PSYCHOMETRIKA—VOL. 62, NO. 4, 579–598
DECEMBER 1997

THREE-MODE ORTHOMAX ROTATION

HENK A. L. KIERS

UNIVERSITY OF GRONINGEN

Factor analysis and principal components analysis (PCA) are often followed by an orthomax rotation to rotate a loading matrix to simple structure. The simple structure is usually defined in terms of the simplicity of the columns of the loading matrix. In Three-mode PCA, rotational freedom of the so called core (a three-way array relating components for the three different modes) can be used similarly to find a simple structure of the core. Simple structure of the core can be defined with respect to all three modes simultaneously, possibly with different emphases on the different modes. The present paper provides a fully flexible approach for orthomax rotation of the core to simple structure with respect to three modes simultaneously. Computationally, this approach relies on repeated (two-way) orthomax applied to supermatrices containing the frontal, lateral or horizontal slabs, respectively. The procedure is illustrated by means of a number of exemplary analyses. As a by-product, application of the Three-mode Orthomax procedures to two-way arrays is shown to reveal interesting relations with and interpretations of existing two-way simple structure rotation techniques.

Key words: varimax, quartimax, three-mode principal components analysis, simple structure.

In factor analysis and principal components analysis (PCA) the loading matrix is determined up to a rotation. To exploit this rotational freedom, the loading matrix is often rotated to "simple structure" so as to simplify the interpretation of the factors or components. A well known class of such simple structure rotation procedures is the "Orthomax" family (Crawford & Ferguson, 1970; Jennrich, 1970). Varimax (Kaiser, 1958) and quartimax (Carroll, 1953; Ferguson, 1954; Neuhaus & Wrigley, 1954; Saunders, 1953) are the best known members of the orthomax family. Orthomax rotation finds a rotation matrix $\mathbf{T}$ ($r \times r$) of a loading matrix $\mathbf{\Lambda}$ ($m \times r$) such that $\tilde{\mathbf{\Lambda}} = \mathbf{\Lambda T}$ has optimal simple structure in terms of the orthomax criterion employed. Specifically, the orthomax criterion is m^{-1} times the function

$$\text{ORMAX}(\tilde{\mathbf{\Lambda}}, \gamma) = \sum_{l=1}^{r} \left(\sum_{i=1}^{m} \tilde{\lambda}_{il}^4 - \frac{\gamma}{m} \left(\sum_{i=1}^{m} \tilde{\lambda}_{il}^2 \right)^2 \right), \tag{1}$$

where $\tilde{\lambda}_{il}$ denotes the element (i, l) of $\tilde{\mathbf{\Lambda}} = \mathbf{\Lambda T}$, and γ is the parameter monitoring the choice of the orthomax criterion (e.g., $\gamma = 0$ yields the quartimax criterion, and $\gamma = 1$ yields the varimax criterion). Orthomax rotation consists of maximizing (1) over orthonormal matrices $\mathbf{T}$, for some selected γ.

Orthomax rotation is a procedure for simple structure rotation of a *matrix*, that is, a two-way array. It exploits the rotational freedom of two-mode PCA or factor analysis. In Three-mode PCA, a similar rotational freedom is encountered. The present paper offers a class of methods for exploiting this type of rotational freedom. Before describing these, we give a brief description of Three-mode PCA and its rotational freedom.

Three-mode PCA (Kroonenberg & de Leeuw, 1980; Tucker, 1966) can be described as follows. Let $\mathbf{X}$ denote an I $\times$ J $\times$ K data array. Then the Three-mode PCA model is given by

This research has been made possible by a fellowship from the Royal Netherlands Academy of Arts and Sciences to the author. The author is obliged to Jos ten Berge and two anonymous reviewers for useful comments on an earlier version of this paper. Requests for reprints should be sent to Henk A. L. Kiers, Department of Psychology, Grote Kruisstraat 2/1, 9712 TS Groningen, THE NETHERLANDS.

0033-3123/97/1200-94247$00.75/0

$$\hat{x}_{ijk} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} a_{ip} b_{jq} c_{kr} g_{pqr}, \qquad (2)$$

where $\hat{x}_{ijk}$ denotes the estimate for the element (i, j, k) of **X**; **A**, **B**, and **C** (with elements a_{ip}, b_{jq}, and c_{kr}, respectively) are component matrices of orders I × P, J × Q, and K × R, respectively, and **G** is a P × Q × R three-way array denoted as the *core*, with elements g_{pqr}, $i = 1, \ldots, I, j = 1, \ldots, J, k = 1, \ldots, K, p = 1, \ldots, P, q = 1, \ldots, Q$, and $r = 1, \ldots, R$. The matrices **A**, **B**, and **C** can be considered component scores for "A-mode entries" (in **A**), "B-mode entries" (in **B**), and "C-mode entries" (in **C**), respectively. The elements of the core indicate how the components from the different modes interact. Three-mode PCA consists of fitting model (2) to a data array by minimizing the sum of squared residuals, $\Sigma_i \Sigma_j \Sigma_k (x_{ijk} - \hat{x}_{ijk})^2$, over **A**, **B**, **C**, and **G**. Kroonenberg and de Leeuw (1980) have proposed an alternating least squares (ALS) algorithm for this.

To give some insight in the role of the core array, we give the following tensorial description of the Three-mode PCA model. The array $\hat{\mathbf{X}}$, with elements $\hat{x}_{ijk}$, see (2), can be written as

$$\hat{\mathbf{X}} = g_{111}(\mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1) + g_{112}(\mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_2) + g_{113}(\mathbf{a}_1 \otimes \mathbf{b}_1 \otimes \mathbf{c}_3) + \cdots$$
$$+ g_{211}(\mathbf{a}_2 \otimes \mathbf{b}_1 \otimes \mathbf{c}_1) + \cdots + g_{PQR}(\mathbf{a}_P \otimes \mathbf{b}_Q \otimes \mathbf{c}_R), \qquad (3)$$

where $(\mathbf{a}_i \otimes \mathbf{b}_j \otimes \mathbf{c}_k)$ denotes the triple tensor product of column i of **A**, column j of **B**, and column k of **C**. Each triple tensor product is defined here as the (P × Q × R) three-way array that comes about by calculating all possible products of elements from the three different vectors. It is clear from (3) that the elements of the core indicate the importance of the different tensor product terms. It is also clear from (3) that reducing the number of "significant" core elements would facilitate the interpretation of the Three-mode PCA solution, because it would involve fewer terms to take into account. Therefore, in the present paper we consider procedures that rotate the core such that it has a simple structure in the sense that many elements are so small that we may ignore them. First, however, we will study the rotational freedom in more detail.

Tucker (1966) already described the rotational freedom of the Three-mode PCA model. Specifically, he has shown that postmultiplication of matrices **A**, **B** and **C** by nonsingular matrices can always be compensated by applying the inverse of these matrices to the core array. Specifically, it can be verified that, if we choose $\tilde{\mathbf{A}} = \mathbf{AS}^{-1}$, $\tilde{\mathbf{B}} = \mathbf{BT}^{-1}$, $\tilde{\mathbf{C}} = \mathbf{CU}^{-1}$ and $\tilde{\mathbf{G}}$ with elements

$$\tilde{g}_{ijk} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} s_{ip} t_{jq} u_{kr} g_{pqr}, \qquad (4)$$

$i = 1, \ldots, P, j = 1, \ldots, Q, k = 1, \ldots, R$, then $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{G}}$ give the same estimates for $\hat{\mathbf{X}}$ as **A**, **B**, **C** and **G** do (see, e.g., Kiers, 1992, for a derivation). As a consequence, when we have obtained a solution, we may always transform this in three directions such that the core becomes "simple". Suggestions to rotate the core to a simpler structure have been made by Cohen (1974, 1975), MacCallum (1976), Kroonenberg (1983, chap. 5), and Kiers (1992). The first three authors proposed procedures for making the frontal core slabs as diagonal as possible. Kiers proposed to rotate the core such that it would be optimally "superdiagonal" (i.e., large core elements on the three-way diagonal of the core, and small elements elsewhere). All these approaches have in common that they rotate the core to an a priori structure. However, in certain cases such an a priori structure may turn out to be unfeasible. For instance, in cases where the core is not "cubic" (i.e., with equal dimensions) such a superdiagonal structure is impossible by definition. When an a priori simple structure cannot be obtained it is still conceivable that other simple structures do exist. To the author's knowledge, only two attempts have been made to find such data dependent (rather than a priori) simple structures. These are Murakami's (1983) proposal to apply varimax to a supermatrix of frontal core planes, and Kruskal's (1988) method for "tri-quartimax" rotation. Murakami's method is limited to rotation in only one of the three directions. Whereas this is sufficient for his particular purposes, in other situations it will be desirable to further simplify the core by rotation in the other directions as well. Kruskal's method does use the full rotational freedom of the core. His tri-quartimax procedure maximizes a combination of normalized quartimax functions applied to the supermatrices consisting of the frontal, lateral and horizontal planes, respectively, of the core. Kruskal proposed to maximize this function over oblique rotations, thus ignoring the (as we will see) much simpler possibility of using orthogonal rotations. No details on either the implementation, or the usefulness of his procedure have been published.

