Multivariate Behavioral Research, 36 (1), 111-150

An Overview of Analytic Rotation in Exploratory Factor Analysis

Michael W. Browne
The Ohio State University

The use of analytic rotation in exploratory factor analysis will be examined. Particular attention will be given to situations where there is a complex factor pattern and standard methods yield poor solutions. Some little known but interesting rotation criteria will be discussed and methods for weighting variables will be examined. Illustrations will be provided using Thurstone's 26 variable box data and other examples.

Introduction

The rotation of a factor matrix is a problem that dates back to the beginnings of multiple factor analysis. In the early days before electronic computers were available, the process of rotation had to be carried out by hand. It was tremendously time consuming, lasting weeks or even months. During this arduous process, however, there was opportunity for considerable subjective input. Each decision as to the next change in orientation of factor axes could be guided, not only by the current configuration of points, but by background knowledge concerning the manifest variables.

My first acquaintance with factor analysis occurred after computers had arrived on the scene and reliable computerized methods for orthogonal rotation had been developed. Four authors (Carroll, 1953; Saunders, 1953; Neuhaus & Wrigley, 1954; Ferguson, 1954) had independently proposed rotation criteria with different rationales that yielded the same solution in orthogonal rotation. This solution was known as quartimax (see Harman, 1976, Section 13.3). A little later, Kaiser (1958, 1959) had proposed his

I am indebted to Krishna Tateneni and to Carina Tornow for their help in the development of the rotation algorithms used for providing the results reported here. Part of the research reported here was carried out during my sabbatical leave at the University of Virginia. My thanks go to Jack McArdle for suggesting the analysis of the WJR-WAIS data and for several valuable discussions. The work reported here has benefited, also, from helpful comments by Norman Cliff, Bob Cudeck, William Meredith, John Nesselroade, Jim Steiger and Michael Walker.

Support for this research was provided by NIMH grant N4856118101 and NIA grant AG-04704.

varimax criterion and associated algorithm. This method was found to be less prone to producing a general factor than quartimax and became very popular. Subsequently, generalizations of quartimax and varimax were proposed and were shown to belong to a one parameter family of orthogonal rotation criteria known as the orthomax family (Harman, 1976, p. 299).

Still, effective computerized methods for orthogonal rotation did not constitute a final solution. Thurstone (1935, 1947) had emphasized oblique rotation and it was generally felt that correlated factors were a more plausible representation of reality. The development of an effective computerized method for oblique rotation was somewhat of a struggle. In the days of hand rotation it had been customary to first carry out a rotation to a simple reference structure (Thurstone, 1947; Mulaik, 1972, pp. 219-224 ; Yates, 1987, pp. 18-20). This rotation of the reference structure was feasible and a method for rotating to a simple factor pattern was not available. It did not seem to matter as the columns of a simple reference structure were rescaled to yield a primary factor pattern with a similar configuration. Difficulties were encountered, however, when computerized methods were used to rotate the reference structure. There was a strong tendency for factor collapse, that is for correlations between factors to approach one as the iterative procedure proceeded. This approach of optimizing a function of the reference structure was indirect. It did not involve the direct optimization of a function of the factor pattern loadings *that are inspected and used for interpretive purposes.*

Two suggestions were made on how to by-pass the factor collapse associated with the reference structure by using an orthogonal rotation to detect a simple pattern of loadings. Harris and Kaiser (1964) suggested the orthoblique method which involves an orthogonal rotation of a column-scaled principal axes factor matrix and a second column rescaling on the result to yield the oblique factor pattern. Hendrickson and White (1964) proposed the two stage promax method that first derives a target from an orthogonal rotation, and then obtains the factor pattern from an oblique target rotation. These two methods do not optimize prespecified functions of the factor pattern loadings to be interpreted and are also indirect, but in a different manner to the reference structure approach. Furthermore, each involves a power parameter that must be chosen anew in each application. This choice noticeably affects the quality of the solution.

The problems of oblique rotation were solved by Jennrich and Sampson (1966) who discovered a way of rotating directly to a simple factor pattern, thereby eliminating the intermediate rotation to a reference structure. This direct approach eliminates the problem of factor collapse if the rotation criterion to be minimized increases in value whenever elements of the factor pattern increase in magnitude (Jennrich & Sampson, 1966, p. 318; Yates, 1987, p. 55). Jennrich and Sampson employed the quartimin criterion (Carroll, 1953) and their direct quartimin procedure finally provided a usable computerized approach to oblique rotation. Direct quartimin, however, took longer to gain wide acceptance than varimax did, possibly because of the more or less simultaneous availability of effective methods for confirmatory factor analysis (Jöreskog, 1969)

The development of effective computerized methods for rotation had a number of consequences. First of all the time consuming aspect of factor rotation was eliminated. Rotating factor matrices became quick and easy. Secondly the opportunity for use of background knowledge concerning the variables during the rotation process was eliminated. Some regarded this as a desirable change of direction to greater objectivity, since the rotation process was no longer influenced by the investigator and depended only on the choice of rotation algorithm. Also, eliminating the need to learn complicated hand rotation procedures made rotation available to many who had little training in factor analysis and accepted the output of a rotation program without question. Varimax, in particular became universally used and alternative rotation methods had difficulty in being accepted by users and by journal reviewers. Perfect cluster solutions were handled effectively by varimax and direct quartimin, and were easy to interpret, so that little effort was made to seek the more complex patterns originally regarded by Thurstone as representative of reality.

Currently, orthogonal rotation with varimax still predominates in articles published in prestigious psychological journals (Fabrigar, Wegener, MacCallum & Strahan, 1999). The varimax-based promax method for oblique rotation (Hendrickson & White, 1964) is still included in some packages and is fairly frequently employed. Some informed users employ direct quartimin.

Confirmatory factor analysis procedures are often used for exploratory purposes. Frequently a confirmatory factor analysis, with prespecified loadings, is rejected and a sequence of modifications of the model is carried out in an attempt to improve fit. The procedure then becomes exploratory rather than confirmatory (see Nesselroade & Baltes 1984, pp. 272-273). In this situation the use of exploratory factor analysis, with rotation of the factor matrix, appears preferable. All statistical information produced by any confirmatory factor analysis program, including standard errors for rotated factor loadings, is currently also provided in a readily accessible exploratory factor analysis program (Browne, Cudeck, Tateneni, & Mels, 1998). The discovery of misspecified loadings, however, is more direct through rotation of the factor matrix than through the examination of model modification indices.

In this article, attention will be limited to single stage rotation methods that minimize a smooth function of factor pattern coefficients to attain simplicity. Both orthogonal and oblique rotation will be considered, although oblique rotation is probably more appropriate in most practical situations. Particular emphasis will be given to situations where a perfect cluster solution is inappropriate and more complex patterns are required. An overview of some rotation criteria will first be presented. Most of these are virtually unknown but are very interesting. Methods for row standardization of a factor matrix prior to rotation in order to improve the solution will also be examined. Some illustrative analyses will be carried out and general conclusions will be drawn as to what can and what cannot be accomplished in rotation.

Preliminaries

It will be convenient to provide a brief summary of salient features of rotation before proceeding to the main purpose of the article.

Analytic Rotation

Analytic rotation methods involve the postmultiplication of an input $p \times m$ factor matrix, $\mathbf{A}$, by a $m \times m$ matrix, $\mathbf{T}$, to yield a rotated primary factor pattern matrix,

$$\mathbf{\Lambda} = \mathbf{AT}$$

that minimizes a continuous function, $f(\mathbf{\Lambda})$, of its factor loadings. This function is intended to measure the complexity of the pattern of loadings in $\mathbf{\Lambda}$. By minimizing the complexity function, $f(\mathbf{\Lambda})$, the rotation procedure yields a rotated matrix $\mathbf{\Lambda}$ with a simple pattern of loadings.

Let the factor correlation matrix after rotation be represented by $\mathbf{\Phi}$. In orthogonal rotation the transformation matrix is required to satisfy the $m(m - 1)/2$ constraints

$$\mathbf{\Phi} = \mathbf{T}'\mathbf{T} = \mathbf{I} \tag{1}$$

defining factors that are uncorrelated and have unit variances. In direct oblique rotation (Jennrich & Sampson, 1966) a complexity function, $f(\mathbf{\Lambda})$, is also minimized but now a smaller number, m, of constraints

$$\text{Diag}(\mathbf{\Phi}) = \text{Diag}(\mathbf{T}^{-1}\mathbf{T}^{-1\prime}) = \mathbf{I} \tag{2}$$

is imposed, defining factors that are correlated but still have unit variances. This process defines a factor pattern, $\mathbf{\Lambda}$, that directly minimizes (Jennrich & Sampson, 1966) the complexity criterion.

Thus orthogonal and oblique rotation involve the same problem of minimizing a complexity criterion, and only the constraints imposed differ. It is appealing to make use of complexity functions that are suitable for both orthogonal and oblique rotation. Since fewer constraints are imposed in oblique rotation, it is generally possible to obtain a lower value of the complexity function and consequent greater simplicity of the factor pattern than in orthogonal rotation.

Simplicity of a Factor Pattern

Thurstone (1947) provided five rules concerning the ideal positioning of zero loadings to aid the identification of a simple structure at a time when it was routine to carry out oblique rotation on the reference structure. Since the factor pattern is obtained by scaling columns of the reference structure, and the positions of zeros do not change, the same rules may be applied to the factor pattern. In orthogonal rotation the reference structure and factor pattern coincide.

Thurstone's (1947) rules for simple structure of a factor matrix with *m* columns are:

1. Each row should contain at least one zero.
2. Each column should contain at least *m* zeros.
3. Every pair of columns should have several rows with a zero in one column but not the other.
4. If $m \geq 4$, every pair of columns should have several rows with zeros in both columns.
5. Every pair of columns of $\mathbf{\Lambda}$ should have few rows with nonzero loadings in both columns.

Yates (1997, p. 34) pointed out that Thurstone originally intended the first condition to define simple structure and the rest were intended to yield overdetermination and stability of the configuration of factor loadings. For example, m - 1 zero loadings per column are sufficient to identify the factor matrix in oblique rotation, provided that certain rank conditions are met (e.g. Algina, 1980), whereas Thurstone's second condition requires *m* zeros. The requirement of more zeros than the minimum number is intended to yield greater stability of the configuration.

