



# Three-mode analysis of multimode covariance matrices

Pieter M. Kroonenberg<sup>1\*</sup> and Frans J. Oort<sup>2</sup>

<sup>1</sup>Department of Education, Leiden University, Leiden, The Netherlands

<sup>2</sup>Department of Medical Psychology, AMC, University of Amsterdam, Amsterdam, The Netherlands

Multimode covariance matrices, such as multitrait-multimethod matrices, contain the covariances of subject scores on variables for different occasions or conditions. This paper presents a comparison of three-mode component analysis and three-mode factor analysis applied to such covariance matrices. The differences and similarities between the non-stochastic and stochastic approaches are demonstrated by two examples, one of which has a longitudinal design. The empirical comparison is facilitated by deriving, as a heuristic device, a statistic based on the maximum likelihood function for three-mode factor analysis and its associated degrees of freedom for the three-mode component models. Furthermore, within the present context a case is made for interpreting the core array as second-order components.

## 1. Introduction

Tucker (1966) introduced three-mode analysis to handle three-way data arrays. Best known is his work on three-mode component models for raw data matrices, but he also proposed a common factor variant to be used on multivariable-multioccasion (or multimode) covariance matrices. In this paper, we discuss and review the major three-mode developments for such matrices, compare component and factor models and present two extended examples. Recent overviews of other methods for multimode covariance matrices can be found in Wothke (1996), and Bagozzi, Yi, and Nassen (1999). The major aim is to show that these models can be fruitfully applied in realistic situations, especially because very few applications exist in the literature. C. W. Snyder and colleagues seem to have been the only authors seriously interested in

\*Request for reprints should be addressed to Pieter M. Kroonenberg, Department of Education, Leiden University, Wassenaarseweg 52, 2333 AK, Leiden, The Netherlands (e-mail: [kroonenb@fsw.leidenuniv.nl](mailto:kroonenb@fsw.leidenuniv.nl)).

applying Tucker's model for multimode covariance matrices (especially Snyder, 1976; Snyder, Bridgman, & Law, 1981), and they also presented (non-least-squares) programs to perform the basic analysis (e.g. Snyder & Law, 1979). To facilitate the comparison of component and factor models, we present and discuss measures of fit, especially for the component case, and propose to employ Lohmöller and Wold's (1982) interpretation of the core array as a second-order component matrix.

## 2. Component models

In order to appreciate the form the three-mode covariance model for multimode covariance matrices takes, we will first reiterate the basic three-mode model for the observed scores and develop the covariance model from this formulation. Moreover, this section serves to introduce the terminology and some preliminaries that are necessary for understanding several details of the covariance form of the three-mode component model.

### 2.1. Tucker3 model for observed scores

#### 2.1.1. Model description

The best-known three-mode component model is the Tucker3 model (Tucker, 1966). The basic assumption in the model is that each mode of the data has its own components and that the relationships between the components are described by the core array. The standard interpretation of the elements of the core array is that they are weights which indicate how much each of the combinations of components of the three different modes contributes to the (implied) data, or alternatively as scores of idealized subjects on latent variables under prototype conditions (see Tucker, 1966; Kroonenberg, 1983, pp. 157ff.). The observed score form of the model in matrix form, using combination-mode matrices (Tucker, 1966, p. 281), is

$$\mathbf{X} = \mathbf{A}\mathbf{G}(\mathbf{C}' \otimes \mathbf{B}') + \mathbf{E}, \quad (1)$$

where the  $I \times J \times K$  three-data array  $\underline{\mathbf{X}}$ , the  $I \times J \times K$  three-way array with residuals  $\underline{\mathbf{E}}$ , and the  $P \times Q \times R$  three-way core array  $\underline{\mathbf{G}}$  are written as combination-mode two-way matrices  $\mathbf{X}$ ,  $\mathbf{E}$ ,  $\mathbf{G}$  of order  $I \times JK$ ,  $I \times JK$ , and  $P \times QR$ , respectively. The three-way arrays are underlined and bold to distinguish them from their corresponding (two-way) matrices in which within each of the matrices the frontal slices have been placed next to each other, so that, for instance,  $j$  denotes the inner index moving fastest and  $k$  the outer index moving slowest.  $\mathbf{A}$  is the  $I \times P$  matrix with the coefficients of the subjects of the first mode on the subject components.  $\mathbf{B}$  is the  $J \times Q$  matrix with coefficients of the variables on the variable components, and  $\mathbf{C}$  the  $K \times R$  matrix with the coefficients of the conditions on the condition components. In the original  $\underline{\mathbf{X}}$  every element of the matrix,  $x_{ijk}$ , represents the value of a specific combination of levels of the original modes. In a similar manner, each element  $g_{pqr}$  of the core array  $\underline{\mathbf{G}}$  represents the value, (mutual) weight, or interaction of a specific combination of the components of the modes. In contrast with Oort (1999), we will not explicitly model a mean structure but assume that the columns of the combination mode of  $\mathbf{X}$  are centred across the subjects. All values in  $\mathbf{X}$  are divided by  $1/\sqrt{N}$ , so that  $\mathbf{X}'\mathbf{X}$  is a covariance matrix.

### 2.1.2. Identification

The model in (1) uniquely identifies the subspaces of the specified dimensionality (given unequal eigenvalues for the last components and the first one not included) as they are the maximum variance projections ( $\hat{\mathbf{X}} = \mathbf{A}\mathbf{A}'\mathbf{X}(\mathbf{C}\mathbf{C}' \otimes \mathbf{B}\mathbf{B}')$ ), but the coordinate axes of the component spaces are arbitrary and thus in that respect the model is under-identified. Any component matrix can be postmultiplied by a non-singular matrix without changing the fit of the model provided the inverse transformation is applied to the core array. Thus, as an example, any  $P \times P$  non-singular matrix  $\mathbf{U}$  may be applied to  $\mathbf{A}$ ,  $\mathbf{A}^* = \mathbf{A}\mathbf{U}$ , provided  $\mathbf{G}^* = \mathbf{U}^{-1}\mathbf{G}$ . This also means that the columnwise orthonormal  $\mathbf{A}$  does not have  $IP$  independent parameters, but only  $IP - P^2$ . During the alternating least-squares estimation procedure, the identifiability is realized by imposing orthonormality restrictions on the component matrices which leads automatically to three-way orthogonality of the core array (see Weesie and Van Houwelingen, 1983).

Note that the  $P^2$  undetermined parameters may be split into  $P^2 = \frac{1}{2}P(P+1) + \frac{1}{2}P(P-1)$ , where the number of parameters of the first term can be used to orthogonally transform  $\mathbf{A}$ , and the number of parameters in the second term to transform the core orthogonally along one of its dimensions.

### 2.1.3. Estimation

The parameters for the Tucker3 model are estimated using the least-squares discrepancy function

$$\min_{\theta} \|\mathbf{X} - \hat{\mathbf{X}}(\theta)\|^2 \quad (2)$$

where  $\theta$  is the vector of parameters consisting of the elements of  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ , and  $\mathbf{G}$ . Details can be found in Kroonenberg and De Leeuw (1980).

## 2.2. Covariance form

### 2.2.1. Model description

Given the column centring of the combination mode of  $\mathbf{X}$ , the multimode covariance form of the Tucker3 model is

$$\Sigma = \mathbf{X}'\mathbf{X} = (\mathbf{C} \otimes \mathbf{B})\mathbf{G}'\mathbf{A}'\mathbf{A}\mathbf{G}(\mathbf{C}' \otimes \mathbf{B}') + \mathbf{E}'\mathbf{E} \quad (3)$$

or, using  $\Phi = \mathbf{A}'\mathbf{A}$  and  $\Theta = \mathbf{E}'\mathbf{E}$ ,

$$\Sigma = \mathbf{X}'\mathbf{X} = (\mathbf{C} \otimes \mathbf{B})\mathbf{G}'\Phi\mathbf{G}(\mathbf{C}' \otimes \mathbf{B}') + \Theta, \quad (4)$$

with  $\Theta$  the unstructured covariance matrix of the residuals  $\mathbf{E}$ , and  $\Phi$  the covariance matrix of the subject components. To simplify the model, we may use the transformational freedom for  $\mathbf{A}$  and assume that the subject components are (columnwise) orthonormal, so that (4) can be simplified to

$$\Sigma = \mathbf{X}'\mathbf{X} = (\mathbf{C} \otimes \mathbf{B})\mathbf{G}'\mathbf{G}(\mathbf{C}' \otimes \mathbf{B}') + \Theta, \quad (5)$$

which shows that the covariance matrix does not explicitly depend on  $\mathbf{A}$  but only on the parameters  $\mathbf{B}$ ,  $\mathbf{C}$ , and  $\mathbf{G}$ . Equation (5) can be simplified even further to

$$\Sigma = \mathbf{X}'\mathbf{X} = (\mathbf{C} \otimes \mathbf{B})\Psi(\mathbf{C}' \otimes \mathbf{B}') + \Theta, \quad (6)$$

where  $\Psi = \mathbf{G}'\mathbf{G} = \mathbf{G}'\Phi\mathbf{G} = \mathbf{G}'\mathbf{A}'\mathbf{A}\mathbf{G}$  is the cross-product array of  $\mathbf{A}^* = \mathbf{A}\mathbf{G}$ . However, in the Kiers, Kroonenberg, and ten Berge (1992) approach (see below)  $\mathbf{G}$  is estimated and not  $\Psi$ .  $\Psi$  has been called the *core covariance matrix* by Lohmöller (1979), but it

may also be called the *component covariance matrix* as it is the matrix with covariances between the *QR* combination components.

As pointed out by Lohmöller (1979; see also Tucker, 1966; Bentler & Lee, 1978; Browne, 1984), when  $\Psi$  is a diagonal matrix, all combination components are uncorrelated, and  $\Sigma = \mathbf{C}\mathbf{C}' \otimes \mathbf{B}\mathbf{B}'$ . In the full-dimensional model,  $\Sigma = \mathbf{R}_C \otimes \mathbf{R}_B$ , which can be considered a *null three-mode model* for the multimode covariance matrix. This model is referred to by Browne (1984) as Swain's direct-product model. The implication is that, on the one hand, the covariances between the combination variables can be modelled in a multiplicative fashion by covariances between (components of) the variables,  $\mathbf{R}_B$ , which are constant across occasions, and, on the other hand, the covariances between (components of) occasions,  $\mathbf{R}_C$ , which are constant across variables. Any deviation from the diagonal structure of  $\Psi$  is an indication of non-constant correlation between either the variables or the occasions, or both. Lohmöller and Wold (1982) present a Monte Carlo study in which they examine how deviations from this constant situation affect the core array (for a summary, see Kroonenberg, 1983, chap. 13).

### 2.2.2. Estimation

The parameters of the covariance form of the Tucker3 model are estimated by rearranging calculations within (2), and not by minimizing the least-squares difference  $\|\Sigma - \hat{\Sigma}\|^2$ . In particular, Kiers *et al.* (1992) noted, after examining the estimation of the least-squares discrepancy function (2), that the component matrix  $\mathbf{A}$  is only involved in the minimizing process in combination with  $\mathbf{X}$ , via  $\Lambda = \mathbf{X}'\mathbf{A}$ , and that  $\mathbf{X}$  only occurs via  $\mathbf{X}'\mathbf{X}$ . Thus instead of using raw data directly, the iterations can be based on the covariances  $\mathbf{X}'\mathbf{X}$  and the matrix  $\Lambda$ . Because of this, the algorithm for the Tucker3 model can be redesigned so that it only uses the multimode covariance matrix rather than the original observed scores (Kiers *et al.*, 1992; see also Murakami, 1983, for the Tucker2 model). The only drawback from the point of view of observed score modelling is that the subject scores  $\mathbf{A}$  are not available but only the covariances of the original variables with the subject components in  $\Lambda = \mathbf{X}'\mathbf{A}$ . However, if the observed scores are available,  $\mathbf{A}$  can be recovered from  $\Lambda$  and  $\mathbf{X}$  after convergence (for details see Kiers *et al.*, 1992). Because the revised algorithm operates directly on multimode covariance matrices, three-mode component analysis becomes a useful alternative for covariance structure analysis via structural equation modelling in very difficult circumstances, such as few observations, very large numbers of variables, or inadequate fulfilment of distributional assumptions.