Kruskal's (1988) idea has been an important point of departure for the present paper. However, in the present paper we consider orthogonal rather than oblique rotations, and we consider orthomax rather than quartimax rotation. The reason to use orthogonal rotation is that, in addition to the simplifications obtained by using orthogonal rotations, it has an interest in its own right: In certain applications it is desired that components remain orthogonal after rotation, for instance, because this allows a simple fit partitioning (ten Berge, de Leeuw & Kroonenberg, 1987). The reason to use orthomax rather than quartimax is that, in the two-way case, quartimax is known for its tendency to yield a "general factor", and varimax (or more generally orthomax) has been proposed as an attempt to avoid this problem. The present paper will therefore start by generalizing Kruskal's tri-quartimax criterion to Three-mode Orthomax criteria.

The above announced generalization of Kruskal's (1988) criterion is based on a combination of two-way orthomax criteria applied to supermatrices consisting of vectorized slabs of the core array. A detailed account of this generalized criterion and the ensuing class of methods will be given in the present paper. It will be shown that the generalized criterion offers great flexibility, allowing for (combinations of) different kinds of orthomax criteria, and for differential weighting of the orthomax criteria involved. By means of differential weighting of the orthomax criteria, it is possible to attach different weights to the simplicity in the different directions of the core. In the extreme case of setting certain weights to zero, attention will be limited exclusively to one or two modes. This is interesting in practice, for instance when one or two of the component matrices have itself been rotated to simple structure. A case in point is Murakami's (1983) method, which can be viewed as a special case of our method.

In the next section, we describe the Three-mode Orthomax criterion. In a subsequent section, it will be shown how this function can be maximized iteratively. After the description of the algorithm, the performance of the ensuing methods is studied by a number of exemplary analyses, and a small simulation study.

Since a two-way array can be considered a three-way array with only one slab, we can apply the general combined orthomax criterion to two-way matrices as well. Although this does not lead to new methods, it does lead to new interpretations of existing methods. These new perspectives are discussed at the end of the paper.

Three-Mode Orthomax Rotation

In the present paper, Three-mode Orthomax rotation is developed as a method to maximize a weighted sum of orthomax criteria, see (1), applied to the core. Before defining the Three-mode Orthomax criterion, we have to explain our notation: $\tilde{\mathbf{G}}^1$, $\tilde{\mathbf{G}}^2$, and $\tilde{\mathbf{G}}^3$ denote the matrices whose columns consist of the vectorized horizontal, lateral and frontal slabs, respectively, of the core $\tilde{\mathbf{G}}$; $\tilde{\mathbf{G}}_1, \ldots, \tilde{\mathbf{G}}_R$ denote the frontal slabs of $\tilde{\mathbf{G}}$. In Three-mode Orthomax rotation, we maximize

$$f(\mathbf{S}, \mathbf{T}, \mathbf{U}) = \sum_{l=1}^{3} w_l \text{ORMAX}(\tilde{\mathbf{G}}^l, \gamma_l) = w_1 \text{ORMAX}\left(\begin{pmatrix} \tilde{\mathbf{G}}_1' \\ \cdot \\ \cdot \\ \tilde{\mathbf{G}}_R' \end{pmatrix}, \gamma_1\right)$$

$$+ w_2 \text{ORMAX}\left(\begin{pmatrix} \tilde{\mathbf{G}}_1 \\ \cdot \\ \cdot \\ \tilde{\mathbf{G}}_R \end{pmatrix}, \gamma_2\right) + w_3 \text{ORMAX}((\text{Vec}(\tilde{\mathbf{G}}_1) \ldots \text{Vec}(\tilde{\mathbf{G}}_R)), \gamma_3), \quad (5)$$

over orthonormal matrices **S**, **T** and **U**, where w_1, w_2 and w_3 are fixed prespecified weights, and γ_1, γ_2 and γ_3 are prespecified values of the γ parameter; here $\tilde{\mathbf{G}}$ is the core rotated by **S**, **T** and **U**, as defined in (4), and Vec( ) denotes the vector obtained by stringing out a matrix columnwise. Maximization of (5) comes down to maximizing the weighted sum of orthomax criteria applied to supermatrices that consist of the vectorized horizontal, lateral and frontal slabs, respectively, of the core. By maximizing the weighted sum of these criteria, we in fact maximize a weighted sum of simplicity measures applied to each of the slabs obtained in each of the three directions. That is, we aim at finding rotations such that, on the average, the elements vary maximally within the slabs. The terms of (5) can be interpreted as in ordinary orthomax. Specifically, by choosing $w_1 = (\text{QR})^{-1}$, $w_2 = (\text{PR})^{-1}$, $w_3 = (\text{PQ})^{-1}$, the three constituent parts of (5) all refer to original orthomax functions (which come about by dividing the ORMAX function by the number of rows of the matrix to which it is applied, see (1)). For instance, when $w_1 = (\text{QR})^{-1}$ and $\gamma_1 = 1$ (hence orthomax reduces to varimax), the first term sums the variances of the squared core elements in the horizontal slabs. More generally, the Three-mode version of varimax is obtained by taking $w_1 = (\text{QR})^{-1}$, $w_2 = (\text{PR})^{-1}$, $w_3 = (\text{PQ})^{-1}$, and $\gamma_1 = \gamma_2 = \gamma_3 = 1$.

The Three-mode Orthomax criterion is a generalization of Kruskal's (1988) tri-quartimax criterion. In the orthogonal case, his criterion reduces to the sum of fourth powers of all core elements, multiplied by a constant. In the Three-mode Orthomax criterion this is obtained as the special case with $\gamma_1 = \gamma_2 = \gamma_3 = 0$. In the present paper, this method is denoted as "Three-mode Quartimax".

Clearly, it is possible to set one or two of the weights in (5) equal to 0, thus deleting the associated terms from the criterion. For instance, setting $w_1 = 0$ in (5) implies that the orthomax criterion is applied only to the lateral and the frontal slabs (even though the core is rotated in *all three* directions). Simplicity of the horizontal slabs is thus not aimed at. In this way a flexible approach is offered in which different weights for different modes can monitor the importance attached to the modes. Further flexibility is implied by the fact that the γ parameters in each orthomax function can be specified at wish. Thus, each choice of weights and parameters would seem to lead to a different model, and it may be quite hard to choose among the multitude of possibilities. Fortunately, however, it turns out that there is a dependency in the choice of the γ parameters and the weights, which can be seen as follows. Since the optimization problem does not alter by multiplication by a constant, we may use the normalization $(w_1 + w_2 + w_3) = (\text{QR})^{-1} + (\text{PR})^{-1} + (\text{PQ})^{-1}$, and we can write criterion (5) as

$$f(\mathbf{S}, \mathbf{T}, \mathbf{U}) = w_1 \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} \tilde{g}_{pqr}^4 - w_1 \frac{\gamma_1}{QR} \sum_{p=1}^{P}\left(\sum_{q=1}^{Q}\sum_{r=1}^{R}\tilde{g}_{pqr}^2\right)^2 + w_2 \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} \tilde{g}_{pqr}^4$$

$$- w_2 \frac{\gamma_2}{PR}\sum_{q=1}^{Q}\left(\sum_{p=1}^{P}\sum_{r=1}^{R}\tilde{g}_{pqr}^2\right)^2 + w_3 \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R}\tilde{g}_{pqr}^4 - w_3 \frac{\gamma_3}{PQ}\sum_{r=1}^{R}\left(\sum_{p=1}^{P}\sum_{q=1}^{Q}\tilde{g}_{pqr}^2\right)^2$$