The *complexity* of a variable in a factor pattern refers to the number of nonzero elements in the corresponding row of the factor matrix. A variable with complexity one will be referred to as a *perfect indicator*. If all variables are perfect indicators the factor matrix is said to have a perfect cluster configuration as in the following example

$$\mathbf{L} = \begin{bmatrix} \mathrm{x} & 0 & 0 \\ \mathrm{x} & 0 & 0 \\ \mathrm{x} & 0 & 0 \\ \mathrm{x} & 0 & 0 \\ 0 & \mathrm{x} & 0 \\ 0 & \mathrm{x} & 0 \\ 0 & 0 & \mathrm{x} \\ 0 & 0 & \mathrm{x} \\ 0 & 0 & \mathrm{x} \end{bmatrix}$$

where x refers to a nonzero quantity. As will be seen, this configuration is the one sought by most available complexity criteria. Thurstone's first rule, however, is less stringent and is satisfied by a more complex configuration where each variable has a complexity of at most m - 1, for example:

$$\mathbf{L} = \begin{bmatrix} \mathrm{x} & 0 & \mathrm{x} \\ 0 & \mathrm{x} & 0 \\ \mathrm{x} & \mathrm{x} & 0 \\ \mathrm{x} & 0 & 0 \\ 0 & \mathrm{x} & \mathrm{x} \\ \mathrm{x} & \mathrm{x} & 0 \\ 0 & 0 & \mathrm{x} \\ 0 & \mathrm{x} & \mathrm{x} \\ \mathrm{x} & 0 & \mathrm{x} \end{bmatrix}$$

Some Complexity Functions for Analytic Rotation

Thurstone (1947, pp. 140-142) constructed a data set to demonstrate the utility of factor analysis as an approximation in situations where relationships between factors are nonlinear, and also to illustrate his principles of simple structure. This data set was constructed from the height, length and weight of thirty hypothetical boxes. Twenty six nonlinear functions of these three characteristics were calculated for each box and then correlated. Thurstone carried out a factor analysis extracting three factors, followed by a subsequent hand rotation to demonstrate that it was possible to recover the simple structure of the problem despite the nonlinearity of the measurements. On the whole the variables are complex, and only three are perfect indicators.

Although a simple structure is known to exist, and can be recovered making use of prior knowledge, Thurstone's box data pose problems to blind rotation procedures (Butler, 1964; Eber, 1966; Cureton & Mulaik, 1971). Well known methods, such as varimax and direct quartimin, that are available in statistical software packages, fail with these data. This is due to the complexity of the variables rather than to their nonlinearity. Other artificial data can be constructed to yield similar problems (e.g. Rozeboom, 1992) without any involvement of nonlinearity.

In the present section we shall review some rotation criteria that have been eclipsed by well known methods and are not generally available in software packages. Some are virtually unknown. Their characteristics will be demonstrated in subsequent sections by applying them in some selected examples. Particular attention will be paid to their performance when applied to the box problem.

All rotation criteria to be considered are expressed as complexity functions to be minimized to yield a simple pattern of loadings. All of these complexity functions have a greatest lower bound (GLB) of zero. It will be instructive to consider the hypothetical situations in which this GLB is attained. Sometimes these situations are unrealistic and would not occur with real data. One example is where there is only one non zero element per row. Nevertheless, knowledge concerning the GLB conveys insight into the type of configuration that the rotation criterion seeks.

Crawford-Ferguson Family of Rotation Criteria

Let **s** be a (row or column) vector consisting of nonnegative elements, $s_j \geq 0$, $j = 1,2,$. This vector is regarded as being simple if it has few nonzero elements and complex if it has many. A measure of complexity of **s**, due to Carroll (1953), is

$$
\begin{aligned}
c(\mathbf{s}) &= \sum_{j} \sum_{\ell \neq j} s_j s_\ell \\
&= s_1 s_2 + s_1 s_3 + s_1 s_4 + \ldots + s_2 s_1 + s_2 s_3 + s_2 s_4 + \ldots \\
&\quad + s_3 s_1 + s_3 s_2 + s_3 s_4 + \ldots
\end{aligned}
\tag{3}
$$

It is immediately apparent from this definition that $c(\mathbf{s}) \geq 0$, and that its GLB of zero is attained if and only if **s** has at most one nonzero zero element. Also $c(\mathbf{s})$ increases if any zero element of **s** is replaced by a nonzero element (assuming $\mathbf{s} \neq \mathbf{0}$). In this manner $c(\mathbf{s})$ measures the complexity of **s**. Now consider the $p \times m$ matrix **S** of squared factor loadings, $s_{ij} = \lambda^2_{ij}$, $i = 1, ..., p$, $j = 1, ..., m$. Let $\mathbf{s}_i$ be the $1 \times m$ vector formed from the i^{th} row of **S** and $\mathbf{s}_{.j}$ be the $p \times 1$ vector formed from the j^{th} column of **S**.

Crawford and Ferguson (1970) suggested a family of complexity functions based on $c(\mathbf{s})$ in Equation 3. This family is indexed by a single parameter, κ ($0 \le \kappa \le 1$), and its members are of the form:

$$
\begin{aligned}
f(\mathbf{L}) &= (1-\kappa)\sum_{i=1}^{p} c(\mathbf{s}_{i.}) &&+ \quad \kappa \sum_{j=1}^{m} c(\mathbf{s}_{.j}) \\
&= (1-\kappa)\sum_{i=1}^{p}\sum_{j=1}^{m}\sum_{\ell \ne j}^{m} \lambda^2_{ij}\lambda^2_{i\ell} &&+ \quad \kappa \sum_{j=1}^{m}\sum_{i=1}^{p}\sum_{k \ne i}^{p} \lambda^2_{ij}\lambda^2_{kj} \\
&\quad \text{Row (variable) complexity} && \quad \text{Column (factor) complexity}
\end{aligned}
\tag{4}
$$

Thus the Crawford-Ferguson criterion is a weighted sum of a measure of complexity of the p rows of $\mathbf{\Lambda}$ and a measure of complexity of the m columns.

The first term, or measure of row (variable) complexity, is Caroll's (1953) quartimin criterion which is motivated by Thurstone's rules three, four and five. Although it attains its GLB for a perfect cluster configuration it also attains the GLB when all nonzero loadings are in the first column and zeros occur elsewhere. Consequently the first term is insensitive to a general factor. The second term measures column (factor) complexity and will attain the GLB if each column has a single nonzero element. It reflects the intent of Thurstone's second rule by penalizing too many nonzero elements in a column.

In orthogonal rotation the Crawford-Ferguson (CF) family is equivalent to the orthomax family (Crawford & Ferguson, 1970, pp. 324-326). Values of κ that yield particular members of the orthomax family are shown in Table 1. Quartimax, varimax and equamax were previously known members of the orthomax family. Parsimax and factor parsimony were suggested by Crawford and Ferguson (1970). Parsimax results in equal contributions from

Table 1
<u>The Orthomax Family of Rotation Criteria</u>

$\kappa = 0$	$\kappa = \frac{1}{p}$	$\kappa = \frac{m}{2p}$	$\kappa = \frac{m-1}{p+m-2}$	$\kappa = 1$
Quartimax (Quartimin)	Varimax	Equamax	Parsimax	Factor Parsimony

variable complexity and factor complexity when all elements of $\boldsymbol{\Lambda}$ are equal. Factor parsimony consists of the factor complexity term in Equation 4 alone and is primarily of theoretical interest. The complexity function in Equation 4 resulting from a particular choice of κ will be distinguished from the corresponding member of the orthomax family by its name prefixed by CF. Thus the varimax simplicity function is the original function maximized by Kaiser (1958) while CF-varimax will refer to the complexity function that is given by Equation 4 with

$$\kappa = \frac{1}{p},$$

and is to be minimized in the rotation process. In orthogonal rotation varimax and CF-varimax yield the same solution. Their equivalence is a result of the invariance of within-row sums of squared factor loadings in orthogonal rotation (Crawford & Ferguson, 1970, pp. 324-325). This invariance is no longer true in oblique rotation, so that the equivalence no longer holds. Oblique varimax can result in factor collapse whereas oblique CF-varimax cannot.

A positive characteristic of the CF family is that none of its members can result in factor collapse under direct oblique rotation of the factor pattern (Crawford, 1975).

Yates' Geomin

This criterion employs an adaptation of a measure of row complexity first suggested by Thurstone [1935, p. 163; 1947, p. 355, (32)]. Again, let the $m \times 1$ vector $\mathbf{s}$ consist of nonnegative elements, s_j, $j = 1,...,m$. Thurstone's measure of complexity of the vector $\mathbf{s}$ is

$$c(\mathbf{s}) = s_1 s_2 s_3 \ldots s_m = \prod_{j=1}^{m} s_j, \tag{5}$$

which will be zero if at least one element of $\mathbf{s}$ is zero. Thurstone's measure of complexity of $\mathbf{s}$ (Equation 5) differs from Carroll's measure (Equation 3) in that Equation 5 is zero if only one element of $\mathbf{s}$ is zero whereas Equation 3 requires m - 1 zeroes. Also, unlike Equation 3, Equation 5 does not increase if there are two or more zeroes and one of these becomes nonzero. Consequently, in unmodified form, Equation 5 results in indeterminacy since it remains at its minimum if one element of $\mathbf{s}$ is zero, no matter how the other elements change.

Thurstone [1935, p. 163; 1947, p. 355, (32)] proposed the first complexity function for rotation to simple structure. It is applied to the elements, ℓ_{ij}, of the reference structure, **L**, and consists of a sum of *p* row complexity measures Equation 5 defining

$$\mathbf{s}_i = \left(\ell_{i1}^2\,\ell_{i2}^2 \ldots \ell_{im}^2\right)$$

This complexity function is

$$f(\mathbf{L}) = \sum_{i=1}^{p} c(\mathbf{s}_i) = \sum_{i=1}^{p} \prod_{j=1}^{m} \ell_{ij}^2 \quad (6)$$

It clearly attains its GLB of zero if Thurstone's first rule for simple structure is satisfied; that is each row of **L** contains at least one zero. Thurstone (1935, pp. 185-197) gave details of an algorithm for minimizing *f*(**L**) in Equation 6, but it was not successful (Thurstone, 1935, p. 197) so that his complexity function had little impact at the time. It was modified by Yates [1987, p. 46, (40a)] by *replacing* the sum of within row products of squared *reference structure* elements by a sum of within row *geometric means* of squared *factor pattern* coefficients

$$f(\boldsymbol{\Lambda}) = \sum_{i=1}^{p} c(\mathbf{s}_i)^{\frac{1}{m}} = \sum_{i=1}^{p} \left(\prod_{j=1}^{m} \lambda_{ij}^2 \right)^{\frac{1}{m}} \quad (7)$$

where $\mathbf{s}_i$ is now redefined as a vector of squared factor pattern coefficients

$$\mathbf{s}_i = \left(\lambda_{i1}^2\,\lambda_{i2}^2 \ldots \lambda_{im}^2\right)$$

The complexity function in Equation 7, named "geomin" by Yates, also attains its GLB of zero if Thurstone's first rule for simple structure is satisfied; that is each row of $\boldsymbol{\Lambda}$ contains at least one zero. It has the indeterminacy of Equation 5 in that, if any element in a row of $\boldsymbol{\Lambda}$ is zero, the values of the other elements in the same row have no influence on the value of the function. Yates (1987, pp. 67-74) devised a "soft-squeeze" algorithm in an attempt to bypass this problem. Here no attempt is made to provide an algorithm that chooses one of several local minima. Rather, Yates' complexity function is altered slightly to reduce the indeterminacy of its minimizer. The complexity function in Equation 7 is modified by adding a small positive quantity, ϵ to λ_{ij}^2 to obtain

(8) $$f(\Lambda) = \sum_{i=1}^{p} \left[\prod_{j=1}^{m} \left(\lambda_{ij}^2 + \epsilon \right) \right]^{\frac{1}{m}}$$

A zero loading no longer results in difficulties and $f(\Lambda)$ is not affected greatly if ϵ is small. A value of ϵ = .01 seems satisfactory for three or four factors. It may need to be increased slightly for more factors. Although the complexity function in Equation 8 is a minor modification of Yates' geomin, it will be referred to subsequently as geomin for simplicity.

Yates intended geomin for oblique rotation only. It can also be used for orthogonal rotation but may not be optimal for this purpose.

