varimax solution) and ten random additional starts; the same convergence criterion was used as in the analyses of the artificial data. The optimal function values were .0074, .0107, .0149, .1202, .1628, and .1969, respectively, where it may be of interest to mention that none of these were found after using the rational start. Clearly, a large jump in function values occurs after the third analysis (corresponding to $p = 27$). On the basis of these results, it seems rational to choose the solution with $p = 27$ as the one striking the best compromise between parsimony (as many near zeros as possible) and fit (average squared deviation from zero for the alleged near zeros). The $p = 27$ solution is displayed in Table 3 (with elements larger than .30 or smaller than −.30 set in bold face), and it can be verified that, in contrast to the solutions from the five well-known simple structure rotation techniques, the present solution does indeed agree completely with the original constituting rules used to construct the data. The correlation between the components 1 and 2 is .20, between components 1 and 3 is .27 and between 2 and 3 is .28. For comparison, Table 3 also displays the pattern resulting from $\mathrm{HKM}^{1/2}$ rotation, and it can be seen that differences are substantial indeed. For this solution, the correlations between the components are .16, .21 and .18, respectively.

To give some insight in computation times for the SIMPLIMAX method, we measured, for each value of p, the computation times on a 80386/80387 machine for the 20 randomly started runs (which constitute the major part of a full SIMPLIMAX analysis). For p ranging from 25 to 30, we found computation times of 23.5, 22.3, 23.2, 21.8, 21.7, and 20.8 seconds, respectively, which seems acceptable from a practical point of view.

Fixing the Variable Complexities in the Target

In the above described SIMPLIMAX method, only the average complexity of the variables needed to be specified in advance. In certain cases, however, the user may wish to specify the complexity of each variable a priori. For instance, the user may be interested in obtaining that rotation that gives the best solution assigning each variable to one and only one factor (complexity one for each variable). Other situations are those where each variable is to be assigned to (at most) two factors, or where, for each variable, the complexity can be specified in advance, on the basis of substantive reasonings. Although such procedures limit the range of possible solutions and may hence miss a solution which is more parsimonious than the one obtained, they also have the advantage of limiting the number of analyses to be done and avoiding the possibly difficult decision as to which is the "right" value for p. Therefore, it will now be described how a rotation to a truly simple pattern with the complexity specified per variable can be obtained.

As above, the problem is to minimize the function (3) over W for given T. In contrast to the above, now for the matrix W it has been specified how many zeros it should have per row. Given this specification, we can obtain the optimal matrix W updating one row of W at a time. Let $\mathbf{w}_i'$ denote row i of W, and assume that $\mathbf{w}_i$ is constrained to have at least p_i zero elements. The elements of W should again be found such that they minimize the function (3), hence the elements of $\mathbf{w}_i$, $i = 1, \ldots, m$, should be chosen so that they minimize

$$f_i(T, \mathbf{w}_i) = \sum_{j=1}^{r} (1 - w_{ij})a_{ij}^2. \tag{4}$$

The solution is given at once by setting those p_i elements of $\mathbf{w}_i$ to zero that correspond to the p_i smallest elements of $\{a_{i1}^2, \ldots, a_{im}^2\}$, while the other elements are set to one. Thus, one can obtain the complete solution for W, given T. Combining this with (one cycle of) Browne's (1972b) procedure for updating T, one has an algorithm that decreases the function monotonically, until the function value stabilizes. For instance, in the case where each variable is specified to have complexity one, the unit elements in W are assigned to the positions corresponding to the rowwise highest (in the absolute sense) values of PT. In this case, the method minimizes the sum of $(r - 1)$ rowwise smallest squares of the elements of the transformed pattern matrix.

Analysis of Twenty-four Psychological Tests Data by SIMPLIMAX with Complexity One

To illustrate the SIMPLIMAX method with the constraint that each variable in the target has complexity one, we analyzed an empirical data set that has been subjected to various simple structure rotations in the literature. The data set (Holzinger & Swineford, 1939; also, see Harman, 1976, p. 124) consists of scores on 24 psychological tests. The correlation matrix was analyzed by PCA, and the resulting four-components solution was rotated by SIMPLIMAX with complexity one. The results are displayed in Table 4. For comparative purposes we also used the HKIC procedure (as this technique was explicitly designed to reveal a simple structure of complexity one). The HKIC results are also reported in Table 4. All elements larger than .35 (in the absolute sense) are set in bold face. It can be seen that the solutions are slightly different. The function values for (4), the sums of rowwise smallest squares, are 1.99 for the HKIC solution, and 1.68 for the optimal oblique complexity one solution. The latter solution was obtained from 3 of the 9 random starts used in this analysis. When the HKIC solution was used as a start, a local minimum was found (f = 1.71); local minima were also found for the six other starts (f ranging from 1.71 to 2.01).