### 2.3. Interpretation of the core array: Second-order component loadings

The core array can be interpreted in several ways (see Kroonenberg, 1983, pp. 157ff.), but here we will concentrate on the interpretation as second-order components. Starting from (1), we obtain  $\mathbf{X} = \mathbf{A}^*\mathbf{F}' + \mathbf{E}$  with  $\mathbf{A}^* = \mathbf{A}\mathbf{G}$  and  $\mathbf{F} = (\mathbf{C} \otimes \mathbf{B})$ . The covariance form is  $\mathbf{X}'\mathbf{X} = \mathbf{F}\mathbf{A}^*\mathbf{A}^*\mathbf{F}' + \mathbf{\Theta} = \mathbf{F}\Psi\mathbf{F}' + \mathbf{\Theta}$ , where  $\mathbf{F}$  is the first-order component loading matrix and  $\mathbf{A}^*$  the first-order component score matrix, with  $\Psi$  as the first-order component covariance matrix. Now  $\Psi = \mathbf{G}'\mathbf{A}'\mathbf{A}\mathbf{G} = \mathbf{G}'\Phi\mathbf{G}$ ,  $\mathbf{G}$  being the (transposed) second-order loading matrix and  $\mathbf{A}$  the second-order component score matrix, with  $\Phi$  the second-order covariance matrix (see Lohmöller, 1979, for an earlier exposition of this point).

## 2.4. Software

The component models have been fitted by a program (T3Covar) based on the Kiers *et al.* (1992) algorithm which will be included in the next release of the 3WayPack (Kroonenberg, 1996) package for three-mode data analysis. For further information, see <http://three-mode.leidenuniv.nl>.

## 3. Factor models

### 3.1. Tucker–Bloxom models

Tucker (1966, pp. 301ff.) also described a three-mode common-factor model with unique variances for combination variables. Snyder (1968) produced an unpublished master's thesis under the direction of Tucker in which unique variances were also defined for the measures and the conditions. Bloxom (1968) first formulated a truly stochastic three-mode data model,

$$\mathbf{x}' = \xi' \mathbf{G}(\mathbf{C}' \otimes \mathbf{B}') + \boldsymbol{\varepsilon}', \quad (7)$$

where  $\mathbf{x}$  and  $\boldsymbol{\varepsilon}$  are  $JK \times 1$  random vectors of the observed scores and the unobserved residuals respectively, and  $\xi$  the  $P \times 1$  random vector of unobserved scores with their covariance matrices given by  $\Sigma$ ,  $\Theta$ , and  $\Phi$  respectively, so that the multimode covariance matrix is modelled as

$$\Sigma = (\mathbf{C} \otimes \mathbf{B})\mathbf{G}'\Phi\mathbf{G}(\mathbf{C}' \otimes \mathbf{B}') + \Theta, \quad (8)$$

where  $\Theta$  contains the residual variances on the diagonal and non-zero off-diagonal elements in case of correlated residuals. Note that we have two fixed modes and one random mode. In general, the first (random) mode is associated with the subjects and the second and third (fixed) modes with variables and occasions, respectively. The latter modes are characterized by the terms *slow* and *fast*, where slow and fast refer to the index of the mode. This distinction corresponds to Tucker's (1966) usage of outer mode and inner mode, respectively. In (7),  $\mathbf{B}$  is the fast mode and  $\mathbf{C}$  is the slow mode because of the way the data are organized, with the  $\mathbf{B}$  mode nested within the  $\mathbf{C}$  mode. The three-mode factor model (8) can be written as a special case of the ordinary common factor model,

$$\Sigma = \Lambda\Phi\Lambda' + \Theta, \quad (9)$$

with the matrix of factor loadings  $\Lambda$  restricted to

$$\Lambda = (\mathbf{C} \otimes \mathbf{B})\mathbf{G}', \quad (10)$$

so that standard programs for structural equation modelling can be used to estimate the parameters of the three-mode factor model. Contributions to three-mode modelling of multimode covariance matrices can be found in Bentler and Lee (1978, 1979), Bentler, Poon, and Lee (1988), Bloxom (1968), McDonald (1980; especially p. 165), Lee and Fong (1983), Verhees and Wansbeek (1990), and Oort (1999, 2001).

### 3.2. Restrictions

The basic three-mode common factor model (3MFM) is generally underidentified, so that restrictions need to be imposed. Moreover, models with specific restrictions can be more easily interpreted, such as those with simple structures for  $\mathbf{B}$  and  $\mathbf{C}$ . Another way to restrict the three-mode factor model (8) is to specify not only a Kronecker (or direct-product) structure for the factor loadings via  $(\mathbf{B} \otimes \mathbf{C})$ , but also similar structures for the

covariance matrix  $\Phi$  and the residual covariance matrix  $\Theta$  (see Oort, 1999)—for example,

$$\Theta = \Theta_C \otimes \Theta_B, \tag{11}$$

where  $\Theta_C$  and  $\Theta_B$  are  $K \times K$  and  $J \times J$ . Such a restriction, employed by Browne (1984) for direct-product models, saves a considerable number of degrees of freedom.

On the other hand, sometimes it is necessary to increase the number of model parameters due to covarying residual factors. That is, residual factors of the same variables may be correlated across occasions, and one or more lags for the auto-correlation may be included. This could also be accomplished with the Kronecker product restriction, for example,  $\Theta_C$  symmetric and  $\Theta_B$  diagonal.

Another way to model the covariance matrix of residual factors  $\Theta$ , especially in the case of repeated measurements, as in our second example, is to impose an autoregressive (AR) structure, that is,

$$\Theta = (\mathbf{I}_\Theta - \mathbf{B}_\Theta)^{-1} \Psi_\Theta (\mathbf{I}_\Theta - \mathbf{B}'_\Theta)^{-1}, \tag{12}$$

where  $\mathbf{I}_\Theta$  is an identity matrix,  $\mathbf{B}_\Theta$  is a subdiagonal matrix containing parameters for the autoregressive effects, and  $\Psi_\Theta$  is a diagonal matrix of residual variances, also called variances of the innovation factors. Matrices  $\mathbf{I}_\Theta$ ,  $\mathbf{B}_\Theta$ , and  $\Psi_\Theta$  are of the same order as  $\Theta$ , that is,  $JK \times JK$ . In some applications, one might even combine the structures of (11) and (12) by specifying the autoregressive structure for  $\Theta_B$  or  $\Theta_C$  rather than for  $\Theta$  itself.

### 3.3. Identification

A necessary condition for achieving the identification of a three-mode factor model is that all common factors  $\xi$  have a scale. Assuming that the core array  $\underline{\mathbf{G}}$  does not contain free parameters, the factors  $\xi$  can be given scales either by fixing  $Q + R$  factor loadings at a non-zero value, one element in each column of  $\mathbf{B}$  and  $\mathbf{C}$ , or by fixing  $P$  factor variances, that is, the diagonal elements of the  $\Phi$  matrix. In the latter case, there is an indeterminacy in the constraint of (10) that must be removed by fixing one of the elements of either matrix  $\mathbf{B}$  or matrix  $\mathbf{C}$  at a non-zero value. The same goes for other Kronecker product constraints, such as the  $\Theta$  restriction of (11)—for details, see Oort (1999, Table 1).

When three-mode factor models are used in an exploratory context, that is, without explicitly imposing an hypothesized structure, the number of components has to be determined through a series of trials, and (subject) factors are generally specified as orthogonal so that their covariance matrix  $\Phi$  is diagonal. The  $\mathbf{B}$  and  $\mathbf{C}$  matrices can be chosen to be of echelon form, while the core array  $\underline{\mathbf{G}}$  can then be chosen to be either diagonal or of echelon form as well. For interpretation, however, the matrices often have to be transformed—see Bentler and Lee (1978, 1979) or Oort’s (1999) summary.

### 3.4. Estimation

Assuming a multivariate normal distribution for the observed variables  $\mathbf{x}$ , and assuming that the model is identified, maximum likelihood estimates for all parameters can be obtained by minimizing the maximum likelihood function

$$F(\mathbf{S}, \hat{\Sigma}) = \ln |\hat{\Sigma}| - \ln(\mathbf{S}) + \text{tr}(\mathbf{S}\hat{\Sigma}^{-1}) - JK, \tag{13}$$

where  $\hat{\Sigma}$  is the implied covariance matrix,  $\mathbf{S}$  the observed covariance matrix and  $JK$  the number of variables. Minimization of the maximum likelihood function gives estimates

for all model parameters. The estimation procedure also provides estimates for the asymptotic standard errors of the model parameters, and a chi-square test of overall goodness of fit. Even if the assumption of normality is not met, we still estimate the three-mode factor model parameters through the maximum likelihood function (13). However, the standard errors should not be interpreted and the chi-square distribution should not be used rigorously when evaluating the goodness of fit index. But there is no reason to believe that point estimates of the model parameters are seriously biased (Bollen, 1989, pp. 415ff.), and we can still use the root mean square error of approximation and Akaike's information criterion (see below) to compare the fit of different models to the same data.

### 3.5. Fit statistics

We will report the chi-square measure of overall goodness of fit ( $\chi^2$ ), the root mean square error of approximation,  $RMSEA = \sqrt{(\chi^2/df - 1)/(N - 1)}$ , and Akaike's information criterion,  $AIC = \chi^2 - 2df = df\{(\chi^2/df) - 2\}$ . The RMSEA and the AIC can be used to compare the fit of different models to the same data. As a rule of thumb, RMSEA values smaller than 0.05 are indicative of close fit but values smaller than 0.08 are still considered reasonable (Browne & Cudeck, 1992). Simulation studies suggest that under non-normality  $\chi^2$  will be overestimated (Curran, West & Finch, 1996), so that the RMSEA and AIC values will turn out too high. We will therefore apply the RMSEA rule of thumb conservatively.

### 3.6. Software

Three-mode factor models can be fitted with standard software. McDonald (1980) and Bentler *et al.* (1988) have shown that their three-mode factor models can reparameterized in such a way that the model parameters can be estimated with commercial computer programs such as LISREL (Jöreskog & Sörbom, 1996; Jöreskog, Sörbom, du Toit, & du Toit, 1999; <http://www.ssicentral.com/lisrel/mainlis.htm>) and EQS (Bentler, 1995; <http://www.mvsoft.com>). However, neither program is as versatile as Mx (Neale, 1997; <http://griffin.vcu.edu/mx>), which is freely available through the internet. In fact, the fitting of three-mode factor models with Mx can be done in a straightforward fashion without reparameterization. We have used Mx throughout for all of the three-mode factor models of interest.

## 4. Component versus factor models

The choice between component and factor models depends on several aspects of a study, such as the type of research questions, the assumptions one is willing to make about the data, the number of subjects, variables and conditions in the data set, the amount of prior knowledge, etc. It is not uncommon that not all aspects will favour one particular approach and that one has to make do with a component model when a factor model was intended.

### 4.1. Sample versus population

In factor analysis, it is explicitly assumed that the subjects at hand are a (random) sample from a specific population, and the parameters of the model are estimated on the basis

of this sample. If the model is valid, then the parameter estimates approximate the population values. In principal components analysis (PCA), the analysis of the sample is carried out without explicitly assuming that the sample is a random one. The component model is not considered to be an explicit model for reality, but a parsimonious description of the relationships in the sample. However, the researcher will expect to find similar results in similar samples.

#### **4.2. Estimation and distributional assumptions**

The parameters of three-mode component models are estimated through a least-squares discrepancy function (2). With three-mode factor models, it is also possible to use the maximum likelihood function (13) through which standard errors and a chi-square measure of overall goodness of fit can be obtained. However, the use of the maximum likelihood estimation method requires distributional assumptions not required by the least-squares method. If these distributional assumptions are not met, we must revert to least-squares estimates for the factor model parameters as well.

#### **4.3. Interpretation and poor fit**

One important difference is that factor models are commonly only interpreted when a good or reasonable fit is achieved. Lack of fit is seen as far less a problem in component models because a maximum variance projection onto a lower-dimensional space is always possible and valid. In both models high values of fit are obviously preferred, but when the data are very noisy, that is, contain a large amount of error, a component solution may still describe the majority of the systematic information even when the fit of the model is relatively modest.

#### **4.4. Exploratory and description versus confirmatory and testing**

The choice between three-mode component and factor models also depends upon on whether one wants to reduce the dimensionality of the data in order to describe and explore the patterns present or whether one aims to confirm and test explicitly restricted models for the data. Given that one wants to build models and has a theory for them, one could use the component approach as a preliminary check on the dimensionality and to find adequate starting values. One could also use the component models to form an idea of what the structure in the data is like and use the confirmatory factor analysis approach to test the adequacy of the model 'gleaned' from the component model. However, three-mode factor analysis can also be used as an exploratory technique (Bentler & Lee, 1978, 1979; Bentler, Poon, & Lee, 1988; Lee & Fong, 1983).

