British Journal of Mathematical and Statistical Psychology (1999), **52**, 243–272 *Printed in Great Britain*

Stochastic three-mode models for mean and covariance structures

Frans J. Oort*

Leiden University, The Netherlands

With three-mode models, the three modes are analysed simultaneously. Examples are the analysis of multitrait-multimethod data where the modes are traits, methods and subjects, and the analysis of multivariate longitudinal data where the modes are variables, occasions and subjects. If we consider the subjects mode as random, and the other modes as fixed, such data can be analysed using stochastic three-mode models. Three-mode factor analysis models and composite direct product models are special cases, but they are models for the covariance structure only. Stochastic three-mode models for mean and covariance structures are presented, and the identification, estimation and interpretation of the model parameters are discussed. Interpretation is facilitated by introducing a new terminology and by considering various special cases. Analyses of real data from the field of economic psychology serve as an illustration.

1. Introduction

Three-mode data are collected in many disciplines. Examples are multitrait-multimethod data where the modes are traits, methods and subjects, and multivariate longitudinal data where the modes are variables, occasions and subjects. Such data should be analysed with models that take the three-mode structure into account. In three-mode models components are defined for each of three modes.

Suppose the first mode has N levels and P components, the second mode has J levels and Q components, and the third mode has K levels and R components. Tucker's (1966) three-mode principal components model is then given by

$$\mathbf{X} = \mathbf{AG}(\mathbf{C}' \otimes \mathbf{B}') + \mathbf{E}, \tag{1}$$

where $\mathbf{X}$ is an $N \times JK$ matrix of observed scores, $\mathbf{A}$ is an $N \times P$ matrix of coefficients relating first mode levels to first mode components, $\mathbf{B}$ is a $J \times Q$ matrix of coefficients relating second mode levels to second mode components, $\mathbf{C}$ is a $K \times R$ matrix of coefficients relating third mode levels to third mode components, $\mathbf{G}$ is a $P \times QR$ matrix of coefficients relating the components of the three different modes to each other, and $\mathbf{E}$ is an $N \times JK$ matrix of approximation errors. The symbol $\otimes$ denotes the (right) Kronecker product or direct product. The coefficients in matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are called component loadings, which can be

* Requests for reprints should be addressed to F. J. Oort, Department of Medical Psychology, Academic Medical Centre, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands.

interpreted as scores or weights of levels on components. The **G** matrix is called the core matrix. There are several ways to interpret the **G** coefficients, but the substantive meaning is often unclear in practice.

There are several algorithms for obtaining least squares estimates of **A**, **B**, **C** and **G**. Some of these algorithms have been implemented in a computer program for carrying out three-mode principal components analysis (Kroonenberg, 1994).

In Tucker's principal components model (1) all levels of all modes are considered fixed. However, in most applications in the social and behavioural sciences, one of the modes refers to subjects, which are often considered random. Taking the levels of the first mode as random, we have a three-mode model for the observed scores of a randomly chosen subject:

$$\mathbf{x}' = \xi'\mathbf{G}(\mathbf{C}' \otimes \mathbf{B}') + \varepsilon', \tag{2}$$

where the fixed matrices **X**, **A** and **E** of (1) have been replaced by random vectors $\mathbf{x}'$, ξ', and ε'; **x** is a $JK \times 1$ random vector of observed scores for an arbitrary subject, ξ is a $P \times 1$ random vector of unobserved scores of this subject on common factors, and ε is a $JK \times 1$ random vector of unobserved scores on residual factors. Model (2) has been described previously by Bloxom (1968). It can be seen as a stochastic version of the three-mode principal components model, and so the analysis of (2) is called three-mode factor analysis. As an aside, it should be noted that Tucker had considered a three-mode factor analysis model as well, though he estimated the model parameters through a non-stochastic procedure (Tucker, 1966, pp. 299ff.).

In three-mode factor analysis, as described by Bloxom, the covariance structure of the observed variables is analysed, but the mean structure is disregarded. Bentler & Lee (1978, 1979), extending Bloxom's work, present models that do feature intercepts. However, from their assumption that these intercepts are equal to the population means it follows that they have implicitly chosen all factor means equal to zero and that the mean structure is not analysed.

Here we present three-mode models for the analysis of covariance and mean structures. That is, we investigate three-mode structures not only in the factor loadings, as is done in the three-mode factor analyses of Bloxom (1968) and Bentler & Lee (1978, 1979), but also in the intercepts. Moreover, in addition to looking for three-mode structures in the measurement parameters (factor loadings, intercepts), we look for possible three-mode structures in other structural parameters: the means, variances and covariances of the common factors, and residual factors. This increases the flexibility of the model and its usefulness in applications.

Furthermore, we pay special attention to the interpretation of the parameters that feature in three-mode models. Bentler & Lee (1978, 1979) only consider exploratory use of three-mode models. Therefore, the only restrictions on the parameter matrices in their models are minimal identification constraints. As a consequence, the resulting parameter estimates are difficult to interpret. The understanding of three-mode models and the interpretation of their parameters are easier when three-mode models have parameter matrices with simple structures. Further insight may be gained by looking into a few special cases of three-mode models, such as models for multivariate longitudinal data and Browne's (1984) composite direct product models.

The three-mode models presented in this paper are referred to as stochastic three-mode models (S3MMs). In all S3MMs one of the three-modes is considered random. S3MMs are described as special cases of the linear latent variable model used in structural equation modelling (SEM). This means that the general SEM theory applies to the S3MMs and that, in principle, S3MMs can be fitted with standard SEM software. Bentler, Poon & Lee (1988) have already noted that the models described by Bentler & Lee (1978, 1979) are special cases of the structural equation model of Bentler & Weeks (1980) and have presented one possible reparameterization of these models as a standard structural equation model.

Compared to the usual linear model in SEM, the S3MM approach to three-mode data has the advantage that the three-mode structure is taken into account explicitly. As a consequence, S3MMs are far more parsimonious. The advantage of S3MMs over three-mode principal components analysis models is that *a priori* knowledge of the data can be used, and tested, by fitting confirmatory S3MMs. Moreover, by using hierarchically related S3MMs, all kinds of specific hypotheses can be tested.

Another advantage of the present approach to modelling three-mode data is that it unifies three-mode models that have appeared in the literature. The S3MMs are very general, and various other three-mode models can be described as special cases. Still, the general interpretation of the S3MM parameters applies to the parameters of the special models as well. Because of this, our interpretation of S3MMs may enhance the understanding of these special cases.

The discussion of S3MMs and their interpretation is facilitated by some new terminology which is introduced first. In S3MMs, one mode is random and two modes are fixed. Three-mode data are usually organized in such a way that the levels of the first mode are the subjects. With multivariate longitudinal data, the second mode is usually associated with occasions and the third mode with variables. With multitrait-multimethod data, the second mode is usually associated with methods and the third mode with traits. However, this all depends on how the data set is organized. To keep things general, and for the sake of clarity, we will assume the first mode to be the random mode, associated with subjects. The two fixed modes will be called the 'fast mode' and 'slow mode'. The fast mode is the mode whose levels change fastest in the data set, and the slow mode is the mode whose levels change slower. With multivariate longitudinal data sets, the scores of the subjects are usually given first for all variables on the first occasion, then for all variables on the second occasion, and so on, thus associating the slow mode with the occasions and the fast mode with the variables. With multitrait-multimethod data the slow mode and fast mode usually refer to methods and traits, respectively. Slow and fast modes could also be termed 'outer' and 'inner' modes (as suggested by a referee). These latter terms may further clarify the distinction between slow and fast modes for people who are familiar with nested loops in computer programming languages.

