PSYCHOMETRIKA—VOL. 56, NO. 2, 197–212
JUNE 1991

SIMPLE STRUCTURE IN COMPONENT ANALYSIS TECHNIQUES FOR MIXTURES OF QUALITATIVE AND QUANTITATIVE VARIABLES

HENK A. L. KIERS

UNIVERSITY OF GRONINGEN

Several methods have been developed for the analysis of a mixture of qualitative and quantitative variables, and one, called PCAMIX, includes ordinary principal component analysis (PCA) and multiple correspondence analysis (MCA) as special cases. The present paper proposes several techniques for simple structure rotation of a PCAMIX solution based on the rotation of component scores and indicates how these can be viewed as generalizations of the simple structure methods for PCA. In addition, a recently developed technique for the analysis of mixtures of qualitative and quantitative variables, called INDOMIX, is shown to construct component scores (without rotational freedom) maximizing the quartimax criterion over all possible sets of component scores. A numerical example is used to illustrate the implication that when used for qualitative variables, INDOMIX provides axes that discriminate between the observation units better than do those generated from MCA.

Key words: multiple correspondence analysis, INDSCAL, varimax, quartimax, orthomax, discrimination between objects.

Introduction

In principal component analysis (PCA), as well as in factor analysis, the solution for the loadings of variables on components is determined only up to a rotation, and therefore, the loading matrix is typically rotated to simple structure. Kaiser (1958) has discussed a number of simple structure criteria, all defined in terms of optimal patterns of small and large loadings (in an absolute sense). For a detailed discussion of the rationale behind simple structure rotation, the reader is referred to Harman (1976).

In PCA, the rotation of a loading matrix is paralleled by a rotation of the component score matrix. Explicitly, if Z denotes the $n \times m$ matrix of standardized scores of n observation units (objects) on m variables, and X ($n \times r$) the matrix with standardized component scores on r components, then the $m \times r$ loading matrix is given by $A = n^{-1}Z'X$. Thus, rotating the component scores matrix X by an orthogonal matrix T implies that the loadings corresponding to these rotated component scores are given by $AT = n^{-1}Z'XT$, and the loadings corresponding to the rotated component scores can be obtained by applying the same rotation to the original loadings.

Strictly speaking, ordinary PCA can be applied only to quantitative variables; however, as an analogue for the multivariate analysis of a set of qualitative variables, multiple correspondence analysis (MCA) is one of the best known alternatives (see, e.g., Tenenhaus & Young, 1985) that also constructs coordinates for the observation units (objects) in a low-dimensional space. When each qualitative variable is represented by means of a set of binary indicator variables specifying for each category whether an object belongs to it (1) or not (0), MCA can be formulated as a PCA of the total set of these indicator variables with respect to some predefined metrics. Thus, just

The Netherlands organization for scientific research (NWO) is gratefully acknowledged for funding this project. This research was conducted while the author was supported by a PSYCHON-grant (560-267-011) from this organization. The author is obliged to Jos ten Berge for his comments on an earlier version.

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

0033-3123/91/0600-89148$00.75/0

as in PCA, object coordinates can be seen as component scores determined up to a rotation only.

Apart from techniques for the analysis of solely qualitative or quantitative variables, several methods have been proposed for the analysis of mixtures. One type has been proposed independently by several authors (de Leeuw, 1973; de Leeuw & van Rijckevorsel, 1980; Escofier, 1979; Nishisato, 1980, pp. 103–107), and although the suggestions differ slightly in the way in which quantitative variables are transformed, essentially the same approach is used to handle qualitative variables. Of these, the one method of concern here is equivalent to the approach of de Leeuw and van Rijckevorsel (1980) and will be called PCAMIX. Explicitly, PCAMIX can be described as follows: Suppose n is the number of objects and m is the number of variables; if the j-th variable is qualitative, let G_j denote the $n \times m_j$ indicator matrix for the j-th variable, where m_j is the number of categories of variable j. D_j is defined as the diagonal matrix of frequencies of the categories of this variable, and the $n \times n$ matrix J as the centering operator $J = (I - \mathbf{1}\mathbf{1}'/n)$, where $\mathbf{1}$ is the vector of order n with unit elements. Given this notation, a so-called quantification matrix for a qualitative variable is defined as

$$S_j = JG_jD_j^{-1}G_j'J. \tag{1}$$

When the j-th variable is quantitative and the column vector $\mathbf{z}_j$ contains the standardized scores of the n objects on variable j, a quantification matrix is defined by

$$S_j = n^{-1}\mathbf{z}_j\mathbf{z}_j'. \tag{2}$$

In terms of (1) and (2), PCAMIX can be described as the method that maximizes

$$\mathrm{f}(X) = \sum_{j=1}^{m} \operatorname{tr} X'S_jX, \tag{3}$$

over X, subject to $X'X = nI_r$, where X ($n \times r$) contains (standardized) object coordinates (or component scores). The solution is given by the first r eigenvectors of $\Sigma_j S_j$. Clearly, given only qualitative variables, the quantification matrices are all of the form in (1), and the PCAMIX solution for the object coordinates given by the eigenvectors of $\Sigma_j JG_jD_j^{-1}G_j'J$, which is the well-known solution for MCA. Alternatively, given only quantitative variables, the quantification matrices are all of the form in (2) and the PCAMIX solution for the component scores is given by the eigenvectors of $n^{-1}\Sigma_j\mathbf{z}_j\mathbf{z}_j'$, which is equivalent to the well-known PCA solution.

From (3), it is evident that rotating X does not change the function value since for any T for which $T'T = TT' = I_r$, $\mathrm{f}(XT) = \mathrm{f}(X)$. The choice of such a rotation may depend on many different criteria, but it is standard practice in PCAMIX as well as in the special case of MCA, to use as components those that successively account for the maximum inertia, given by the terms $\mathbf{x}_l'\Sigma_jS_j\mathbf{x}_l$, where $\mathbf{x}_l$ denotes the l-th column of X, $l = 1, \ldots, r$. As is the case in ordinary PCA, this may yield components that are difficult to interpret. The present paper describes techniques for rotation of the component scores such that certain criteria possibly leading to a good interpretation of the components are optimized.