#### **4.5. Restricting the numbers of parameters**

Component analysis, with its search for low-dimensional subspaces, typically favour fewer components than factor models while by limiting the number of coefficients per factor, the latter increase the number of factors without drastically increasing the number of parameters. Parsimony is thus achieved in component models by limiting the number of components, and in factor models by using larger numbers of sparse factors. Recently, there have been developments to introduce restrictions into the model part of component models that blur some of these differences—see, for



instance, for two-mode PCA, Takane, Kiers, and De Leeuw (1995) and Kiers, Takane, and ten Berge (1996).

#### 4.6. Fitting residuals and (co)variances

Irrespective of the imposition of restrictions, a fundamental difference remains that in factor analysis the residuals are explicitly modelled, while this is not the case in component analysis, and that the former aims to model the covariances as well as possible while the latter only aims to explain as much variance as possible.

### 5. Model comparison

Component and factor models are based on conceptually different points of view. However, the model part of the covariance form of the Tucker3 model is equal to the three-mode factor model, so that it is possible to use it as an alternative in particular difficult cases, such as (i) covariance matrices with relatively small numbers of subjects or large number of variables, (ii) covariance matrices for which no clear *a priori* structure is known or hypothesized, (iii) when the implied covariance matrix does not fit the observed covariance matrix very well, or (iv) when the distributional assumptions are clearly untenable.

#### 5.1. Measures of fit

In order to evaluate the relative performance of different models on the same data, the root mean square residual (RMR) for the off-diagonal elements of the covariance matrix can be used. For comparisons across data sets the correlation matrix is more appropriate. The advantage of the RMR is that, in either model, the covariances are estimated on the basis of the model used and differences between the estimates and the observed values can be assessed in both cases irrespective of the method by which the estimates were derived. Another measure which can be computed for both models is the proportion of variance explained,  $V = \text{tr}(\Lambda\Phi\Lambda')/\text{tr}\mathbf{S}$ , which is equal to the proportion of explained variance from the alternating least-squares procedure for the Tucker3 model, albeit that this is the criterion explicitly optimized in the component case.

Further, purely heuristically, comparisons between component and factor models can also be based on statistics calculated from the maximum likelihood discrepancy function  $F$  in (13). Generally in maximum likelihood fitting, the diagonals of the observed and the implied covariance matrices are equal and therefore this equality is employed for the component case as well. This also solves the problem that the implied covariance matrix  $\hat{\Sigma}$  is rank-deficient due to the reduced rank of the component matrices in the model. Thus for component models we also used  $F$  to obtain the common measures of fit of the stochastic approach such as the RMSEA and the AIC. Of course, no distributional assumptions are made in the component case. The values for the 'maximum likelihood' statistic for the component models will probably be comparatively higher, as the maximum likelihood discrepancy function was not used for minimization. On the other hand, the residual covariances are unstructured, which might lead to lower fit compared to factor models.

The number of degrees of freedom to be used for the covariance form of the Tucker3 model is determined as the difference between the number of independent values in the sample covariance matrix,  $JK/(JK + 1)/2$ , and the number of estimated parameters. In

the observed score algorithm, not only are  $\mathbf{B}$ ,  $\mathbf{C}$ , and  $\mathbf{G}$  estimated but also  $\Lambda = \mathbf{X}'\mathbf{A}$ . However,  $\Lambda$  is not involved in calculating  $\hat{\Sigma}$ ; moreover, because of the orthonormality of  $\mathbf{A}$ ,

$$\Lambda = \mathbf{X}'\mathbf{A} = (\mathbf{C} \otimes \mathbf{B})\mathbf{G}'\mathbf{A}'\mathbf{A} = (\mathbf{C} \otimes \mathbf{B})\mathbf{G}', \quad (14)$$

and thus  $\Lambda$  does not include independent parameters. Therefore, the number of parameters (first discussed by Weesie & Van Houwelingen, 1983, for the observed score Tucker3 model) becomes  $(JQ - Q^2) + (KR - R^2) + (PQR - \frac{1}{2}P(P - 1))$ . The first two terms need corrections because  $\mathbf{B}$  and  $\mathbf{C}$  are determined up to non-singular transformations; the correction for the third term referring to  $\mathbf{G}$  is introduced because we may transform  $\mathbf{G}$  orthogonally along its first mode without affecting the fit of the model. The other  $\frac{1}{2}P(P + 1)$  restrictions on  $\mathbf{A}$  were used for the orthonormality of  $\mathbf{A}$  when we went from (3) to (4).

Timmerman and Kiers (2000; see also Kiers & der Kinderen, 2003) suggested a model-selection procedure analogous to Cattell's scree plot for two-mode component analysis. In particular, they based the selection on choosing the model with the highest proportion of fitted sums of squares,  $V_S$ , within the class of models with the same sum of numbers of components ( $S = P + Q + R$ ). To compare classes with different  $S$ , they computed  $dif_S = V_S - V_{S-1}$ . Due to the non-nestedness of three-mode component models,  $dif_S$  cannot always be compared for successive values of  $S$ . Only those  $dif_S$  were considered which were sequentially highest. Timmerman and Kiers defined a *salience value*,  $b_S = dif_S/dif_{S^*}$ , where  $dif_{S^*}$  has the highest value after  $dif_S$ . They proposed to select the model for which  $b_S$  has the highest value, and they call this the *DIFFIT* criterion. Finally, the authors defined a lower bound for  $dif_S$  to be taken into account. The  $dif_S$  should be greater than the average proportion of explained variability taken over all feasible values of  $S$  ( $s_{\max} = \min(I, JK) + \min(J, IK) + \min(K, IJ) - 3$ ; in Timmerman & Kiers, 2000 'max' is inadvertently printed instead of 'min'). Note, by the way, that  $s_{\max}$  has to be corrected in case a mode has been centred. For instance, in the first example,  $J$  should be replaced by  $J - 1$ . The Timmerman and Kiers approach cannot be used for confirmatory factor models, because there is no provision for restrictions on configurations as  $S$  is only based on the number of components. It would probably be worthwhile to supplement their study using the *degrees of freedom* rather than  $S$ . (In our two examples the correlation between  $df$  and  $S$  was  $-.85$  and  $-.92$ , respectively.)

## 5.2. Model selection

It does not always make sense to search for the true or best model. Often several models are acceptable from the fitting point of view, and, especially in component models, the choice of a model is often determined by the amount of detail one wants to consider. Generally, substantive considerations may guide the selection as well. Our model search will primarily be one of exclusion, in the sense of excluding models from further consideration because other models fit better with fewer or equal numbers of parameters. Three types of plots can be used to evaluate both the models and the measures themselves, as they allow insight into the behaviour of both of them: fitted sums of squares  $V$  versus RMR ( $V$ -RMR plot), fitted sums of squares  $V$  versus sums of numbers of components  $S$  (three-mode scree plot), and deviance plots with  $\chi^2$  versus degrees of freedom  $df$  (deviance plot). In these plots, curves are drawn by connecting component models with the same  $(Q, R)$  combinations and increasing  $P$ . Such models are nested with respect to both their fitted sums of squares and  $df$  (or  $S$ ). Furthermore, for the last two plots, a convex hull can be drawn to connect favoured models. The

general idea is that models within the hull are disfavoured compared to the 'hull models' that have similar or better fit and equal  $df$  (or  $S$ ), equal fit with more  $df$  (or smaller  $S$ ), or a combination of both.

The least controversial measures are  $V$  and RMR, because they can be calculated straightforwardly for both fitting methods. A plot of these two quantities gives insight in how both the differences in variances and covariances are reduced in increasingly complex models. Moreover, it can be discerned which of the two measures gains more and whether the gain of one is at the expense of the other. The plot of  $V$  versus  $S$  gives insight in the behaviour of Timmerman and Kiers' proposal for model selection. In the deviance plots of  $\chi^2$  versus  $df$ , model evaluation is aided by drawing straight lines between the loci of models with equal  $\chi^2/df = k$ . If the origin is included in such a plot, these lines can be seen to fan out from the origin with slope  $k$ . These lines also represent models with same RMSEA, as, for constant  $k$ ,  $RMSEA = \sqrt{\{(k - 1)/(N - 1)\}}$ . The loci of models with equal AIC can be drawn in such plots as  $AIC = \chi^2 - 2 \times df$  or  $\chi^2 = AIC + 2 \times df$ , so that they are straight lines with a constant slope of 2 and a different intercept for each AIC. Thus in a single plot the relative performance of (non-nested) models can be assessed with respect to  $\chi^2/df$ , RMSEA, and AIC. Verbeek (1983) gives a statistical discussion of deviance plots; Fowlkes, Freeny, and Landwehr (1988) give some practical considerations for its use.

## 6. Example: Economic activities data

### 6.1. Data

The first example is based on a study on similarities and differences in perceptions of economic activities in Hungary, the Netherlands, Poland, and the United Kingdom. The original three-mode data set consists of 390 individuals (mode 1) scoring 20 economic activities (mode 3) on 12 bipolar seven-point rating scales (mode 2); for details, see Antonides, Farago, Ranyard, and Tyszka (1997) and Veldscholte, Kroonenberg, and Antonides (1998). Here we use the same subset Oort (1999) selected to illustrate his stochastic three-mode models. In particular, he selected seven of the rating scales (see Table 2) and nine of the economic activities (see Table 3) on the basis of context and earlier analyses such that the rating-scale mode and the economic activity mode had a clear structure and interpretation. In particular, the scales were chosen to represent social values, economic value, and expected risk (see Table 2), while the activities were chosen to represent being employed, being an owner, uncontroversial financial activities, and controversial financial activities (see Table 3). A selection was necessary because analysing the covariance structure of 240 variables with structural equation models is not feasible with present-day personal computers.

The nature of the data and the way the subset was selected clearly call for an analysis with factor models rather than component models. We are dealing with random variables and we are looking for the confirmation of a specific structure in the loadings.

### 6.2. Three-mode factor analysis

Two types of models were fitted, exploratory factor models and confirmatory ones. The former followed Bentler and Lee (1978, 1979) by specifying echelon forms for the  $7 \times 3$  factor matrix  $\mathbf{B}$  and the  $9 \times 4$  factor matrix  $\mathbf{C}$  (taking care that not all the zeros are associated with the variables of one factor) and with either an echelon form for the

**Table 1.** Economic activities data: Fit measures for three-mode factor and component models (390 subjects × 7 scales × 9 economic activities)

Model	S = P + Q + R	SS (Fit)			$\chi^2$	df <sup>a</sup>	$\chi^2/df$	RMSEA <sup>b</sup>	AIC <sup>c</sup>	RMR <sup>d</sup>
		V	dif <sub>S</sub>	b <sub>S</sub>						
<i>Exploratory factor models</i>										
E7.3.4	14	.25			5522	1858	2.97	.071	1806	.25
<b>E12.3.4</b>	<b>19</b>	<b>.30</b>			<b>5496</b>	<b>1906</b>	<b>2.88</b>	<b>.069</b>	<b>1684</b>	<b>.27</b>
<i>Confirmatory factor models</i>										
C7.3.4	14	.25			5964	1911	3.12	.074	2142	.26
C12.3.4 <sup>e</sup>	19	.29			5678	1866	3.04	.072	1946	.25
<b>C12.3.4a<sup>f</sup></b>	<b>19</b>	<b>.29</b>			<b>5985</b>	<b>1929</b>	<b>3.05</b>	<b>.073</b>	<b>2127</b>	<b>.28</b>
C12.3.4b <sup>g</sup>	19	.27			6660	1977	3.37	.078	2706	.29
<i>Component models</i>										
1 × 1 × 1	3	.0936	.0936	1.45	8852	2001	4.42	.094	4850	.34
2 × 2 × 1	5	.1582	.0646	1.08	7989	2016	4.00	.088	3999	.31
2 × 2 × 2	6	.1630	.0048	–	7957	1985	4.01	.088	3987	.31
<b>3 × 2 × 2</b>	<b>7</b>	<b>.2226</b>	<b>.0596</b>	<b>2.14</b>	<b>7030</b>	<b>1983</b>	<b>3.55</b>	<b>.081</b>	<b>3064</b>	<b>.28</b>
4 × 2 × 2	8	.2404	.0278	1.03	7107	1982	3.59	.082	3143	.28
4 × 3 × 2	9	.2675	.0271	1.11	6539	1972	3.32	.077	2595	.27
5 × 3 × 2	10	.2864	.0189	–	6644	1970	3.37	.078	2704	.27
5 × 3 × 3	11	.3101	.0237	–	6298	1951	3.23	.076	2396	.25
6 × 3 × 3	12	.3344	.0243	–	6192	1947	3.18	.075	2298	.25
7 × 3 × 3	13	.3532	.0188	–	6279	1944	3.23	.076	2391	.25
<b>7 × 4 × 3</b>	<b>14</b>	<b>.3776</b>	<b>.0244</b>	<b>1.21</b>	<b>5857</b>	<b>1923</b>	<b>3.05</b>	<b>.073</b>	<b>2011</b>	<b>.23</b>
8 × 4 × 3	15	.3978	.0202	–	5909	1918	3.08	.073	2072	.23
9 × 4 × 3	16	.4151	.0173	–	6025	1914	3.15	.074	2197	.23
9 × 4 × 4	17	.4333	.0182	–	5581	1876	2.98	.071	1829	.22
<i>Additional component models</i>										
7 × 3 × 4	14	.37			6059	1921	3.15	.074	2217	.24
12 × 3 × 4	19	.44			6649	1906	3.49	.080	2837	.25
13 × 5 × 4	22	.52			5514	1804	3.06	.073	1906	.20
13 × 6 × 8	27	.56			4167	1456	2.86	.069	1255	.14
48 × 6 × 8	62	.90			4487	826	5.43	.107	2835	.10
63 × 7 × 9	79	1.00			0	0	0.00	.000	0	.00

<sup>a</sup>  $df = JK(JK + 1)/2 - [(J \times Q - Q^2) + (K \times R - R^2) + (P \times Q \times R - \frac{1}{2}P(P - 1))]$ .