$$= (w_1 + w_2 + w_3) \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R}\tilde{g}_{pqr}^4 - \frac{\gamma_1}{QR} w_1 \sum_{p=1}^{P}\left(\sum_{q=1}^{Q}\sum_{r=1}^{R}\tilde{g}_{pqr}^2\right)^2$$

$$- \frac{\gamma_2}{PR} w_2 \sum_{q=1}^{Q}\left(\sum_{p=1}^{P}\sum_{r=1}^{R}\tilde{g}_{pqr}^2\right)^2 - \frac{\gamma_3}{PQ} w_3 \sum_{r=1}^{R}\left(\sum_{p=1}^{P}\sum_{q=1}^{Q}\tilde{g}_{pqr}^2\right)^2$$

$$= (QR)^{-1}\,\text{ORMAX}\left(\begin{pmatrix}\tilde{\mathbf{G}}_1' \\ \cdot \\ \cdot \\ \tilde{\mathbf{G}}_R'\end{pmatrix}, \tilde{\gamma}_1\right) + (PR)^{-1}\,\text{ORMAX}\left(\begin{pmatrix}\tilde{\mathbf{G}}_1 \\ \cdot \\ \cdot \\ \tilde{\mathbf{G}}_R\end{pmatrix}, \tilde{\gamma}_2\right)$$

$$+ (PQ)^{-1}\,\text{ORMAX}((\text{Vec}(\tilde{\mathbf{G}}_1) \ldots \text{Vec}(\tilde{\mathbf{G}}_R)), \tilde{\gamma}_3),$$

where $\tilde{\gamma}_1 \equiv \gamma_1 w_1 QR$, $\tilde{\gamma}_2 \equiv \gamma_2 w_2 PR$ and $\tilde{\gamma}_3 \equiv \gamma_3 w_3 PQ$. It is now clear that the Three-mode Orthomax criterion is governed only by the three parameters $\tilde{\gamma}_1$, $\tilde{\gamma}_2$ and $\tilde{\gamma}_3$. Hence, one may fix the weights to arbitrary nonzero values without loss of flexibility, because, with fixed nonzero weights, any parameter set $\{\tilde{\gamma}_1, \tilde{\gamma}_2, \tilde{\gamma}_3\}$ can still be obtained by appropriate choices of γ_1, γ_2 and γ_3. We therefore choose to take $w_1 = (QR)^{-1}$, $w_2 = (PR)^{-1}$ and $w_3 = (PQ)^{-1}$ in the sequel (unless specified otherwise), and choose different criteria by only varying the γ parameters.

It should be noted that fixing the weights as above is not always useful. For instance, if we wish to obtain the criterion corresponding to a particular set of values for w_1, w_2, γ_1, γ_2 and γ_3 with $w_3 = 0$ in (5), we can indeed get this, alternatively, by taking $w_1^* = (QR)^{-1}$, $w_2^* = (PR)^{-1}$ and $w_3^* = (PQ)^{-1}$, and $\gamma_1^* = \gamma_1 w_1 QR/c$, $\gamma_2^* = \gamma_2 w_2 PR/c$ and $\gamma_3^* = 0$, where $c \equiv (w_1 + w_2)/((QR)^{-1} + (PR)^{-1} + (PQ)^{-1})$. However, this is not a very insightful way of defining a criterion that takes into account only orthomax simplicity in the first two modes, and not in the third.

Maximization of the Three-Mode Orthomax Criterion

To maximize (5) over **S**, **T** and **U**, it is proposed here to alternately maximize (5) over **S** while keeping **T** and **U** fixed, over **T** while keeping **S** and **U** fixed, and over **U** while keeping **S** and **T** fixed. Due to the symmetry of the problem, these three maximization problems are fully equivalent. In fact, the procedure for maximization over **S** can be used for maximization over **T** and over **U** as well, after appropriately permuting the current core array, and permuting the values for the weights and the γ parameters accordingly. Therefore, in the present section, we only consider the problem of maximizing

$$f(\mathbf{S},*) = w_1\text{ORMAX}\left(\begin{pmatrix}\mathbf{H}_1'\mathbf{S}' \\ \cdot \\ \cdot \\ \mathbf{H}_R'\mathbf{S}'\end{pmatrix}, \gamma_1\right) + w_2\text{ORMAX}\left(\begin{pmatrix}\mathbf{SH}_1 \\ \cdot \\ \cdot \\ \mathbf{SH}_R\end{pmatrix}, \gamma_2\right)$$

$$+ w_3\text{ORMAX}((\text{Vec}(\mathbf{SH}_1) \ldots \text{Vec}(\mathbf{SH}_R)), \gamma_3), \quad (6)$$

where **H** denotes **G** multiplied by the current **T** and **U** matrices. We will elaborate each of the terms in (6), but before doing so, we rewrite the ORMAX function (1) as

$$\text{ORMAX}(\tilde{\boldsymbol{\Lambda}}, \gamma) = \sum_{i=1}^{m} \sum_{l=1}^{r} \tilde{\lambda}_{il}^4 - \frac{\gamma}{m} \sum_{l=1}^{r} (\tilde{\boldsymbol{\lambda}}_l' \tilde{\boldsymbol{\lambda}}_l)^2, \tag{7}$$

where the first part simply gives the sum of the fourth powers of all elements in $\tilde{\boldsymbol{\Lambda}}$, and the second term is expressed in terms of the vectors representing the columns $\tilde{\boldsymbol{\lambda}}_l$ ($l = 1, \ldots, r$) of $\tilde{\boldsymbol{\Lambda}}$. Now, we define $\mathbf{H}_F = (\mathbf{H}_1 \mid \ldots \mid \mathbf{H}_R)$, the supermatrix with frontal planes of **H** next to each other, and elaborate the first term of (6) as

$$\text{ORMAX}\left(\begin{pmatrix} \mathbf{H}_1'\mathbf{S}' \\ \cdot \\ \cdot \\ \mathbf{H}_R'\mathbf{S}' \end{pmatrix}, \gamma_1\right) = \text{ORMAX}(\mathbf{H}_F'\mathbf{S}', \gamma_1) = \sum_{q=1}^{Q} \sum_{r=1}^{R} \sum_{l=1}^{P} \left(\sum_{p=1}^{P} h_{pqr} s_{lp}\right)^4$$

$$- \frac{\gamma_1}{QR} \left(\sum_{p=1}^{P} (\mathbf{s}_p' \mathbf{H}_F \mathbf{H}_F' \mathbf{s}_p)^2\right), \tag{8}$$

where $\mathbf{s}_p'$ is the p-th row of **S**. Writing $\mathbf{h}_q^r$ for the q-th column of $\mathbf{H}_r$, we can elaborate the second term of (6) as

$$\text{ORMAX}\left(\begin{pmatrix} \mathbf{SH}_1 \\ \cdot \\ \cdot \\ \mathbf{SH}_R \end{pmatrix}, \gamma_2\right) = \sum_{i=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} \left(\sum_{p=1}^{P} s_{ip} h_{pqr}\right)^4 - \frac{\gamma_2}{PR} \left(\sum_{q=1}^{Q} \left(\sum_{r=1}^{R} \mathbf{h}_q'^r \mathbf{S}'\mathbf{S}\mathbf{h}_q^r\right)^2\right)$$

$$= \sum_{l=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} \left(\sum_{p=1}^{P} s_{lp} h_{pqr}\right)^4 - \text{constant}, \tag{9}$$

where the second line of (9) is obtained by replacing the index i by l in the first term and by noting that the second term is constant because $\mathbf{S}'\mathbf{S} = \mathbf{I}$. Clearly the first term in (9) is equal to the first term in (8). Finally, the third term of (6) can be elaborated as

$$\text{ORMAX}((\text{Vec}(\mathbf{SH}_1) \ldots \text{Vec}(\mathbf{SH}_R)), \gamma_3) = \sum_{i=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} \left(\sum_{p=1}^{P} s_{ip} h_{pqr}\right)^4$$

$$- \frac{\gamma_3}{PQ} \sum_{r=1}^{R} ((\text{Vec}(\mathbf{SH}_r))'(\text{Vec}(\mathbf{SH}_r)))^2 = \sum_{l=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} \left(\sum_{p=1}^{P} s_{lp} h_{pqr}\right)^4$$