The pattern matrices differ mainly in terms of a few variables, with variables 20 and 22 belonging mainly to component 1 in the HKIC solution, and to component 2 in the SIMPLIMAX solution. Other differences can be observed as well. For instance, if we consider only values higher than .35 (bold face), which is the lowest value such that each variable is associated to at least one component, the SIMPLIMAX solution yields exactly complexity one for all variables, whereas in the HKIC solution, variables 13, 16, and 21 belong to two components.

Discussion

As reported by Cureton and Mulaik (1975), the Box Data have a long tradition of withstanding attempts to recover the original constituting rules by analytic simple structure procedures, as was confirmed by the analyses of these data by the five well-known simple structure techniques employed here. The fact that the SIMPLIMAX procedure proposed in the present paper does succeed in recovering the underlying structure demonstrates that SIMPLIMAX can indeed perform better than the most popular existing techniques. There is at least one alternative technique that also recovered the underlying structure for the Box Data: Cureton and Mulaik's weighted varimax procedure. However, their procedure is based on assuming certain symmetries in a pattern,

TABLE 4

Pattern Matrices Resulting from HKIC Rotation and SIMPLIMAX With Complexity One Rotation of the Solution for the 24 Psychological Tests Data

	HKIC				SIMPLIMAX With Complexity One			
Variable	1	2	3	4	1	2	3	4
1	**.82**	- .09	.05	- .04	**.65**	.04	.17	.09
2	**.71**	- .09	- .06	- .11	**.59**	.02	.04	.00
3	**.78**	- .04	- .23	- .02	**.65**	.09	- .12	.10
4	**.70**	.09	- .07	- .14	**.58**	.20	.03	- .04
5	- .04	**.83**	.11	- .04	- .04	**.82**	.13	- .07
6	- .01	**.85**	- .09	.09	- .01	**.85**	- .06	.06
7	- .04	**.93**	.03	- .11	- .03	**.91**	.06	- .13
8	.20	**.61**	.11	- .07	.16	**.63**	.16	- .05
9	- .04	**.90**	- .10	.07	- .03	**.90**	- .07	.04
10	- .30	.08	**.94**	.00	- .30	.04	**.90**	- .03
11	- .10	.02	**.62**	.28	- .14	.03	**.61**	.28
12	.13	- .18	**.86**	- .12	.06	- .16	**.87**	- .07
13	**.42**	- .02	**.59**	- .18	.31	.04	**.65**	- .10
14	- .18	.14	- .09	**.76**	- .21	.16	- .09	**.72**
15	- .04	- .02	- .10	**.76**	- .09	.02	- .09	**.73**
16	**.46**	- .18	- .17	**.58**	.33	- .07	- .09	**.64**
17	- .15	.00	.13	**.73**	- .19	.02	.13	**.70**
18	.30	- .28	.29	**.45**	.18	- .19	.34	**.50**
19	.20	- .03	.07	**.39**	.12	.03	.11	**.42**
20	**.38**	.31	- .10	.19	.30	**.39**	- .03	.23
21	**.42**	- .03	**.43**	.00	.30	.04	**.49**	.08
22	**.36**	.30	- .05	.18	.28	**.37**	.02	.22
23	**.50**	.26	.08	.03	**.40**	.34	.16	.11
24	.00	.29	**.50**	.12	- .05	.30	**.52**	.12

which happen to be present in the current data set, but for which there is no reason to expect their presence in real-life data sets. If the symmetry is not present, their procedure may well fail to recover a truly simple pattern, as was found after some experimentation with artificial data.

The SIMPLIMAX method generally yields oblique rotations. In some cases, however, one might want to restrict the class of rotations to orthogonal rotations. For such cases, one can modify the SIMPLIMAX method by minimizing (1) subject to the constraint that T is orthonormal. To do so, we could simply replace the Browne (1972b) procedure by the Browne (1972a) procedure, using his Formula (3) (on p. 116). Alternatively, we could construct a possibly simpler algorithm by exploiting the fact that there now is an explicit solution (Cliff, 1966) for the problem of minimizing $\sigma(T, G)$ over T (considering G fixed); the problem of minimizing σ over G (considering T fixed) is solved by setting g_{ij} equal to 0 for the pairs (i, j) corresponding to the p smallest values of a_{ij}^2, and setting $g_{ij} = a_{ij}$ for the other values of (i, j). Alternating these procedures of updating G and T, we also have a monotonically convergent algorithm for the present minimization problem.