<sup>b</sup>  $RMSEA = \sqrt{[(\chi^2/df - 1)/(N - 1)]}$ .

<sup>c</sup>  $AIC = \chi^2 - 2df$ .

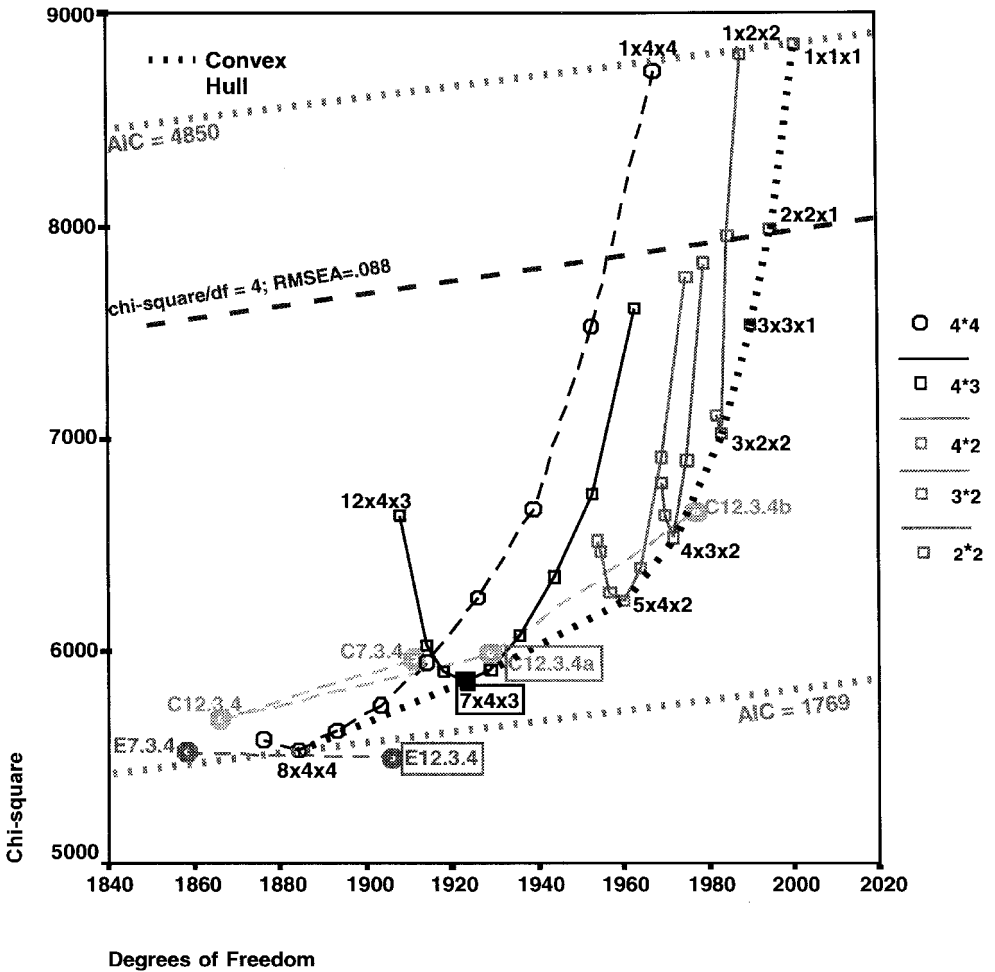
<sup>d</sup> RMR = root mean square residual (off-diagonal elements); RMR of the original covariance matrix is 0.40.

<sup>e</sup> **G** diagonal.

<sup>f</sup> Additional  $\Phi$  restriction.

<sup>g</sup> Additional  $\Theta$  restriction.

7 × 12 core array ( $P = 7$ ;  $P = Q + R$ ) or a 12 × 12 diagonal matrix ( $P = 12$ ;  $P = QR$ ). In the latter case, there is a one-to-one relationship between all  $Q \times R$  combinations of the first-order factors and the second-order factors as their numbers are equal and the core array is diagonal. The confirmatory models had simple structures for the two fixed models **B** and **C** (see Tables 2 and 3, respectively). Two more restricted confirmatory models were derived from the  $P = 12$  case: one with a Kronecker (or direct-product)



**Figure 1.** Economic activities data: Deviance plot. Component models with a constant number of components in the second mode ( $Q$ ) and in the third mode ( $R$ ) are connected for increasing number of second-order components ( $P$ ). The  $7 \times 4 \times 3$  model is the one selected for interpretation. The dashed line represents the convex hull of models which pair lowest chi-squares with highest degrees of freedom. Examples of lines of both constant  $\chi^2/df$  and constant RMSEA are also indicated, as well as lines of constant AIC. The boxed models are the ones selected for interpretation.

structure imposed on the subject-factor covariance matrix  $\Phi$  (C12.3.4a) (similar to (11)), and the other with an additional Kronecker structure on the residual covariance matrix  $\Theta$  (C12.3.4b) (see (9)). In terms of measures of fit, the 'best' exploratory model is E12.3.4. With respect to the confirmatory models, the C12.3.4a model, with the factor covariance matrix having a Kronecker form, is the favoured candidate because of its attractive interpretational characteristics, albeit that C12.3.4 has a better fit (see Table 1).

Neither  $V$ -RMR plots nor three-mode scree plots seem helpful in model selection for factor models. The  $V$ -RMR plot is based on the RMR and the fitted sum of squares, and no attention is paid to the number of parameters that need to be estimated for a model.

In the three-mode scree plot only the numbers of factors are counted and thus it is irrelevant how many of the elements in a factor are constraints and how many are free to be estimated. The deviance plot (Fig. 1) shows the relative positions of the three-mode factor models; it illustrates the superiority of the C12.3.4a and C12.3.4b factor models as both lie more or less on the convex hull, while the other factor models lie well within it. The exploratory factor models fare very well, but their interpretation appears difficult.

The interpretations of the factors in the confirmatory factor model, both for the rating scales (Table 2) and for the economic activities (Table 3) were already clear from the outset, but it is satisfying to see that all variables have high loadings on their designated factors.

**Table 2.** Economic activities data: Factors from the 12 × 3 × 4 factor model (C12.4.3a) and varimax-rotated rating scale components from the 7 × 4 × 3 component model

Rating scales	Factor analysis factors			Component analysis rotated components			
	Social values	Economic values	Expected risk	Social values	Economic values	Well-known	Non-risky
1 Moral	<b>.89</b>	–	–	<b>.86</b>	–.09	–.06	–.06
2 Beneficial for society	<b>.61</b>	–	–	<b>.49</b>	.12	.06	.06
3 Requires great effort	–	<b>.79</b>	–	.00	<b>.62</b>	–.03	–.05
4 Requires much knowledge	–	<b>.78</b>	–	–.05	<b>.67</b>	.02	.17
5 Requires a lot of financial resources	–	<b>.48</b>	–	.12	<b>.38</b>	.00	–.14
6 Well-known	–	–	<b>1.00</b>	.02	–.01	<b>1.00</b>	–.01
7 Non-risky	–	–	<b>.31</b>	.05	–.05	–.01	<b>.97</b>
Proportion of explained variance				.10	.13	.10	.06

Note: Factor correlations are  $r$  (Social Values, Economic Values) = .40;  $r$  (Social Values, Expected Risk) = .36;  $r$  (Economic Values, Expected Risk) = .20.

**6.3. Three-mode PCA: Model selection**

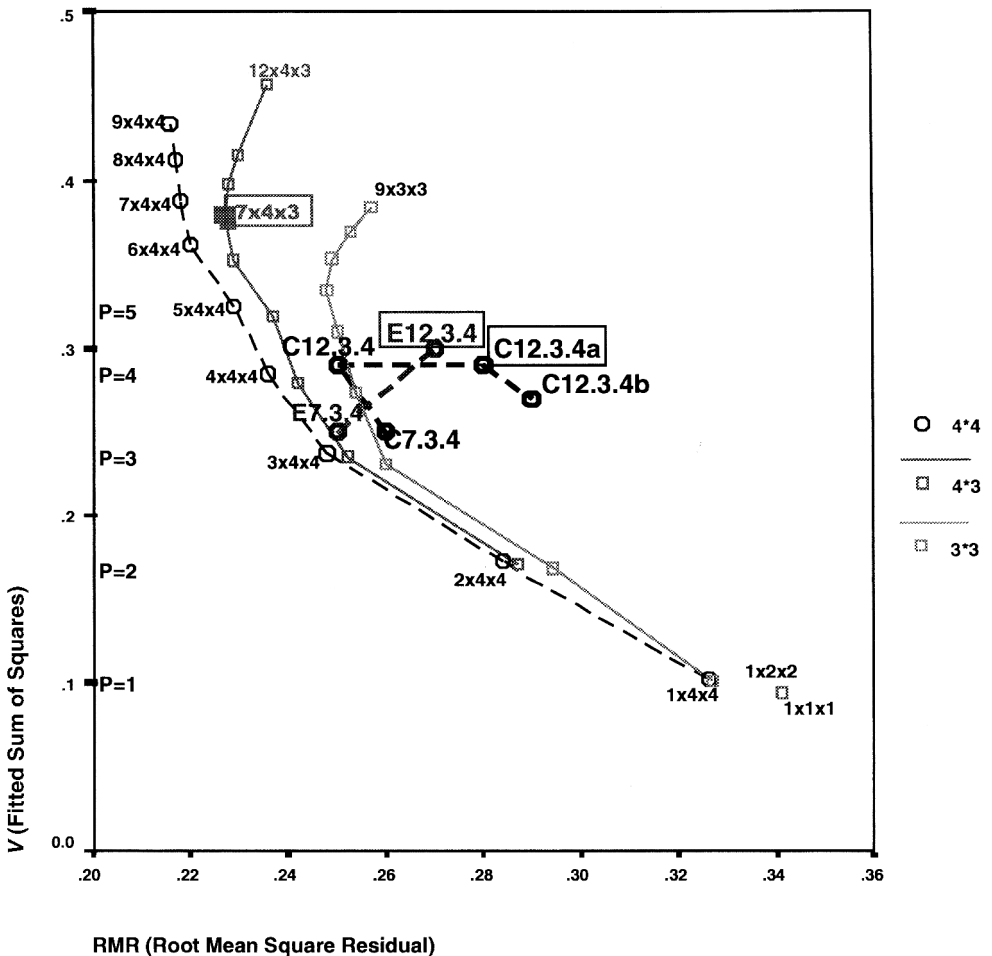
Because of the differences in factor and component models, it is not *a priori* self-evident that the number of factors and the number of components are the same. Therefore, a full model search for small to medium numbers of components was carried out. Using standard scree plots, a first indication of the number of components followed from two-mode PCAs on the original multimode covariance matrix without regard to its structure (7 components), the average scales covariance matrix (3 components) and the average activities covariance matrix (2–3 components). Table 1 lists the fit measures of selected component and factor models for the economic activities data. The RMSEA measures suggest a more or less reasonable fit is possible (0.07–0.08), but the proportion of variance explained is low for reasonable numbers of components (between 0.09 (1 × 1 × 1) and 0.44 (12 × 3 × 4)). Veldscholte *et al.* (1998) noticed that there was a large amount of noise in the raw data. However, this does not mean that there is no recoverable structure in the data, and in the covariances in particular.