In PCAMIX the relation between the loadings and component scores known for PCA, $A = n^{-1}Z'X$, does not generally hold, when A contains loadings (defined in some way) for the qualitative and quantitative variables on the axes. As a consequence, rotating the component scores is no longer equivalent to rotating the loading matrix, and if one wishes to rotate the component scores so the loadings of the variables on the components have optimal simple structure, the standard simple structure rotation techniques cannot be used. The techniques to be developed are for the rotation of the *component scores* so the *loadings* of the variables have optimal simple structure, and will be considered generalizations of orthogonal simple structure rotations for PCA.

Recently, Kiers (1988, 1989c, pp. 61–70) proposed an alternative to PCAMIX, denoted as "INDORT for a mixture of qualitative and quantitative variables", or briefly, INDOMIX. Like PCAMIX, this method yields object coordinates, but does so by optimizing a criterion that differs from that of PCAMIX. We will show that INDOMIX optimizes the quartimax criterion, one of the simple structure criteria to be described shortly. INDOMIX and the PCAMIX quartimax rotation technique differ, however, since the latter optimizes the quartimax criterion only over rotations of the PCAMIX component scores, while INDOMIX optimizes the quartimax criterion over all possible sets of component scores. Therefore, the INDOMIX solution always attains a quartimax value at least as high as that attained by the quartimax rotation of the PCAMIX component scores. This explains why the loadings of the variables obtained by INDOMIX are more clearly clustered than those of PCAMIX, and when applied to sets of solely qualitative variables, INDOMIX tends to yield solutions with clusters of objects that, per axis, are more clearly separated and denser than those possibly found in an MCA solution. Apart from giving a formal explanation of this phenomenon, it will also be discussed by means of an illustrative data set.

A Definition of Squared Loadings in PCAMIX

Like PCA, PCAMIX finds component scores for objects on several components. In ordinary PCA, the loadings of variables on components are given by the correlations between the variables and the components. In PCAMIX it is possible to define loadings for the quantitative variables in the same way; explicitly, the loading of the quantitative variable j on component l can be given by $a_{jl} = n^{-1}\mathbf{z}_j'\mathbf{x}_l$, the product-moment correlation between variable j and component l, where $\mathbf{x}_l$ is the l-th column of the component score matrix X. For qualitative variables, however, the product-moment correlation cannot be used, and instead, another coefficient has to be chosen to express the correlation between a qualitative variable and a (quantitative) component. One such index is the discrimination measure (Gifi, 1990) defined as the contribution of a component to the inertia of a variable that is accounted for, and formally by $c_{jl} \equiv n^{-1}\mathbf{x}_l'S_j\mathbf{x}_l$, with S_j defined as in (1). This measure can be interpreted as the squared correlation between variable j when it is optimally quantified and component l (Gifi, 1990, p. 96), or alternatively, as the well-known correlation ratio η^2. In both interpretations, the measure c_{jl} is considered a squared correlation, and thus, the squared loading of variable j on component l.

To have the same notation for qualitative and quantitative variables, c_{jl} is defined for a quantitative variable as the squared loading of variable j on component l: $c_{jl} = a_{jl}^2 = n^{-2}(\mathbf{z}_j'\mathbf{x}_l)^2 = n^{-1}\mathbf{x}_l'(n^{-1}\mathbf{z}_j\mathbf{z}_j')\mathbf{x}_l$. Using the definition of S_j in (1) and (2), $c_{jl} = n^{-1}\mathbf{x}_l'S_j\mathbf{x}_l$ for both qualitative and quantitative variables. It is of interest to note that PCAMIX can be formulated as the method that maximizes $n\Sigma_{jl}c_{jl}$, with c_{jl} defined as above, over X subject to $X'X = nI_r$.

Simple Structure Rotations for PCA

Kaiser (1958) described several simple structure criteria, as well as procedures to optimize these criteria over orthogonal rotations of the loading matrix, included in the orthomax family of orthogonal rotations (Clarkson & Jennrich, 1988; Crawford & Ferguson, 1970; Jennrich, 1970) to be discussed here. The orthomax family of simple structure rotations for PCA can be described as the set of techniques maximizing the orthomax criterion (denoted by the acronym ORMAX) expressible in terms of the squared loadings of the variables on the axes. If the loading for variable j on axis l is given by a_{jl}, $j = 1, \ldots, m$, $l = 1, \ldots, r$, the ORMAX criterion is given by

$$\text{ORMAX} = \sum_{j=1}^{m} \sum_{l=1}^{r} a_{jl}^4 - \frac{\gamma}{m} \sum_{l=1}^{r} \left(\sum_{j=1}^{m} a_{jl}^2 \right)^2 . \tag{4}$$

Although in principle γ can be any scalar, we assume $0 \leq \gamma \leq 1$. The lower value of $\gamma = 0$ yields the quartimax criterion (QMAX)

$$\text{QMAX} = \sum_{j=1}^{m} \sum_{l=1}^{r} a_{jl}^4,$$

originally proposed by Ferguson (1954); the higher value $\gamma = 1$ yields the varimax criterion (VMAX) proposed by Kaiser (1958):

$$\text{VMAX} = \sum_{j=1}^{m} \sum_{l=1}^{r} a_{jl}^4 - \frac{1}{m} \sum_{l=1}^{r} \left(\sum_{j=1}^{m} a_{jl}^2 \right)^2 .$$

Techniques for optimizing these criteria have been discussed by various authors, but only the method for maximizing the orthomax criterion given by ten Berge, Knol, and Kiers (1988) is mentioned here. If A is the $m \times r$ matrix of component loadings, $\mathbf{a}_j'$ the j-th row of A, and $E_j = (\delta A'A - m\mathbf{a}_j\mathbf{a}_j')$, for $j = 1, \ldots, m$, with δ defined so $\gamma = \delta(2 - \delta)$, the problem of maximizing the orthomax function is equivalent to simultaneously diagonalizing the set of E_j matrices in the least squares sense, or equivalently, maximizing Σ_j tr(Diag $T'E_jT)^2$ over orthonormal matrices T. For simultaneously diagonalizing a set of matrices, an algorithm proposed by de Leeuw and Pruzansky (1978) can be used.