$$- \frac{\gamma_3}{PQ} \sum_{r=1}^{R} (\text{tr}(\mathbf{SH}_r)'(\mathbf{SH}_r))^2 = \sum_{l=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} \left(\sum_{p=1}^{P} s_{lp} h_{pqr}\right)^4 - \text{constant}, \tag{10}$$

where the first term is again obtained via re-indexing, and it is again used that $\mathbf{S}'\mathbf{S} = \mathbf{I}$ to show that the second term is constant. Combining (8), (9) and (10) with (6), we obtain

$$\mathrm{f}(\mathbf{S},{}^*) = (w_1 + w_2 + w_3) \sum_{q=1}^{Q} \sum_{r=1}^{R} \sum_{l=1}^{P} \left(\sum_{p=1}^{P} h_{pqr} s_{lp} \right)^4 - w_1 \frac{\gamma_1}{QR} \left(\sum_{p=1}^{P} (\mathbf{s}_p' \mathbf{H}_F \mathbf{H}_F' \mathbf{s}_p)^2 \right) + c, \tag{11}$$

where c denotes a constant. Hence we have to maximize

$$\tilde{\mathrm{f}}(\mathbf{S}) = \sum_{q=1}^{Q} \sum_{r=1}^{R} \sum_{l=1}^{P} \left(\sum_{p=1}^{P} h_{pqr} s_{lp} \right)^4 - w_1 \frac{\gamma_1}{QR} (w_1 + w_2 + w_3)^{-1} \sum_{p=1}^{P} (\mathbf{s}_p' \mathbf{H}_F \mathbf{H}_F' \mathbf{s}_p)^2$$

$$= \mathrm{ORMAX}(\mathbf{H}_F' \mathbf{S}', \gamma_1 w_1 (w_1 + w_2 + w_3)^{-1}). \tag{12}$$

It follows that the problem of maximizing (6) over **S** is equivalent to orthomax applied to the supermatrix $\mathbf{H}_F'$ with the γ parameter chosen as $\gamma_1 w_1 (w_1 + w_2 + w_3)^{-1}$. Hence, for updating **S** we do not have to devise a new algorithm, but we can use any orthomax algorithm available, for instance the planar algorithm proposed by Jennrich (1970), which updates two columns of $\mathbf{S}'$ at a time, and iterates over every pair of columns. Cycling through all pairs of columns the algorithm increases the orthomax function value monotonically until it stabilizes. Here, we used Clarkson and Jennrich's (1988, p. 255) formula (15) for efficient computations of the planar updates, supplemented with a procedure for suppressing permutations (see ten Berge, 1995).

Above, we have described how we can maximize f over **S**, while keeping **T** and **U** fixed. As mentioned above, maximizing f over **T** and over **U**, respectively, can be done by applying the procedure for maximizing f over **S**, after permuting the three-way array such that **S** takes the role of **T** and of **U**, respectively; the weights and the γ parameters must be permuted accordingly. After having updated **S**, **T** and **U**, we evaluate the function value. If it has increased, we start another round of updatings, and we continue this procedure until the function value stabilizes. It should be noted that the above sketched algorithm contains a major cycle (updating **S**, **T** and **U**, respectively), and many inner cycles (planar rotations for updating **S**, **T** or **U**). Since each planar rotation increases the function value, one could, for each update of **S** (and similarly of **T** and **U**), use only one complete cycle of planar rotations, rather than as many planar rotations as are necessary for convergence. This will not affect the monotonical convergence of the algorithm, and may increase the efficiency of the algorithm.

To consider some special choices of the γ parameters, we first consider the choices $\gamma_1 = \gamma_2 = \gamma_3 = 1$, which we denote as "Three-mode Varimax" because with these choices (5) combines three varimax functions. It should be noted that the algorithm for maximizing (5) *does not* reduce to alternately applying varimax to certain supermatrices, as one might have expected intuitively. Instead, the algorithm alternates over three orthomax procedures with γ taken as P/(P + Q + R), Q/(P + Q + R) and R/(P + Q + R), respectively. On the other hand, choosing γ_1 = (P + Q + R)/P, γ_2 = (P + Q + R)/Q and γ_3 = (P + Q + R)/R does lead to an algorithm that alternates over three varimax procedures (because the orthomax parameter in (12) then equals 1), which are applied to certain supermatrices consisting of rotated elements of the core. We call this method "Alternating Varimax" here.

In certain cases it is not desirable to rotate the core in all three directions. For instance, it may be useful to rotate the matrix **A** in the Three-mode PCA solution to simple structure, which implies that the matrix **S** must remain fixed. In such cases the function f must be maximized over **T** and **U** only, which can be done by updating only **T** and **U**. Thus, besides flexibility in the choice of the γ parameters, a second kind of flexibility emerges: One may decide to rotate the core in any combination of directions, corresponding to rotation of the component matrices **A**, **B** and **C**. In fact, Murakami's (1983) procedure is the case where only the matrix **B** is rotated (hence **S** and **U** are kept fixed) and where two of the three γ parameters are fixed to zero.

Performance of Three-Mode Orthomax

Above, we have described a class of procedures for Three-mode Orthomax rotation. In the present section we study the performance of the methods in this class. For this purpose we programmed the general Three-mode Orthomax approach in PCMATLAB—the program ORTHMAX3 is available from the author—and applied our procedures to an empirically obtained (three-way) core array, as well as to contrived core arrays. As with many iterative algorithms, the procedures may miss the global maximum of the criterion function. This is because one set of parameters is updated given another set; at particular points, it is possible that a set of parameters can no longer be updated given the others, even though jointly updating the parameters sets would increase the function value. This well-known phenomenon for iterative procedures is usually dealt with by choosing rational starting positions, or using a number of random starts and taking the best solution as "the solution". In the present paper, we choose to combine these procedures: The algorithms were started with $\mathbf{S} = \mathbf{I}$, $\mathbf{T} = \mathbf{I}$ and $\mathbf{U} = \mathbf{I}$ (which for various applications may be considered "rational"), as well as with five sets of random starts (i.e., starts for **S**, **T** and **U** based on orthonormalized versions of matrices with elements drawn randomly from the standard normal distribution). The iterative process was stopped whenever the function value changed by less than .0001%. The solutions reported are the best of the six runs.

Meaudret Data: An Example of Rotation of all Three Modes

The first data set to be analyzed here is the $6 \times 10 \times 4$ data set with measurements on the river Meaudret (Doledec & Chessel, 1987; see Kiers, 1991). We used the same preprocessing method as Kiers (1991) and performed a Three-mode PCA with $P = Q = R = 3$. As mentioned by Kiers (p. 466) this solution explains considerably more of the inertia (77.6%) than does PARAFAC/CANDECOMP (73.5%). Therefore, the core matrix apparently involves important interactions between components that are not covered by the PARAFAC model. Hence, it will not be possible to rotate the core to a nearly superdiagonal core (which corresponds to the PARAFAC model): Some nonnegligible off-superdiagonal elements will remain.

To obtain the Three-mode PCA solution, we used the Kroonenberg, ten Berge, Brouwer and Kiers (1989) Gram-Schmidt based algorithm, which finds a principal axes orientation for the component matrices (see p. 84). This implies that the component matrix **A** contains eigenvectors of $\Sigma_k \hat{\mathbf{X}}_k\hat{\mathbf{X}}_k'$ (where $\hat{\mathbf{X}}_k$ denotes the k-th frontal plane of $\hat{\mathbf{X}}$), and **B** and **C** contain eigenvectors of analogously derived matrices. The core obtained with this "unrotated" solution is given in the first panel of Table 1. It can be seen that this core shows some simplicity (it has six elements higher than .5 (bold face) and 12 elements smaller than .2 (small font), in the absolute sense), but further simplicity is desirable. To find a simpler core, we applied the Three-mode Orthomax rotations described in the present paper. Specifically, we obtained two solutions by Three-mode Quartimax (Three-mode Orthomax with $\gamma_1 = \gamma_2 = \gamma_3 = 0$) and Three-mode Varimax ($\gamma_1 = \gamma_2 = \gamma_3 = 1$), respectively. The Three-mode Quartimax and Three-mode Varimax rotated cores are reported in the second and third panels of Table 1. It can be seen that these two solutions are very similar to each other and considerably more simple than the unrotated core in that they have more "extreme" and fewer medium sized values.