Another more complex rotation scheme that involves iteratively respecified weights was also suggested by Yates (1987, Chapter 6).

McCammon's Minimum Entropy Criterion

McCammon (1966) suggested a rotation criterion based on the entropy function of information theory. For completeness, the entropy function will first be defined.

Consider n nonnegative quantities, $x_i \geq 0$, $i = 1,\ldots,n$ that sum to one,

$$\sum_{i=1}^{n} x_i = 1,,$$

and let $\mathbf{x} = (x_1, x_2,\ldots,x_n)'$. The entropy function is defined by

(9) $$\text{Ent}(\mathbf{x}) = -\sum_{i=1}^{n} e(x_i)$$

where

(10) $$\begin{aligned} e(x_i) &= x_i \ln(x_i) \quad \text{if } x_i > 0 \\ &= 0 \qquad\qquad\quad \text{if } x_i = 0 \end{aligned}$$

This entropy function has a GLB of 0 which is attained if a single x_i is equal to 1 and the rest are equal to 0. It has a least upper bound (LUB) of ln *n* which is attained if all x_i are equal, $x_i = 1/n$, $i = 1,...,n$. If the elements of **x** in Equation 9 are obtained from elements of **s** by rescaling them to have a sum of one,

$$x_i = \frac{s_i}{\sum_{k=1}^{n} s_k}$$

then $c(\mathbf{s}) = \text{Ent}(\mathbf{x})$ is a measure of the complexity of **s** that, like the Carroll measure (Equation 3), implies that there is a single nonzero element of **s** at its GLB of zero. However Ent (**x**), unlike the Carroll measure, is simultaneously a measure of equality of elements of **s**. If Ent(**x**) attains it LUB of ln *n* then all elements of **s** are equal.

Let **S** be the $p \times m$ matrix of squared factor loadings with typical element $s_{ij} = \lambda_{ij}^2$. Consider the following sums of squared factor loadings:

$$S_{i.} = \sum_{j=1}^{m} s_{ij} \quad S_{.j} = \sum_{i=1}^{p} s_{ij} \quad S = \sum_{j=1}^{m} S_{.j} = \sum_{i=1}^{p} \sum_{j=1}^{m} s_{ij} \tag{11}$$

McCammon's minimum entropy complexity function is defined by

$$f(\boldsymbol{\Lambda}) = \frac{-\sum_{j=1}^{m} \sum_{i=1}^{p} e\left(\frac{s_{ij}}{S_{.j}}\right)}{-\sum_{j=1}^{m} e\left(\frac{S_{.j}}{S}\right)} = \frac{\sum_{j=1}^{m} \sum_{i=1}^{p} e\left(\frac{s_{ij}}{S_{.j}}\right)}{\sum_{j=1}^{m} e\left(\frac{S_{.j}}{S}\right)} \tag{12}$$

The numerator consists of the sum of *m* within column entropy functions. It attains its GLB of 0 if each column has a single nonzero element. This suggests that the numerator encourages configurations with columns that contain few large and many small elements. The denominator is the entropy function based on column sums of **S**. It attains its LUB, thereby reducing the criterion optimally, if all column sums, $S_{.j}$, are equal. Consequently the minimum entropy criterion seeks solutions with simple columns and with column sums of squares that do not differ widely.

McCammon intended his minimum entropy criterion for use in orthogonal rotation only. As will be seen subsequently, it is unsatisfactory in oblique rotation.

McKeon's Infomax

McKeon (1968), in an unpublished manuscript, treated a matrix of squared factor loadings as analogous to a two way contingency table and derived a number of simplicity functions based on tests for association. The one he found most effective was based on the likelihood ratio test for association (Agresti, 1990, p. 48, Equation 3.13), which is maximized for maximum simplicity. McKeon also pointed out that, if the squared factor loadings are interpreted as frequencies, his criterion may be regarded as a measure of information about row categories conveyed by column categories and, simultaneously, as a measure of information about column categories conveyed by row categories. He consequently named it infomax.

Here McKeon's infomax criterion is subtracted from its LUB to yield a complexity function. This infomax complexity function is given by

$$f(\mathbf{\Lambda}) = \ln m - \sum_{j=1}^{m} \sum_{i=1}^{p} e\left(\frac{s_{ij}}{S_{.j}}\right) + \sum_{i=1}^{p} e\left(\frac{S_{i.}}{S}\right) + \sum_{j=1}^{m} e\left(\frac{S_{.j}}{S}\right)$$

using notation defined in Equations 10 and 11. This complexity function attains its GLB of zero when the factor matrix has a perfect cluster configuration and the m within column sums of squared loadings are equal ($S_{.1} = S_{.2} = \ldots = S_{.m}$). Thus infomax favors a perfect cluster configuration and simultaneously discourages a general factor.

McKeon's infomax criterion gives good results in both orthogonal and oblique rotation.

Rotation to a Partially Specified Target

The first methods of this type were suggested by Tucker (1940, 1944) and Horst (1941) for use specifically in locating *reference axes* in exploratory hand rotation. Ideally eigenvectors would have been required but, at that time, their exact evaluation was not feasible, and approximations were necessary. When computers, and effective algorithms for computing eigenvectors, were available, Lawley and Maxwell (1964) and Jöreskog (1965) provided algorithms for rotating *reference structures* to partially specified targets in confirmatory rotation. After the Jennrich-Sampson (1966) breakthrough, Gruvaeus (1970) and Browne (1972 a, b) provided algorithms for directly rotating the *factor pattern* to a partially specified target.

Use of this approach to rotation requires the specification of target values for selected factor pattern coefficients. A $p \times m$ target matrix, **B**, with some specified elements and some unspecified elements is required, for example:

$$\mathbf{B} = \begin{bmatrix} ? & 0 & 0 \\ ? & ? & 0 \\ ? & 0 & ? \\ 0 & ? & 0 \\ ? & ? & 0 \\ 0 & ? & ? \\ 0 & 0 & ? \\ 0 & ? & ? \\ ? & 0 & ? \end{bmatrix}$$

This target matrix reflects partial knowledge as to what the factor pattern should be. In the example all specified elements are zero, as is usually the case in practical applications, and unspecified elements are indicated by the ? symbol. Nonzero values can be employed for specified elements, b_{ij}, but it is usually easier to specify zeros for small elements than it is to specify precise values for larger elements. No information is provided by the unspecified b_{ij} = ? and, after rotation, the corresponding rotated loadings, λ_{ij}, may turn out to be large, moderate, or small.

Represent the set of subscripts for specified target loadings, b_{ij}, in column j by I_j. A suitable complexity function for minimization that yields λ_{ij} values that are close to the specified b_{ij} is

(13) $$f(\mathbf{L}) = \sum_{j=1}^{m} \sum_{i \in I_j} (\lambda_{ij} - b_{ij})^2$$

or sum of squared differences between loadings after rotation and specified target values. It is a suitable complexity function for both orthogonal (Browne, 1972a) and oblique (Browne, 1972b) rotation.

Rotation to a partially specified target has similarities to confirmatory factor analysis (Jöreskog, 1969) as values for some factor loadings must be specified in advance. There is a salient difference, however. In confirmatory factor analysis, specified factor loadings are forced to assume the specified values of zero. Misspecified elements may only be detected indirectly through examination of the overall measure of fit supplemented by modification indices. In target rotation, corresponding elements of the rotated factor pattern matrix are only made as close to the specified zeros as possible. Differences can be large so that misspecified zeros are easily detected.

The target may then be changed. Previously misspecified elements of **B** are now left unspecified. Furthermore, any previously unspecified b_{ij} may now be specified to be zero if the corresponding λ_{ij} is near zero. The altered **B** may now be employed in a new target rotation. This procedure may be repeated until the investigator is satisfied with the outcome. When a sequence of targets is employed, the process ceases to be confirmatory and becomes a non-mechanical exploratory process, guided by human judgment. This sequential procedure is a modernization of Tucker's (1944) "semi-analytical method of factorial rotation."

Standardization of Factor Loadings

The simplicity of the pattern of a rotated solution can sometimes be improved by carrying out an initial standardization on rows of the initial factor matrix. Thus, if **A** is the initial factor matrix, the initial standardization is of the form,

$$\mathbf{A}^* = \mathbf{D}_v \mathbf{A} \tag{14}$$

where $\mathbf{D}_v$ is a positive definite diagonal matrix. The complexity function, $f(\mathbf{A}^*\mathbf{T})$, is minimized with respect to **T**, subject to the constraints of Equations 1 or 2, and $\boldsymbol{\Lambda}^* = \mathbf{A}^*\mathbf{T}$ is restandardized,

$$\boldsymbol{\Lambda} = \mathbf{D}_v^{-1} \boldsymbol{\Lambda}^*$$

to yield the final simple pattern matrix, $\boldsymbol{\Lambda}$.

The two standardization procedures to be considered here were both originally derived with varimax in mind, but appear to be applicable to other members of the Crawford-Ferguson family and to some other rotation criteria. They do not seem to be appropriate for rotation to a partially specified target.

Kaiser Standardization

Kaiser (1958) noted that rows of **A** that yield low communalities have little effect on the final varimax solution. In order to ensure that all variables have the same influence on the rotated solution he recommended that the standardization should yield an $\mathbf{A}^*$ with equal row sums of squares. His main motivation for this was to improve generalizability of the rotated solution across batteries, where communalities for variables change.

In the Kaiser standardization, the weights are chosen to be inverse square roots of communalities:

$$\mathbf{D}_v = \mathbf{D}_h^{-1/2}$$

where

$$\mathbf{D}_h = \mathrm{Diag}(\mathbf{A}\mathbf{A}') \qquad (15)$$

This standardization is frequently employed, not only with varimax, but also with other rotation criteria, both in orthogonal and oblique rotation.

Cureton-Mulaik Standardization

Cureton and Mulaik (1975) demonstrated that orthogonal varimax rotation yields an unsatisfactory solution when applied to the Thurstone box data. Since these data were artificially constructed, the optimal factor pattern is known. Twenty three of the twenty six variables are complex in that they have non-negligible loadings on at least two of the three factors. The remaining three variables are pure indicators with single non-negligible loadings. One serves as an indicator for each of the three factors.

Varimax is effective when a perfect orthogonal cluster solution exists but gives poor results with complex factor patterns. The aim of the Cureton-Mulaik (CM) standardization is to provide a weighting system that downweights complex variables and emphasizes pure indicators with single non-negligible loadings. This will improve the varimax solution. Detection of pure indicators and of complex variables must, however, be accomplished without the use of a simple pattern, since this is unknown prior to rotation. Two assumptions allow the forecasting of pure indicators before knowing the optimal simple pattern. The first is the assumption of a positive manifold (Thurstone, 1947, pp. 341-343; Yates, 1987, pp. 87-89). This assumption was originally formulated in the language of multidimensional geometry, but is *equivalent* to assuming that it is possible to find an orthogonal rotation of the factor matrix where all non-negligible factor loadings for each variable (i.e. in each row of $\mathbf{\Lambda}$) have the same sign. The second assumption is that the set of test points is scattered on the positive manifold implying that a substantial proportion of the variables will load on more than on factor.