The SIMPLIMAX method seems to have two disadvantages compared to the well-known oblique simple structure rotation methods. The first is its severe local minimum problem. The second is the possibly problematic choice of the number p. As to the local minimum problem, it seems important to note that it can be mitigated by taking a sufficient number of random restarts (as was illustrated for the artificial data), where only 3 out of 40 complete analyses resulted in a locally optimal solution. Therefore, it should be emphasized that a user employs many different (usually random) restarts for one SIMPLIMAX analysis. With modern computational facilities and the introduction of parallel processing computers (which are particularly suitable for handling independent analyses of the same data as is the case with multistart procedures), this limitation may be of minor importance in the near future. Moreover, the SIMPLIMAX algorithm usually converges fast and requires little computational effort, as indicated by the reasonable computation times for the Box Data. The second disadvantage (the possibly problematic decision with respect to the value of p) seems somewhat more disturbing. It resembles the decision problems the user is faced with in the orthoblique, oblimin and Promax methods. In all these methods the user also has to specify a certain parameter (viz., the power of M in orthoblique, γ in oblimin, and the power parameter in Promax) which may affect the solution considerably. For those methods some more or less standard choices have been suggested (based on results from simulation studies), but there are no choices that are good for *all* situations. So, although for such methods the decision problem is usually ignored, it is nevertheless present. In the SIMPLIMAX method the decision problem cannot be ignored as easily. Without the use of external substantive information on the data at hand, the choice for the value of p must be made on the basis of comparing different results for different values of p. Fortunately, this comparison can be based on the fit values after rotation of the pattern matrix at hand. These values are straightforward measures for simplicity of the solution, and are comparable across solutions with different values for p. In this respect, one can view the fact that the value for p in SIMPLIMAX has to be chosen on the basis of results for the data at hand as an advantage rather than a disadvantage over the parameter choice procedures in orthoblique, oblimin and Promax: For the latter, no choice procedures seem to be available that are based on results for the data at hand.

References

Browne, M. W. (1972a). Orthogonal rotation to a partially specified target. *British Journal of Mathematical and Statistical Psychology*, *25*, 115–120.

Browne, M. W. (1972b). Oblique rotation to a partially specified target. *British Journal of Mathematical and Statistical Psychology, 25*, 207–212.

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

Clarkson, D. B., & Jennrich, R. I. (1988). Quartic rotation criteria and algorithms. *Psychometrika, 53*, 251–259.

Cliff, N. (1966). Orthogonal rotation to congruence. *Psychometrika, 31*, 33–42.

Cureton, E. E., & Mulaik, S. A. (1975). The weighted varimax rotation and the Promax rotation. *Psychometrika, 40*, 183–195.

Digman, J. M. (1966). *The procrustes class of factor-analytic transformations*. Unpublished manuscript, University of Hawaii.

Gruvaeus, G. T. (1970). A general approach to Procrustes pattern rotation. *Psychometrika, 35*, 493–505.

Harman, H. H. (1976). *Modern factor analysis* (3rd ed.). Chicago: University of Chicago Press.

Harris, C. W., & Kaiser, H. F. (1964). Oblique factor analytic solutions by orthogonal transformations. *Psychometrika, 29*, 347–362.

Hendrickson, A. E., & White, P. O. (1964). PROMAX: A quick method for rotation to oblique simple structure. *British Journal of Statistical Psychology, 17*, 65–70.

Holzinger, K. J., & Swineford, F. (1939). *A study in factor analysis: The stability of a bi-factor solution* (Supplementary Educational Monographs No. 48). Chicago: University of Chicago, Department of Education.

Jennrich, R. I., & Sampson, P. F. (1966). Rotation for simple loadings. *Psychometrika, 31*, 313–323.

Jöreskog, K. G. (1965). *On rotation to a specified simple structure* (Research Bulletin). Princeton, NJ: ETS.

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

Kashiwagi, S. (1989). *Orthogonal and oblique Procrustes factor rotation methods based on the procedure of single plane assuming the multiple classification of test vectors, Japanese psychological monographs 19*. Tokyo: University Press. (in Japanese)

Lawley, D. N., & Maxwell, A. E. (1964). Factor transformation methods. *British Journal of Statistical Psychology, 17*, 97–103.

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

Shiba, S. (1972). *Factor Analysis*. Tokyo: University of Tokyo Press. (in Japanese)

Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.

Manuscript received 10/19/92
Final version received 10/25/93