**Table 3.** Economic activities data: Economic activity factors from the  $12 \times 3 \times 4$  factor model and varimax-rotated economic activity components from a  $7 \times 4 \times 3$  component model

Economic activities	Factor analysis: Economic activity factors				Component analysis: Economic activity components			
	Being employed	Being an owner	Uncontroversial financial activities	Controversial financial activities	Work—earning money	Uncontroversial financial activities	Controversial financial activities	
1 Being employed as a teacher	<b>1.00</b>	—	—	—	<b>.47</b>	.02	— .08	
2 Being employed as a house painter	<b>.61</b>	—	—	—	<b>.48</b>	.04	— .04	
3 Being the owner of a newspaper	—	<b>1.24</b>	—	—	<b>.53</b>	— .04	.07	
4 Being the owner of a night club	—	<b>1.33</b>	—	—	<b>.46</b>	— .03	.20	
5 Buying a life insurance	—	—	<b>0.99</b>	—	— .05	<b>.60</b>	— .10	
6 Buying government bonds	—	—	<b>1.23</b>	—	— .09	<b>.60</b>	.18	
7 Paying taxes	—	—	<b>1.30</b>	—	.13	<b>.51</b>	.01	
8 Gambling in a casino	—	—	—	<b>1.03</b>	.12	.10	<b>.40</b>	
9 Receiving bribes	—	—	—	<b>1.71</b>	— .13	— .10	<b>.86</b>	
Proportion of explained variance					.14	.12	.13	
Factor correlations								
Being employed	1.00							
Being employer	.63	1.00						
Uncontroversial financial activities	.56	.46	1.00					
Controversial financial activities	.17	.38	.42	1.00				

The plot of the fitted sums of squares versus RMR (Fig. 2) shows that the major gains in fit depend on the number of second-order components  $P$ , and given  $P$ , increasing numbers of components for  $Q$  and  $R$  lead to decreases in RMR. Given models with equal ( $Q, R$ ), the gains in fitted sums of squares are at the expense of increasing RMR when  $P$  keeps on increasing beyond a certain value. As one would expect in PCA, fit in variances is given preference over fit in covariances. Given that we are looking at covariance matrices, models above the turning point are less attractive. For choosing a particular model the plot is less suitable, because there is no control over the number of parameters. The only thing that can clearly be seen is that the gain for increasing  $P$  gradually diminishes. The gains between maxima for a given  $P$  are 7.1% ( $1 \rightarrow 2$ ), 6.5% ( $2 \rightarrow 3$ ), 4.7% ( $3 \rightarrow 4$ ), 4.1% ( $4 \rightarrow 5$ ), 3.6% ( $5 \rightarrow 6$ ), so that no more than five second-order components seems indicated given the models examined.

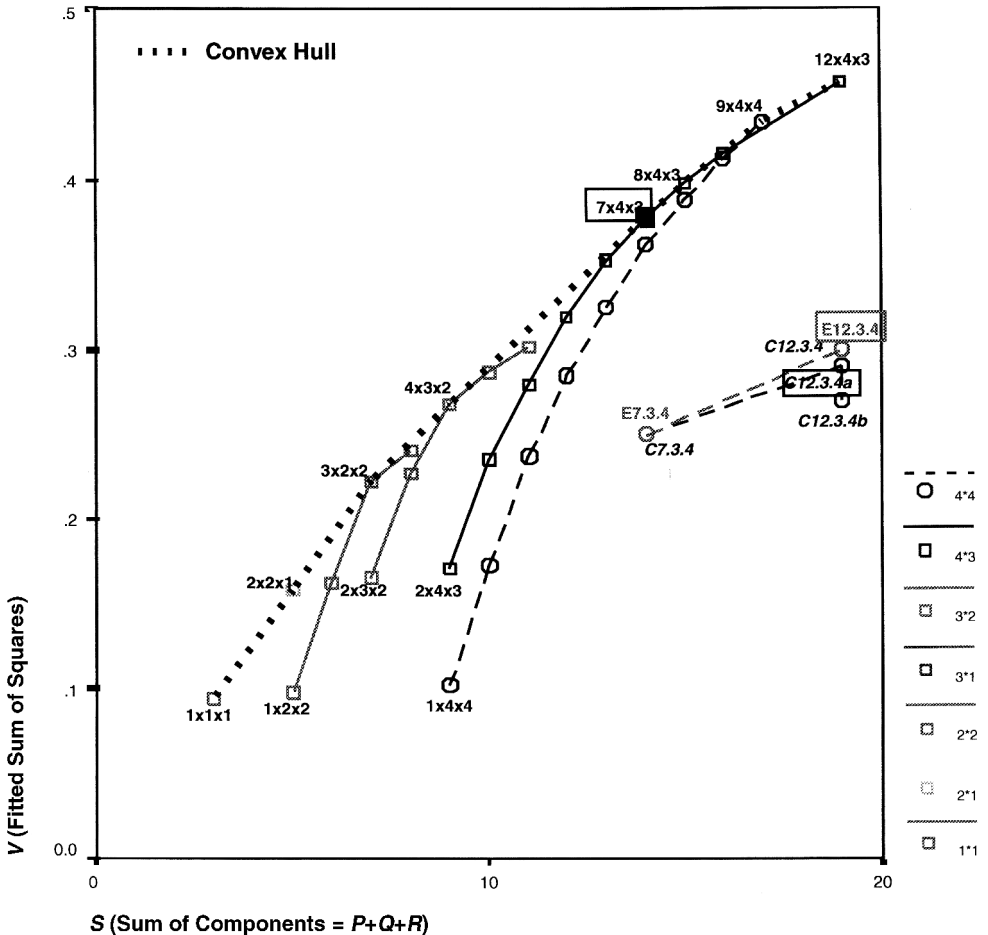
According to Timmerman and Kiers' (2000) DIFFIT proposal (see Table 1), a 3



**Figure 2.** Economic activities data: V-RMR plot. Component models with a constant number of components in the second mode ( $Q$ ) and in the third mode ( $R$ ) are connected for increasing number of second-order components ( $P$ ). Also all confirmatory common factor models are connected, as well as their exploratory versions. The boxed models are the ones selected for interpretation.



(second order)  $\times 2$  (scales)  $\times 2$  (activities) solution is indicated, which, however, only explains 22% of the variation. The lower bound ( $1/s_{\max}$ ) after correction for centring = 0.0145, and this bound is still not reached for  $S = 19$ . More insight into the working of the DIFFIT criterion is obtained by examining the three-mode scree plot (Fig. 3). The strong dependence of the fit on the number of second-order components is again very clear. Further inspection of the fit shows that for models with the same  $S$ , the  $Q > R$  models always show better fit. The curvature in  $V$  versus  $S$  for models with equal  $(Q, R)$  is indicative of the diminishing contribution to the fit of additional second-order components. The dashed curve connects all points with maximal value of  $S$  within their class provided the curve remains convex. The Timmerman-Kiers criterion thus comes



**Figure 3.** Economic activities data: Three-mode scree plot. Component models with a constant number of components in the second mode ( $Q$ ) and in the third mode ( $R$ ) are connected for increasing number of second-order components ( $P$ ). The  $7 \times 4 \times 3$  model is the one selected for interpretation. The dashed line represents the convex hull of models with highest fitted sums of squares given  $S$ . Factor models with equal numbers of factors are located on vertical lines irrespective of restrictions on the factors; they have comparatively low fitted sums of squares. The boxed models are the ones selected for interpretation.

down to selection of models which lie on the (dashed) convex hull. From Fig. 3, one might decide to select the  $4 \times 3 \times 2$  model ( $V = .2675$ ) or  $7 \times 4 \times 3$  model ( $V = .3776$ ) as alternative models to adopt if one prefers more detail than the  $3 \times 2 \times 2$  model.

The deviance plot of  $\chi^2$  versus  $df$  (Fig. 1) shows U-shaped curves for models with the same  $(Q, R)$  and increasing  $P$ . The reason for the curvature is that, for increasing  $P$ , models with the same  $(Q, R)$  have increasing fit in their variances, which is not represented in the  $\chi^2$ , coupled with decreasing fit of the covariances on which the  $\chi^2$  is based. Thus models at the lower end of the curve are the preferred ones in any curve. On the convex hull we find several of the previous favourably mentioned models (such as  $3 \times 2 \times 2$ ,  $4 \times 3 \times 2$ ,  $7 \times 4 \times 3$ ), but not all of them. The primary cause for this is that  $S$  does not take the number of levels in a mode into account. One component of a mode with 390 levels has the same weight as one with only 7 levels. The effect in this example seems to be mitigated by the fact that for the mode with a large number of levels, the contribution to the fit of components beyond the first few is small.

From a data reduction point of view, balancing fit with interpretability, we prefer the  $7 \times 4 \times 3$  solution (38% explained variability). The basis for this decision is that the model has an acceptable RMSEA, lies on two convex hulls, is the best in its  $S$ -class, has a decent  $\chi^2/df$  ratio, and allows for a comprehensible substantive interpretation without being too general or too detailed. Based on the deviance plot, one could also choose the  $4 \times 3 \times 2$  or  $5 \times 4 \times 2$  model because both are in a region where there is a change from large gains in  $\chi^2$  for a few  $df$  to small gains in  $\chi^2$  for large  $df$ .

#### 6.4. Three-mode PCA: Interpretation

Table 2 also presents four varimax-rotated rating-scale components; all varimax rotations were performed on the orthonormal component matrices, making it in fact an oblique cluster rotation (see Harris & Kaiser, 1964). The components can be characterized as social values, economic values, well-known, and non-risky. Table 3 presents three varimax-rotated activity components, which can be characterized as work—(earning money), uncontroversial financial activities, and controversial financial activities. During the selection separate components were foreseen for being employed and being an owner, but they did not emerge here.

The core array (Table 4) indicates how the components of the activities and rating components modes go together for each subject component. The results of the analysis will be interpreted with the second-order component interpretation of the core array. From Table 4 we see that the (rotated) second-order components explain more or less the comparable amounts of the variability (between .039 (4th) and .081 (5th)) of the combination components. By only paying attention to the high (above 1.3) and medium values (above 1.0) of the core array (a value of 1.3 corresponds roughly to 1% explained variability for a combination of components) the following statements can be made.

The first four second-order components are primarily based on the rating scale components, which can be seen from the diagonal structure in each of the activity sections, albeit that this is rather weak for controversial financial activities. The *first* second-order component is primarily determined by the well-known scale, the *second* second-order component is primarily influenced by economic values, the *third* second-order component is concerned with social values, and the *fourth* second-order component relates to the lack of risk of various activities. On the other hand, the *fifth* second-order component is related to one type of activity, i.e. uncontroversial financial activities, rather than on specific scales. The *sixth* second-order component is again

**Table 4.** Economic activities data: Core array (after Varimax rotation of all three modes) for the  $7 \times 4 \times 3$  component model

Second-order (subject) component no.	Proportion of explained variability	Economic activity components											
		Work—earning money				Uncontroversial financial activities				Controversial financial activities			
		Well-known	Economic values	Social values	Non-risky	Well-known	Economic values	Social values	Non-risky	Well-known	Economic values	Social values	Non-risky
1	.053	<b>2.77</b>	-.02	.38	-.02	<b>1.15</b>	.01	.29	.25	.42	.06	-.03	-.21
2	.053	.15	<b>2.41</b>	<b>1.11</b>	-.09	-.38	<b>1.17</b>	-.64	-.30	-.11	.51	-.18	.13
3	.059	.14	.28	<b>2.11</b>	.14	.25	-.06	<b>1.18</b>	-.33	.18	.07	<b>2.09</b>	-.03
4	.039	-.04	-.10	.28	<b>2.15</b>	.12	-.40	.06	<b>1.19</b>	.01	.00	-.16	.85
5	.081	-.00	-.41	-.30	-.27	<b>1.12</b>	<b>2.37</b>	<b>2.34</b>	<b>1.33</b>	-.11	.13	.12	-.05
6	.050	.22	.32	-.02	-.05	-.12	<b>1.12</b>	-.08	.14	.12	<b>2.73</b>	.20	-.30
7	.042	-.01	.00	.04	-.01	-.47	.24	.13	.12	-.43	.04	-.38	<b>-1.09</b>

Note: Bold face is used where the explained variance of a component combination exceeds 0.07%.

determined by economic values, but in a reversed pattern from the second second-order component with respect to importance of type of activities. Finally the *seventh* second-order component is only determined by attention to controversial financial activities and stresses their unfamiliarity and their risk.