These two assumptions will now be illustrated graphically considering the special case, where m = 2, that is illustrated in Figure 1. Suppose that $\mathbf{A}^* = D_h^{-1/2}\mathbf{A}$ so that each row of $\mathbf{A}^*$ has length one. When the points corresponding to rows of $\mathbf{A}^*$ are plotted they will fall on the circumference of a unit circle. Points corresponding to rows of **A** would lie within the unit circle. Consequently the points plotted in the figure are referred to as *test points extended to the unit circle*. With a positive manifold *it is possible to find* orthogonal axes, I and II in Figure 1, such that all extended test points, marked by crosses, lie either on the first quadrant (between points marked by dots I and II) or on the third quadrant (between points marked by dots I- and II-). Points lying on the first quadrant represent tests with two non-negative factor loadings. Those on the third quadrant represent tests with two non-positive factor loadings. All points on the third quadrant may be reflected to the first quadrant by multiplying the appropriate row of $\mathbf{A}^*$ by -1. Consequently there is no loss of generality in assuming that all test points lie on the first quadrant as is the case in Figure 1.

These extended test points are scattered fairly evenly on the first quadrant thereby satisfying the second assumption. Points coinciding with dot I represent tests with a single nonzero loading on the first factor and a zero loading on the second factor. Those coinciding with dot II have a single

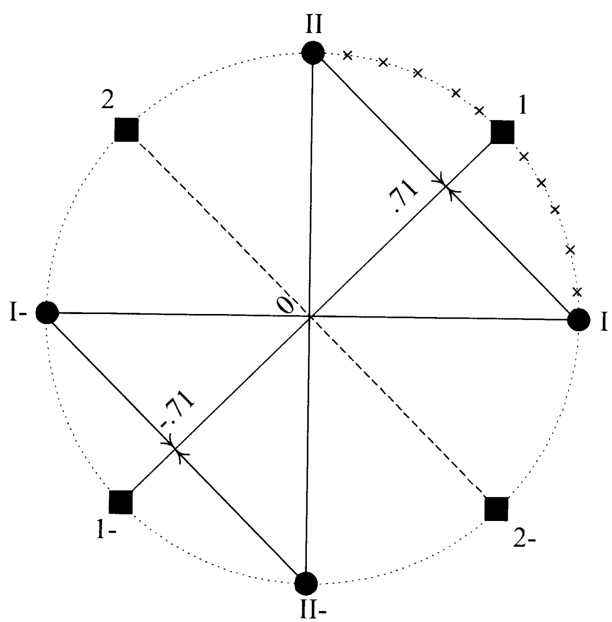

Figure 1
Cureton-Mulaik Standardization: Scatter of Points on Quadrant. $m = 2$ $m^{-1/2} = .71$

nonzero loading on the second factor and a zero loading on the first factor. Points in the center of the quadrant (coinciding with square 1) represent tests with equal loadings on both factors. Most tests have complexity two (two nonzero loadings, not necessarily of the same size) as most test points coincide neither with dot I nor dot II.

The location of the optimal orthogonal axes I and II is not known in advance. Cureton and Mulaik (1975) devised an ingenious method for forecasting the complexity of each test from its loading on the first principal axis alone. Let **a** represent the $p \times 1$ vector of loadings of tests on the first principal axis or, equivalently the first principal component of $\mathbf{AA}'$. Also let $\mathbf{a}^* = D_h^{-1/2}\mathbf{a}$ be the vector of principal component loadings divided by square roots of communalities. Elements, a^*_i, of $\mathbf{a}^*$ will lie between -1 and 1. If the absolute value $|a^*_i| = 1$ then the i-th test will have a nonzero loading only on the first principal axis. Tests with $a^*_i = 1$ will coincide with the square labeled 1 in Figure 1; those with $a^*_i = -1$ will coincide with the square labeled 1-. Other elements of $\mathbf{a}^*$ will be orthogonal projections of test points onto the axis joining the two squares labeled 1 and 1-. If this assumption of evenly scattered test points is true the principal axis, $\mathbf{a}^*$, joining 1 and 1- will pass through the center of the scattering of extended test points on the positive manifold (see Figure 1). Good indicators with a single substantial loading will be close to dots I, II, I- or II- and will yield $|a^*_i| \approx 1/\sqrt{m} = 1\sqrt{2} \approx .71$. Complex tests with substantial loadings of the same sign will be close to the squares 1 and 1- in Figure 1 and will yield $|a^*_i| \approx 1$. Complex tests with substantial loadings of different signs will be close to the squares 2 and 2- in Figure 1 and will yield $|a^*_i| \approx 0$.

This information may be used to provide a weighting system that assigns weights between zero and one to tests, or rows of $\mathbf{A}^*$, based only on the absolute values $0 \leq |a^*_i| \leq 1$ of elements of $\mathbf{a}^*$. Weights of 1 to are assigned to good indicators with elements of $\mathbf{a}^*$ near to.71 in absolute value. Weights of zero are assigned to complex tests with elements of $\mathbf{a}^*$ that are near 0 or near 1 in absolute value.

The reasoning presented here generalizes to $m > 2$ factors. This only requires the replacement of $1/\sqrt{2}$ by $1/\sqrt{m}$ as the value of a^*_i where a maximum weight should be assigned. Thus the weight assigned to the i-th test should be $w_i = 1$ if $|a_i|^* = 1/\sqrt{m}$ and $w_i = 0$ if either $|a_i|^* = 1$ or if $|a_i|^* = 0$.

Cureton and Mulaik (1975) proposed a weight assignment scheme for this purpose giving two functions, one to be used if $|a_i|^* \geq 1/\sqrt{m}$ and the other if $|a_i|^* < 1/\sqrt{m}$. This weight assignment scheme is expressed in an algebraically equivalent form here, involving a single function. The CM weight for test i is given by

$$w_i = \cos^2\left[\frac{\operatorname{acos}\left(m^{-1/2}\right) - \operatorname{acos}\left(|a_i^*|\right)}{\operatorname{acos}\left(m^{-1/2}\right) - \delta\left(a_i^*, m\right)} \cdot \frac{\pi}{2}\right] \tag{16}$$

where

$$\delta(a_i^*, m) = \begin{cases} \frac{\pi}{2} & \text{if } |a_i^*| < m^{-1/2} \\ 0 & \text{otherwise} \end{cases}$$

Since these weights apply to rows of $\mathbf{A}^*$ the combined weight matrix $\mathbf{D}_v$ in Equation 14 with weights for rows of $\mathbf{A}$ is given by

$$\mathbf{D}_v = \mathbf{D}_w \mathbf{D}_h^{-1/2}$$

where diagonal elements of $\mathbf{D}_w$ and $\mathbf{D}_h$ are given by Equations 16 and 15 respectively.

The shape of the function, $w(a^*_i)$ defining the w_i in Equation 16 when $m = 2$ is shown in Figure 2. It yields a zero weight at $a^* = 0$, increases steadily to yield a unit weight at $a^* = .71$, and decreases again to yield a zero weight at $a^* = 1$. It is of interest that, although the function is continuous, there is a discontinuity in its first derivative at $a^* = m^{-1/2}$. Alternative functions of a^* were suggested by Yates (1987, Chapter 5) for providing standardization weights.

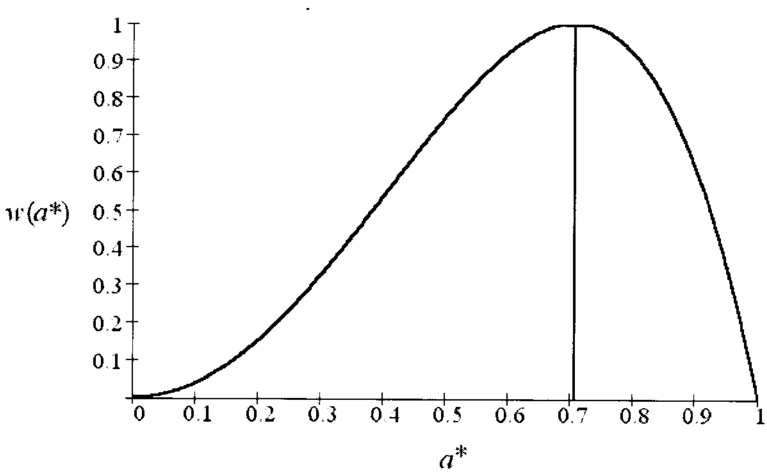

Figure 2
Cureton-Mulaik Weight Function. $m = 2$ $m^{-1/2} = .71$

Although it was originally intended for orthogonal varimax, it will be seen subsequently that the CM standardization can be helpful with other rotation criteria, both in orthogonal and in oblique rotation.

It is worth bearing in mind that both the CM and the Kaiser standardizations may have undesirable consequences in small samples since they increase the influence of tests that have small communalities and consequently yield unstable factor loadings (cf. MacCallum, Widaman, Zhang, & Hong, 1999).

Computational Considerations

All computations reported subsequently were carried out with the Comprehensive Exploratory Factor Analysis (CEFA) program (Browne, Cudeck, Tateneni, & Mels, 1998). Some facilities of this program that are relevant to the present work will be described briefly.

A General Method for Rotation

In CEFA an arbitrary complexity function may be tried out in both orthogonal and oblique rotation with minimal algebraic development and minimal additional programming effort, involved only in computation of the complexity function. Kaiser's (1959) algorithm for orthogonal rotation and Jennrich and Sampson's (1966) algorithm for direct oblique rotation of the factor pattern were adapted for use with arbitrary complexity functions. This was accomplished by replacing the closed form solution of a specific nonlinear equation for obtaining the angle of rotation by Brent's derivative free (Brent, 1973, Chapter 5; Press, Teukolsky, Vetterling & Flannery, 1986, Section 10.2) unidimensional search algorithm. While the methods employed may not be as rapid as special purpose algorithms, they were found to be accurate and sufficiently efficient for routine use on the fast personal computers currently available.

Detection of Multiple Local Minima of the Complexity Function

In some situations complexity functions will yield multiple local minima and consequent multiple alternative rotations. Some complexity functions are more prone to local minima than others, but all dealt with here can yield multiple local minima at least occasionally. This instability will not be detected by the user if a single rotation is carried out. Special steps to seek out local minima must therefore be taken.

Gebhart (1968) investigated the presence of local maxima in orthogonal rotation and suggested the use of multiple starting points, generated by means of random orthogonal rotations of the initial factor matrix. This enabled him to show that, in certain circumstances, even the generally stable varimax criterion is susceptible to multiple solutions. Multiple starting points have also been used by others. Random orthogonal rotations were employed to give starting points by Kiers (1994) for alternative oblique simplimax solutions. Rozeboom (1992) has strongly advocated the use of random initial oblique rotations.

CEFA can carry out a random orthogonal rotation to yield a random starting point, followed by a particular orthogonal or oblique rotation. This process may be repeated a specified number of times and is convenient for detecting multiple solutions. It has been applied repeatedly in this article.

Numerical Examples

The best known rotation methods, available in most commercial software packages, are varimax (Kaiser, 1958) in orthogonal rotation and direct quartimin (Jennrich & Sampson, 1966) or promax (Hendrickson & White, 1966) in oblique rotation. A number of promising but virtually unknown rotation criteria have been surveyed in earlier sections of this article. In order to try these out initially, two well-known data sets were used. The first is a well known correlation matrix presented by Harman (1976, Table 7.4) for twenty four psychological tests based on 145 cases. It was obtained from part of a data set collected by Holzinger and Swineford (1939). The derived factor matrix has been used repeatedly by Harman and many other authors to demonstrate the effectiveness of various rotation procedures. The second is Thurstone's (1947) artificially constructed box data set described earlier. A substantial number of different types of rotation were applied to these matrices. Some rotated factor matrices will be reported but some results will be described without presentation of factor matrices because of space considerations. The computer program that was used is available on the world wide web (Browne, Cudeck, Tateneni, & Mels, 1998) so that readers may replicate results described if they wish to do so.