Simple Structure Rotations for PCAMIX

As noted earlier, PCAMIX rotated component scores do not correspond to a loading matrix that can be found by rotating the original loading matrix, and thus it does not suffice to express the simple structure criteria for ordinary PCA in terms of the original loadings and a rotation matrix. Instead, these criteria will be expressed in terms of the squared loadings of the variables on the rotated components.

The orthomax criterion given in (4) can be expressed in terms of the squared PCAMIX loadings by replacing a_{jl}^2 by c_{jl} for all j and l:

$$\text{ORMAX} = \sum_{j=1}^{m} \sum_{l=1}^{r} c_{jl}^2 - \frac{\gamma}{m} \sum_{l=1}^{r} \left(\sum_{j=1}^{m} c_{jl} \right)^2 . \tag{5}$$

Substituting $n^{-1}\mathbf{x}_l'S_j\mathbf{x}_l$ for c_{jl} in (5), gives

$$f_{or}(X) \equiv n^2\text{ORMAX} = \sum_{j=1}^{m} \sum_{l=1}^{r} (\mathbf{x}_l'S_j\mathbf{x}_l)^2 - \frac{\gamma}{m} \sum_{l=1}^{r} \left(\sum_{j=1}^{m} \mathbf{x}_l'S_j\mathbf{x}_l \right)^2$$

$$= \sum_{j=1}^{m} \sum_{l=1}^{r} (\mathbf{x}_l' S_j \mathbf{x}_l)^2 - \frac{\gamma}{m} \sum_{l=1}^{r} (\mathbf{x}_l' \Sigma_j S_j \mathbf{x}_l)^2. \tag{6}$$

The problem of maximizing $f_{or}(FT)$ over orthonormal matrices T, where F $(n \times r)$ is the unrotated PCAMIX component score matrix, will be translated as the simultaneous diagonalization of a set of matrices, as in ten Berge et al. (1988) for the orthomax rotation of a PCA loading matrix. Letting the l-th column of T be $\mathbf{t}_l$,

$$f_{or}(FT) = \sum_{j=1}^{m} \sum_{l=1}^{r} (\mathbf{t}_l' F' S_j F \mathbf{t}_l)^2 - \frac{\gamma}{m} \sum_{l=1}^{r} (\Sigma_j \mathbf{t}_l' F' S_j F \mathbf{t}_l)^2$$

$$= \sum_{j=1}^{m} \sum_{l=1}^{r} (\mathbf{t}_l' (F' S_j F - \delta m^{-1} \Sigma_k F' S_k F) \mathbf{t}_l)^2, \tag{7}$$

with δ chosen so $2\delta - \delta^2 = \gamma$. For $\delta = \gamma = 1$, the right-hand side of (7) follows since the variance can be written as an average squared deviation from the mean. If $\delta \neq 1$, the second equality in (7) follows from

$$\sum_{j=1}^{m} \sum_{l=1}^{r} (\mathbf{t}_l' (F' S_j F - \delta m^{-1} \Sigma_k F' S_k F) \mathbf{t}_l)^2$$

$$= \sum_{j=1}^{m} \sum_{l=1}^{r} (\mathbf{t}_l' F' S_j F \mathbf{t}_l)^2 + \sum_{j=1}^{m} \sum_{l=1}^{r} (\mathbf{t}_l' \delta m^{-1} \Sigma_k F' S_k F \mathbf{t}_l)^2$$

$$- 2 \sum_{j=1}^{m} \sum_{l=1}^{r} (\mathbf{t}_l' F' S_j F \mathbf{t}_l)(\mathbf{t}_l' \delta m^{-1} \Sigma_k F' S_k F \mathbf{t}_l)$$

$$= \sum_{j=1}^{m} \sum_{l=1}^{r} (\mathbf{t}_l' F' S_j F \mathbf{t}_l')^2 + m \sum_{l=1}^{r} \delta^2 m^{-2} (\mathbf{t}_l' \underline{\Sigma_k F'} S_k F \mathbf{t}_l)^2$$

$$- 2\, \delta m^{-1} \sum_{l=1}^{r} (\mathbf{t}_l' \Sigma_j F' S_j F \mathbf{t}_l)(\mathbf{t}_l' \Sigma_k F' S_k F \mathbf{t}_l)$$

$$= \sum_{j=1}^{m} \sum_{l=1}^{r} (\mathbf{t}_l' F' S_j F \mathbf{t}_l)^2 + \delta^2 m^{-1} \sum_{l=1}^{r} (\mathbf{t}_l' \Sigma_j F' S_j F \mathbf{t}_l)^2 - 2\, \delta m^{-1} \sum_{l=1}^{r} (\mathbf{t}_l' \Sigma_j F' S_j F \mathbf{t}_l)^2$$

$$= \sum_{j=1}^{m} \sum_{l=1}^{r} (\mathbf{t}_l' F' S_j F \mathbf{t}_l)^2 + m^{-1} \underline{(\delta^2 - 2\delta) \sum_{l=1}^{r} (\mathbf{t}_l' \Sigma_j F' S_j F \mathbf{t}_l)^2}.$$

Because F contains the PCAMIX solution, the columns of F are the first r eigenvectors of $\Sigma_j S_j$, normalized to sums of squares of n; hence, $F' \Sigma_j S_j F = n\Lambda$, where Λ is the diagonal matrix with the first r eigenvalues of $\Sigma_j S_j$ on its diagonal. Using this result, (7) can be rewritten as

$$\mathrm{f}_{or}(FT) = \sum_{j=1}^{m} \sum_{l=1}^{r} (\mathbf{t}_l'(F'S_jF - \delta m^{-1}n\Lambda)\mathbf{t}_l)^2. \tag{8}$$

If $\tilde{E}_j$ is defined by

$$\tilde{E}_j \equiv (mF'S_jF - \delta n\Lambda),$$

for $j = 1, \ldots, m$, (8) can be rewritten as

$$\mathrm{f}_{or}(FT) = m^{-2} \sum_{j=1}^{m} \mathrm{tr}\ T'\tilde{E}_jT(\mathrm{Diag}\ T'\tilde{E}_jT).$$

As ten Berge (1984, p. 348) has shown, maximizing this function over orthonormal matrices T is equivalent to the problem of simultaneously diagonalizing the set of matrices $\tilde{E}_1, \ldots, \tilde{E}_m$ in the least squares sense, and thus, the algorithm proposed by de Leeuw and Pruzansky (1978) can be used, or any other that simultaneously diagonalizes a set of symmetric matrices.