## 6.5. Conclusion

When comparing the interpretations of the component and factor solutions, the similarities between the components and factors of both the ratings and the economic activities are obvious, their different numbers notwithstanding. The most striking difference between the two solutions is the way in which the relationships between scores are portrayed: via correlations in the factor model and via the core array with the second-order subject components in the component model.

The correlations between the first and second factors of the economic activities are in accordance with their joint loading in the component solution. The values for the rating scales are also fairly similar to those of the component solution.

## 7. Example: Complex tracking behaviour data

### 7.1. Data

The complex tracking behaviour experiment by Parker and Fleishman (1960) was a forerunner of today's complex flight simulators. A tracking device was constructed 'to simulate roughly the display characteristics and control requirements of an airborne radar intercept mission' (Tucker, 1967, pp. 140-141). A control system involving a control stick and rudder pedals as in a standard aircraft control system was manipulated by the subject to control a target dot in the centre of a cathode ray oscillograph, and the needle of a potentiometer was used as a side-slip indicator. The subject's task was to defeat an induced movement of the dot so as to keep the dot in the centre of the oscillograph. Side-slip error occurred when the subject did not co-ordinate the movement of the stick and the rudder pedals. Measures of horizontal error, vertical error, side-slip error and time on target were obtained. The component error measures were summed to give absolute displacements of the target dot in the two dimensions and of the side-slip indicator from the central position of the potentiometer needle. The time-on-target measure was the time the dot and side-slip indicator needle were simultaneously within the tolerance limits.

Ten segments of the learning curve were selected by Parker and Fleishman (1960, Fig. 9) to represent successive stages in learning, and scores of the 203 subjects on each measure on each stage of practice were converted to stanine form for the correlations presented as Appendix F in Parker and Fleishman (1960). The error scores were reversed to produce accuracy measures. There were four measurements for each of ten stages of practice. The correlations in Parker and Fleishman's Appendix F were only given to two decimal places, causing the correlation matrix to be negative definite with one small negative eigenvalue. This value was arbitrarily set at .01, and the correlation matrix was recomputed with the original positive eigenvalues and the amended one so that it became positive definite and therefore amenable to treatment with structural equation models. With three-mode PCA, which is not influenced by the negative definiteness, it could be checked that the results for limited numbers of components were not influenced by the procedure. Subsequently, the correlation matrix was

transformed to a covariance matrix using the variances supplied by Parker and Fleishman, although the effect was relatively minor as nearly all variances were equal to 2.

The purpose of the analysis of these data was to investigate the usefulness of the two techniques *vis-à-vis* a well-structured longitudinal data set. Interestingly, this was in fact the first multimode data set analysed with three-mode models (Tucker, 1967). As indicated and worked out in Section 8.2 below, longitudinal data should in general be modelled taking the longitudinal character of the data into account. However, for the main analyses of the tracking data, we followed the approach taken by Tucker (1967), as we wanted to redo his analysis and evaluate it with present-day improvements in three-mode component analysis.

For this data set, there was no clear *a priori* structure other than that from Tucker's (1967) analysis. Therefore, a model search with three-mode component analysis was carried out first, before attempting a three-mode factor analysis.

### 7.2. Three-mode PCA: Model selection

A full model search for small to medium numbers of components was carried out. Using standard scree plots, a first indication of the number of components followed from two-mode PCAs on the original multimode covariance matrix (7 components), the average stages covariance matrix (3 components) and the average measurement covariance matrix (2 components). Substantive considerations suggest that no more than two components should be chosen for the measures as there are measures for direction control and one for side-slip control, and the time on target is clearly dependent on both. Table 5 lists the fit measures of selected component and factor models for the tracking data. For the models under consideration an (inadequate) RMSEA of 0.127 for the  $4 \times 3 \times 2$  model was the best value possible given a maximum of two measures components, but the proportion of variance explained is high for reasonable numbers of components (between 0.51 ( $3 \times 2 \times 2$ ) and 0.67 ( $8 \times 4 \times 2$ )).

The plot of the fitted sums of squares versus RMR (not shown here) shows again that major gains in fit depend on the number of second-order components,  $P$ , and given  $P$ , increasing numbers of components for  $Q$  and  $R$  lead to decreases in RMR.

According to Timmerman and Kiers' (2000) DIFFIT proposal (see Table 5), a 2 (second order)  $\times$  2 (stages of practice)  $\times$  1 (measurement) solution is indicated, which explains 40% of the variation. The lower bound ( $1/s_{\max}$ ) is 0.0196, which implies that no inspected model after  $7 \times 4 \times 2$  should be taken into consideration. From Table 5, one might decide to select the  $3 \times 2 \times 2$  model ( $V = .51$ ) or Tucker's  $7 \times 4 \times 2$  model ( $V = .66$ ) as alternative models to adopt if one prefers more detail than the  $2 \times 2 \times 1$  model, which provides a too simplistic explanation of the data. The single measurement component, for instance, does not fit the side-slip data.

The deviance plot of  $\chi^2$  versus  $df$  (Fig. 4) again shows U-shaped curves. On the convex hull we find for larger numbers of components a somewhat different set of models (such as  $2 \times 2 \times 1$ ,  $3 \times 2 \times 2$ ,  $4 \times 2 \times 3$ ,  $6 \times 4 \times 3$ ,  $7 \times 4 \times 3$ ,  $7 \times 4 \times 4$ ) compared to previously mentioned models; however, the number of components for the third mode is generally too large for sensible interpretation. Tucker's  $7 \times 4 \times 2$  model does not stand out on any of the measures.

From a data reduction point of view, balancing fit with interpretability, we decided to report the  $4 \times 3 \times 2$  model. This model has the best RMSEA of all models with no more than two measure components (0.127 compared to 0.136 for Tucker's model) and the

**Table 5.** Complex tracking behaviour data: Fit of three-mode component and factor models (203 subjects × 10 stages of practice × 4 measurements)

Model	Sum	SS (Fit)	dif <sub>S</sub>	b <sub>S</sub>	χ <sup>2</sup>	df <sup>a</sup>	χ <sup>2</sup> /df	RMSEA <sup>b</sup>	AIC <sup>c</sup>	RMR <sup>d</sup>
<i>Component models</i>										
1 × 1 × 1	3	.2728	.2728	2.13	7634	807	9.46	.168	6020	.79
<b>2 × 2 × 1</b>	<b>5</b>	<b>.4007</b>	<b>.1279</b>	<b>2.16</b>	<b>5816</b>	<b>798</b>	<b>7.29</b>	<b>.145</b>	<b>4220</b>	<b>.66</b>
2 × 2 × 2	6	.4598	.0591	1.08	4988	793	6.29	.133	3402	.37
3 × 2 × 2	7	.5143	.0545	1.49	4790	791	6.06	.130	3208	.34
4 × 2 × 2	8	.5509	.0366	1.05	4750	790	6.01	.129	3170	.35
<b>4 × 3 × 2</b>	<b>9</b>	<b>.5635</b>	<b>.0126</b>	–	<b>4527</b>	<b>777</b>	<b>5.83</b>	<b>.127</b>	<b>2973</b>	<b>.31</b>
5 × 3 × 2	10	.5983	.0348	1.38	4586	775	5.92	.128	3036	.31
6 × 3 × 2	11	.6235	.0252	1.01	4687	774	6.06	.130	3139	.31
6 × 4 × 2 <sup>e</sup>	12	.6355	.0120	–	7559	759	9.96	.173	6041	.28
<b>7 × 4 × 2<sup>f</sup></b>	<b>13</b>	<b>.6604</b>	<b>.0249</b>	<b>1.67</b>	<b>4984</b>	<b>757</b>	<b>6.58</b>	<b>.136</b>	<b>3470</b>	<b>.28</b>
8 × 4 × 2	14	.6753	.0149	–	5229	756	6.92	.140	3717	.29
8 × 4 × 3	15	.6839	.0086	–	4101	725	5.66	.124	2651	.26
9 × 4 × 3	16	.6978	.0139	–	4392	721	6.09	.130	2950	.26
9 × 4 × 4	17	.7062	.0084	–	3729	688	5.42	.121	2353	.23
<i>Factor models (6 × 3 × 2)</i>										
C-diag	12	.54			2792	749	3.73	.116	1299	.40
C-block	12	.57			2303	569	4.05	.123	1165	.37
C-lag3	12	.57			2400	653	3.68	.115	1094	.39
C-ar2	12	.51			2507	681	3.68	.115	1145	.42
CR-block	12	.52			2675	731	3.66	.114	1213	.38
<b>CR-lag3</b>	<b>12</b>	<b>.53</b>			<b>2699</b>	<b>752</b>	<b>3.59</b>	<b>.113</b>	<b>1195</b>	<b>.39</b>
CR-ar2	12	.52			2728	759	3.59	.113	1210	.40

<sup>a</sup>  $df = (J \times Q - Q^2) + (K \times R - R^2) + (P \times Q \times R - \frac{1}{2}P(P - 1))$ .

<sup>b</sup>  $RMSEA = \sqrt{((\chi^2/df - 1)/(N - 1))}$ .

<sup>c</sup>  $AIC = \chi^2 - 2df$ .

<sup>d</sup> RMR = root mean square residual (off-diagonal elements); RMR of the original covariance matrix is 1.26.

<sup>e</sup> This model has an unexpected and unexplained extremely high value for χ<sup>2</sup> of 7559.

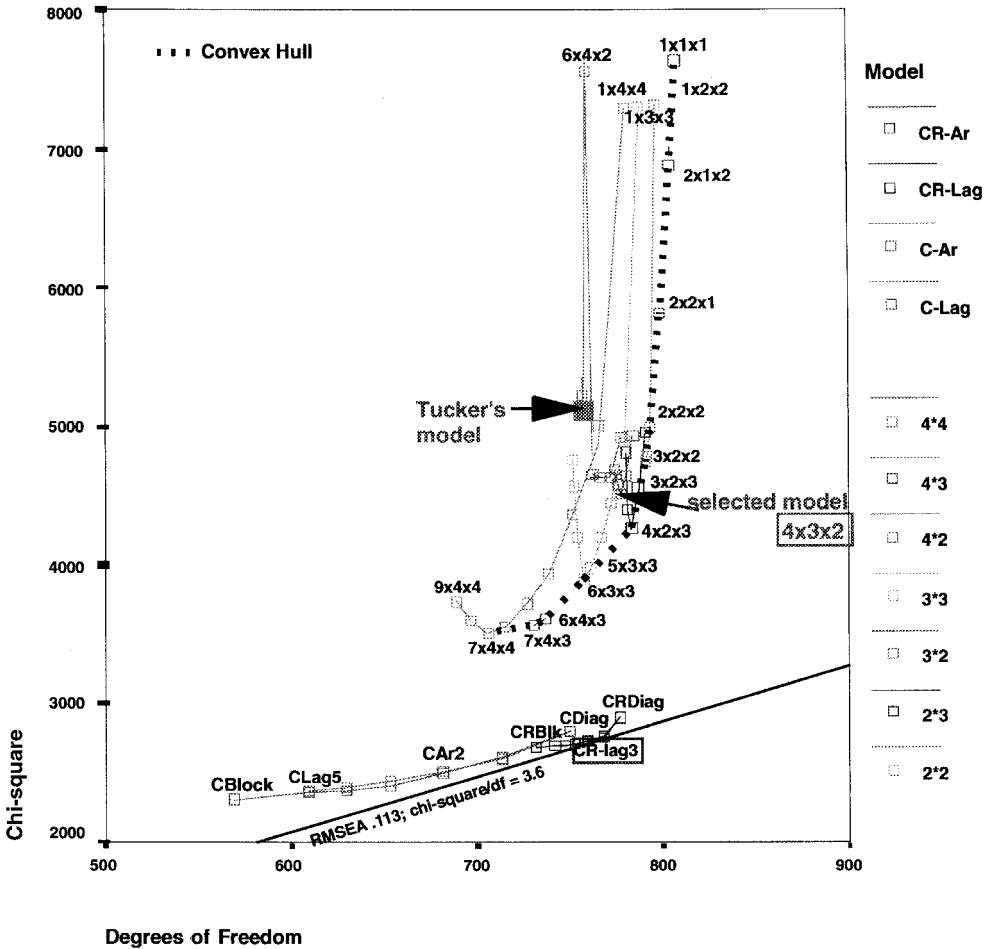
<sup>f</sup> Tucker's original solution.

best AIC (2973 versus 3470 for Tucker's model). The RMSEA is very high, but no component model has a conventionally acceptable RMSEA. It is also the best in its S-class (S = 9), but it misses the convex hull in Fig. 4, albeit only just. It allows for a comprehensible substantive interpretation without being too general or too detailed.