Twenty Four Psychological Tests

Standard rotation methods give good results when applied to the twenty four psychological tests. They yield rotations in which about two thirds of the variables are pure indicators. Perfect clusters predominate.

The twenty four psychological tests were employed here to verify that the rotation criteria considered are satisfactory in typical situations. Four factors were extracted by maximum likelihood. The following exploratory rotation criteria were examined: CF-quartimax (equivalent to direct quartimin in oblique rotation), CF-varimax, geomin, minimum entropy and infomax. These were applied with no standardization in both orthogonal and oblique rotation and using 20 random starting points. Although oblique rotation is frequently to be recommended, there can be situations where orthogonal rotation is required so both are considered here. CM standardization was also applied in conjunction with all orthogonal and oblique rotation methods.

In orthogonal rotation it was found that each complexity function resulted in convergence to a unique minimum with the exception of geomin. This yielded two local minima which differed somewhat from the other solutions in that there was a greater incidence of variables with substantial loadings on more than one factor. It is worth noting that Yates (1987) did not intend geomin for orthogonal rotation. The infomax orthogonal rotation is reported in Table 2. It is typical of the rotations obtained and does not differ greatly from the three orthogonal rotations reported in Harman (1976, Table 13.7).

CM standardization had only a slight effect in most cases. This standardization did, however, reduce orthogonal geomin's tendency, mentioned earlier, to yield more than one substantial loading per row of the rotated factor matrix.

In oblique rotation minimum entropy alone proved unsatisfactory. The algorithm failed to converge within 500 iterations from all random starting points. When iteration was terminated, factor collapse was evident in that two pairs of factors had correlations of one. Since the minimum entropy complexity function (Equation 12) need not tend to infinity as any interfactor correlation coefficient tends to one, factor collapse is possible (cf. Jennrich & Sampson, 1966, p. 318; Yates, 1987, p. 55). McCammon did not intend the minimum entropy complexity function to be used for oblique rotation.

The other rotation criteria, including geomin, all resulted in convergence to a single minimum from all starting points. As an example the infomax oblique rotation is also shown in Table 2. Other rotation criteria resulted in similar configurations and all are reasonably similar to the direct quartimin solution reported by Harman (1976, Table 14.10).

The main point demonstrated by these trials is that most complexity functions are reasonably satisfactory in situations where a high proportion of variables are perfect indicators.

Thurstone's 26 Variable Box Data

As was pointed out by Yates (1987), the current tendency to select pure indicators for factor analysis may be influenced by the fact that the generally

Table 2
Holzinger and Swineford, 24 Psychological Tests: Infomax Rotations

	Orthogonal				Oblique			
	Fac1	Fac2	Fac3	Fac4	Fac1	Fac2	Fac3	Fac4
Var 1	0.71	0.18	0.10	0.12	0.74	-0.03	0.05	0.02
Var 2	0.44	0.13	0.03	0.07	0.47	0.01	-0.01	0.01
Var 3	0.56	0.15	-0.09	0.08	0.63	0.01	-0.15	0.02
Var 4	0.53	0.25	0.04	0.05	0.55	0.13	-0.01	-0.06
Var 5	0.20	0.75	0.20	0.11	0.00	0.77	0.12	-0.04
Var 6	0.20	0.78	0.05	0.20	0.01	0.80	-0.07	0.09
Var 7	0.20	0.81	0.13	0.03	0.02	0.88	0.06	-0.14
Var 8	0.36	0.58	0.20	0.09	0.22	0.53	0.14	-0.06
Var 9	0.19	0.82	0.03	0.19	0.00	0.86	-0.10	0.08
Var10	-0.01	0.17	0.84	0.14	-0.29	0.04	0.93	0.03
Var11	0.19	0.19	0.50	0.35	-0.03	0.00	0.48	0.31
Var12	0.30	0.02	0.69	0.06	0.16	-0.17	0.77	-0.07
Var13	0.50	0.20	0.47	0.04	0.42	0.01	0.50	-0.11
Var14	0.08	0.22	0.09	0.54	-0.12	0.07	-0.04	0.63
Var15	0.14	0.14	0.07	0.51	-0.01	-0.02	-0.05	0.59
Var16	0.43	0.10	0.03	0.51	0.35	-0.15	-0.11	0.56
Var17	0.11	0.16	0.23	0.56	-0.11	-0.03	0.12	0.63
Var18	0.35	0.05	0.31	0.43	0.21	-0.21	0.24	0.45
Var19	0.27	0.17	0.14	0.35	0.16	0.00	0.06	0.35
Var20	0.41	0.40	0.08	0.27	0.32	0.27	-0.02	0.21
Var21	0.44	0.19	0.39	0.19	0.33	-0.01	0.38	0.09
Var22	0.41	0.39	0.08	0.27	0.31	0.25	-0.02	0.21
Var23	0.53	0.39	0.19	0.20	0.44	0.23	0.12	0.09
Var24	0.22	0.38	0.48	0.27	-0.01	0.24	0.45	0.18

Factor Correlations

	Fac1	Fac2	Fac3	Fac4	Fac1	Fac2	Fac3	Fac4
Fac1	1.00				1.00			
Fac2	0.00	1.00			0.51	1.00		
Fac3	0.00	0.00	1.00		0.41	0.41	1.00	
Fac4	0.00	0.00	0.00	1.00	0.50	0.52	0.47	1.00

available rotation procedures give reasonable results only in this situation. This deviates from Thurstone's original concept of simple structure where variables of complexity at most *m* - 1 were regarded as admissible. His box data (Thurstone, 1947, p. 369) illustrate his principles of simple structure and twenty one of his constructed variables are of complexity two. These data will be employed to investigate the extent to which the rotation criteria being considered here can provide a complex solution when it exists.

Because of the manner in which they were constructed, Thurstone's box data involve virtually no error of measurement. As a result of this and the effect of rounding to two places, the correlation matrix given by Thurstone (1947, p. 370) is indefinite. Cureton and Mulaik (1975, Table 4) reported the first three principal components obtained from Thurstone's correlation matrix. They employed this matrix to demonstrate that their standardization procedure enables varimax to reproduce Thurstone's simple structure. The same matrix was used here to investigate the capabilities of the complexity functions used earlier with the twenty four psycholgical tests. CF-quartimax (equivalent to direct quartimin in oblique rotation), CF-varimax, geomin and infomax were applied to the box data in both orthogonal and oblique rotation. Minimum entropy was employed only in orthogonal rotation because of its tendency for factor collapse in oblique rotation.

At an early stage some authors (Butler, 1964; Eber, 1966; Cureton & Mulaik, 1971) found that the box data yield more than one simple structure. Consequently 100 random starts were used for each trial rotation in order to yield a reasonable chance of discovering multiple solutions, or local minima. When no standardization was applied, it was found that minimum entropy in orthogonal rotation, and geomin and infomax in both orthogonal and oblique rotation, were capable of reproducing Thurstone's simple structure reasonably well at one of their local minima. The best orthogonal minimum entropy solution and best oblique geomin solution are shown in Table 3. Row headings show the manner in which the variable was constructed from the measurement of height, *h*, length, *l*, and width, *w*. Column headings name the height factor, *H*, length factor, *L*, and width factor, *W*. The oblique geomin solution shown in Table 3 and the best oblique infomax solution were very close to the factor pattern obtained by Thurstone (1947, p. 371) using hand rotation.

The three columns under Standardization-None in Table 4 give details of the occurrence of local minima when analyzing the box data. The number of different local minima that occurred after 100 random starts is shown in the column headed by #LM. The column headed by %√ shows the percentage occurrence of the particular local minimum that was judged to yield a satisfactory solution, in that the pattern of loadings reflected the

Table 3
Thurstone, Box Problem: Orthogonal and Oblique Rotations

	Orthogonal Minimum Entropy			Oblique Geomin (ϵ = .01)		
	H	*L*	*W*	*H*	*L*	*W*
h	0.98	0.04	0.12	0.99	-0.02	-0.01
l	0.19	0.93	0.21	0.06	0.94	0.05
w	0.14	0.15	0.96	0.00	0.06	0.97
hl	0.71	0.66	0.18	0.64	0.64	-0.01
h^2l	0.88	0.42	0.18	0.84	0.38	0.01
hl^2	0.49	0.82	0.22	0.39	0.81	0.03
$2h+2l$	0.63	0.72	0.17	0.55	0.71	-0.02
h^2+l^2	0.62	0.71	0.18	0.54	0.70	-0.01
hw	0.68	0.09	0.72	0.60	0.00	0.65
h^2w	0.84	0.06	0.51	0.79	-0.02	0.42
hw^2	0.57	0.13	0.91	0.44	0.03	0.86
$2h+2w$	0.65	0.07	0.75	0.56	-0.02	0.69
h^2+w^2	0.62	0.08	0.74	0.53	-0.01	0.68
lw	0.15	0.66	0.73	-0.02	0.61	0.64
l^2w	0.15	0.79	0.57	-0.03	0.77	0.45
lw^2	0.14	0.50	0.84	-0.03	0.44	0.78
$2l+2w$	0.16	0.66	0.72	-0.01	0.62	0.63
l^2+w^2	0.18	0.66	0.69	0.02	0.62	0.60
h/l	0.63	-0.77	-0.03	0.75	-0.83	0.01
l/h	-0.63	0.77	0.03	-0.75	0.83	-0.01
h/w	0.69	-0.02	-0.70	0.82	0.01	-0.83
w/h	-0.69	0.02	0.70	-0.82	-0.01	0.83
l/w	-0.01	0.76	-0.65	-0.01	0.85	-0.80
w/l	0.01	-0.76	0.65	0.01	-0.85	0.80
hlw	0.58	0.54	0.60	0.45	0.48	0.47
$h^2+l^2+w^2$	0.47	0.58	0.61	0.34	0.53	0.49

Factor Correlations

	H	*L*	*W*	*H*	*L*	*W*
H	1.00			1.00		
L	0.00	1.00		0.21	1.00	
W	0.00	0.00	1.00	0.28	0.27	1.00

Table 4
Box Data: Random Start Results

		Standardization				
		None			CM	2Stage
		#LM	%√	GMS	√	√
Orthogonal	CF-Quartimax	1	0	*N*	*Y*	*N*
	CF-Varimax	1	0	*N*	*Y*	*N*
	Infomax	3	15	*N*	*Y*	*Y*
	Geomin (ϵ = .01)	4	27	*N*	*Y*	*Y*
	Minent	2	56	*N*	*Y*	*Y*
Oblique	CF-Quartimax	1	0	*N*	*Y*	*N*
	CF-Varimax	?	6	*N*	*Y*	*Y*
	Infomax	6	12	*N*	*Y*	*Y*
	Geomin (ϵ = .01)	4	16	*Y*	*Y*	*Y*

manner in which the data were constructed from the attributes *h*, *l* and *w*. It is of interest that the standard methods, quartimax and varimax in orthogonal rotation and direct quartimin (CF-quartimax) in oblique rotation, converged consistently to a single minimum, but this minimum was unsatisfactory. In oblique rotation, CF-varimax showed great difficulty in converging and the iterative procedure was usually terminated prematurely by the program after a maximum of 500 iterative cycles had been reached. In 6 instances, however, fairly satisfactory solutions were obtained, but all occurred after premature termination and they differed from each other. Infomax and geomin (orthogonal and oblique) and minimum entropy (orthogonal) yielded several local minima but in each instance one of the local minima was satisfactory. In some cases the percentage of cases in which the satisfactory solution was attained was fairly small. The satisfactory solution would probably be missed if random starts were not used.