We will study the Three-mode Varimax solution in a little more detail. Inspecting the core elements, we see that there are five elements higher than .5 (four of which are even higher than .8) and 20 smaller than .2, in the absolute sense. Only two elements are

TABLE 1

Unrotated and Varimax Rotated 3×3×3 Core Arrays for the Meaudret Data

Unrotated Core Array (Frontal Planes)

	C1			C2			C3		
	B1	B2	B3	B1	B2	B3	B1	B2	B3
A1	**2.07**	0.06	0.21	0.07	0.14	–0.26	0.04	0.34	–0.25
A2	-0.19	**-0.60**	–0.39	**0.99**	0.08	**-0.60**	–0.38	**0.62**	–0.36
A3	-0.12	**0.59**	0.29	0.13	-0.05	-0.13	-0.14	0.22	-0.16

Three-mode Quartimax Rotated Core Array (Frontal Planes)

	C1			C2			C3		
	B1	B2	B3	B1	B2	B3	B1	B2	B3
A1	**2.11**	0.00	-0.04	0.00	0.34	-0.06	0.01	0.00	0.00
A2	0.03	-0.03	0.40	0.14	**0.96**	-0.06	**0.98**	0.13	**0.58**
A3	0.00	-0.15	**0.83**	0.11	0.04	-0.02	0.15	0.11	0.07

Three-mode Varimax Rotated Core Array (Frontal Planes)

	C1			C2			C3		
	B1	B2	B3	B1	B2	B3	B1	B2	B3
A1	**2.11**	0.01	-0.03	0.00	0.33	-0.06	0.01	0.01	0.00
A2	0.03	-0.02	0.38	0.14	**0.97**	-0.06	**0.97**	0.13	**0.59**
A3	0.00	-0.14	**0.84**	0.12	0.06	-0.02	0.17	0.11	0.09

"Orthomax with γ_1=γ_2=γ_3=6" Rotated Core Array (Frontal Planes)

	C1			C2			C3		
	B1	B2	B3	B1	B2	B3	B1	B2	B3
A1	**2.10**	0.05	0.05	0.00	0.30	-0.07	0.04	-0.02	-0.02
A2	0.05	0.00	0.26	0.16	**0.97**	-0.06	**0.90**	0.10	**0.63**
A3	0.01	-0.12	**0.88**	0.14	0.20	-0.02	0.27	0.12	0.20

somewhat moderate in size. Clearly, this solution is much simpler than the unrotated one, that had nine elements of moderate size. Upon comparison of the Three-mode Quartimax and Varimax solutions, we can see that in the Three-mode Varimax solution the two medium sized core elements are somewhat smaller than in the Three-mode Quartimax solution. To see if taking larger values for γ_1, γ_2 and γ_3 will strengthen this effect, we also computed the Three-mode Orthomax solution with $\gamma_1 = \gamma_2 = \gamma_3 = 3$ ("Alternating Varimax Rotations"), and with $\gamma_1 = \gamma_2 = \gamma_3 = 6$. We only report the latter solution, and note that the Alternating Varimax solution was intermediate between the Three-mode Varimax solution and the solution for $\gamma_1 = \gamma_2 = \gamma_3 = 6$. It can be seen that the elements that were medium sized in the Three-mode Varimax solution indeed become smaller in the Three-mode Orthomax with $\gamma_1 = \gamma_2 = \gamma_3 = 6$ solution. However, in the latter solutions some elements that were small in the Three-mode Varimax solution, now have grown to medium size. For this reason, for the present data set, we prefer the Three-mode Varimax solution.

The full results of the Three-mode PCA with Three-mode Varimax rotation of the core can now be interpreted by studying the most important interactions of the three sets of components. Hence, we should focus on the five interactions with highest core elements ((A1, B1, C1), (A3, B3, C1), (A2, B2, C2), (A2, B1, C3), and (A2, B3, C3)), pay some attention to the interactions (A2, B3, C1) and (A1, B2, C2), and we may ignore the other terms; here "A1", "B2", etc. indicate Component 1 for mode A, Component 2 for mode B, etc. To give a complete, substantive interpretation, we study the component matrices as well. These are given in Table 2 (for the labels we refer to Kiers, 1991). We only consider the interpretation of the five most important interaction terms. From Table 2 we see that A1 can be interpreted as contrasting Station 2 to the others, B1 indicates the pollutedness of the water and C1 has high values for the second part of the year (say Fall). Thus, the term (A1, B1, C1) describes that especially in the Fall, Station 2 is relatively more polluted than the other stations. This contribution is very similar to the contribution of the first ORTCP-A dimension, reported by Kiers (1991, p. 466). Component A3 contrasts the midstream Stations 3, 4 and 5 to Stations 2 and 6, B3 measures mainly NO_3. Hence, the second term, (A3, B3, C1), represents the relatively large NO_3 concentrations at Stations 3, 4 and 5 in the Fall. Component A2 is a general component related to all stations, B2 mainly represents the temperature, and C2 contrasts Summer versus Winter. So the term (A2, B2, C2) represents the (quite obvious) fact that temperature is higher in Summer than in Winter, and that this holds for all stations. C3 describes a gradual decrease from August, via November and February to June, so the term (A2, B1, C3) indicates that, at all stations, the water is most polluted in August, gradually gets better in November and February, and is least polluted in June, for all stations. Finally, (A2, B3, C3) represents the same cyclic decrease for the NO_3 concentrations (again for all stations).

Above, the Three-mode Varimax rotated core has been compared to the unrotated core and to some other Three-mode Orthomax rotated cores. An alternative (and rather common) procedure is not to rotate the core to simple structure, but to rotate the matrices **A**, **B** and **C** to simple structure. This approach has the advantage that it aims at easily interpretable dimensions, but has the disadvantage that the core may not be simple, and that hence many interactions may have to be taken into account. For the present data set, the solution resulting from varimax of **A**, **B** and **C** is given in Table 3. It can be seen that the component matrices **C** and to some extent **A** are indeed more simple than those associated with the Three-mode Varimax rotated core, but, the core is considerably more complex, and many more interactions need to be interpreted than in the Three-mode Varimax rotated solution. Moreover, the matrices **B**, which play the most important role in the substantive interpretation, hardly differ (except for a permutation). Furthermore, it is doubtful if the varimax rotated version of **C** is more easily interpretable than the one

TABLE 2

Rotated Component Matrices for the Three-mode Varimax Rotated Solution of the Meaudret Data

Matrix **A** (Measurement Stations)			
Station 1	−.19	.33	−.14
Station 2	.89	.24	−.31
Station 3	.23	.40	.56
Station 4	−.03	.48	.35
Station 5	−.20	.39	.20
Station 6	−.28	.55	−.63
Matrix **B** (Chemical Variables)			
Temp.	−.01	.91	.13
Flux	−.13	−.12	−.31
PH	−.31	−.12	−.17
Cond.	.34	−.20	.25
O_2	−.33	−.24	.01
BOD	.41	−.05	−.15
COD	.42	−.05	−.22
NH_4	.41	.02	−.04
NO_3	−.11	−.15	.82
PO_4	.36	−.10	.23
Matrix **C** (Occasions)			
June	.21	.48	−.83
August	.58	.60	.41
November	.78	−.51	−.02
February	.13	−.39	−.38

reported in Table 2, which showed an interesting contrast between Summer and Winter and a cyclic seasonal effect. It can be concluded that, in the present example, the loss of simplicity in **A**, **B** and **C**, seems to be amply compensated by a considerable gain in simplicity of the core using Three-mode Varimax.