This paper is organized as follows. First we give a general description of the stochastic three-mode models for mean and covariance structures, and discuss the identification and the estimation of the model parameters. Next, we consider the interpretation of the model parameters, and some special cases of S3MMs. Finally, the models are illustrated by analysing a data set from the field of economic psychology.

2. Model

Stochastic three-mode models are models for mean and covariance structures. They are introduced by writing them as special cases of the mean and covariance structures of the linear latent variable model (LLVM).

The LLVM for the observed scores of an arbitrary subject is

$$\mathbf{x} = \tau + \Lambda\xi + \Delta\varepsilon, \tag{3}$$

where $\mathbf{x}$ is a $JK \times 1$ random vector of observed scores of an arbitrary subject, ξ is a $P \times 1$ random vector of unobserved scores of this subject on common factors (or latent constructs), ε is a $JK \times 1$ random vector of unobserved scores of this subject on residual factors, Λ is a $JK \times P$ matrix of constants, to be interpreted as factor loadings, τ is $JK \times 1$ vector of constants, to be interpreted as intercepts, and Δ is a $JK \times JK$ diagonal matrix of constants, to be interpreted as loadings of the observed variables on the residual factors.

We assume

$$E(\varepsilon) = \mathbf{0}, \tag{4}$$

$$\mathrm{Cov}(\xi, \varepsilon) = \mathbf{0}, \tag{5}$$

and write

$$E(\xi) = \kappa, \tag{6}$$

$$\mathrm{Cov}(\xi, \xi') = \Phi, \tag{7}$$

$$\mathrm{Cov}(\varepsilon, \varepsilon') = \boldsymbol{\Theta}, \tag{8}$$

where both Φ and $\boldsymbol{\Theta}$ are symmetric. It follows that the mean structure of the LLVM is

$$E(\mathbf{x}) = \mu = \tau + \Lambda\kappa, \tag{9}$$

and the covariance structure equation is

$$\mathrm{Cov}(\mathbf{x}\mathbf{x}') = \Sigma = \Lambda\Phi\Lambda' + \Delta\boldsymbol{\Theta}\Delta'. \tag{10}$$

The mean structure and covariance structure in (9) and (10) feature six different parameter matrices. We distinguish between measurement parameters and other structural parameters. Measurement parameters, in Λ, τ and Δ, represent characteristics of the variables (loadings, intercepts). Other structural parameters, given by Φ, κ and $\boldsymbol{\Theta}$, represent characteristics of the population of subjects (means, variances, covariances).

2.1. Restrictions

S3MMs are special cases of the LLVM. The LLVM becomes a three-mode model if at least one of the following six restrictions is imposed. These restrictions concern all parameter matrices that feature in the mean structure and covariance structure in (9) and (10)—i.e. Λ, τ, Δ, Φ, κ and $\boldsymbol{\Theta}$. The restrictions are:

$$\Lambda = (\Lambda_S \otimes \Lambda_F)\Gamma, \tag{11}$$

where Λ_S is a $J \times Q$ matrix of coefficients relating slow mode levels to slow mode components, Λ_F is a $K \times R$ matrix of coefficients relating fast mode levels to fast mode components, and Γ is a $QR \times P$ matrix relating combinations of slow and fast mode components to factors;

$$\tau = \tau_S \otimes \tau_F, \tag{12}$$

where τ_S is a $J \times 1$ vector of coefficients associated with the slow mode levels, and τ_F is a

$K \times 1$ vector of coefficients associated with the fast mode levels;

$$\Delta = \Delta_S \otimes \Delta_F, \tag{13}$$

where Δ_S is a $J \times J$ diagonal matrix of coefficients relating slow mode levels to slow mode residual components, and Δ_F is a $K \times K$ diagonal matrix of coefficients relating fast mode levels to fast mode residual components;

$$\mathbf{\Theta} = \mathbf{\Theta}_S \otimes \mathbf{\Theta}_F, \tag{14}$$

where $\mathbf{\Theta}_S$ is a $J \times J$ symmetric matrix of coefficients relating the slow mode residual components to each other, and $\mathbf{\Theta}_F$ is a $K \times K$ symmetric matrix of coefficients relating the fast mode residual components to each other.

Similar restrictions of Φ and κ are meaningful only if the transformation matrix Γ is chosen in such a way that each possible combination of slow and fast mode components transforms into a separate factor. The number of factors then equals the product of the numbers of slow and fast mode components, $P = QR$. In that case, the following restrictions can be imposed:

$$\Phi = \Phi_S \otimes \Phi_F, \tag{15}$$

where Φ_S is a $Q \times Q$ symmetric matrix of coefficients relating the slow mode components to each other, and Φ_F is an $R \times R$ symmetric matrix of coefficients relating the fast mode components to each other;

$$\kappa = \kappa_S \otimes \kappa_F, \tag{16}$$

where κ_S is a $Q \times 1$ vector of coefficients associated with the slow mode components, and κ_F is an $R \times 1$ vector of coefficients associated with the fast mode components.

All six restrictions given by (11)–(16) contain the Kronecker product as a distinctive characteristic, and we refer to them as the Kronecker-product restrictions. If all six restrictions are imposed, then the mean and covariance structures, given earlier for the LLVM by (9) and (10), become

$$E(\mathbf{x}) = \mu = \tau_S \otimes \tau_F + ((\Lambda_S \otimes \Lambda_F)\Gamma)(\kappa_S \otimes \kappa_F) \tag{17}$$

and

$$\begin{aligned}\text{Cov}(\mathbf{x}, \mathbf{x}') = \Sigma = {} & ((\Lambda_S \otimes \Lambda_F)\Gamma)(\Phi_S \otimes \Phi_F)(\Gamma'(\Lambda_S' \otimes \Lambda_F')) \\ & + (\Delta_S \otimes \Delta_F)(\mathbf{\Theta}_S \otimes \mathbf{\Theta}_F)(\Delta_S' \otimes \Delta_F').\end{aligned} \tag{18}$$

But, typically, not all six restrictions are imposed simultaneously.

What are the minimal requirements of a three-mode model? That is, which of the six Kronecker-product restrictions have to be imposed on the LLVM, in order to call it a three-mode model? One may contend that all measurement parameters should conform to a three-mode structure. This would mean that at least the first three restrictions, concerning Λ, τ and Δ, have to be imposed. This requirement is very strict. It is not met by, for example, the three-mode factor models of Bentler & Lee (1978, 1979), which do not comply with the τ restriction. A less strict requirement is that the three-mode model at least comply with the Λ restriction. All published three-mode models meet this requirement. If we want to be lenient, we could limit ourselves to just any one of the six Kronecker-product restrictions. Or, even more lenient, any LLVM with only a pattern of free and fixed parameters that conforms to the pattern produced by the Λ restriction could be called a three-mode model.

The above description of parameters featuring in S3MMs has been kept neutral. The interpretation of the S3MM parameters will be discussed at length, after the treatment of their identification and estimation.

3. Identification

Equations (9) and (10) express all variances, covariances and means of the observed variables as functions of model parameters. If, conversely, all free (to be estimated) model parameters can be independently expressed as functions of the variances, covariances or means of the observed variables, then the model is identified. Bollen (1989) gives general guidelines for identification. These guidelines yield identification conditions that are either necessary or sufficient.

Bekker, Merckens & Wansbeek (1994) give an identification rule that is both necessary and sufficient. If and only if the kernel of the Jacobian matrix for a particular model is empty, then that model is identified. The Jacobian matrix contains the first-order derivatives of the equations implied by (9) and (10) with respect to the model parameters that are free to be estimated. The kernel (or null space) is the set of vectors that maps the linear transformations defined by the Jacobian matrix onto the zero matrix. If the model is not identified, then the kernel contains expressions featuring the non-identified model parameters. So, by inspecting the Jacobian's kernel, one can learn which parameters need to be (further) restricted for the model to be identified. Computation of the Jacobian matrix and its kernel is laborious. It can be done with computer programs for symbolic computation like Maple (Char, Geddes, Gonnet, Leong, Monagan & Watt, 1991) and Mathematica (Wolfram, 1991). Bekker *et al.* (1994) also supply a suite of computer programs to evaluate the identification of various covariance models.