As mentioned above, ten Berge et al. (1988) have shown that the problem of maximizing the orthomax function over orthogonal rotations of a PCA loading matrix is equivalent to the problem of simultaneously diagonalizing the matrices $E_j = (\delta A'A - m\mathbf{a}_j\mathbf{a}_j')$, where A is the $m \times r$ PCA loading matrix, and $\mathbf{a}_j'$ is the j-th row of A. In the special case where PCAMIX is applied to a set of solely quantitative variables, the procedure for the orthomax rotation of the PCAMIX solution is equivalent to the simultaneous diagonalization of a set of $\tilde{E}_j$ matrices proportional to the E_j matrices. Explicitly, when all variables are quantitative, $S_j = n^{-1}\mathbf{z}_j\mathbf{z}_j'$ and $\tilde{E}_j = (mn^{-1}F'\mathbf{z}_j\mathbf{z}_j'F - \delta n\Lambda)$, where Λ is the diagonal matrix with eigenvalues of $n^{-1}\Sigma_j\mathbf{z}_j\mathbf{z}_j'$. Thus, $\mathbf{a}_j' = n^{-1}\mathbf{z}_j'F$, and $mn^{-1}F'\mathbf{z}_j\mathbf{z}_j'F = mn\mathbf{a}_j\mathbf{a}_j'$. Because $A'A = \Sigma_j\mathbf{a}_j\mathbf{a}_j' = n^{-2}F'\Sigma_j\mathbf{z}_j\mathbf{z}_j'F = n^{-1}F'\Sigma_jS_jF = \Lambda$, matrix $\tilde{E}_j$ can be written as $\tilde{E}_j = (mn\mathbf{a}_j\mathbf{a}_j' - \delta nA'A) = -nE_j$, which shows that the orthomax rotation procedure for ordinary PCA is a special case of the orthomax rotation procedure for PCAMIX.

Because the quartimax criterion and the varimax criterion are special case of the orthomax criterion, we immediately have a quartimax and varimax procedure for rotating the PCAMIX solution by setting γ to 0 or 1, respectively. The $\tilde{E}_j$ matrices do not explicitly contain γ, but depend on γ because δ depends on γ. For the quartimax procedure, we obtain $\gamma = 0$ when $\delta = 0$ (or $\delta = 2$, which is less convenient and therefore ignored); for the varimax procedure, taking $\delta = 1$, we have $\gamma = 1$.

INDOMIX

Although PCAMIX is the best-known method for the analysis of a set of qualitative and quantitative variables, Kiers (1988, 1989c) has recently suggested INDOMIX as an alternative approach for the analysis of such sets of variables, as a compromise between PCAMIX and another method for the analysis of mixtures of qualitative and quantitative variables developed by Saporta (1976). The latter strategy analyzes mixtures of qualitative and quantitative variables by means of an ordinary PCA on certain correlation measures between the variables, and yields loadings for the variables (optimally representing the variables) without giving component scores for the objects. The compromise is directed at optimally representing the variables (as Saporta's method does) while at the same time providing component scores for the objects (as

PCAMIX does), and as shown below, optimizes the quartimax criterion over all possible sets of component scores.

INDOMIX is based on applying a constrained variant of INDSCAL (Carroll & Chang, 1970) to the quantification matrices S_j defined above; explicitly, we minimize the INDSCAL loss function

$$\sigma(X,W_j) = \sum_{j=1}^{m} \|S_j - XW_jX'\|^2,$$

over X ($n \times r$) and diagonal matrices $W_1, \ldots, W_m$, where X is constrained such that $X'X = nI_r$ (see Kiers, 1989c; also see Kiers, 1989a, for details on an efficient INDOMIX algorithm for large data sets). For arbitrary but fixed X, the W_j matrices that minimize $\sigma(X,W_j)$ are given by $W_j = \text{Diag}(X'S_jX)$; thus, minimizing $\sigma(X,W_j)$ reduces to maximizing

$$g(X) = \sum_{j=1}^{m} \text{tr}\,(\text{Diag}\; X'S_jX)^2 = \sum_{j=1}^{m}\sum_{l=1}^{r} (\mathbf{x}_l'S_j\mathbf{x}_l)^2 = n^2\sum_{j=1}^{m}\sum_{l=1}^{r} c_{jl}^2, \tag{9}$$

subject to $X'X = nI_r$ (see Kiers, 1989b). Since maximizing g(X) is equivalent to maximizing the quartimax function, INDOMIX maximizes the quartimax function over component score matrices X, subject to $X'X = nI_r$. INDOMIX differs from quartimax *rotation* in that the latter maximizes the quartimax function over rotations T of the component score matrix F. The class of matrices X over which INDOMIX maximizes the quartimax function consists not only of all rotated versions FT of F, but also all matrices X with columns outside the column-space of F for which $X'X = nI_r$. Thus, the maximum of g(X) is always at least as large as the maximum of g(FT) since INDOMIX maximizes the quartimax criterion over *all* possible component score matrices, yielding a quartimax value at least as high as the maximum possible quartimax value obtained by rotation of the PCAMIX solution. In turn, INDOMIX yields solutions that have a simpler structure than the optimally rotated PCAMIX solutions when simple structure is considered in the quartimax sense.

The quartimax criterion was one of the first analytic simple structure criteria, and is not the most prevalent in ordinary PCA. As Kaiser (1958) pointed out, the quartimax criterion tends to give a solution with one general component, quite contrary to the purpose of achieving maximum simple structure. This tendency does not seem to be present in INDOMIX, however, and in practice it is found that even though INDOMIX maximizes only the quartimax function, INDOMIX tends to provide solutions with a high amount of simple structure in terms of the varimax criterion as well. Nevertheless, it does seem useful to also discuss methods that maximize the orthomax criterion for *any* γ over all possible sets of component scores, in the same way as INDOMIX maximizes the quartimax criterion.