### 7.3. Three-mode PCA: Interpretation

Table 6 shows the stages-of-practice components for the 4 × 3 × 2 solution after a varimax rotation on the unit-length components. The first component (early) is determined by the first two stages, the second component (middle) by stages 3–6, and the third by stages 7–10. The two components for the measurements (Table 7) can be labelled 'direction control' and 'side-slip control', and total time on target is influenced especially by direction control and much less by side-slip control.

The patterns in the second-order components can be deduced from the core array



**Figure 4.** Complex tracking behaviour data: Deviance plot. Component models with a constant number of components in the second mode ( $Q$ ) and in the third mode ( $R$ ) are connected for increasing number of second-order components ( $P$ ). The  $4 \times 3 \times 2$  model is the one selected for interpretation. The  $6 \times 4 \times 2$  model has an unexplained disproportionately high  $\chi^2$ . From its location in the plot it can be seen that the  $4 \times 3 \times 2$  model has the best RMSEA and AIC given two measurement components. The confirmatory common factor models have very much better RMSEA and AIC values than the component models. The boxed models are the ones selected for interpretation.

(Table 8) which portrays the relationships in a very succinct way. One of the difficulties in interpreting core arrays is that one has to take into account the nature of the components of both the stages and the measurements. To support the interpretation of the core array, we can look at the structured loading matrices (see (14)) which combine the component and core information (Fig. 5). The first second-order component (Fig. 5(a)) shows positive weights for direction control and zero and negative weights for side-slip control. In the middle period very good direction control is achieved to the detriment of side-slip control, and achieving control in both is only possible via an

**Table 6.** Complex tracking behaviour data: stages-of-practice components for  $4 \times 3 \times 2$  solution of Tucker3 model after varimax transformation (of orthonormal **B**)

Stages of practice	Components		
	Early	Middle	Late
1	<b>.70</b>	-.04	-.02
2	<b>.67</b>	.08	-.03
3	.15	<b>.45</b>	-.11
4	-.04	<b>.55</b>	-.03
5	-.02	<b>.52</b>	.03
6	-.03	<b>.39</b>	.19
7	-.05	.18	<b>.38</b>
8	-.05	.06	<b>.46</b>
9	-.03	-.05	<b>.55</b>
10	.21	-.17	<b>.54</b>
Proportion of explained variance	.25	.22	.09

**Table 7.** Complex tracking behaviour data: Measurement components for  $4 \times 3 \times 2$  solution of Tucker3 model after varimax transformation

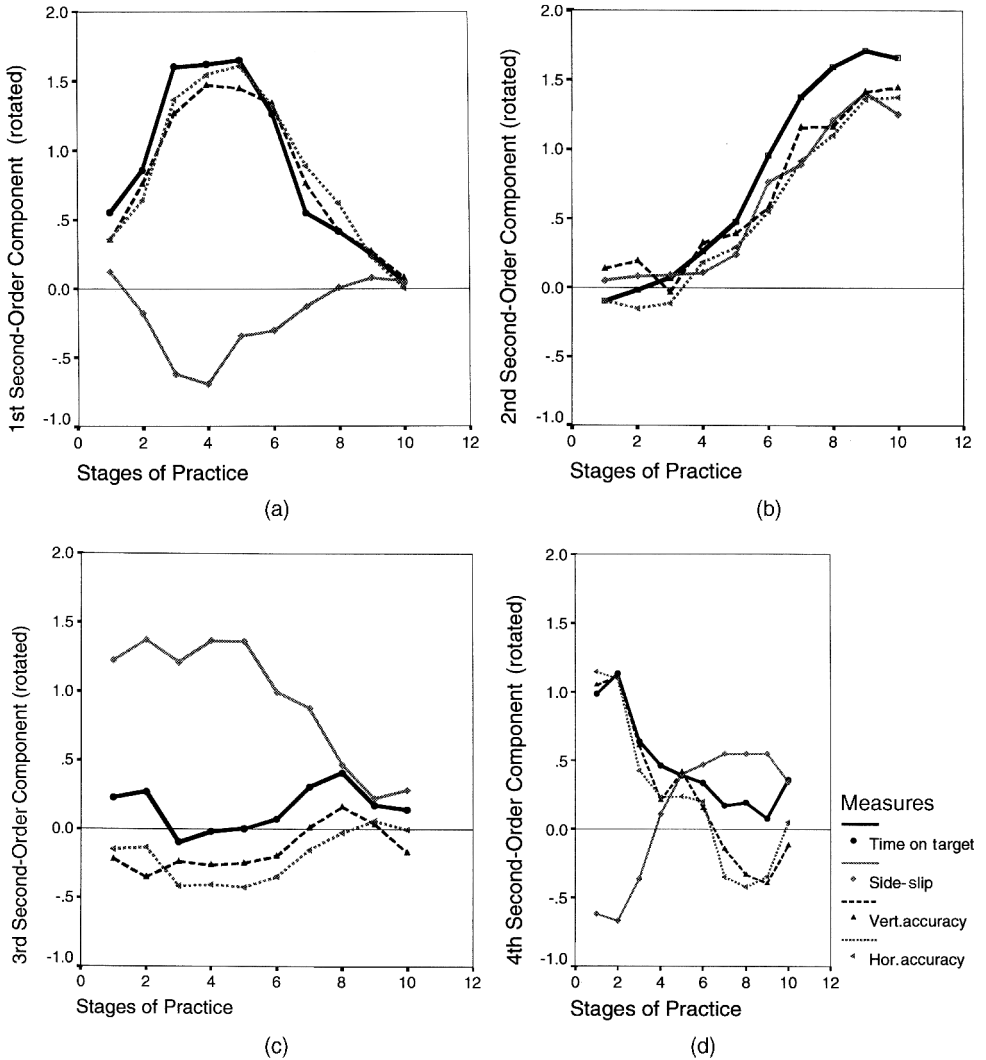
Measurements	Direction control	Side-slip control
Horizontal accuracy	<b>.56</b>	-.12
Vertical accuracy	<b>.56</b>	-.05
Side-slip accuracy	-.03	<b>.97</b>
Total time on target	<b>.61</b>	<b>.21</b>
Proportion of explained variance	.40	.12

**Table 8.** Complex tracking behaviour data: Core array after transformation of component matrices **B** and **C**

Second-order components	Direction control			Side-slip control		
	Early	Middle	Late	Early	Middle	Late
1	<b>1.41</b>	<b>5.21</b>	<b>1.26</b>	.01	-.88	.02
2	.14	.94	<b>4.80</b>	.16	.51	<b>2.63</b>
3	.24	-.76	.23	<b>1.88</b>	<b>2.62</b>	.89
4	<b>2.75</b>	<b>1.21</b>	-.47	-.89	.38	<b>1.24</b>



average performance on both (i.e. a zero score). The second second-order component (Fig. 5(b)) shows, after an initial average grappling of the tasks, a gradual increase in all measures with clear mastery in the end in both types of control, thus achieving a high time on target. The third and fourth second-order components (Fig. 5(c) and 5(d)) show good control for one aspect at the expense of the other in the beginning, ending up somewhere near the average. Interestingly enough, the fourth component shows initial good direction control at the cost of side-slip control, then in the middle a reversal takes place, and in the end both types of control are reasonably balanced.



**Figure 5.** Complex tracking behaviour data: Structured loading matrix for each second-order (subject) component: (a) first component; (b) second component; (c) third component; (d) fourth component. Time on target is represented by solid dots (●), side-slip control by diamonds (◆), vertical accuracy by up-pointing triangles (▲), and horizontal accuracy by left-pointing triangles (◄).

#### 7.4. Three-mode factor analysis

On the basis of Tucker's original analysis, we chose six subject factors ( $P = 6$ ), three stages-of-practice factors ( $Q = 3$ ) and two measurement factors ( $R = 2$ ). An additional consideration was that we wanted a not too large  $P = QR$  model because of its straightforward interpretability without second-order factors. We first used Bentler and Lee's (1978, 1979) identification constraints including echelon-form factor matrices, but the resulting models were uninterpretable. Therefore, we will only report confirmatory models in which we specify a simple structure for both component matrices  $\mathbf{B}$  and  $\mathbf{C}$ , even though their fit is worse than that of the exploratory models. With the simple-structure  $\mathbf{B}$  and  $\mathbf{C}$  matrices, setting the  $6 \times 6$  core array  $\mathbf{G}$  diagonal, and the  $6 \times 6$  factor correlation matrix  $\Phi$  symmetric, we get the confirmatory model C-diag, where 'diag' refers to the diagonal matrix of residual covariances  $\Theta$ . The fit of the confirmatory model is poor (second part of Table 5), probably due to covarying residual factors. That is, residual factors of the same variables are probably correlated across occasions. The C-diag model may thus also be seen as a zero-lagged residual covariance matrix. With 4 variables and 10 occasions, there are  $4 \times 45$  parameters representing residual covariances. Setting these  $\Theta$  parameters free to be estimated, which is equivalent to considering all nine lags (C-lag9), one obtains a block-diagonal  $\Theta$  matrix. The C-lag9 model (= C-block) gives a worse RMSEA but a better AIC than the C-diag model (Table 5). However, the fit is still poor. We need to find better fitting models that (i) take into account the covariances of residual factors of the same variables on different occasions, and (ii) are more parsimonious than the C-block model.

In the C-lag1 model only  $4 \times 9$  parameters in  $\Theta$ , representing covariances between residual factors of adjacent trials, are set free to be estimated. This does improve the fit somewhat. Adding correlated residuals for longer lags than the third lag does not improve the fit any further. Although the fit of these latter models is considerably better than the fit of the models C-diag, the fit is still poor.

Another way to satisfy requirements (i) and (ii) is to impose an autoregressive (AR) structure on the covariance matrix of residual factors  $\Theta$  (see (12)). The fit of such a model with only first-order autoregressive effects is still poor (C-ar1). Addition of second-order autoregressive effects does improve the fit a little (C-ar2; Table 5), but not sufficiently. Yet another way of saving degrees of freedom is to impose a Kronecker structure on the residual covariances (11). In the CR-block model, the  $\Theta$  restriction of (9) is imposed with  $\Theta_C$  diagonal and  $\Theta_B$  symmetric, so that the resulting  $\Theta$  matrix has the same block diagonal structure as the C-block model. To gain some more degrees of freedom we imposed autoregressive structures on the  $\Theta_B$  matrix (e.g. CR-ar2; Table 5). According to the AIC, the fit of these restricted models is worse than the fit of the same models without the  $\Theta$  restriction, but according to the RMSEA the restricted models fit relatively better. This is also true for models with lagged structures for the  $\Theta_B$  matrix (e.g. CR-lag3; Table 5). We imposed a Kronecker structure on the factor correlation matrix  $\Phi$ , but this gave worse fit in all cases.

Figure 4 summarizes the fit of all series of models discussed. Overall there is little difference between the lagged and autoregressive models of one type of series as the convex hulls more or less coincide. The major decision would be whether or not to impose the Kronecker structure on  $\Theta$ . If one decides to model it, the CR-lag3 model ( $\chi^2(752) = 2699$ ; RMSEA = 0.113; AIC = 1195) seems to be the best to choose, as is also indicated by the AIC (see Table 5).

We have to conclude that none of the three-mode factor models provides a satisfactory fit. In view of this, one should not try to interpret the parameter estimates.

That said, with the simple structure of the component matrices  $\mathbf{B}$  and  $\mathbf{C}$ , the diagonal core array  $\mathbf{G}$ , and the standardized  $\Phi$ , the interpretation would have been straightforward and in fact would have led largely to the same conclusions as those reached from the component models. In addition, it would have provided information about the dependence of the variables via the lag information in the residual covariance matrix  $\Theta$ .

Figure 4 and Table 5 show that in terms of  $\chi^2$ , RMSEA, and AIC the factor models perform better than the component models, but, as we have seen, this is still not good enough to allow interpretation of the factor models. This is partly due to the restriction to two measurement components, but with reasonable numbers of components, the fit to the covariances for the component models is never as good as for the factor models.

### 7.5. Conclusions

The complex tracking behaviour data clearly show that an adequate explained variance is not necessarily indicative of a good fit to the covariances, emphasizing that component models and factor models are fundamentally different in their orientation. The example also illustrates that even though no well-fitting model could be found for the factor models, a very interpretable component model existed. The irony is that the poorly fitting factor model (CR-lag3) has a similar and clear interpretation to the component model. From a methodological point of view, it was shown that the core array can be seen as a succinct summary of the major relations in the data, especially via a comparison with the structured loadings.