TABLE 3

Solution for the Meaudret Data After Varimax Rotation Component Matrices

Matrix A (Measurement Stations)			
Station 1	−.06	.13	.37
Station 2	.97	.02	.04
Station 3	.12	.69	−.19
Station 4	−.04	.58	.09
Station 5	−.17	.40	.20
Station 6	−.05	−.05	.88
Matrix B (Chemical Variables)			
Temp.	−.01	−.01	.92
Flux	−.16	−.28	−.17
PH	−.33	−.12	−.14
Cond.	.37	.24	−.16
O_2	−.33	.08	−.23
BOD	.40	−.19	−.07
COD	.40	−.25	−.08
NH_4	.40	−.09	.02
NO_3	−.02	.84	−.02
PO_4	.38	.20	−.06
Matrix C (Occasions)			
June	−.03	.03	.98
August	.08	.92	.04
November	.92	.08	−.04
February	.37	−.38	.18

Associated Core Array (Frontal Planes)

	C1			C2			C3	
B1	B2	B3	B1	B2	B3	B1	B2	B3
1.55	−0.27	−0.27	**1.12**	−0.28	0.33	0.27	−0.22	0.21
0.03	**0.76**	−0.44	**0.68**	**0.51**	**0.63**	**−0.61**	−0.23	0.13
−0.69	−0.16	−0.25	−0.15	−0.05	0.36	**−0.64**	−0.32	0.17

Artificial Core Array 1

In the above example, a simple core was obtained by various Three-mode Orthomax procedures described in the present paper. One may wonder to what extent similar results could have been found by using other procedures, such as Kiers' (1992) rotation to su-

TABLE 4

Superdiagonalization of Artificial Core Array 1

Unrotated Core (Frontal Planes)								
1	0	0	0	0	0	0	0	0
0	0	0	0	0	1	1	0	0
0	1	0	0	0	0	0	0	0
Superdiagonalized Core (Frontal Planes)								
.87	.00	.00	.00	−.50	.00	−.50	.00	.00
−.50	.00	.00	.00	.87	.00	−.87	.00	.00
.00	.00	−.50	.00	.00	.00	.00	.00	.87

perdiagonality. In fact, for this data set it turned out that the results obtained by rotation to superdiagonality were quite similar to those found by Three-mode Varimax. To see if this is always the case, some very simple artificial core arrays were constructed. One of these is reported in the first panel of Table 4. Applying Kiers' superdiagonalization procedure to this core resulted in the rotated core displayed in the second panel. It can be seen that the core has in fact become considerably more complex than it was before (it has eight nonnegligible elements rather than the four it had before). Instead, the Three-mode Varimax algorithm applied to this core left the core completely unchanged, thus respecting the given simplicity of this core. The same happened for some other Three-mode Orthomax variants tested. This illustrates that superdiagonalization is not guaranteed to yield a simple core, even if there is one. It certainly will, if the simple core is superdiagonal, but there is no reason to expect that it will be simple as soon as full superdiagonality cannot be attained. The Three-mode Orthomax procedures described in the present paper perform better in that respect.

Artificial Core Array 2

For the two-way case, it has been argued that varimax might be preferred over quartimax, because quartimax has a tendency to yield a "general factor". To see if this preference generalizes to the three-way case, we constructed a second artificial core matrix (reported in Table 5), which had a first slab that, in the two-way case, would lead to a general factor when rotated by quartimax. When applying Three-mode Quartimax to this array, it turned out that no general factor appeared. That is, in the case where rotations in all three modes were performed, the general factor phenomenon was not observed.

In practice, we do not always rotate over all three modes. Therefore, it is interesting to see if rotation over less than three modes sometimes leads to a general factor. As an example, the core in Table 5 was rotated over **T** and **U** only. It turned out that Three-mode Quartimax and Varimax left the core unchanged (thus yielding the "general factor" present in this core). On the other hand, the Alternating Varimax approach (Three-mode Orthomax with $\gamma_1 = 2$, $\gamma_2 = 4$ and $\gamma_3 = 4$) did rotate the core to a solution without a

TABLE 5

Artificial Core Array 2

Unrotated Core (Frontal Planes)			
.8	.3	.16	.06
.8	.3	−.16	−.06
.8	−.3	−.16	.06
.8	−.3	.16	−.06
Alternating Varimax Rotated Core (Frontal Planes)			
.35	.78	−.07	−.16
.35	.78	.07	.16
.78	.35	.16	.07
.78	.35	−.16	−.07

general factor (see second panel of Table 5). It can be concluded that, at least when less than three modes are rotated, Three-mode Quartimax and Varimax may find a general factor (in a similar way as in the two-way case). The results also suggest that Alternating Varimax is more similar to ordinary two-way varimax than is Three-mode Varimax itself.

The present example has shown that Three-mode Orthomax results in certain cases depend strongly on the choice for γ_1, γ_2 and γ_3. To see how different the outcomes of the Three-mode Orthomax approaches with some different choices for γ_1, γ_2 and γ_3 are "in practice", we conducted a small simulation study, which is reported in the next subsection.

A Small Simulation Study: Empirical Differences Between Four Approaches for Three-mode Orthomax

To gain some experience with the Three-mode Orthomax procedures described in the present paper, we constructed 40 $3 \times 3 \times 3$ test core arrays in the following way. Each core was, to some extent, based on the simple core with frontal planes

$$\mathbf{H}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \mathbf{H}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } \mathbf{H}_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{13}$$

Next, to obtain a test core array, the above simple core was rotated by random orthonormal matrices **S**, **T** and **U**, and to this rotated core we added a random $3 \times 3 \times 3$ array. For construction of the first 20 test core arrays, the elements of the random array were drawn from N(0, .1); the last 20 test core arrays were based on the same structural and random parts, but the random part was now multiplied by $(10)^{1/2}$ before it was added to the structural part. The first 20 cores are denoted as "Low Noise Core Arrays" (the structural part has sum of squares 5, whereas the random part has an expected sum of squares of 2.7), the others are called "High Noise Core Arrays" (the random part has an expected sum of squares of 27). The first set of test arrays was constructed to simulate situations where a simple core is basically available. The second set of arrays simulates situations where the a priori simple structure is very weak. Thus the latter arrays can be considered as almost random.

TABLE 6

Frequencies of Finding Global Optima (out of 200) for Each Condition in the Simulation Study

	Three-mode Quartimax	Three-mode Varimax	Alternating Varimax	Orthomax with γ=6
Low Noise	187	187	183	188
High Noise	184	184	177	179

The 40 test core arrays were rotated over **S**, **T** and **U**, by four Three-mode Orthomax procedures: Three-mode Quartimax, Three-mode Varimax, Alternating Varimax (Orthomax with $\gamma_1 = \gamma_2 = \gamma_3 = 3$), and Three-mode Orthomax with $\gamma_1 = \gamma_2 = \gamma_3 = 6$. Each analysis involved ten randomly started runs. In each run, the algorithm was considered converged when the function values of consecutive iterations differed less than .0001%. The best of the ten runs was considered to yield the globally optimal solution. A solution was considered to have missed the global optimum if the function value was more than .1% smaller than the (alleged) global optimum.

For each analysis, we checked how many runs missed the global optimum. In Table 6, total frequencies of finding the global optimum for the four methods, for the two sets of 20 core arrays are reported. Considering that each cell pertains to 200 runs, it can be concluded that the methods do not miss the global optimum frequently. Although for the present data we do not know the global optima, the fact that the best value was always found at least four times (in ten runs), and usually more often (on average, it was found in more than nine out of ten runs), suggests that in most cases the global optimum has indeed been found, and that using ten random starts will usually be more than sufficient to find the global optimum.

We also kept track of the computation times of each series of ten runs for rotating one test array. Because the procedures have all been programmed in PCMATLAB, the algorithms are not programmed as efficient as they could be. Computation times (with a Pentium 100 MHz processor) can therefore be considered an upper bound. The computation times ranged from 5 to 85 seconds. On average, Three-mode Quartimax used 15.3 seconds, Three-mode Varimax 19.6, Alternating Varimax 21.4 and Three-mode Orthomax with $\gamma_1 = \gamma_2 = \gamma_3 = 6$ used 27.5 seconds for ten runs. Thus, computation times tend to increase with increasing γ's, but they never become prohibitive. Computation times did not seem to differ systematically for High Noise data and for Low Noise data.