The column of Table 4 headed by GMS indicates whether (*Y*) or not (*N*) the global minimum (smallest local minimum) yielded a satisfactory solution. In all cases, except for oblique geomin, a local minimum that was not the global minimum was the one judged satisfactory. There were no cases where more than one local minimum could be judged satisfactory, bearing in mind that oblique CF-varimax solutions that occur after premature termination cannot be regarded as yielding local minima.

The results of the box data trials where no standardization was used may be summarized as follows. A complexity function that had a single minimum never yielded a satisfactory solution. A complexity function with several local minima always had one that yielded a satisfactory solution. This satisfactory solution usually did not correspond to the global minimum. Consequently, there was no way of choosing the satisfactory solution without using knowledge of Thurstone's data generation process.

The column of Table 4 headed CM indicates the result of the rotation using the CM standardization. In every case, in orthogonal rotation or oblique rotation, using CF-varimax or other complexity functions, a satisfactory solution was obtained. The success Cureton and Mulaik (1975) experienced when using their standardization procedure in conjunction with orthogonal varimax was experienced again here using a variety of other criteria in both orthogonal and oblique rotation.

The final column headed 2Stage gives the results obtained when using an orthogonal rotation with the CM standardization to obtain a starting point and then carrying out an orthogonal or oblique rotation using the same complexity function with no standardization. In the three situations, orthogonal CF-quartimax, orthogonal CF-varimax and oblique CF-quartimax, where no satisfactory solution could be obtained using random starting points, use of an orthogonal rotation with the CM standardization to obtain a starting point did *not* result in a satisfactory solution. It seems clear that no satisfactory solution exists in these three situations. On the other hand, in all situations where a satisfactory solution was found to exist, it was found that using the CM standardization to obtain a starting point for a rotation with no standardization always resulted in the satisfactory solution being obtained.

Recapitulation

After these results had been obtained, it seemed that a final solution to the rotation problem had been found in the CM standardization. This standardization gave good results with the twenty four psycholgical tests and some other classical data sets where other approaches also gave good results. It also gave good results with the Thurstone box data where other approaches either had failed, or could only provide a satisfactory solution from specific starting points. If a satisfactory solution did exist, the CM standardization could be used to provide a good starting point. This obviated the need for random starting points and the choice of a rotated factor matrix from several alternatives.

It was subsequently found, however, that one can construct situations where the CM standardization will fail. This will now be considered.

Possible Failure of the CM Standardization

Consider Figure 1 once more. The first principal axis passes through the center of test points evenly scattered on the first quadrant of the circle. Both optimal orthogonal axes make angles of 45° with the principal axis. The CM weight function shown in Figure 2 assigns high weights to points making an angle of 45° with the principal axis and low weights to points close to the first principal axis or to the second principal axis.

In the configuration of test points shown in Figure 1 many tests have complexity greater than one. Now consider Figure 3. The variables form two orthogonal and essentially perfect clusters. (Perfect clusters are not shown in the figure to avoid superimposition of test points. Points adjoining the squares marked 1 and 2 should be regarded as coinciding with them.) The first principal axis passes through the denser configuration of points. It connects squares 1 and 1- and has the worst possible location. Low weights are assigned to the simple tests in the vicinity of squares 1, 1-, 2, 2- and high

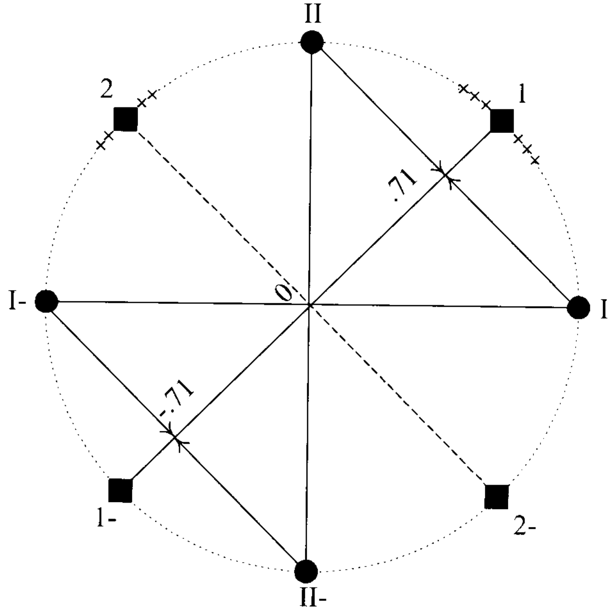

Figure 3
Cureton-Mulaik standardization: Orthogonal Perfect Clusters. $m = 2$ $m^{-1/2} = .71$

weights to any complex tests in the vicinity of dots I, I-, II, II-. Thus in this situation the CM procedure encourages the most complex configuration of points possible and avoids the attainable perfect orthogonal cluster configuration.

Similar reasoning will extend to more factors but is more difficult to illustrate graphically. The first factor matrix shown in Table 5 was constructed to illustrate failure of the CM standardization in an example with 12 variables and 3 factors. The first 10 variables exhibit a perfect orthogonal cluster configuration with three factors. If these 10 variables alone were included in the battery, the principal axes would coincide with the perfect cluster configuration shown and all CM weights would be zero. The last two variables are of complexity two but have very small communalities compared to the first ten tests. Consequently their inclusion in the battery has minimal effect on the orientation of the principal axes. These two complex variables, however, determine the orientation of the rotated axes after CM standardization.

Application of orthogonal CF-varimax with no standardization to the constructed factor matrix yields a matrix that agrees with it to three decimal places. This solution is stable under random restarts.

The last column shows the Kaiser weights ($h_{ii}^{-1/2}$)Although they apply considerable emphasis to the last two rows which have complexity two, CF-varimax with Kaiser standardization yields a matrix with elements that differ from those of the constructed matrix by not more than.01. The solution is stable under random restarts.

The Cureton-Mulaik weights w_i, are also shown in Table 5. They weight the perfect indicators out of the rotation process so that it is influenced almost entirely by the last two variables of complexity two. As a result the CM standardized varimax solution shown in the second part of Table 2 shows no semblance of a simple pattern. Whereas CM standardized varimax gives good results for the box data where the best configuration is complex it gives poor results in the present example where a near perfect orthogonal cluster configuration is possible.

This artificial example has two characteristics. Firstly a solution exists where most tests are perfect indicators and there is a small number of complex additional tests with low communalities. Secondly, at the rotation where most tests are perfect indicators, the factors are *uncorrelated*. In order to illustrate the necessity for this second condition two additional examples were constructed. The factor matrix given in Table 5 was used again but, in one case all factor intercorrelations were taken to be.3, and in the other they were taken to be .5. Oblique CF-varimax with CM standardization was now carried out with random starts in all three situations.

Table 5
Illustration of a Failure of Cureton-Mulaik Standardization in Orthogonal Varimax Rotation

Constructed Matrix

	Fac1	Fac2	Fac3	K Wts. $h_{ii}^{-1/2}$
Var1	0.90	0.00	0.00	1.11
Var2	0.90	0.00	0.00	1.11
Var3	0.90	0.00	0.00	1.11
Var4	0.90	0.00	0.00	1.11
Var5	0.00	0.80	0.00	1.25
Var6	0.00	0.80	0.00	1.25
Var7	0.00	0.80	0.00	1.25
Var8	0.00	0.00	0.70	1.43
Var9	0.00	0.00	0.70	1.43
Var10	0.00	0.00	0.70	1.43
Var11	0.10	0.10	0.00	7.07
Var12	0.10	0.00	0.10	7.07

Cureton-Mulaik Weighted Varimax

	Fac1	Fac2	Fac3	CM Wts. w_i
Var1	0.52	0.52	0.52	0.00
Var2	0.52	0.52	0.52	0.00
Var3	0.52	0.52	0.52	0.00
Var4	0.52	0.52	0.52	0.00
Var5	-0.46	0.63	-0.17	0.00
Var6	-0.46	0.63	-0.17	0.00
Var7	-0.46	0.63	-0.17	0.00
Var8	-0.40	-0.15	0.55	0.00
Var9	-0.40	-0.15	0.55	0.00
Var10	-0.40	-0.15	0.55	0.00
Var11	0.00	0.14	0.04	0.92
Var12	0.00	0.04	0.14	0.92

Results are shown in Table 6. When all factor correlations were equal to ρ = 0 oblique CF-varimax gave a poor solution that differs a little from the corresponding orthogonal rotation in Table 5. With ρ = .3 the same approach yielded a reasonably good approximation to a perfect cluster solution and this improved further with ρ = .5. The unfortunate effect of CM standardization in this situation disappeared as ρ increased.

The conclusions to be drawn from this investigation is that CM standardization can be very helpful under some circumstances and can result in poor solutions in others. It should form part of the kit of tools to be used in rotation, but its results should be evaluated against alternative solutions before being accepted.

Table 6
CM Fail Data: Effect of Interfactor Correlation Coefficient on Oblique CF-Varimax with CM Standardization

	ρ = 0			ρ = .3			ρ = .5		
	Fac1	Fac2	Fac3	Fac1	Fac2	Fac3	Fac1	Fac2	Fac3
Var 1	0.52	0.44	0.34	0.97	-0.08	-0.15	0.94	-0.03	-0.06
Var 2	0.52	0.44	0.34	0.97	-0.08	-0.15	0.94	-0.03	-0.06
Var 3	0.52	0.44	0.34	0.97	-0.08	-0.15	0.94	-0.03	-0.06
Var 4	0.52	0.44	0.34	0.97	-0.08	-0.15	0.94	-0.03	-0.06
Var 5	-0.47	0.74	-0.31	0.05	0.78	0.02	0.06	0.76	0.03
Var 6	-0.47	0.74	-0.31	0.05	0.78	0.02	0.06	0.76	0.03
Var 7	-0.47	0.74	-0.31	0.05	0.78	0.02	0.06	0.76	0.03
Var 8	-0.41	-0.35	0.72	0.05	0.02	0.68	0.06	0.04	0.66
Var 9	-0.41	-0.35	0.72	0.05	0.02	0.68	0.06	0.04	0.66
Var10	-0.41	-0.35	0.72	0.05	0.02	0.68	0.06	0.04	0.66
Var11	0.00	0.14	0.00	0.11	0.09	-0.01	0.11	0.09	0.00
Var12	0.00	0.00	0.14	0.11	-0.01	0.08	0.11	0.00	0.09

Factor Correlations

	Fac1	Fac2	Fac3	Fac1	Fac2	Fac3	Fac1	Fac2	Fac3
Fac1	1.00			1.00			1.00		
Fac2	0.04	1.00		0.34	1.00		0.47	1.00	
Fac3	0.13	0.50	1.00	0.39	0.21	1.00	0.48	0.37	1.00

A Practical Example: Investigation of the Simple Structure of the WAIS

The examples considered up to this point either were artificial or well-known in the literature for showing off rotation procedures to advantage. An application of the rotation methods discussed earlier in a practical situation will now be discussed.