The orthomax function (6) can be rewritten as

$$\begin{aligned} f_{or}(X) &= \sum_{j=1}^{m}\sum_{l=1}^{r} (\mathbf{x}_l'S_j\mathbf{x}_l)^2 - \frac{\gamma}{m}\sum_{l=1}^{r} (\mathbf{x}_l'\Sigma_jS_j\mathbf{x}_l)^2 \\ &= \sum_{j=1}^{m}\sum_{l=1}^{r} (\mathbf{x}_l'(S_j - \delta m^{-1}\Sigma_kS_k)\mathbf{x}_l)^2, \end{aligned} \tag{10}$$

with δ again chosen so $2\delta - \delta^2 = \gamma$. The last step in the derivation of (10) is based on a reasoning similar to that used in deriving (8). Clearly, choosing $\delta = 0$ (or $\delta = 2$), we have $f_{or}(X) = g(X)$. However, (10) can be defined for any other γ between 0 and 1 (corresponding to $\delta = 1 \pm (1 - \gamma)^{1/2}$), and in particular, the varimax function ($\gamma = 1$) is maximized over all component score matrices X by maximizing

$$h(X) = \sum_{j=1}^{m} \sum_{l=1}^{r} (\mathbf{x}_l'(S_j - m^{-1}\Sigma_k S_k)\mathbf{x}_l)^2,$$

over X subject to $X'X = nI_r$. Thus, this method applies orthogonally constrained INDSCAL to the matrices $(S_j - m^{-1}\Sigma_k S_k)$, which are centered with respect to the mean of these matrices.

Ten Berge et al. (1988) have provided an algorithm for maximizing function (9) that converges monotonically if the S_j are positive semidefinite. Their algorithm might also be used for maximizing (10), with S_j replaced by $(S_j - \delta m^{-1}\Sigma_k S_k)$, but monotone convergence is no longer guaranteed. An algorithm for which monotone convergence is guaranteed has been described by Kiers (1990), and is a slightly adapted version of the ten Berge et al. algorithm.

Given the above discussion, it is apparent that several methods are now available for finding components for mixtures of qualitative and quantitative variables maximizing the orthomax simple structure criteria (for the loadings of the variables on the components). One approach is to perform PCAMIX first and then rotate the component scores to optimize the simple structure criteria over orthogonal rotations of the PCAMIX component scores. Sets of components are found that account best for the inertia as measured in PCAMIX, and among these components, that set yielding the greatest simple structure value. The other approach, the generalization of INDOMIX maximizing $f_{or}(X)$ in (10), seeks components that have the best possible simple structure at the cost of a possible loss (which tends to be rather small in practice) in explained inertia.

We might note briefly that generalized INDOMIX (maximizing (10)) applied to a set of solely quantitative variables provides an alternative to ordinary PCA. This method will provide orthomax values at least as high as the orthomax values of rotated PCA loadings, and therefore, when the main objective is to find components with a clear simple structure in terms of any of the orthomax criteria, and maximally accounting for the inertia is less important, generalized INDOMIX might be a useful alternative to ordinary PCA. Although the explained inertia is no longer maximized by generalized INDOMIX, it cannot be very small either, as can be seen in the case of varimax ($\delta = 1$): Function $h(X)$ is proportional to the sum of column-variances of the elements of C, $m^{-1}\Sigma_j\Sigma_l(c_{jl} - m^{-1}\Sigma_k c_{kl})^2$, which cannot be maximal if the explained inertia, $n\Sigma_{jl}c_{jl}$, is very small.

Relations Between INDOMIX and Simple Structure Rotations of PCAMIX

The generalization of INDOMIX to maximize the orthomax function over all possible component score matrices always yields an orthomax function value at least as high as the one obtained by the orthomax rotation of the PCAMIX solution. As we indicate in Results 1 to 3 below, INDOMIX also yields values of other orthomax criteria at least as high as the ones attained by optimally rotated PCAMIX solutions. As notation, the component score matrix of the INDOMIX solution is denoted as X_I, the unrotated PCAMIX solution as F, and the optimally rotated PCAMIX solution as FT_0 where T_0 is the rotation matrix optimizing the simple structure criterion at hand.

Result 1. ORMAX(X_I) $\geq$ ORMAX(F).

Result 2. VMAX(X_I) $\geq$ VMAX(F).

Result 3. ORMAX(X_I) $\geq$ ORMAX(FT_0) = ORMAX(F), if

(a) $r = 1$, or

(b) $m = 2$, and both variables are qualitative (i.e., when PCAMIX is equivalent to correspondence analysis).

Proof 1. The ORMAX function (5) can be rewritten as

$$\mathrm{ORMAX}(X) = n^{-2}\mathrm{g}(X) - n^{-2}\frac{\gamma}{m}\,\mathrm{k}(X),$$

where $0 \leq \gamma \leq 1$, and k(X) is defined as

$$\mathrm{k}(X) = \sum_{l=1}^{r} (\mathbf{x}_l'\Sigma_j S_j \mathbf{x}_l)^2.$$

If an eigendecomposition of $\Sigma_j S_j$ is given by $\Sigma_j S_j = K\Lambda K'$, then k($X$) can be rewritten as

$$\mathrm{k}(X) = \sum_{l=1}^{r} (\mathbf{x}_l' K\Lambda K'\mathbf{x}_l)^2.$$

From the Cauchy-Schwarz inequality, it follows that

$$(\mathbf{x}_l' K\Lambda K'\mathbf{x}_l)^2 = [(\mathbf{x}_l' K\Lambda K')(\mathbf{x}_l)]^2 \leq n(\mathbf{x}_l' K\Lambda^2 K'\mathbf{x}_l),$$

and

$$\sum_{l=1}^{r} (\mathbf{x}_l' K\Lambda K'\mathbf{x}_l)^2 \leq n \sum_{l=1}^{r} (\mathbf{x}_l' K\Lambda^2 K'\mathbf{x}_l) = n \;\mathrm{tr}\; X'K\Lambda^2 K'X. \tag{11}$$

It is readily verified (ten Berge, 1983) that the right-hand side in (11) is smaller than or equal to the sum of the first r values in $n^2\Lambda^2$. If Λ_r denotes the diagonal matrix with the first r values in Λ, then

$$\mathrm{k}(X) = \sum_{l=1}^{r} (\mathbf{x}_l' K\Lambda K'\mathbf{x}_l)^2 \leq n^2 \mathrm{tr}\;\Lambda_r^2. \tag{12}$$