Although the evaluation of the Jacobian's kernel provides a way of checking a model's identification conclusively, general guidelines are still needed for specifying a model in the first place. A necessary condition for achieving the identification of a model is that all latent variables have a scale and an origin. Scales and origins can be imposed through the measurement parameters (in Λ, τ and Δ), or through the other structural parameters (in Φ, κ and $\boldsymbol{\Theta}$).

In the LLVM, latent variables are the common factors ξ and the residual factors ε. Scales for the ξ factors can be provided either by fixing P factor loadings at a non-zero value, one element in each column of Λ, or by fixing the factor variances, that is, the diagonal elements of the Φ matrix. Scales for the ε factors can be provided either by fixing the residual factor loadings (the Δ diagonal) or by fixing the residual factor variances (the $\boldsymbol{\Theta}$ diagonal) at a non-zero value. Origins for the ξ factors can be provided either by fixing P intercepts, one intercept per factor (in τ), or by fixing all factor means (in κ). There is no need explicitly to provide origins for the ε factors, as they have zero means by assumption ((4)).

In S3MMs one or more Kronecker-product restrictions ((11)–(16)) are imposed. The scale and origin restrictions must then be applied to the constituent matrices. However, application of a Kronecker-product restriction to a matrix that is not restricted to provide scales and origins introduces a new indeterminacy that must be removed by fixing one of the elements of one of the constituent matrices at a non-zero value.

Table 1 summarizes the different ways of imposing scales and origins on the latent variables under various combinations of Kronecker-product restrictions. In confirmatory context, where the Λ matrix (or the Λ_S, Λ_F and Γ matrices) has some simple structure, the

Table 1. Minimum numbers of fixed elements for providing scales and origins of the latent variables in three-mode models

Restrictions		Providing scales and origins through measurement parameters	Providing scales and origins through structural parameters
Λ	Φ	Scales for ξ through Λ	Scales for ξ through Φ
un	*un*	*P*, one *el* per column of Λ	*P* diag *el*s of Φ
re	*un*	*Q*, one *el* per column of Λ_S *R*, one *el* per column of Λ_F	*P* diag *el*s of Φ 1 *el* of either Λ_S or Λ_F
un	*re*	*P*, one *el* per column of Λ 1 diag *el* of either Φ_S or Φ_F	*Q* diag *el*s of Φ_S *R* diag *el*s of Φ_F
re	*re*	*Q*, one *el* per column of Λ_S *R*, one *el* per column of Λ_F 1 diag *el* of either Φ_S or Φ_F	*Q* diag *el*s of Φ_S *R* diag *el*s of Φ_F 1 *el* of either Λ_S or Λ_F
τ	κ	Origins for ξ through τ	Origins for ξ through κ
un	*un*	*P els* of τ, one per factor	*P els* of κ
re	*un*	*Q els* of τ_S, one per component *R els* of τ_F, one per component	*P els* of κ 1 *el* of either τ_S or τ_F
un	*re*	*P els* of τ, one per factor 1 *el* of either κ_S or κ_F	*Q els* of κ_S *R els* of κ_F
re	*re*	*Q els* of τ_S, one per component *R els* of τ_F, one per component 1 *el* of either κ_S or κ_F	*Q els* of κ_S *R els* of κ_F 1 *el* of either τ_S or τ_F
Δ	$\mathbf{\Theta}$	Scales for ε through Δ	Scales for ε through $\mathbf{\Theta}$
un	*un*	*JK* diag *el*s of Δ	*JK* diag *el*s of $\mathbf{\Theta}$
re	*un*	*J* diag *els* of Δ_S *K* diag *els* of Δ_F	*JK* diag *els* of $\mathbf{\Theta}$ 1 *el* of either Δ_S or Δ_F
un	*re*	*JK* diag *el*s of Δ 1 diag *el* of either $\mathbf{\Theta}_S$ or $\mathbf{\Theta}_F$	*J* diag *els* of $\mathbf{\Theta}_S$ *K* diag *els* of $\mathbf{\Theta}_F$
re	*re*	*J* diag *els* of Δ_S *K* diag *els* of Δ_F 1 diag *el* of either $\mathbf{\Theta}_S$ or $\mathbf{\Theta}_F$	*J* diag *els* of $\mathbf{\Theta}_S$ *K* diag *els* of $\mathbf{\Theta}_F$ 1 *el* of either Δ_S or Δ_F

Note: re = restricted (as in equations (11)–(16)), *un* = unrestricted, *el* = element, diag = diagonal. All elements mentioned have to be set at some non-zero value (or zero by exception; see text). Λ is assumed to have some simple structure, Φ is symmetric, Δ and $\mathbf{\Theta}$ are diagonal. With restricted Λ, assume all Γ columns contain at least one fixed element. If Γ does not contain fixed elements, then both the 'scales for ξ through Λ' and the 'scales for ξ through Φ' regulations must be followed.

directions in Table 1 will generally yield identified models. Still, the directions concern necessary, not sufficient, conditions for identifying all model parameters. Only the Jacobian's kernel offers absolute certainty about a model's identification.

3.1. Identification in exploratory context

The S3MMs presented here are meant for confirmatory model fitting. In exploratory research it is arguably better to carry out three-mode principal component analyses with models like Tucker's (1966) in equation (1). However, it is of course possible to use confirmatory models in an exploratory way. A problem is that in the exploratory context the Table 1 directions will generally not yield identified models.

Bentler & Lee (1978, 1979) considered the exploratory use of their three-mode model. They only modelled the covariance structure, and only imposed the Λ restriction of (11)—matrices Δ and $\boldsymbol{\Theta}$ are diagonal, one of them an identity matrix. In exploratory research, the numbers of factors, slow mode components and fast mode components must be determined through series of trials. The factors are specified as orthogonal, so matrix Φ is diagonal. Matrix Γ can then be chosen to be diagonal, or to be of echelon form (e.g. (51) below). With Γ diagonal, Λ_S and Λ_F can be of echelon form. With Γ of echelon form, identification can be achieved by having Λ_S and Λ_F both contain a fixed diagonal top. That is, the top $Q \times Q$ part of Λ_S and the top $R \times R$ part of Λ_F are both diagonal with fixed non-zero elements (e.g. (50)).

It should be noted that, although echelon form will generally secure identifiability, it is theoretically possible that this will not be the case. Identification will not be achieved in the event of an independent factor (or component) corresponding to a set of variables that all have zero loadings in one of the echelon columns. Moreover, it is practically possible that, during the optimization of the discrepancy function, parameter estimates get close to an irregular point (Shapiro, 1986), which is sometimes called 'empirical underidentification'. Such problems can be solved by rearranging the variables, by trial and error, or on substantive grounds.

3.2. Ways of setting scales and origins are not always arbitrary

In the LLVM, the choice of setting scales and origins through the measurement parameters or through the other structural parameters is arbitrary. That is, whether scales and origins are provided by measurement parameters (in Λ, τ, Δ) or by other structural parameters (in Φ, κ, $\boldsymbol{\Theta}$), models contain the same numbers of parameters to be estimated, the same degrees of freedom, and yield the same fit. This is not always the case with S3MMs, as is apparent from Table 1. For example, with the Λ restriction imposed ((11)), the scales of the P factors ξ can be set either by fixing Q elements in Λ_S and R elements in Λ_F, or by fixing P elements in Φ and one element in Λ_S or Λ_F. So, unless accidentally $Q + R = P + 1$, the two ways of providing scales for the ξ factors yield models with different numbers of free parameters, degrees of freedom and fit. Moreover, three-mode models with equal numbers of parameters are not always simple reparameterizations of each other either.