Over the years it has become apparent that the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955) measures a number of intellectual abilities. Because the battery consists of only eleven tests and is therefore not suitable for a factor analysis extracting seven or more factors, Dr. J. J. McArdle planned and executed a study in which the WAIS was supplemented by tests from the Woodcock-Johnson Revised (WJR) battery (Woodcock & Johnson, 1989). He selected a subset of 16 WJR subtests intended to measure eight Gf-Gc factors (Cattell, 1963; Horn & Cattell, 1966). The names of the 11 WAIS tests and the 16 WJR tests are listed in Table 7. A correlation matrix of the 16 WJR tests based on a sample of size 763 was factor analyzed extracting eight factors. An initial maximum likelihood analysis yielded two Heywood cases. Multiple Heywood cases frequently indicate unstable solutions with multiple minima so that the ordinary least squares solution with no Heywood cases was chosen for further investigation. Oblique rotations of the 16×8 factor matrix were carried out using several complexity functions. It was found that they yielded similar perfect cluster solutions with two substantial loadings per factor. The CF-varimax and geomin solutions are shown in Table 8. Each was obtained consistently from 20 random starting points. The similarity of the two configurations obtained using totally different complexity functions suggests that it is reasonable to regard the pairs of variables indicated by underlined loadings as indicators of the same factor.

Another correlation matrix involving the 11 WAIS variables, 16 WJR indicators and an additional set of 6 WJR variables was factor analyzed extracting eight factors. This correlation matrix had been obtained by maximum likelihood using the EM algorithm (Marcantonio & Pechnyo, 1999) from a data set collected by Dr. McArdle with data missing by design to avoid burdening individual subjects with 33 tests. The maximum number of subjects involved for any block of correlation coefficients was 294. Again an ordinary least squares solution was obtained for consistency with the preliminary factor analysis and because it is not clear that maximum Wishart likelihood should be preferred for analyzing a correlation matrix based on incomplete data.

Table 7
Tests in McArdle's Battery

WAIS tests		**Original WJ-R tests**		**Additional WJ-R tests**	
IN	Information	PV	Picture Vocabulary	SC	Science Knowledge
CO	Comprehension	OV	Oral Vocabulary	SS	Social Studies
AR	Arithmetic	AS	Analysis-Synthesis	HU	Humanities
SI	Similarities	CF	Concept Formation	PLS	Power Letter Series
DSP	Digit Span	MN	Memory for Names	PNS	Power Number Series
VO	Vocabulary	VAL	Visual-Auditory Learn	MA	Matrices
DSY	Digit Symbol	MS	Memory for Sentences		
PC	Picture Completion	MW	Memory for Words		
BD	Block Design	VM	Visual Matching		
PA	Picture Arrangement	COU	Cross-Out		
OA	Object Assembly	IW	Incomplete Words		
		SB	Sound Blending		
		VC	Visual Closure		
		PR	Picture Recognition		
		CA	Calculation		
		AP	Applied Problems		
WJ-R Factors					
Gc		Gsm		Gv	
Gf		Gs		Gq	
Glr		Ga			

Oblique CF-varimax and geomin rotations were obtained once more. The 16 × 8 WJR submatrices extracted from the 33 × 8 rotated factor matrices are shown in Table 9. Results are no longer as similar as they were in Table 8 and the configuration appears to have changed particularly for factors Gf and Glr. This leads to some doubt concerning the matching of factors between the two data sets. Consequently a target rotation was carried out. The only loadings that were specified (= 0) in the target were those in the 16 × 8 WJR indicator block that do *not* correspond to loadings that are underlined in Table 9. Target elements corresponding to underlined loadings and all those in the 11 × 8 WAIS and 6 × 8 additional WJR blocks were left unspecified. Results of the oblique target rotation are shown in

Table 8
WJR Data - 16 Tests, 8 Factors, *N* = 763.

Oblique CF-Varimax Factor Matrix

	Gc	Gf	Glr	Gsm	Gs	Ga	Gv	Gq
PV	<u>0.86</u>	0.01	0.02	0.00	0.04	0.03	0.06	0.02
OV	<u>0.53</u>	0.08	0.06	0.15	-0.03	0.08	-0.04	0.25
AS	-0.04	<u>0.48</u>	0.00	0.10	0.04	0.07	0.23	0.18
CF	0.06	<u>0.64</u>	0.08	0.02	0.11	0.08	0.01	0.04
MN	0.00	-0.01	<u>0.98</u>	-0.01	0.01	0.01	-0.01	-0.01
VAL	0.00	0.18	<u>0.35</u>	0.11	-0.06	0.07	0.32	0.17
MS	0.09	0.19	0.06	<u>0.63</u>	0.00	0.03	-0.09	0.00
MW	-0.03	-0.09	0.01	<u>0.67</u>	0.09	0.08	0.10	0.07
VM	0.04	0.02	0.04	0.06	<u>0.83</u>	0.01	0.00	0.06
COU	-0.03	0.12	0.06	0.00	<u>0.61</u>	0.12	0.14	0.04
IW	0.17	0.10	0.03	0.24	0.11	<u>0.39</u>	0.08	-0.15
SB	0.00	0.01	0.02	0.00	0.00	<u>0.88</u>	0.00	0.03
VC	0.18	0.07	0.04	-0.04	0.18	0.15	<u>0.48</u>	-0.04
PR	0.13	0.12	0.19	0.12	0.12	-0.02	<u>0.36</u>	-0.02
CA	0.01	0.07	0.05	0.02	0.15	0.09	0.05	<u>0.66</u>
AP	0.21	0.10	0.04	0.09	0.04	0.05	-0.03	<u>0.61</u>

Oblique Geomin Factor Matrix (ϵ = .02)

	Gc	Gf	Glr	Gsm	Gs	Ga	Gv	Gq
PV	<u>0.88</u>	-0.02	-0.01	-0.02	0.03	0.01	0.08	0.04
OV	<u>0.57</u>	0.06	0.04	0.12	-0.03	0.05	-0.02	0.28
AS	-0.04	<u>0.45</u>	-0.01	0.05	0.03	0.03	0.30	0.20
CF	0.06	<u>0.65</u>	0.06	-0.01	0.12	0.05	0.05	0.04
MN	0.00	-0.02	<u>0.96</u>	-0.02	0.04	0.02	0.01	0.00
VAL	-0.01	0.13	<u>0.33</u>	0.07	-0.07	0.04	0.40	0.20
MS	0.09	0.24	0.04	<u>0.62</u>	-0.02	0.01	-0.06	0.01
MW	-0.04	-0.06	-0.01	<u>0.66</u>	0.07	0.06	0.14	0.08
VM	0.04	0.01	0.01	0.07	<u>0.87</u>	0.00	-0.02	0.04
COU	-0.04	0.09	0.04	0.00	<u>0.64</u>	0.10	0.15	0.04
IW	0.15	0.11	0.02	0.25	0.09	<u>0.37</u>	0.10	-0.13
SB	0.00	0.00	0.02	0.02	0.00	<u>0.85</u>	-0.01	0.08
VC	0.15	-0.01	0.02	-0.06	0.16	0.12	<u>0.56</u>	-0.02
PR	0.11	0.07	0.18	0.09	0.11	-0.04	<u>0.43</u>	-0.01
CA	0.05	0.02	0.03	-0.02	0.17	0.04	0.06	<u>0.70</u>
AP	0.26	0.07	0.02	0.05	0.05	0.00	-0.01	<u>0.65</u>

Table 10. The correspondence with Table 8 is clearer, albeit not perfect, and the interpretation of the WAIS tests in a framework provided by the WJR markers has been facilitated.

The last column of Table 10 shows the CM weights. There is no tendency for the factorially simple tests in the WJR marker battery to have higher weights than the complex WAIS variables.

Concluding Observations

Some diverse rotation criteria have been tried out in several different situations and this has led to a number of observations. Little has been proved irrefutably but a number of plausible working hypotheses or conjectures have been suggested.

It appears that in cases where a near perfect cluster configuration exists, most complexity functions considered here will have a single minimum and yield an acceptable solution. In situations where the best factor pattern is complex it seems that those complexity functions that are stable and yield a single global minimum can also be expected to yield a poor solution. Complexity functions that are capable of yielding a minimum accompanied by a good solution can be expected to be unstable and to also yield other local minima accompanied by poor solutions. Unfortunately, in situations where there are several minima, the lowest local minimum, or global minimum, need not be accompanied by the best solution. The choice of the best solution therefore *cannot* be made automatically and without human judgment.

The CM weighting scheme can result in an acceptable solution using a stable complexity function that would otherwise be unable to locate it. When used in conjunction with a complexity function that does yield an acceptable solution at one of several local minima, the CM standardization can result in a single global minimum which is accompanied by an acceptable solution. These outcomes can be expected when there is at least one perfect indicator of each factor and a substantial number of variables of higher complexity. The solution is made possible by weighting out the complex variables so that the solution is determined primarily by the perfect indicators. Unfortunately the CM approach can also result in an unacceptable solution when a good unweighted solution is available. This can be expected when most variables are perfect indicators of uncorrelated factors and there are a few complex variables with low communalities. The CM standardization then weights out the perfect indicators and allows the solution to be determined by the complex variables.

It is clear that we are not at a stage where we can rely on mechanical exploratory rotation by a computer program if the complexity of most

Table 9
WJR +WAIS: 33 Tests, 8 Factors, $N \approx 294$.