Inequality (12) provides an upper bound to k(X) that is attained by choosing X as $n^{1/2}K_r$, the matrix with the r standardized eigenvectors of $\Sigma_j S_j$ belonging to the first r eigenvalues of $\Sigma_j S_j$. This is precisely the unrotated PCAMIX solution F for the component scores. Therefore, $k(F) \geq k(X_I)$, and because $\gamma \geq 0$, $-\gamma k(X_I) \geq -\gamma k(F)$. Combining this result with $g(X_I) \geq g(F)$ proves $\text{ORMAX}(X_I) \geq \text{ORMAX}(F)$. □

Proof 2. Result 2 follows immediately from Result 1 for $\gamma = 1$. □

Proof 3. (a) If $r = 1$, rotation of F reduces to a possible reflection of the one column in F, which does not affect the squared loadings. Hence, $\text{ORMAX}(FT_0) = \text{ORMAX}(F)$, and Result 3 follows at once from Result 1. (b) If $m = 2$ and both variables are qualitative, it can be shown that $F'S_1F = F'S_2F = \frac{1}{2}n\Lambda_r$ (Nishisato & Sheu, 1980, p. 471). Hence, $f_{or}(FT_0)$ is the maximum over T of

$$f_{or}(FT) = \sum_{j=1}^{2} \sum_{l=1}^{r} (\mathbf{t}_l'(F'S_jF - \tfrac{1}{2}\delta n\Lambda_r)\mathbf{t}_l)^2, \tag{13}$$

as follows from (7). Substituting $F'S_1F = F'S_2F = \frac{1}{2}n\Lambda_r$ in (13) yields

$$f_{or}(FT) = \sum_{j=1}^{2} \sum_{l=1}^{r} (\mathbf{t}_l'(\tfrac{1}{2}(1-\delta)n\Lambda_r)\mathbf{t}_l)^2.$$

Analogous to the proof of (12), we have

$$f_{or}(FT) = \sum_{j=1}^{2} \sum_{l=1}^{r} (\mathbf{t}_l'(\tfrac{1}{2}(1-\delta)n\Lambda_r)\mathbf{t}_l)^2 \leq \sum_{j=1}^{2} \tfrac{1}{4}(1-\delta)^2 n^2 \text{ tr } \Lambda_r{}^2$$
$$= \tfrac{1}{2}(1-\delta)^2 n^2 \text{ tr } \Lambda_r{}^2. \tag{14}$$

The right-hand side of (14) gives an upper bound to $f_{or}(FT)$, attained for $T = I$. That is, $f_{or}(FT_0)$, the maximum over T of $f_{or}(FT)$, is equal to $f_{or}(F)$; hence, $\text{ORMAX}(FT_0) = \text{ORMAX}(F)$. With this equality, $\text{ORMAX}(X_I) \geq \text{ORMAX}(FT_0) = \text{ORMAX}(F)$ follows immediately from Result 1. □

It is of interest to mention that for the VMAX criterion, Result 3b implies for $m = 2$, that $\text{VMAX}(F) = \text{VMAX}(FT_0) = 0$. The fact that $\text{VMAX}(FT_0) = 0$ follows at once upon substitution of $\delta = 1$ in (14), and implies that the correspondence analysis solution (MCA with $m = 2$) gives a varimax function that is always zero, and cannot be increased by rotating the solution. In addition, it can be concluded that for $m = 2$, the MCA loadings on a component are the same for both variables.

Result 3 has been included because it proves $\text{VMAX}(X_I) \geq \text{VMAX}(FT_0)$ in some special cases. Although this result does not generally hold for $m > 2$, $r > 1$, it is often found that $\text{VMAX}(X_I) \geq \text{VMAX}(FT_0)$ in practice. Because INDOMIX maximizes the sum of *squares* of the c_{jl} values (which will from now on be called loadings instead of squared loadings), the loadings (c_{jl}) tend to be either large or small for each INDOMIX dimension. Dimensions that do not account for a considerable part of the inertia are usually discarded, and therefore, one typically has at least some high loadings for each INDOMIX dimension. On the other hand, it is very unlikely that *all* variables load highly on a component because this can only happen when all variables are related strongly to each other. As a consequence, the INDOMIX solution typically consists of dimensions with some high and some small loadings. Such a pattern of loadings contributes much to the VMAX value. Since (rotated) PCAMIX typically finds less extreme loadings, the VMAX value tends to be smaller than for the INDOMIX solution.

The above reasoning implicitly indicates how cases with VMAX(X_I) < VMAX(FT_0) can be constructed: Data consisting only of strongly related variables will provide one (or more) components with high loadings throughout and other components with small loadings, both in INDOMIX and unrotated PCAMIX. Clearly, both solutions will have a small VMAX value, but after varimax rotation of the PCAMIX solution, small and large loadings will be distributed more evenly over the dimensions, typically increasing the VMAX value of the PCAMIX solution above that of the INDOMIX solution. Examples constructed in this way can be obtained from the author. Moreover, Result 4 describes a class of cases in which INDOMIX *never* yields a higher VMAX value than rotated PCAMIX.

Result 4. If $r = \text{rank}(\Sigma_j S_j)$, then ORMAX($X_I$) $\leq$ ORMAX(FT_0).

Proof. If $r = \text{rank}(\Sigma_j S_j)$, then both F and X_I span the complete column-space of $\Sigma_j S_j$, which follows at once from the normal equations for the respective maximization problems (Kiers, 1989c, p. 103). Because F and X_I are both columnwise orthonormal, $X_I = FT$ for some orthonormal matrix T. As a consequence, ORMAX(X_I) = ORMAX(FT) $\leq$ ORMAX(FT_0), where the inequality follows from the optimality of T_0. □

Comparison of MCA and INDOMIX for Qualitative Variables with Respect to Discriminatory Capability

INDOMIX attains values of several simple structure criteria at least as high as those attained by optimally rotated PCAMIX solutions. As a consequence to be developed below, MCA (i.e., PCAMIX applied to a set of solely qualitative variables) finds object coordinate axes that appear to discriminate between the objects as well as possible in terms of *all* variables, whereas in INDOMIX (INDOQUAL when applied to solely qualitative variables) each axis tends to discriminate between the objects mainly in terms of a subset of the variables, and for different axes, different subsets may be involved.