Scales of latent variables can only be provided by fixing certain parameters at some non-zero value. Parameters used to provide origins for the latent variables can be fixed at either zero or non-zero values. Setting the origins of the ξ factors by fixing certain parameters at zero can give more leeway when applying the τ or κ restrictions ((12) and (16)). For example, when the τ restriction is applied, origins can often be provided by fixing just two elements of τ_S (instead of a total of $Q + R$ elements in τ_S and τ_F, as indicated in Table 1): one at zero and one non-zero. However, if the ξ factors have been given origins by fixing some τ coefficients or all κ coefficients at zero, then one has to be careful when applying the τ and κ restrictions. For example, if all elements of either κ_S or κ_F are fixed at zero, then the elements of the other vector are not identified, although the resulting κ matrix is (viz. $\kappa = \mathbf{0}_{QR \times 1}$).

There is yet another complication with the analysis of the mean structure of three-mode models. If the τ parameters are unrestricted, then it does not matter how origins are imposed on the ξ factors. The choice of the value at which τ or κ parameters are fixed is of no consequence for the fit of the model. However, with the τ restriction of (12) imposed, the fit of a model does depend on the value at which τ or κ parameters are fixed. Moreover, models with different scales for the ξ factors yield different fit as well, when the τ restriction is imposed (unless $\kappa = \mathbf{0}$; compare (9)). In short, if τ parameters are restricted, then models that only differ in the values of the factor means give different fit. Models that only differ in the values of the factor variances also give different fit, unless the factor means are zero.

4. Estimation

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

$$F(\mathbf{m}, \mathbf{S}, \mu, \Sigma) = (\mathbf{m} - \mu)'\Sigma^{-1}(\mathbf{m} - \mu) + \log|\Sigma| - \log|\mathbf{S}| + \mathrm{trace}\,(\mathbf{S}\Sigma^{-1}) - JK, \quad (19)$$

where $\mathbf{m}$ is the vector of observed sample means of $\mathbf{x}$, and $\mathbf{S}$ is the matrix of observed sample variances and covariances of $\mathbf{x}$. JK is the total number of observed variables (see Bollen, 1989; or Jöreskog & Sörbom, 1996).

If the sample contains subjects of different groups, separate parameter estimates can be obtained for G groups simultaneously, using

$$F_G(\mathbf{m}_1, \mathbf{S}_1, \mu_1, \Sigma_1, \ldots, \mathbf{m}_G, \mathbf{S}_G, \mu_G, \Sigma_G) = \sum_{g=1}^{G} \frac{N_g}{N} F(\mathbf{m}_g, \mathbf{S}_g, \mu_g, \Sigma_g), \quad (20)$$

where the subscripts g refer to group g, N_g is the number of subjects in group g, and N is the total number of subjects in the analysis.

Minimization of the ML function gives estimates for all model parameters, which are unbiased, scale-invariant and scale-free. The ML function also provides estimates of the asymptotic standard errors of the model parameters. Finally, a chi-square test of overall goodness-of-fit is available with ML estimation.

There are weaker assumptions than multivariate normality of $\mathbf{x}$ that still justify parameter estimation through minimization of the ML function (Browne & Shapiro, 1988). Alternatively, parameter estimates can be obtained through other procedures, such as weighted or unweighted least squares estimation. Not all procedures share all of the useful properties of ML estimation (see Bollen, 1989; also Browne & Arminger, 1995, for more information on estimation procedures).

The LLVM can be fitted with standard SEM software. Many computer programs for the estimation of SEM parameters are available. Most prominent are the commercial programs LISREL (Jöreskog & Sörbom, 1996) and EQS (Bentler, 1995), though neither is as versatile as Mx (Neale, 1997), which is freely available through the Internet (http://views.vcu.edu/mx). S3MMs can be fitted straightforwardly with Mx or, with a series of user-specified constraints, with other standard SEM software. Bentler *et al.* (1988), for example, have presented one way to reparameterize their three-mode model as a standard structural equation model.

5. Interpretation

The covariance and mean structure equations (9) and (10) feature six different parameter matrices, each of which can be restricted to conform to a multiplicative structure ((11)–(16)).

The six so-called Kronecker-product restrictions involve a total of 13 different parameter matrices.

When interpreting the parameter matrices, it helps to think of the observed variables $\mathbf{x}$ as in some way composed of $\mathbf{x}_S$ and $\mathbf{x}_F$, through

$$\mathbf{x} = f_{\mathbf{x}}(\mathbf{x}_S, \mathbf{x}_F), \tag{21}$$

with $\mathbf{x}_S$ and $\mathbf{x}_F$ being J and K vectors of unobserved variables corresponding to the slow and fast mode levels in the three-mode data. The common factors ξ can be thought of as composed of ξ_S and ξ_F, through

$$\xi = f_{\xi}(\xi_S, \xi_F), \tag{22}$$

with ξ_S and ξ_F being Q and R vectors of factors corresponding to the slow and fast mode components in the three-mode data. In the same way, the residual factors ε can be thought of as composed of ε_S and ε_F, through

$$\varepsilon = f_{\varepsilon}(\varepsilon_S, \varepsilon_F), \tag{23}$$

with ε_S and ε_F being J and K vectors relating to the slow and fast mode levels in the three-mode data. The composition functions $f_{\mathbf{x}}$, f_{ξ} and f_{ε} are left unspecified.

The S3MMs are models for mean and covariance structures. The observed scores on $\mathbf{x}$ are not explicitly modelled. Note that different underlying score models may yield the same mean and covariance structures (see Browne, 1984). And, although we will look into one possible form of the composition functions later (equations (24), (30) and (33)), mean and covariance models can be studied without knowledge of the underlying score models. For now, the composition functions and their constituent vectors of variables are abstractions, established only for ease of interpretation of the parameter matrices.

5.1. Matrix Λ *and vector* τ

The matrix Λ describes the number of factors and the content or meaning of these factors. The coefficients in Λ show which observed variables are indicative of which factors and to what extent.

In the LLVM (equation (3)), the Λ coefficients are called ‘factor loadings’. Factor loadings relate the observed scores on variables to unobserved scores on factors. Factor loadings are regression weights for the regressions of the observed variables on the factors. The Λ restriction features the matrices Λ_S, Λ_F and Γ. As defined by (11), the coefficients of Λ_S and Λ_F relate ‘levels’ to ‘components’. The Λ_S and Λ_F coefficients are therefore often called ‘component loadings’. It follows that the coefficients in Γ must then be interpreted as coefficients that transform products of component loadings into factor loadings. So the most direct characterization of Γ is as a transformation matrix.

It is important to note that ‘component loadings’ is just a name used to distinguish Λ_S and Λ_F coefficients from the factor loadings in Λ. If the Λ restriction is imposed, we suppose that factors are in some way composed of components, and factor loadings are in some way composed of component loadings. Hence the use of the term ‘components’ (which is different from the use of the same term in ‘principal components analysis’).

Another way of looking at the Λ_S and Λ_F coefficients arises from the idea of the composition functions $f_{\mathbf{x}}$ and f_{ξ}. That is, the supposition of $\mathbf{x}_S$, $\mathbf{x}_F$, ξ_S, and ξ_F variables that in some way constitute composite variables $\mathbf{x}$ and ξ. The Λ_S coefficients can then be seen as loadings of $\mathbf{x}_S$ variables on ξ_S factors, and the Λ_F coefficients as loadings of $\mathbf{x}_F$ variables on ξ_F factors. So we could just as well call the Λ_S coefficients 'slow mode factor loadings' and the Λ_F coefficients 'fast mode factor loadings'. Γ then describes the transformation of combined slow and fast mode factor loadings into loadings on the composite factors.