	Gc	Gf	Glr	Gsm	Gs	Ga	Gv	Gq
	Part of Oblique CF-Varimax Factor Matrix							
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
PV	<u>0.41</u>	-0.04	0.07	0.02	0.05	0.12	0.44	0.12
OV	<u>0.59</u>	-0.02	-0.04	0.10	0.16	0.26	0.11	-0.02
AS	-0.12	<u>0.40</u>	0.09	0.16	0.02	0.39	0.10	0.08
CF	0.02	<u>0.48</u>	0.14	0.18	0.05	0.21	0.02	0.04
MN	0.12	0.42	<u>0.24</u>	0.23	0.09	0.04	0.01	-0.21
VAL	0.08	0.31	<u>0.32</u>	0.07	0.17	0.26	0.05	-0.11
MS	0.15	-0.04	0.29	<u>0.61</u>	-0.01	-0.04	-0.02	0.09
MW	-0.02	0.02	-0.01	<u>0.83</u>	-0.01	0.01	0.04	-0.03
VM	0.03	0.02	0.04	0.10	<u>0.82</u>	-0.02	0.01	0.07
COU	-0.15	0.10	0.22	0.06	<u>0.50</u>	0.06	0.28	0.05
IW	0.03	-0.01	0.00	0.43	0.12	<u>0.20</u>	0.41	-0.08
SB	0.05	0.15	0.04	0.35	0.14	<u>0.34</u>	0.16	-0.26
VC	0.01	0.33	0.26	-0.01	0.23	-0.07	<u>0.42</u>	-0.09
PR	0.04	0.27	0.14	0.15	0.23	0.01	<u>0.17</u>	-0.08
CA	0.11	0.07	0.07	0.04	0.38	0.43	-0.11	<u>0.23</u>
AP	0.00	0.17	0.09	0.09	0.15	0.22	0.09	<u>0.50</u>
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	Part of Oblique Geomin Factor Matrix ($\epsilon = .02$)							
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
PV	<u>0.78</u>	-0.02	0.01	0.01	-0.01	0.08	0.28	0.03
OV	<u>0.78</u>	0.06	-0.07	0.08	0.17	0.16	-0.03	-0.15
AS	0.04	<u>0.74</u>	-0.03	0.12	0.00	0.01	-0.04	0.20
CF	0.03	<u>0.78</u>	0.00	0.13	-0.01	-0.08	-0.02	0.02
MN	-0.11	0.73	<u>0.07</u>	0.19	-0.02	0.01	0.07	-0.16
VAL	0.05	0.71	<u>0.13</u>	0.03	0.11	0.09	0.03	-0.01
MS	0.15	0.01	0.28	<u>0.59</u>	-0.01	-0.06	0.01	-0.06
MW	-0.06	0.02	0.05	<u>0.84</u>	-0.02	0.02	0.02	0.01
VM	0.03	-0.04	0.02	0.10	<u>0.81</u>	-0.04	0.17	-0.04
COU	-0.03	0.21	0.11	0.05	<u>0.43</u>	0.02	0.33	0.14
IW	0.24	0.08	-0.03	0.45	0.06	<u>0.20</u>	0.26	0.14
SB	0.05	0.44	-0.04	0.35	0.09	<u>0.27</u>	0.03	0.04
VC	0.02	0.51	0.05	-0.03	0.05	0.00	<u>0.47</u>	0.01
PR	-0.01	0.42	0.02	0.13	0.13	0.00	<u>0.21</u>	-0.04
CA	0.35	0.26	0.03	0.00	0.47	0.02	-0.21	<u>0.12</u>
AP	0.40	0.21	0.05	0.05	0.21	-0.23	0.00	<u>0.24</u>
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 10
WJR + WAIS 33 Tests 8 Factors, *N* = ± 294

	Oblique Target Rotation: Factor Matrix								
	Gc	Gf	Glr	Gsm	Gs	Ga	Gv	Gq	CM Wts
IN	0.74	0.05	-0.09	0.07	-0.20	-0.09	0.01	0.39	0.58
CO	0.71	0.26	-0.26	0.12	-0.08	-0.34	0.26	0.13	0.62
AR	0.22	0.08	-0.04	0.37	0.08	-0.41	0.09	0.53	0.42
SI	0.45	0.14	0.28	0.00	0.00	-0.05	0.11	0.04	0.13
DSP	-0.04	0.24	-0.17	0.53	0.26	-0.03	-0.13	0.13	0.53
VO	1.14	-0.04	0.13	0.18	-0.17	-0.40	-0.08	-0.07	0.83
DSY	-0.03	-0.07	0.16	-0.03	0.75	0.11	0.02	0.08	0.47
PC	0.14	-0.09	-0.03	0.16	0.04	0.12	0.35	0.45	0.25
BD	-0.19	-0.30	0.62	0.46	0.00	-0.11	0.20	0.41	0.36
PA	0.16	0.32	0.10	0.05	0.14	-0.16	0.34	0.13	0.19
OA	0.03	-0.42	0.56	0.35	0.09	-0.01	0.36	0.16	0.37
PV	<u>0.76</u>	-0.08	-0.05	0.00	-0.08	0.09	0.22	0.13	0.46
OV	<u>0.83</u>	0.03	0.07	0.00	0.01	0.06	-0.14	0.02	0.47
AS	-0.04	<u>0.42</u>	0.15	0.00	-0.05	0.25	0.08	0.31	0.25
CF	0.05	<u>0.43</u>	0.21	0.10	0.01	0.02	0.12	0.17	0.19
MN	0.06	0.26	<u>0.36</u>	0.20	0.07	-0.00	0.15	-0.13	0.36
VAL	0.07	0.12	<u>0.44</u>	0.07	0.08	0.14	0.11	0.13	0.17
MS	0.07	-0.14	0.16	<u>0.79</u>	-0.02	-0.07	0.01	0.12	0.44
MW	-0.09	0.12	-0.10	<u>0.73</u>	0.04	0.20	-0.07	-0.06	0.72
VM	0.02	-0.01	0.00	0.06	<u>0.86</u>	-0.02	0.00	0.09	0.48
COU	-0.09	-0.05	0.11	0.03	<u>0.51</u>	0.19	0.27	0.18	0.32
IW	0.21	0.03	-0.09	0.20	0.08	<u>0.48</u>	0.12	0.00	0.33
SB	0.10	0.15	0.18	0.09	0.07	<u>0.48</u>	-0.05	-0.07	0.38
VC	0.15	0.11	0.17	-0.04	0.23	0.09	<u>0.49</u>	-0.05	0.35
PR	0.08	0.16	0.14	0.09	0.23	0.07	<u>0.21</u>	-0.04	0.21
CA	0.15	0.12	0.14	0.00	0.26	0.09	-0.21	<u>0.47</u>	0.27
AP	0.15	0.23	-0.09	0.15	0.09	-0.10	0.09	<u>0.61</u>	0.26
SC	0.37	0.03	0.26	0.04	-0.29	0.21	0.10	0.40	0.23
SS	0.66	-0.05	-0.03	0.12	-0.15	-0.08	0.08	0.42	0.47
HU	0.61	0.09	0.15	0.08	-0.08	0.08	0.08	0.08	0.18
PLS	0.12	0.21	0.22	0.21	0.08	-0.03	0.08	0.24	0.05
PNS	0.34	0.42	-0.06	-0.10	0.29	-0.15	0.04	0.29	0.19
MA	-0.07	0.58	0.01	-0.01	0.09	-0.13	0.18	0.33	0.41

variables is not close to one. One approach that is open to the user is to try more than one complexity function; for example, a member of the Crawford-Ferguson family, infomax and either minimum entropy (orthogonal) or geomin (oblique). Multiple starting points are also to be recommended in order to detect multiple minima of a complexity function. In situations where some information about the configuration is available it is worth trying a sequence of rotations to partially specified targets. The choice of modifications to yield the next target would be guided by the results of the present rotation. All this involves human thought and judgment, which seems unavoidable if exploration is to be carried out. A choice from several alternatives will be accompanied by capitalization on chance, requiring a follow up confirmatory factor analysis on new data (cf. Nesselroade and Baltes, 1984, pp. 272-273). It does not seem realistic to expect that exploration of the structure of a battery of tests should be accomplished with the analysis of a single sample.

References

Agresti, A. (1990). *Categorical data analysis*. New York: Wiley.

Algina, J. (1980). A note on identification in the oblique and orthogonal factor-analysis models. *Psychometrika, 3*, 393-396.

Brent, R. P. (1973). *Algorithms for minimization without derivatives*. Englewood Cliffs, NJ: Prentice-Hall.

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.

Browne, M. W., Cudeck, R., Tateneni, K. & Mels, G. (1998) *CEFA: Comprehensive Exploratory Factor Analysis*. WWW document and computer program. URL http://quantrm2.psy.ohio-state.edu/browne/

Butler, J. M. (1964). Simplest data factors and simple structure in factor analysis. *Educational and Psychological Measurement, 24*, 755-763.

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

Cattell, R. B. (1963). Theory for fluid and crystallized intelligence: A critical experiment. *Journal of Experimental Psychology, 54*, 1-22.

Crawford, C. B. (1975). A comparison of the direct oblimin and primary parsimony methods of oblique rotation. *British Journal of Mathematical and Statistical Psychology, 28*, 201-213.

Crawford, C. B. & Ferguson, G. A. (1970). A general rotation criterion and its use in orthogonal rotation. *Psychometrika, 35*, 321-332.

Cureton, E. E. & Mulaik, S. A. (1971). On simple structure and the solution to Thurstone's "invariant" box problem. *Multivariate Behavioral Research, 6*, 375-387.

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

Eber, H. W. (1966). Toward oblique simple structure: Maxplane. *Multivariate Behavioral Research, 1*, 112-125.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999), Evaluating the use of factor analysis in psychological research. *Psychological Methods, 4,* 281-290.

Ferguson, G. A. (1954). The concept of parsimony in factor analysis. *Psychometrika, 19,* 347-362.

Gebhardt, F. (1968). A counterexample to two-dimensional varimax rotation. *Psychometrika, 33,* 35-36.

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

Harman, H. H. (1976). *Modern factor analysis, third edition*. Chicago: University of Chicago Press.

Harris, C. W. & Kaiser, H. F. (1964). Oblique factor analytic transformations 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: Department of Education, University of Chicago.

Horn, J. L. & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 57,* 253-270.

Horst, A. P. (1941). A non-graphical method for transforming an arbitrary factor matrix into a simple structure factor matrix. *Psychometrika, 6,* 79-99.

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 Memorandum* 65-13. Princeton, NJ: Educational Testing Service.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika, 34,* 183-202.

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

Kaiser, H. F. (1959). Computer program for varimax rotation in factor analysis. *Educational and Psychological Measurement, 19*, 413-420.

Kiers, H. A. L. (1994). SIMPLIMAX: Oblique rotation to an optimal target with simple structure. *Psychometrika, 59,* 567-579.

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

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84-99.

Marcantonio, R. & Pechnyo, M. (1999). Missing value analysis. In *SYSTAT 9: Statistics II* (pp. 1-46). Chicago: SPSS.

McCammon, R. B. (1966). Principal component analysis and its application in large-scale correlation studies. *Journal of Geology, 74*, 721-733.

McKeon, J. J. (1968). *Rotation for maximum association between factors and tests*. Unpublished manuscript, Biometric Laboratory, George Washington University.

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

Neuhaus, J. O. & Wrigley, C. (1954). The quartimax method: an analytical approach to orthogonal simple structure. *British Journal of Statistical Psychology, 7,* 187-191.

Nesselroade, J. R. & Baltes, P. B. (1984). From traditional factor analysis to structural-causal modeling in developmental research. In V. Sarris & A. Parducci (Eds.) *Perspectives in psychological experimentation: Toward the year 2000.* Hillsdale, NJ: Erlbaum.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1986). *Numerical recipes in Fortran, second edition*. Cambridge, UK: Cambridge University Press.

Rozeboom, W. W. (1992). The glory of suboptimal factor rotation: Why local minima in analytic optimization of simple structure are more blessing than curse. *Multivariate Behavioral Research, 27*, 585-599.

Saunders, D. R. (1953). *An analytic method for rotation to orthogonal simple structure*. Research Bulletin, 53-10, Princeton, NJ: Educational Testing Service.

Thurstone, L. L. (1935). *The vectors of mind*. Chicago: University of Chicago Press.

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

Tucker, L. R. (1940). A rotational method based on the mean principal axis of a subgroup of tests. *Psychological Bulletin, 5*, 289-294.

Tucker, L. R. (1944). A semi-analytical method of factorial rotation to simple structure. *Psychometrika, 9*, 43-68.

Wechsler, D. (1955). *Manual for the Wechsler adult intelligence scale*. New York: The Psychological Corporation.

Woodcock, R. W. & Johnson, M. B. (1989). *Woodcock-Johnson psycho-educational battery-revised*. Allen, TX: DLM Teaching Resources.

Yates, A. (1987). *Multivariate exploratory data analysis: A perspective on exploratory factor analysis*. Albany: State University of New York Press.

Accepted October, 2000.