## 8. Discussion

In this paper, we have shown that multimode covariance matrices can be fruitfully analysed with three-mode models. Within the context of covariance structure analysis, three-mode factor analysis can be seen as a combination of two good ideas. The first is that interpretation can be facilitated by using a Kronecker or direct-product structure to handle the multimode aspect—see also Browne's (1984) direct-product model. The second idea is that a larger number of variables can be described with a small number of factors. Both ideas lead to a reduction in the number of parameters, and their combination can yield spectacular gains of degrees of freedom in accordance with the desire for parsimonious models. Of course, these reductions increase the likelihood of inadequate fitting models, as the complex tracking behaviour example showed.

In terms of interpretation, the great difference between the factor and component techniques lies in the assumptions made about the underlying models, in the same way as for two-mode analysis. For model selection, the various plots served to provide an overview of the relative performance of the various models; however, model selection is only partially guided by 'objective' methods. All model selection procedures used here rather served to exclude models, rather than select the 'true' one.

### 8.1. Numerical issues

The actual choice between component and factor models is further complicated by numerical considerations. Component models are solved by linear algebra and projection using alternating least squares, so that values can always be computed for the parameters. In addition, alternating least squares does not use inversion of the (implied) covariance matrix so that numerical instability and initialization are not problematic.

Moreover, very large covariance matrices can be handled in a relatively short time, in contrast with factor models for which, at least with  $Mx$ , the estimation often requires several hours. When there are relatively few subjects and many variables, covariance structure modelling gets easily into numerical difficulty, but this is not so much the case for the component models.

A case in point is the economic activities data. As mentioned in Section 6.1, the original data consisted of a  $390 \times 12 \times 20$  data block, which was impossible to estimate with covariance structure models. For component models, such data sets can be tackled, even within reasonable time. In fact, the largest amount of time goes into the eigenanalysis of the  $240 \times 240$  sample covariance matrix and the inversion of the implied covariance matrix at the end to compute the maximum likelihood discrepancy function, both of which are not necessary for the analyses themselves. As an example, a  $7 \times 4 \times 3$  analysis of the full data set took a little under 3 minutes for the central iterations and gave the following results:  $\chi^2 = 70\,462$  with  $df = 28\,774$ ; RMSEA = 0.061 and a proportion of explained variance of .13. This is most likely not the optimal solution but serves to indicate that a solution of the entire data set can be computed for the component approach (see also Veldscholte *et al.*, 1998).

### 8.2. Other factor models

There are several other models related to the three-mode factor models discussed here. Oort (1999) gives a general description of stochastic three-mode models for mean and covariance structures. The three-mode factor models of Bloxom (1968) and Bentler and Lee (1978, 1979), and the composite direct-product models of Browne (1984) are mentioned as special cases, but these models are for the covariance structure only. Oort also mentions stochastic three-mode models especially suited for the analysis of multivariate longitudinal data. The covariance structure of such models for the tracking data is given by

$$\Sigma = (\mathbf{C} \otimes \mathbf{I})\Phi(\mathbf{C}' \otimes \mathbf{I}) + \Theta. \quad (15)$$

This model yielded a better fit:  $\chi^2(447) = 906$ , RMSEA = 0.071, AIC = 11.5 (with  $\Theta$  containing covariances for the first four lags). Many further restrictions are possible, and the mean structure can be included as well. Special cases of these longitudinal three-mode models are latent (growth) curve models, and autoregressive models (Oort, 2001). However, longitudinal three-mode models fall outside the scope of the true three-mode models discussed here, as can be seen from (15) which contains neither a (stages-of-practice) component matrix  $\mathbf{B}$  nor a core array  $\mathbf{G}$  (cf. (8)). Another three-mode factor analysis model presented in the literature is McDonald's (1984, 1985) invariant factors model. Oort (2001) shows how this model can be written as a special case of his stochastic three-mode models. For the tracking data, where the occasions are the second mode and the variables the third mode, matrix  $\mathbf{B}$  in (8) and (10) is substituted by an identity matrix, and  $\mathbf{G}$  is a  $R \times JR$  matrix (as  $P = R$  and  $Q = J$ ) transforming the  $JR$  first-order factors into  $R$  invariant second-order factors. The invariant factors model is very restrictive, and its fit to the tracking data turns out to be very poor ( $\chi^2(747) = 3865$ , RMSEA = 0.144).

### 8.3. Conclusion

We have presented ways to use non-stochastic three-mode component models and stochastic three-mode factor models for the analysis of multimode covariance matrices;

through two examples we have illustrated how they can be used, and we have discussed their advantages and limitations. The first example illustrated that a low proportion of explained variance can go together with acceptable values for measures such as the RMSEA, while the second example showed the reverse. Even though the proportion of explained variance typically belongs to component models and the RMSEA to factor models, this does not necessarily mean that in the first example only factor models can be interpreted and in the second example only component models.

## Acknowledgements

This research was conducted at Leiden University, as part of the 'Theory and practice of multivariate longitudinal analysis' research project funded by the Netherlands Organization of Scientific Research (NWO), grant 575-30-005. The authors thank the editor and reviewers for their comments which led to considerable improvements in the manuscript.

## References

- Antonides, G., Farago, K., Ranyard, R., & Tyszka, T. (1997). Perceptions of economic activities: A cross-country comparison. In G. Antonides, W. F. van Raaij, S. Maital (Eds), *Advances in economic psychology* (pp. 39–59). Chichester: Wiley.
- Bagozzi, R. P., Yi, Y., & Nassen, K. D. (1999). Representation of measurement error in marketing variables: Review of approaches and extension to three-facet designs. *Journal of Econometrics*, *89*, 393–421.
- Bentler, P. M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.
- Bentler, P. M., & Lee, S.-Y. (1978). Statistical aspects of a three-mode factor analysis model. *Psychometrika*, *43*, 343–352.
- Bentler, P. M., & Lee, S.-Y. (1979). A statistical development of three-mode factor analysis. *British Journal of Mathematical and Statistical Psychology*, *32*, 87–104.
- Bentler, P. M., Poon, W.-Y., & Lee, S.-Y. (1988). Generalized multimode latent variable models: Implementation by standard programs. *Computational Statistics & Data Analysis*, *6*, 107–118.
- Bloxom, B. (1968). A note on invariance in three-mode factor analysis. *Psychometrika*, *33*, 347–350.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Browne, M. W. (1984). The decomposition of multitrait-multimethod matrices. *British Journal of Mathematical and Statistical Psychology*, *37*, 1–21.
- Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*, 230–258.
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*, 16–29.
- Fowlkes, E. B., Freeny, A. E., & Landwehr, J. M. (1988). Evaluating logistic models for large contingency tables. *Journal of the American Statistical Association*, *83*, 611–622.
- Harris, C. W., & Kaiser, H. F. (1964). Oblique factor analytic solutions by orthogonal transformations. *Psychometrika*, *29*, 347–362.
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8 user's reference guide*. Chicago: Scientific Software International.
- Jöreskog, K. G., Sörbom, D., du Toit, S. H. C., & du Toit, M. (1999). *LISREL 8. New statistical features*. Chicago: Scientific Software International.
- Kiers, H. A. L., & der Kinderen, A. (2003). A fast method for choosing the number of components in Tucker3 analysis. *British Journal of Mathematical and Statistical Psychology*, *56*, 119–125.

- Kiers, H. A. L., Kroonenberg, P. M., & ten Berge, J. M. F. (1992). An efficient algorithm for TUCKALS3 on data with large numbers of observation units. *Psychometrika*, *57*, 415–422.
- Kiers, H. A. L., Takane, Y., & ten Berge, J. M. F. (1996). The analysis of multitrait–multimethod matrices via constrained component analysis. *Psychometrika*, *61*, 601–628.
- Kroonenberg, P. M. (1983). Three-mode principal component analysis: Theory and applications. Leiden: DSWO Press.
- Kroonenberg, P. M. (1996). *3WAYPACK user's manual: A package of three-way programs* (unpublished manuscript). Leiden: Department of Educational Sciences, Leiden University.
- 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.
- Lee, S.-Y., & Fong, W.-K. (1983). A scale invariant model for three-mode factor analysis. *British Journal of Mathematical and Statistical Psychology*, *36*, 217–223.
- Lohmöller, J.-B. (1979). Die trimodale Faktorenanalyse von Tucker: Skalierungen, Rotationen, andere Modelle. *Archiv für die Psychologie*, *131*, 137–166.
- Lohmöller, J.-B., & Wold, H. (1982). *Three-mode path models with latent variables and partial least squares (PLS) parameter estimation*. (Research Report 80.03, revised). Munich: University of the Federal Armed Forces, Department of Education.
- McDonald, R. P. (1980). A simple comprehensive model for the analysis of covariance structures: Some remarks on applications. *British Journal of Mathematical and Statistical Psychology*, *33*, 161–183.
- McDonald, R. P. (1984). The invariant factors model for multimode data. In H. G. Law, C. W. Snyder Jr., J. A. Hattie, & R. P. McDonald (Eds), *Research methods for multimode data analysis* (pp. 285–307). New York: Praeger.
- McDonald, R. P. (1985). *Factor analysis and related methods*. Hillsdale, NJ: Erlbaum.
- Murakami, T. (1983). Quasi three-mode principal component analysis: A method for assessing the factor change. *Behaviormetrika*, *14*, 27–48.
- Neale, M. C. (1997). *Mx: Statistical modeling* (4th Ed.). Richmond, VA: Department of Psychiatry, Virginia Commonwealth University.
- Oort, F. J. (1999). Stochastic three-mode models for mean and covariance structures. *British Journal of Mathematical and Statistical Psychology*, *52*, 243–272.
- Oort, F. J. (2001). Three-mode models for multivariate longitudinal data. *British Journal of Mathematical and Statistical Psychology*, *54*, 49–78.
- Parker, J. F., & Fleishman, E. A. (1960). Ability factors and component performance measures as predictors of complex tracking behavior. *Psychological Monographs: General and Applied*, *74* (16, Whole No. 503).
- Snyder, C. W. (1976). Multivariate analysis of intrinsic individual differences in disjunctive conceptual behavior. *Multivariate Behavioral Research*, *11*, 195–216.
- Snyder, C. W., Bridgman, R. P., & Law, H. G. (1981). Three-mode factor analytic reference curves for concept identification. *Personality and Individual Differences*, *2*, 265–272.
- Snyder, C. W., & Law, H. G. (1979). Three-mode common factor analysis: Procedure and computer programs. *Multivariate Behavioral Research*, *14*, 435–441.
- Snyder, F. W. (1968). *A unique variance model for three-mode factor analysis* (Research Report). Urbana: University of Illinois, Department of Psychology.
- Takane, Y., Kiers, H. A. L., & de Leeuw, J. (1995). Component analysis with different sets of constraints on different dimensions. *Psychometrika*, *60*, 259–280.
- Timmerman, M. E., & Kiers, H. A. L. (2000). Three-mode principal components analysis: Choosing the numbers of components and sensitivity to local optima. *British Journal of Mathematical and Statistical Psychology*, *53*, 1–16.
- Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika*, *31*, 279–311.
- Tucker, L. R. (1967). Three-mode factor analysis of Parker–Fleishman complex tracking behavior data. *Multivariate Behavioral Research*, *2*, 139–152.
- Veldscholte, C. M., Kroonenberg, P. M., & Antonides, G. (1998). Three-mode analysis of

- perceptions of economic activities in Eastern and Western Europe. *Journal of Economic Psychology*, 19, 321–351.
- Verbeek, A. (1983). The geometry of model selection in regression. In T. K. Dijkstra (Ed), *Misspecification analysis* (pp. 20–36). Heidelberg: Springer-Verlag.
- Verhees, J., & Wansbeek, T. J. (1990). A multimode direct product model for covariance structure analysis. *British Journal of Mathematical and Statistical Psychology*, 43, 231–240.
- Weesie, J., & Van Houwelingen, H. (1983). *GEPCAM users' manual: Generalized principal components analysis with missing values* (Technical report). Utrecht: University of Utrecht, Institute of Mathematical Statistics.
- Wothke, W. (1996). Models for multitrait-multimethod matrix analysis. In G. A. Marcoulides & R. E. Schumacker (Eds), *Advanced structural equation modeling: Issues and techniques* (pp. 7–56). Mahwah, NJ: Erlbaum.

Received 11 January 2001; revised version received 30 August 2002