The transformation coefficients in Γ weight the effects of combinations of slow mode and fast mode factors on the composite factors in ξ. The Γ coefficients can thus also be seen as loadings of combinations of slow and fast mode factors ξ_S and ξ_F on composite factors ξ.

In most applications of three-mode models, Γ is used as a selection matrix. By fixing all Γ coefficients at either unity or zero, certain combinations of slow and fast mode factors are selected to be transformed into factors. Or, by fixing some of the Γ coefficients at zero and leaving the rest free to be estimated, one can fix all elements of the diagonal of Φ at unity, so that Φ can be interpreted as a correlation matrix. In this case Γ is used not only for selection, but also to absorb scaling.

Most often, the Λ restriction is imposed with a diagonal Γ matrix, so that all combinations of slow mode components and fast mode components are transformed into separate factors.

The intercepts in vector τ concern the level of factors in the same way that factor loadings concern the content of factors. For example, suppose intelligence is measured through a number of tests, one of which is a verbal test. Then the factor loading for the verbal test shows to what extent the verbal test is indicative of intelligence, and the intercept indicates how difficult the verbal test is. If the variables $\mathbf{x}$ are test items, then factor loadings and intercepts can be interpreted as item discrimination and (the reciprocal of) item difficulty, respectively (Oort, 1996, Chapter 2).

The τ restriction of (12) gives a decomposition of the τ vector. If the τ restriction is imposed, then τ_S and τ_F give the contributions to the intercepts for slow and fast mode components separately. Suppose, for example, that various abilities are measured through different methods—for example, mental tests and written tests. Then the intercept for a mental arithmetic test can be composed of two coefficients, one showing how mentally able one must be, and the other showing how arithmetically able one must be, in order to do well on the mental arithmetic test.

In many applications of S3MMs the measurement parameters in Λ and τ are associated with only one of the two fixed modes. For example, with multivariate longitudinal data, the measurement parameters are characteristics of the variables and should be invariant across occasions.

5.2. Matrices Δ *and* $\boldsymbol{\Theta}$

Δ is a diagonal matrix containing the factor loadings of the observed variables on the residual factors. Of course, there are as many residual factors as observed variables. As the scaling of the unobserved residual factors is arbitrary, the Δ coefficients are arbitrary as well, and cannot usually be interpreted. In fact, because of the identification requirements discussed above, Δ is mostly set to identity.

As Δ will usually be a $JK \times JK$ identity matrix, the Δ restriction of equation (13) is satisfied as a matter of course with $\Delta_S = \mathbf{I}_{J\times J}$ and $\Delta_F = \mathbf{I}_{K\times K}$. However, in some applications, researchers may wish to interpret $\boldsymbol{\Theta}$ as a correlation matrix, so that elements of Δ have to be set free to be estimated. With Δ containing scale-absorbing coefficients, the Δ restriction can be imposed to test for a multiplicative structure.

An attractive interpretation of the Δ_S and Δ_F coefficients is that of factor loadings of $\mathbf{x}_S$ and $\mathbf{x}_F$ variables on the ε_S and ε_F factors of composition function f_ε.

$\mathbf{\Theta}$ contains the variances and covariances of the residual factors. In many applications, residual factors are assumed to be uncorrelated and $\mathbf{\Theta}$ is then diagonal. However, with three-mode data in particular, it is often appropriate also to consider symmetric $\mathbf{\Theta}$ matrices. With multitrait-multimethod data, for example, residual factors of variables measuring the same traits through different methods may be correlated. Or, with multivariate longitudinal data, the residual factors of the same variables on different occasions may be correlated. Still assuming that the residual factors of different traits are uncorrelated, $\mathbf{\Theta}$ will be symmetric, either having a block diagonal (if the traits correspond with slow mode components) or consisting of diagonal blocks (if the traits correspond with the fast mode components).

If the $\mathbf{\Theta}$ restriction (14) is imposed, $\mathbf{\Theta}_S$ and $\mathbf{\Theta}_F$ show the separate contributions of the slow and fast mode components. $\mathbf{\Theta}_S$ and $\mathbf{\Theta}_F$ can be viewed as variance-covariance matrices of the abstract ε_S and ε_F factors that feature in the composition function f_ε.

The decomposition of $\mathbf{\Theta}$ is particularly interesting if one of $\mathbf{\Theta}_S$ and $\mathbf{\Theta}_F$ is non-diagonal. If we allow the slow mode components of the residual factors to correlate, then the $\mathbf{\Theta}$ restriction can be applied with a symmetric $\mathbf{\Theta}_S$ and a diagonal $\mathbf{\Theta}_F$, yielding a $\mathbf{\Theta}$ structure with diagonal blocks. Conversely, with $\mathbf{\Theta}_S$ diagonal and $\mathbf{\Theta}_F$ symmetric, $\mathbf{\Theta}_S \otimes \mathbf{\Theta}_F$ yields a block diagonal $\mathbf{\Theta}$.

With multivariate longitudinal data, slow and fast mode components usually correspond to occasions and variables respectively. So we are more familiar with a $\mathbf{\Theta}$ matrix consisting of diagonal blocks than with a block diagonal $\mathbf{\Theta}$ matrix.

5.3. Matrix Φ and vector κ

Φ contains the variances and covariances of the common factors ξ. Scaling of the unobserved factors is arbitrary. If scales are imposed by fixing the factor variances at one, then Φ can be interpreted as a correlation matrix.

The κ vector contains the factor means. The origins of unobserved factors are arbitrary as well, and can be set through either κ or τ. Because of the arbitrary origins of the common factors, comparison of factor means across groups or across occasions is meaningful only after making sure that the measurement parameters are invariant across groups and occasions.

As noted before, the Φ and κ restrictions ((15) and (16)) can only be meaningfully applied if all combinations of slow and fast mode components have been transformed into factors (i.e. Γ diagonal, so that $P = QR$).

If the Φ restriction is applied, then Φ_S and Φ_F give the separate contributions of slow and fast mode components. Adopting the idea of the composition function f_ξ, Φ_S and Φ_F may be viewed as variance-covariance matrices of the abstract ξ_S and ξ_F factors. The decomposition of Φ makes it easy to test specific hypotheses such as whether slow mode factors are uncorrelated. If so, with Φ_S diagonal, $\Phi_S \otimes \Phi_F$ would result in a block diagonal matrix.

Vectors κ_S and κ_F may be interpreted as the mean vectors of the slow and fast mode factors ξ_S and ξ_F that feature in the composition function f_ξ. In order to meet identification requirements discussed above, many researchers automatically set κ equal to zero, $\kappa = \mathbf{0}_{QR\times 1}$. In that case, the κ restriction is obviously satisfied with either $\kappa_S = \mathbf{0}_{Q\times 1}$ or $\kappa_F = \mathbf{0}_{R\times 1}$. Yet there are applications, for example with multivariate longitudinal data, where it is not permissible to set all κ equal to zero. If may then be interesting to impose the κ restriction to test for a multiplicative structure (Oort, 1999).

6. Models for observed scores

As noted before, S3MMs are models for mean and covariance structures, not for scores. Although, for the purpose of interpretation, we did introduce the idea of composition functions $f_{\mathbf{x}}, f_{\xi}$ and f_{ε} (equations (21)–(23)), we have not specified in what way the composite variables in $\mathbf{x}$, ξ and ε are composed of their constituent variables. Here we will go into some possible models for observed scores and their associated mean and covariance structures.

The main reason for introducing these score models is that they may further help the understanding of three-mode models and the interpretation of their parameters. Additionally, the score models lead to certain special cases of three-mode models that may be of particular interest. Still, it is emphasized that there may be score models other than the ones discussed below, that yield the same mean and covariance structures. Moreover, mean and covariance models may be utilized without contemplating underlying score models.