Each qualitative variable defines a set of disjoint groups of objects that fall in different categories of the qualitative variable. Thus, for each variable and axis, the group averages can be computed on that axis. The variance of these group averages is called the between-groups-variance, or because the groups are defined by the categories, the between-categories-variance. For both MCA and INDOQUAL, the solution for the matrix of object coordinates X is centered columnwise, and as a consequence, the between-categories-variance of $\mathbf{x}_l$ with respect to the categories of variable j can be given as

$$\sigma^2_{B(jl)} = n^{-1}\sum_{g=1}^{m_j} n_g(x_l^{gj})^2 = n^{-1}\mathbf{x}_l'G_jD_j^{-1}G_j'\mathbf{x}_l = n^{-1}\mathbf{x}_l'JG_jD_j^{-1}G_j'J\mathbf{x}_l = c_{jl}, \tag{15}$$

where n_g denotes the number of objects in category g of variable j, x_l^{gj} denotes the average value of $\mathbf{x}_l$ in category g of variable j (see, e.g., Tenenhaus & Young, 1985, p. 98). Thus, the loading of variable j on axis l is equal to the between-categories-variance for axis l with respect to the categories defined by variable j. This between-categories-variance can also be considered as the amount of discrimination, provided by axis l, between the objects that fall in different categories of variable j, and the loading c_{jl} indicates how strongly axis l discriminates between the objects in terms of the categories of variable j.

This interpretation of the loadings provides the basis for our statement that the

INDOQUAL axes discriminate between the objects better than do the MCA axes. As noted below, the difference between INDOQUAL and MCA is an immediate consequence of the fact that INDOQUAL provides a solution with better simple structure than MCA, at least in terms of the QMAX criterion, and typically, also in terms of the VMAX criterion. INDOQUAL loadings usually have a simpler structure than MCA loadings, even after optimal rotation of the MCA solution, and thus, INDOQUAL finds loadings that are overall more diverse than those resulting from MCA. Because the loadings are bounded between zero and one, INDOQUAL loadings tend to the extreme values of zero and one more than MCA loadings. From (15) it also follows that INDOQUAL yields more extreme between-categories-variances than MCA, implying that the INDOQUAL coordinate axes discriminate between objects better in terms of the categories of *certain* variables (those with large between-categories-variances), and at the same time, worse in terms of other variables than MCA. In this way, it can be said that INDOQUAL finds axes each of which seek to discriminate between the objects, to a larger extent than MCA, in terms of (possibly different) subsets of variables. Because INDOQUAL seeks to discriminate the objects in terms of fewer variables than MCA, INDOQUAL will succeed better in actually discriminating between the objects. MCA tries to discriminate between the objects as well as possible in terms of *all* variables. When certain variables define completely different groupings of objects, the MCA axes will tend to make compromises by discriminating between the objects a little worse in terms of both variables. On the other hand, INDOQUAL will optimally discriminate between the objects in terms of either one of these opposite variables, and will possibly discriminate between the objects in terms of the other variable by means of a different axis.

In addition, it can be said that a subset of variables that load highly on an axis consists of variables that, at least in one respect, are rather strongly related to each other. If all variables in a subset load highly on an axis, this axis discriminates the categories of each of these variables well, which is only possible if the partitions (groupings) defined by the different variables overlap highly. This is another way of saying that the qualitative variables involved are highly related with respect to the partition of objects into groups that are best discriminated by the axis. In short, INDOQUAL finds axes that discriminate between the objects better than MCA overall, and does so by discriminating between the objects in terms of the categories of subsets of variables that have highly overlapping partitions.

Although INDOQUAL finds loadings that tend more to zero and one than those in MCA, this does not imply that each axis of INDOQUAL *always* has loadings that are larger than those of MCA. In practice, INDOQUAL often yields a solution with loadings that do not only have a simpler structure than those of MCA, but that also consist of certain values that are higher than the highest MCA loadings for a corresponding MCA axis. As a consequence, INDOQUAL yields a solution in which, with respect to each axis, objects separate more clearly than in MCA into (denser) clusters of objects representing the categories of those highly loading variables.

An Illustrative Analysis of Empirical Data

The empirical data to be analyzed in the present section is from Hartigan (1975, p. 228), and consists of 24 objects such as screws and nails that are classified according to 5 qualitative variables (whether they have a thread, type of head, head indentation, kind of bottom, and whether made of brass). In addition, their length (in half inches) is measured, and considered a qualitative variable with five categories (1 through 5 half inches). Although the data are of little practical interest, they serve to illustrate the

TABLE 1

MCA Loadings Before and After Varimax Rotation

	Before Varimax Rotation			After Varimax Rotation		
	Dimension			Dimension		
	1	2	3	1	2	3
Thread	0.93	0.02	0.00	0.95	0.00	0.00
Head	0.95	0.64	0.73	0.96	0.64	0.72
Head indentation	0.94	0.67	0.07	0.95	0.73	0.00
Bottom	0.55	0.02	0.00	0.50	0.05	0.01
Length	0.29	0.82	0.70	0.24	0.79	0.79
Brass	0.06	0.03	0.47	0.09	0.00	0.47
$\Sigma_j c_{jl}$	3.72	2.20	1.97	3.69	2.20	2.00
explained inertia ($\Sigma_{jl} c_{jl}$)		7.90			7.90	
varimax function value		2.05			2.23	
quartimax function value		5.83			5.97	

clustering phenomenon, because the objects are well-described in terms of predefined clusters (those of screws, bolts, nails, and tacks), whereas this clustering does not refer directly to a qualitative variable in the analysis.

The MCA solutions and the INDOQUAL solutions for $r = 3$ will be compared, with the MCA solution considered both before and after varimax rotation (see Table 1 for the MCA loadings). The varimax rotation only changes the loadings slightly, mainly those of the fifth variable, and these changes lead to increasing simple structure as expressed by the varimax and quartimax function values. The amount of explained inertia is equal in the two solutions. We also applied a quartimax rotation, which provided practically the same results as the varimax rotation (with loadings never deviating more than .1). Table 2 gives the loadings found by INDOQUAL, where obviously the INDOQUAL components and those on the rotated MCA solution have high loadings for the same variables, although the loadings in the INDOQUAL components are higher. This is reflected by the varimax and quartimax function values being higher for INDOQUAL than for the rotated MCA solution, and in the better

TABLE 2

INDOQUAL Loadings

	Dimension		
	1	2	3
Thread	0.99	0.00	0.00
Head	0.99	0.40	0.95
Head indentation	0.99	0.87	0.00
Bottom	0.41	0.04	0.00
Length	0.17	0.77	0.88
Brass	0.06	0.00	0.07
$\Sigma_j c_{jl}$	3.61	2.08	1.90
explained inertia ($\Sigma_{jl} c_{jl}$)		7.59	
varimax function value		2.84	
quartimax function value		6.34	

simple structure of INDOQUAL obtained at the cost of a (small) loss in explained inertia as compared to the MCA solution.