The main question to be answered by our simulation was to what extent solutions of the different methods differ in practice. To measure agreement among solutions, we compared the matrices **S**, **T** and **U** obtained from the last four analyses with those obtained in the first (Three-mode Quartimax) analysis. The rows of the matrices were permuted such that they optimally matched those from the Three-mode Quartimax solution. Next, Tucker's (1951) phi coefficients (which in this case reduce to simple scalar products) were computed between corresponding rows. The nine phi coefficients for matrices **S**, **T** and **U** were next summed to form the overall agreement criterion. The overall agreement values are reported in Table 7, averaged within the two sets of 20 core arrays. Because a value of

TABLE 7

Averaged Values of Overall Agreement with the Three-mode Quartimax Solution

	Three-mode Varimax	Alternating Varimax	Orthomax with γ=6
Low Noise	9.00	8.91	8.92
High Noise	9.00	8.98	8.79

9 indicates perfect agreement, it can be concluded that the solutions of Three-mode Varimax and of Alternating Varimax are, on average, very similar to that of Three-mode Quartimax; the "Three-mode Orthomax with $\gamma_1 = \gamma_2 = \gamma_3 = 6$" solutions tend to deviate a little more from the Three-mode Quartimax solutions. Upon inspection of the actual agreement values, it turned out that the differences were mainly caused by a few solutions: For Three-mode Varimax all values exceeded 8.95, Alternating Varimax gave only three solutions with agreement value below 8.85 (which corresponds to an average phi-coefficient of more than .98, and a minimum phi-coefficient of .85), and Three-mode Orthomax with $\gamma_1 = \gamma_2 = \gamma_3 = 6$ gave 8 such solutions. All other solutions can be considered virtually equal to the Three-mode Quartimax and Varimax solutions. It should be noted that the solutions with agreement values below 8.85 may still be very similar to the Three-mode Quartimax solution. Differences become more appreciable when the agreement value drops below 8.0, as happened only once for Alternating Varimax and four times for Three-mode Orthomax with $\gamma_1 = \gamma_2 = \gamma_3 = 6$. It can be concluded that Three-mode Orthomax and Varimax yielded virtually the same results throughout, Alternating Varimax gave deviating outcomes occasionally, and large differences were observed most often when using Three-mode Orthomax with $\gamma_1 = \gamma_2 = \gamma_3 = 6$.

As far as the technical aspects are concerned, we may conclude that the four Three-mode Orthomax procedures tested here do not miss the global optimum frequently, and are not prohibitively slow (even ten runs for an ordinary sized core array usually takes less than one minute computation time). As far as differences in results are concerned, it can be concluded that the first three approaches (based on γ's in the range 0 to 3) usually give very similar solutions, although exceptions do exist. With $\gamma_1 = \gamma_2 = \gamma_3 = 6$, considerable differences were found in four out of 40 times. Therefore, it is advised to use more than one procedure for finding a simpler core, especially in cases where the core rotation procedure has not led to desirable results.

Application to Two-Way Arrays: Implications for Simple Structure Rotation of Matrices

Above, we have focused on rotation of a three-way array to simple structure. Obviously, a matrix can be seen as a three-way array with only one frontal slab. In the present section, we will study what happens when Three-mode Orthomax is applied to a matrix. In particular, it will be studied to what extent Three-mode Orthomax has ordinary orthomax as a special case.

Simple structure rotation of a (factor loading) matrix usually pertains to rotation in *only one* direction. That is, the $(m \times r)$ matrix $\mathbf{\Lambda}$ is postmultiplied by a rotation matrix $\mathbf{T}$

to obtain a "simple" matrix $\tilde{\Lambda} = \Lambda\mathbf{T}$. We may represent the matrix by an $m \times r \times 1$ three-way array $\mathbf{G}$ with a single frontal plane $\mathbf{G}_1 = \Lambda$, and the rotated matrix by the three-way array with frontal plane $\tilde{\mathbf{G}}_1 = \Lambda\mathbf{T}$. We will now study what happens if the Three-mode Orthomax procedure is applied to such a core. To make the procedure optimally similar to the two-way case, we only rotate over $\mathbf{T}$, and keep $\mathbf{S}$ and $\mathbf{U}$ fixed to identity matrices. For the thus rotated core array the Three-mode Orthomax criterion (5) reduces to

$$f(\mathbf{T}) = w_1\text{ORMAX}(\mathbf{T}'\Lambda', \gamma_1) + w_2\text{ORMAX}(\Lambda\mathbf{T}, \gamma_2) + w_3\text{ORMAX}(\text{Vec}(\Lambda\mathbf{T}), \gamma_3)$$

$$= (w_1 + w_2 + w_3)\text{ORMAX}(\Lambda\mathbf{T}, \gamma_2 w_2(w_1 + w_2 + w_3)^{-1}) + \text{constant}, \quad (14)$$

where the second line is derived analogously to (11); the weights are now taken as $w_1 = 1/m$, $w_2 = 1/r$, and $w_3 = 1/mr$, unless specified otherwise. Clearly, (14) is proportional to the ordinary orthomax function with γ parameter $\gamma_2 w_2(w_1 + w_2 + w_3)^{-1}$, which shows that orthomax is indeed a special case of Three-mode Orthomax. However, this is not the only interesting special case of the present application of Three-mode Orthomax.

Some other interesting special cases of (14) are the following:

1. "Three-mode Quartimax" ($\gamma_1 = \gamma_2 = \gamma_3 = 0$) leads to ordinary quartimax. In fact, as soon as $\gamma_2 = 0$, any choice for the other parameters leads to ordinary quartimax.
2. "Three-mode Varimax" ($\gamma_1 = \gamma_2 = \gamma_3 = 1$) does not lead to ordinary varimax, but to orthomax with $\gamma = m/(m + r + 1)$. To find ordinary varimax, one should choose $\gamma_2 = w_2^{-1}(w_1 + w_2 + w_3)$.
3. "Varimax of the rows of Λ" (based on $w_1 = 1/m$, $\gamma_1 = 1$, and $w_2 = w_3 = \gamma_2 = \gamma_3 = 0$) leads to quartimax. Note that, here, we set w_2 and w_3 equal to zero to warrant the interpretation in terms of varimax of the rows (the first mode) only. In fact, Neuhaus and Wrigley (1954) originally described their criterion as maximizing the sum of rowwise variances of squared factor loadings, which was later found to be equivalent to the quartimax criterion.
4. "Varimax of rows and columns of Λ" (based on $w_1 = 1/m$, $w_2 = 1/r$, $\gamma_1 = \gamma_2 = 1$, and $w_3 = \gamma_3 = 0$) leads to orthomax with $\gamma = m/(m + r)$. Thus, orthomax with $\gamma = m/(m + r)$ is provided with a simple and useful interpretation: It maximizes the sum of rowwise and columnwise variances of the squared loadings. This criterion could be seen as an alternative to Saunders' parsimax criterion (see Mulaik, 1972, p. 263), which also aims at maximizing simplicity in terms of rows and columns.

Discussion

The present paper has described a flexible approach for orthogonal rotation of the core in all three directions aimed at simplicity of the core in all three directions. The main purpose of the present paper was to offer the tools for this type of methods. Some attention has been paid to the effect of the different choices. For instance, although it has been observed that different choices for the γ parameters often lead to similar results, it has also been seen that this is not always the case. Similarly, differences between Three-mode Quartimax and Varimax have often been found to be small, but again, this need not always be true. A rather extreme choice for the γ parameters, for instance that of $\gamma = 3$ or $\gamma = 6$, seems to be worth considering as well. The present paper gives some empirical information that should help in making the different choices. On the basis of this information, it is recommended to use the Three-mode Varimax and/or Alternating Varimax procedures to start with. When results are unacceptable, one could use the full flexibility of Three-mode Orthomax.

In practice, often substantive information will help choosing among the different Three-mode Orthomax methods. For instance, if the presence of a general factor can be expected, one could use relatively high values of the γ parameters to avoid it. Or, as another example, consider the situation where it is desirable to keep one (or even two) of the component matrices fixed, because the orientation of the component matrix has been determined by applying a simple structure rotation to the component matrix itself. Here, we encounter a trade-off relation: We can increase simplicity of the core by decreasing the simplicity of the component matrices and vice versa. Simplicity of the component matrices implies that it is easy to interpret the components, whereas simplicity of the core implies that the set of nonnegligible interactions between components from different modes is relatively small. As an example of the latter simplicity, we can view PARAFAC (Harshman, 1970; Harshman & Lundy, 1984a) as the extreme case where the core is very simple, but where there is no possibility of simplifying the component matrices. Which kind of simplicity is to be preferred, and how the choice between these desiderata is to be made depends on the situation at hand, and could, for instance, be determined by studying the stability of the solution, as suggested by Harshman and Lundy.