First, consider

$$\mathbf{x} = \mathbf{x}_{\mathrm{S}} \otimes \mathbf{x}_{\mathrm{F}}. \tag{24}$$

The mean and covariance structure equations for such a model are given by Graybill (1983, p. 368; note Graybill's use of the *left* Kronecker product). Writing Σ_{SS} for Cov ($\mathbf{x}_{\mathrm{S}}$, $\mathbf{x}_{\mathrm{S}}'$), Σ_{FF} for Cov ($\mathbf{x}_{\mathrm{F}}$, $\mathbf{x}_{\mathrm{F}}'$), Σ_{SF} for Cov ($\mathbf{x}_{\mathrm{S}}$, $\mathbf{x}_{\mathrm{F}}'$), Σ_{FS} for Cov ($\mathbf{x}_{\mathrm{F}}$, $\mathbf{x}_{\mathrm{S}}'$), μ_{S} for E ($\mathbf{x}_{\mathrm{S}}$), and μ_{F} for E ($\mathbf{x}_{\mathrm{F}}$), we can write

$$E(\mathbf{x}) = \mu = \mathrm{Vec}(\Sigma_{\mathrm{SF}}) + (\mu_{\mathrm{S}} \otimes \mu_{\mathrm{F}}) \tag{25}$$

and

$$\begin{aligned} \mathrm{Cov}(\mathbf{x}, \mathbf{x}') &= \Sigma \\ &= (\Sigma_{\mathrm{SS}} \otimes \Sigma_{\mathrm{FF}}) + (\mu_{\mathrm{S}}\mu_{\mathrm{S}}' \otimes \Sigma_{\mathrm{FF}}) + (\Sigma_{\mathrm{SS}} \otimes \mu_{\mathrm{F}}\mu_{\mathrm{F}}') \\ &\quad + \mathrm{Comm}_{JK}((\Sigma_{\mathrm{FS}} \otimes \Sigma_{\mathrm{SF}}) + (\mu_{\mathrm{F}}\mu_{\mathrm{S}}' \otimes \Sigma_{\mathrm{SF}}) + (\Sigma_{\mathrm{FS}} \otimes \mu_{\mathrm{S}}\mu_{\mathrm{F}}')), \end{aligned} \tag{26}$$

where Vec (Σ_{SF}) is a JK vector consisting of stacked Σ_{SF} columns, and Comm_{JK} is a so-called commutation matrix (Graybill, 1983, p. 315). If $\mathbf{x}_{\mathrm{S}}$ and $\mathbf{x}_{\mathrm{F}}$ are independent, then the mean and covariance equations simplify to

$$E(\mathbf{x}) = \mu = (\mu_{\mathrm{S}} \otimes \mu_{\mathrm{F}}), \tag{27}$$

and

$$\mathrm{Cov}(\mathbf{x}, \mathbf{x}') = \Sigma = (\Sigma_{\mathrm{SS}} \otimes \Sigma_{\mathrm{FF}}) + (\mu_{\mathrm{S}}\mu_{\mathrm{S}}' \otimes \Sigma_{\mathrm{FF}}) + (\Sigma_{\mathrm{SS}} \otimes \mu_{\mathrm{F}}\mu_{\mathrm{F}}'). \tag{28}$$

So an observed score model like (24) does not readily fit into the framework of three-mode models described by (9)–(16). If, however, the scores on $\mathbf{x}$ are deviation scores, measured from their means, then all $\mathbf{x}$ variables have zero means, so we can set $\mu_{\mathrm{S}} = \mathbf{0}$ and $\mu_{\mathrm{F}} = \mathbf{0}$, and (28) simplifies to Swain's direct product model for the covariance structure (Browne, 1984),

$$\mathrm{Cov}(\mathbf{x}, \mathbf{x}') = \Sigma = \Sigma_{\mathrm{SS}} \otimes \Sigma_{\mathrm{FF}}, \tag{29}$$

which is a special case of the three-mode models described above (i.e. if $\Lambda = \mathbf{I}_{JK \times JK}$, $\tau = \kappa = \mathbf{0}_{JK}$, and $\mathbf{\Theta} = \mathbf{0}_{JK \times JK}$, then the Φ restriction can be applied with $\Phi_S = \Sigma_{SS}$ and $\Phi_F = \Sigma_{FF}$). Still, as our interest is in modelling mean structures as well as covariance structures, we do not want to assume the scores on the observed variables $\mathbf{x}$ to be deviation scores.

Whether the observed score model given by (24) can be written as a special case of the LLVM (3) depends on further modelling of $\mathbf{x}_S$ and $\mathbf{x}_F$. We will not pursue this here.

It should be noted that if normal distributions are assumed for the unobserved variables $\mathbf{x}_S$ and $\mathbf{x}_F$, then the observed variables $\mathbf{x}$ will not be normally distributed, and minimization of the ML function (19) need not yield asymptotically correct standard errors and chi-square distributed test statistics (Browne, 1984).

We turn to observed score models that result from specifying models for the latent variables ξ and ε. Consider the following model for the common factor scores:

$$\xi = \xi_S \otimes \xi_F. \tag{30}$$

If ξ_S and ξ_F are independent then

$$E(\xi) = \kappa = E(\xi_S) \otimes E(\xi_F), \tag{31}$$

and the κ restriction applies, with $\kappa_S = E(\xi_S)$ and $\kappa_F = E(\xi_F)$. If ξ_S and ξ_F are given origins by setting their means equal to zero, then

$$\mathrm{Cov}(\xi, \xi') = \Phi = \mathrm{Cov}(\xi_S, \xi_S') \otimes \mathrm{Cov}(\xi_F, \xi_F'), \tag{32}$$

and the Φ restriction applies, with $\Phi_S = \mathrm{Cov}(\xi_S, \xi_S')$ and $\Phi_F = \mathrm{Cov}(\xi_F, \xi_F')$. Likewise, consider the residual factor scores to be modelled

$$\varepsilon = \varepsilon_S \otimes \varepsilon_F. \tag{33}$$

Assuming that ε_S and ε_F are independent and have zero means, then

$$\mathrm{Cov}(\varepsilon, \varepsilon') = \mathbf{\Theta} = \mathrm{Cov}(\varepsilon_S, \varepsilon_S') \otimes \mathrm{Cov}(\varepsilon_F, \varepsilon_F'), \tag{34}$$

and the $\mathbf{\Theta}$ restriction applies, with $\mathbf{\Theta}_S = \mathrm{Cov}(\varepsilon_S, \varepsilon_S')$ and $\mathbf{\Theta}_F = \mathrm{Cov}(\varepsilon_F, \varepsilon_F')$.

Substitution of either (30) or (33), or both, into the LLVM of (3) does yield models for the observed scores that imply mean and covariance structures within the framework of three-mode models. With such an observed score model, it is again noted that, if the distributions of ξ_S, ξ_F, ε_S and ε_F are multivariate normal, then the distribution of $\mathbf{x}$ cannot be mutivariate normal. Browne & Shapiro (1988; superseding Browne, 1987) give less strict conditions for the ML estimation method to retain its usual asymptotic properties, but these conditions are not met either, as they do not admit restrictions on the covariance matrices Φ and $\mathbf{\Theta}$.