In both the INDOQUAL and the MCA solutions, the first component is highly correlated with the first three variables that are most important in distinguishing screws and bolts from nails and tacks. The component scores on these axes are pictured in Figure 1 by means of plots for the objects on the first axis of both solutions using a stem and leaf diagram in which the component scores of the objects are divided into 30 intervals. Because it is of interest to see how well the original clustering in the data appears in the solution, the objects are indicated by the letters T (tack), N (nail), S (screw), and B (bolt). It should be noted that the component scores are normalized here to unit sums of squares (instead of sums of squares equal to n). Obviously, the objects are clustered more clearly with respect to the first INDOQUAL component than with respect to the first MCA component, and the original categories appear as partly separated clusters where nails and tacks now form one cluster. The bolts and screws form different clusters, but are not well-separated. With respect to the second and third axes, similar plots could be made; one would again find a clearer clustering with respect to the

	MCA	INDOQUAL
0.28 – 0.30		
0.26 – 0.28		
0.24 – 0.26		
0.22 – 0.24	NN	
0.20 – 0.22	NNNNTTN	TNNNNNTTN
0.18 – 0.20		
0.16 – 0.18	TN	NNN
0.14 – 0.16	N	
0.12 – 0.14		
0.10 – 0.12		
0.08 – 0.10		
0.06 – 0.08		
0.04 – 0.06		
0.02 – 0.04		
0.00 – 0.02		
–0.02 – 0.00		
–0.04 – –0.02		
–0.06 – –0.04		
–0.08 – –0.06		
–0.10 – –0.08		
–0.12 – –0.10	SS	
–0.14 – –0.12	S	
–0.16 – –0.14	S	
–0.18 – –0.16	S	SSSS
–0.20 – –0.18	S	SS
–0.22 – –0.20		B
–0.24 – –0.22		BBBBB
–0.26 – –0.24	BBBBB	
–0.28 – –0.26	B	
–0.30 – –0.28		

FIGURE 1
Stem and leaf diagrams for the scores on the first MCA and INDOQUAL components.

INDOQUAL components than with respect to the MCA components, demonstrating that the INDOQUAL axes have a better discriminatory capability than the MCA axes.

References

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition. *Psychometrika, 35*, 283–319.

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

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

de Leeuw, J. (1973). *Canonical analysis of categorical data*. Unpublished doctoral dissertation, University of Leiden.

de Leeuw, J., & Pruzansky, S. (1978). A new computational method to fit the weighted Euclidean distance model. *Psychometrika, 43*, 479–490.

de Leeuw, J., & van Rijckevorsel, J. L. A. (1980). HOMALS and PRINCALS, some generalizations of principal components analysis. In E. Diday et al. (Eds.), *Data analysis and informatics II* (pp. 231–242). Amsterdam: Elsevier Science Publishers.

Escofier, B. (1979). Traitement simultané de variables qualitatives et quantitatives en analyse factorielle [Simultaneous treatment of qualitative and quantitative variables in factor analysis]. *Cahiers de l'Analyse des Données, 4*, 137–146.

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

Gifi, A. (1990). *Nonlinear multivariate analysis*. New York: Wiley.

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

Hartigan, J. A. (1975). *Clustering algorithms*. New York: Wiley.

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.

Kiers, H. A. L. (1988). Principal components analysis on a mixture of quantitative and qualitative data based on generalized correlation coefficients. In M. G. H. Jansen & W. H. van Schuur (Eds.), *The many faces of multivariate analysis, Vol. I, Proceedings of the SMABS-88 Conference in Groningen* (pp. 67–81). Groningen: Rion.

Kiers, H. A. L. (1989a). A computational short-cut for INDSCAL with orthonormality constraints on positive semi-definite matrices of low rank. *Computational Statistics Quarterly, 2*, 119–135.

Kiers, H. A. L. (1989b). INDSCAL for the analysis of categorical data. In R. Coppi & S. Bolasco (Eds.), *Multiway data analysis* (pp. 155–167). Amsterdam: Elsevier Science Publishers.

Kiers, H. A. L. (1989c). *Three-way methods for the analysis of qualitative and quantitative two-way data*. Leiden: DSWO Press.

Kiers, H. A. L. (1990). Majorization as a tool for optimizing a class of matrix functions, *Psychometrika, 55*, 417–428.

Nishisato, S. (1980). *Analysis of categorical data: Dual scaling and its applications*. Toronto: University Press.

Nishisato, S., & Sheu, W.-J. (1980). Piecewise method of reciprocal averages for dual scaling of multiple-choice data. *Psychometrika, 45*, 467–478.

Saporta, G. (1976). Quelques applications des opérateurs d'Escoufier au traitement des variables qualitatives [Several Escoufier operators for the treatment of qualitative variables]. *Statistique et Analyse des Données, 1*, 38–46.

ten Berge, J. M. F. (1983). A generalization of Kristof's theorem on the trace of certain matrix products. *Psychometrika, 48*, 519–523.

ten Berge, J. M. F. (1984). A joint treatment of varimax rotation and the problem of diagonalizing symmetric matrices simultaneously in the least-squares sense. *Psychometrika, 49*, 347–358.

ten Berge, J. M. F., Knol, D. L., & Kiers, H. A. L. (1988). A treatment of the orthomax rotation family in terms of diagonalization, and a re-examination of a singular value approach to varimax rotation. *Computational Statistics Quarterly, 3*, 207–217.

Tenenhaus, M., & Young, F. W. (1985). An analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis and other methods for quantifying categorical multivariate data. *Psychometrika, 50*, 91–119.

Manuscript received 8/24/89
Final version received 3/20/90