In the exemplary analyses reported in the present paper, we often used a three-mode version of *raw* varimax. Kaiser (1958) advocated the use of "normalized" varimax, which implies that the varimax criterion is not applied to the loading matrix itself, but to the rowwise normalized loading matrix; the resulting rotation matrix is applied to the original, nonnormalized loading matrix. In our applications, we did not use such a normalization procedure. The reason is that such a normalization would have to be applied to the rows of all three supermatrices to which the orthomax criteria are applied (see (5) and (6)). Such a normalization would involve triple fiberwise normalization of **G** (see Harshman & Lundy, 1984b). No closed-form solutions seem available for such triple normalizations, hence one has to resort to iterative procedures, the outcome of which may have lost all resemblance with the original core array. For this reason, we did not consider this possibility.

In the Three-mode Orthomax methods proposed in the present paper, the orthomax criteria are applied to supermatrices consisting of frontal, lateral or horizontal slabs of the core next to each other. Instead, one could apply the orthomax criterion to a supermatrix consisting of all column fibers of the core, and to a supermatrix consisting of all row fibers, as well as to one consisting of all "tubes" (fibers from front to back of the core array) of the core. Thus a "fiberwise" approach can be constructed by combining the three criteria in a similar way as was done for the original ("slabwise") approach. A planar algorithm can be devised along similar lines as for the slabwise approach, but no longer relies on alternately applying orthomax procedures. Some first test analyses pointed out that the method works reasonably, but the algorithm converges much slower than the original Three-mode Orthomax procedure, and is more prone to missing the global optimum.

As in the two-way case, it is not always desirable to impose orthogonality. In the two-way case, oblique rotations have been proposed to overcome artificial orthogonality of the components. In the three-way case the only attempt in this direction has been made by Kruskal (1988). Despite our focus on orthogonality, the results of the present paper can be used for oblique rotation as well. In analogy to the Promax (Hendrickson & White, 1964) approach, the orthogonal simple structure solution of the core can be used to obtain an oblique simple structure solution. The idea is to form a target core array by taking, for instance, fourth powers of all core elements, and to rotate the original core towards this target core by iterative oblique Procrustes rotations in all three directions. Similarly, the present results could be used to generalize Harris and Kaiser's (1964) orthoblique procedure to the three-way case. In both cases orthogonal rotation procedures are used for obtaining oblique solutions, which have, moreover, proven to be among the best oblique solutions in the two-way case.

As an alternative to *rotation* to simple structure, Kiers (1992) and Rocci (1992) proposed methods to simplify the Three-mode PCA models by constraining certain parameters of the core to zero. One of their problems is to decide which elements should be constrained to zero. The Three-mode Orthomax procedures described in the present paper could be helpful here. A procedure could be to constrain those elements of the core to zero that are small in the Three-mode Orthomax rotated core. The matrices **A**, **B** and **C** that correspond to the Three-mode Orthomax solution could be used as start for the iterative algorithms in Kiers and Rocci's methods.

In the present paper, a generalization of ordinary orthomax to the *three-mode* case has been made. Since three-mode methods have been further generalized towards four-way (Lastovicka, 1981) and even N-way (Kapteyn, Neudecker & Wansbeek, 1986; Polit, 1986) methods, generalizations of the Three-mode Orthomax core rotation procedure to the four-way and N-way case are desired as well. For this purpose the criterion in (5) can be generalized straightforwardly, using N (rather than three) terms consisting of ORMAX applied to supermatrices $\tilde{\mathbf{G}}^1, \ldots, \tilde{\mathbf{G}}^N$ the columns of which refer to each of the respective N modes. An algorithm for maximizing this criterion can be constructed completely analogous to the algorithm for Three-mode Orthomax, and will again consist of alternately applying two-way orthomax rotations to certain supermatrices.

References

Carroll, J. B. (1953). An analytic solution for approximating simple structure in factor analysis. *Psychometrika, 18,* 23–38.

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

Cohen, H. S. (1974, March). *Three-mode rotation to approximate INDSCAL structure (TRIAS).* Paper presented at the Psychometric Society Meeting, Palo Alto.

Cohen, H. S. (1975). *Further thoughts on three-mode rotation to INDSCAL structure, with jackknifed confidence regions for points.* Paper presented at U.S.-Japan seminar on Theory, Methods and Applications of Multidimensional Scaling and Related Techniques, La Jolla.

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

Doledec, S., & Chessel, D. (1987). Rythmes saisonniers et composantes stationnelles en milieu aquatique. I.-Description d'un plan d'observation complet par projection de variables [Seasonal rhythms and stationary components in aquatic environments. I.-Description of a complete observation scheme by projection of variables]. *Acta Oecologia/Ocologia Generalis, 8,* 403–406.

Ferguson, G. A. (1954). The concept of parsimony in factor analysis. *Psychometrika, 19,* 281–290.

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

Harshman, R. A. (1970). *Foundations of the PARAFAC procedure: models and conditions for an "explanatory" multi-mode factor analysis* (UCLA Working Papers in Phonetics, 16, 1–84). Los Angeles, CA: University of California at Los Angeles.

Harshman, R. A., & Lundy, M. E. (1984a). The PARAFAC model for three-way factor analysis and multidimensional scaling. In H. G. Law, C. W. Snyder, J. A. Hattie, & R. P. McDonald (Eds.), *Research methods for multimode data analysis* (pp. 122–215). New York: Praeger.

Harshman, R. A., & Lundy, M. E. (1984b). Data preprocessing and the extended PARAFAC model. In H. G. Law, C. W. Snyder, J. A. Hattie, & R. P. McDonald (Eds.), *Research methods for multimode data analysis* (pp. 216–284). New York: Praeger.

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.

Jennrich, R. I. (1970). Orthogonal rotation algorithms. *Psychometrika, 35,* 229–235.

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

Kapteyn, A., Neudecker, H., & Wansbeek, T. (1986). An approach to *n*-mode components analysis. *Psychometrika, 51,* 269–275.

Kiers, H. A. L. (1991). Hierarchical relations among three-way methods. *Psychometrika, 56,* 449–470.

Kiers, H. A. L. (1992). TUCKALS core rotations and constrained TUCKALS modelling. *Statistica Applicata, 4,* 659–667.

Kroonenberg, P. M. (1983). *Three mode principal component analysis: Theory and applications.* Leiden: DSWO press.

Kroonenberg, P. M., & de Leeuw, J. (1980). Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika, 45*, 69–97.

Kroonenberg, P. M., ten Berge, J. M. F., Brouwer, P., & Kiers, H. A. L. (1989). Gram-Schmidt versus Bauer-Rutishauser in alternating least-squares algorithms for three-mode component analysis. *Computational Statistics Quarterly, 5*, 81–87.

Kruskal, J. B. (1988, June). *Simple structure for three-way data: A new method intermediate between 3-mode factor analysis and PARAFAC-CANDECOMP*. Paper presented at the 53rd Annual Meeting of the Psychometric Society, Los Angeles.

Lastovicka, J. L. (1981). The extension of component analysis to four-mode matrices. *Psychometrika, 46*, 47–57.

MacCallum, R. C. (1976). Transformations of a three-mode multidimensional scaling solution to INDSCAL form. *Psychometrika, 41*, 385–400.

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

Murakami, T. (1983). Quasi three-mode principal component analysis—A method for assessing factor change. *Behaviormetrika, 14*, 27–48.

Neuhaus, J. O., & Wrigley, C. (1954). The quartimax method: An analytic approach to orthogonal simple structure. *British Journal of Mathematical and Statistical Psychology, 7*, 81–91.

Polit, E. (1986). *Une n-acp d'un hypercube de données* [N-way principal components analysis for a hypercube of data]. Unpublished doctoral thesis, Université de Sciences Sociales de Grenoble.

Rocci, R. (1992). Three-mode factor analysis with binary core and orthonormality constraints. *Journal of the Italian Statistical Society, 3*, 413–422.

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

ten Berge, J. M. F. (1995). Suppressing permutations or rigid planar rotations: A remedy against nonoptimal varimax rotations. *Psychometrika, 60*, 437–446.

ten Berge, J. M. F., de Leeuw, J., & Kroonenberg, P. M. (1987). Some additional results on principal components analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika, 52*, 183–191.

Tucker, L. R. (1951). *A method for synthesis of factor analysis studies*. (Personnel Research Section Report No. 984) Washington DC: Department of the Army.

Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika, 31*, 279–311.

Manuscript received 10/21/94
Final version received 11/5/96