6.1. Some special cases

If all measurement parameter matrices of the LLVM (3) are restricted according to the Kronecker-product restrictions given by (11)–(13), then the model for observed scores becomes

$$\mathbf{x} = (\tau_S \otimes \tau_F) + ((\Lambda_S \otimes \Lambda_F)\Gamma)\xi + (\Delta_S \otimes \Delta_F)\varepsilon, \tag{35}$$

with mean and covariance structure equations

$$E(\mathbf{x}) = \mu = (\tau_S \otimes \tau_F) + ((\Lambda_S \otimes \Lambda_F)\Gamma)\kappa \tag{36}$$

and

$$\begin{aligned}\mathrm{Cov}(\mathbf{x}, \mathbf{x}') &= \Sigma \\ &= ((\Lambda_S \otimes \Lambda_F)\Gamma)\Phi(\Gamma'(\Lambda_S' \otimes \Lambda_F')) + (\Delta_S \otimes \Delta_F)\boldsymbol{\Theta}(\Delta_S' \otimes \Delta_F').\end{aligned} \tag{37}$$

Φ, κ and $\boldsymbol{\Theta}$ can be further restricted according to models inspired by the content of the particular applications at hand, such as some psychological theory. For example, imposing a factor structure on Φ turns (37) into a second-order factor model. Imposition of structural regression relations between factors yields models like the one mentioned by Bentler *et al.* (1988, p. 112, eq. 14).

In many applications of three-mode analysis the measurement parameters are associated with only one of the two fixed modes. For example, with multivariate longitudinal data, the measurement parameters should be invariant across the occasions mode. Assuming the occasions to coincide with slow mode levels, measurement invariance can be modelled by further restricting (35)–(37) by setting $\Lambda_S = \mathbf{I}_{J \times J}$, $\Gamma = \mathbf{I}_{JR \times JR}$, $\tau_S = \mathbf{u}_{J \times 1}$, and $\Delta_S = \mathbf{I}_{J \times J}$ ($\mathbf{u}$ (or $\mathbf{U}$) denotes a vector (or a matrix) with all elements equal to unity). These restrictions yield a longitudinal three-mode model described by Oort (1999). This model can be further restricted. For example, imposing the Φ and κ restrictions with $\Phi_S = \mathbf{U}_{J \times J}$ and $\kappa_S = \mathbf{u}_{J \times 1}$ yields the invariant factors model of McDonald (1984). Other restrictions on Φ, κ and $\boldsymbol{\Theta}$ of the longitudinal three-mode model include restrictions that yield latent curve models and restrictions that yield autoregressive models for multivariate longitudinal data (Oort, 1999). In these longitudinal three-mode models the slow mode levels are not summarized by a smaller number of slow mode components ($J = R$).

If neither slow nor fast mode levels are summarized by smaller numbers of components (and $\Gamma = \mathbf{I}_{JK \times JK}$), then Λ can be fixed at identity (as well as Δ: $\Lambda = \Delta = \mathbf{I}_{JK \times JK}$). If we additionally substitute equation (30) into the LLVM (3), we have

$$\mathbf{x} = \tau + \Lambda(\xi_S \otimes \xi_F) + \varepsilon. \tag{38}$$

If we assume ξ_S and ξ_F independent and provide origins by setting their means equal to zero, then we obtain mean and covariance structure equations

$$E(\mathbf{x}) = \mu = \tau \tag{39}$$

and

$$\mathrm{Cov}(\mathbf{x}, \mathbf{x}') = \Sigma = (\Phi_S \otimes \Phi_F) + \boldsymbol{\Theta}. \tag{40}$$

With the mean structure of (39), we could use the τ restriction to test whether the vector of observed variables $\mathbf{x}$ can be decomposed into two independent $\mathbf{x}_S$ and $\mathbf{x}_F$ vectors as in (24) (compare (27), setting $\tau_S = \mu_S$ and $\tau_F = \mu_F$). However, this would imply a covariance structure that is not adequately described by (40) (compare (28); we might set $\Phi_S = \Sigma_{SS}$ and $\Phi_F = \Sigma_{FF}$, leaving a non-diagonal $\boldsymbol{\Theta}$ for $(\mu_S\mu_S' \otimes \Sigma_{FF}) + (\Sigma_{SS} \otimes \mu_F\mu_F')$, which is not parsimonious because $\boldsymbol{\Theta}$ does not really contain new parameters and should be restricted to equal $\boldsymbol{\Theta} = (\tau_S\tau_S' \otimes \Phi_F) + (\Phi_S \otimes \tau_F\tau_F')$).

The covariance structure in (40), with diagonal $\boldsymbol{\Theta}$, is the basic variant of Browne's (1984) composite direct product model. If we assume the residual factors ε to be modelled as in equation (33) (with ε_S and ε_F independent, and having zero means), then the $\boldsymbol{\Theta}$ restriction can be imposed as well (Browne, 1984, eq. 5.2):

$$\mathrm{Cov}(\mathbf{x}, \mathbf{x}') = \Sigma = (\Phi_S \otimes \Phi_F) + (\boldsymbol{\Theta}_S \otimes \boldsymbol{\Theta}_F). \tag{41}$$

Setting $\Phi_S = \mathbf{U}_{J\times J}$ and $\boldsymbol{\Theta}_S = \mathbf{I}_{J\times J}$ yields the model for parallel batteries of tests (Browne, 1984, eq. 5.1). Scale-invariant versions of composite direct product models are obtained by using Λ and Δ to absorb scaling, so that Φ_S and Φ_F can be interpreted as correlation matrices (additionally setting $\Lambda = \Delta$ yields Browne's Equation 5.7). Subsequent imposition of the Λ and Δ restrictions (11) and (13) cancels the scaling invariance, rendering a reparameterization of (41).

It is once more stressed that particular mean and covariance structures can be implied by different underlying score models. This is underlined by Browne (1984, Section 8) who presents two observed score models, one multiplicative and one additive, that both imply the covariance structure given by (41).

7. Illustration

As an example, S3MMs are applied to data from the field of economic psychology. Antonides, Farago, Ranyard & Tyszka (1996) carried out a study of how people from different countries perceive various economic activities. About 800 males (between 30 and 50 years of age, in full-time employment) were asked to rate 20 economic activities on 12 bipolar 7-point scales. There were two versions of the questionnaire. So, per version, only the data of about 400 subjects can be used for the analysis of the questionnaire. Each subject provided scores on 240 (20×12) variables. Computational limitations prevent the analysis of mean and covariance structures of 240 variables. Therefore nine economic activities and seven perception scales have been selected for our illustrative example; see Table 2. Seven perceptions of nine activities yield 63 (9×7) variables. The number of subjects without missing data is 390.

The data set consists of 63 scores of 390 subjects. The data are arranged as follows: for each subject, first the scores on the seven perception scales for the first economic activity, then the scores on the seven perception scales for the second economic activity, and so on, until the ninth economic activity. So, the first mode is the subjects mode, and can be considered as random. The economic activities and perception scales are the fixed modes. In the data set, the levels of the economic activities change more slowly than the levels of the perception scales, so the economic activities correspond with the slow mode levels and the perception scales correspond with the fast mode levels.

It follows that there are nine slow mode levels ($J = 9$), and seven fast mode levels ($K = 7$). The idea is that the nine activities can be summarized by four slow mode components ($Q = 4$), and the seven perceptions by three fast mode components ($R = 3$). See Table 2 for a description of the components.

In the first analyses, we fit a series of hierarchically related models where all (4×3) combinations of components are transformed into 12 separate factors. Then, we fit a series of models with seven factors, four coinciding with the four slow mode components and three coinciding with the three fast mode components. Finally, we look at the exploratory use of S3MMs, with varying numbers of factors. But first we consider the choice of a suitable estimation method.

Table 2. Economic activities and perception scales

Economic activities	Components
1 Being employed as a teacher	1 Being employed (EMPD)
2 Being employed as a house painter	
3 Being the owner of a newspaper	2 Being employer (EMPR)
4 Being the owner of a night club	
5 Buying life insurance	3 Uncontroversial financial activities (UFIN)
6 Buying government bonds	
7 Paying taxes	
8 Gambling in the casino	4 Controversial financial activities (CFIN)
9 Receiving bribes	
Perception scales	**Components**
1 Moral	1 Social values (SOCV)
2 Beneficial for society	
3 Requires great effort	2 Economic values (ECOV)
4 Requires much knowledge	
5 Requires a lot of financial resources	
6 Well known	3 Expected risk (RISK)
7 Risky	

Note: The selected economic activities and perception scales make up 63 items (9×7). The full questionnaire of Antonides *et al.* (1996) has 240 items (20×12).

7.1. Choice of estimation method and fit statistics

The data of Antonides *et al.* (1996) are well suited to illustrate the models discussed, but they are not multivariate normal. Firstly, scores on seven-point scales are not really continuous. Secondly, supposition of an S3MM makes multivariate normality uncertain, given the multiplicative nature of S3MMs and many of the possibly underlying score models. Univariate tests on the normality of the 63 variables (390 observations) reveal that about half of the variables are not distributed even approximately normally. This makes the choice of a suitable estimation method difficult. We could turn to an estimation method that does not require multivariate normality, such as the weighted least squares method (WLS; see Bollen, 1989). Or we could retain the ML estimation method but use a corrected (scaled) test statistic (see Bentler & Dudgeon, 1996). Both options are available in EQS (Bentler, 1995) and recent versions of LISREL (as of version 8.20; Jöreskog, Sörbom, Du Toit & Du Toit, 1999). WLS estimation requires the calculation of the asymptotic variances and covariances of all elements of the covariance matrix $\mathbf{S}$. In our case, with 63 observed variables, the weight matrix contains more than 2 million elements and is unstable when based on only 390 observations.

Given our relatively small sample size we cannot safely use the WLS estimation method. The use of corrected chi-square statistics may alleviate the problem of non-normality somewhat. However, the current version of Mx (version 1.44; Neale, 1997), which we would like to use to estimate our S3MMs, does not provide a chi-square that takes account of non-normality (besides the WLS chi-square). We will therefore use the normal theory ML estimation method and accompanying test statistic (i.e. based on the discrepancy function of equation (19)). As the assumption of multivariate normality is violated, the resulting test statistic may not have a chi-square distribution and the standard errors may not be correct. Yet there is little reason to believe that misapplying the ML estimator will result in seriously biased point estimates of the parameters (Bollen, 1989).

Mx computes three measures of fit (Neale, 1997): the chi-square measure of overall goodness of fit (CHISQ), the root mean square error of approximation (RMSEA), and Akaike's information criterion (AIC). When the ML discrepancy function is used, the AIC is linearly related to the expected cross-validation index (ECVI) of Browne & Cudeck (1992). We prefer the ECVI to the AIC because of its attractive interpretation, and because Browne & Cudeck (1992) provide formulae to calculate confidence intervals for both the ECVI and the RMSEA. They also provide 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. (For the personal computer, small computer programs for computing the confidence intervals for RMSEA and ECVI are freely available through the Internet. M. W. Browne's FITMOD.EXE can be obtained at http://quantrm2.psy.ohio-state.edu/browne (included in the MUTMUM.ZIP package), and P. Dudgeon's RMSEA.ZIP can be obtained at http://www.mhri.edu.au/~pld.)

Although we cannot use the chi-square distribution to interpret the (non-central) chi-square statistic, and although the derived statistics RMSEA and ECVI depend on normality assumptions as well, we can still use the RMSEA and ECVI to compare the fit of different models to the same data. Simulation studies show that under non-normality, CHISQ values will be overestimated (Curran, West & Finch, 1996), so that RMSEA values will turn out too high. We will apply the rule of thumb conservatively. For comparison we will also present the ECVI values, but for judging the fit of the models we will focus mainly on the RMSEA. Tables 3–5 give the fit indices for all models to be discussed below.

7.2. Models with 12 factors

We first consider a model (model 1.0) without the imposition of any of the Kronecker-product restrictions (11)–(16). Next, we consider models with only one of the restrictions imposed. Finally, we consider models with combinations of several restrictions.

Model 1.0 imposes none of the Kronecker-product restrictions, but its pattern of fixed and free parameters does conform to a three-mode model. It has the mean and covariance structures of the LLVM given by (9) and (10), with $J = 9$, $K = 7$, $Q = 4$, $R = 3$ and $P = QR = 12$. The 12 factors are named 'social value (SOCV) of being employed (EMPD)', 'economic value (ECOV) of EMPD', 'expected risk (RISK) of EMPD', 'SOCV of being employer (EMPR)', 'ECOV of EMPR', 'RISK of EMPR', 'SOCV of uncontroversial financial activities (UFIN)', 'ECOV of UFIN', 'RISK of UFIN', 'SOCV of controversial financial activities (CFIN)', 'ECOV of CFIN' and 'RISK of CFIN' (compare Table 2). Λ has a simple structure. Each variable loads on one factor only. Fitting this LLVM, with identification restrictions

$$\lambda_{1,1} = \lambda_{3,2} = \lambda_{6,3} = \lambda_{15,4} = \lambda_{17,5} = \lambda_{20,6} = \lambda_{29,7} = \lambda_{31,8} = \lambda_{34,9} = \lambda_{50,10} = \lambda_{52,11} = \lambda_{55,12} = 1, \tag{42}$$

$$\tau_1 = \tau_3 = \tau_6 = \tau_{15} = \tau_{17} = \tau_{20} = \tau_{29} = \tau_{31} = \tau_{34} = \tau_{50} = \tau_{52} = \tau_{55} = 0, \tag{43}$$

$\Delta = \mathbf{I}_{JK \times JK}$, and $\boldsymbol{\Theta}$ diagonal, yields CHISQ(1824) = 5474.0 and RMSEA = 0.072. According to the RMSEA rule of thumb, the fit of model 1.0 is not unreasonable. For illustrative purposes we continue to investigate the various Kronecker-product restrictions.

Models 1.1–1.5 all have one of the Kronecker-product restrictions imposed. Model 1.1 has

Table 3. Fit results for three-mode models with twelve factors

Model	Restrictions	df	CHISQ	RMSEA	ECVI
1.0	none	1824	5474.0	0.072 (0.070–0.074)	15.38 (14.82–15.96)
1.1	Λ	1866	5678.3	0.073 (0.070–0.075)	15.69 (15.12–16.28)
1.2	τ	1866	5989.2	0.075 (0.073–0.078)	16.49 (15.90–17.10)
1.3	Φ	1887	5824.8	0.073 (0.071–0.075)	15.96 (15.38–16.56)
1.4	κ	1830	5924.5	0.076 (0.074–0.078)	16.51 (15.92–17.12)
1.5	$\boldsymbol{\Theta}$	1872	6114.8	0.076 (0.074–0.079)	16.78 (16.19–17.40)
1.6	Λ, τ	1908	6710.8	0.080 (0.078–0.083)	18.13 (17.50–18.78)
1.7	Λ, τ, Φ	1971	7020.7	0.081 (0.079–0.083)	18.60 (17.96–19.27)
1,8	Λ, τ, Φ, κ	1977	7169.2	0.082 (0.080–0.084)	18.95 (18.30–19.63)
1.9	Λ, τ, Φ, κ, $\boldsymbol{\Theta}$	2025	7797.9	0.086 (0.084–0.088)	20.32 (19.64–21.03)
1.10	Λ, τ, $\boldsymbol{\Theta}$	1956	7373.1	0.084 (0.082–0.086)	19.59 (18.92–20.27)
1.11	Λ, Φ	1929	5984.6	0.074 (0.071–0.076)	16.16 (15.57–16.77)
1.12	Λ, Φ, $\boldsymbol{\Theta}$	1977	6659.5	0.078 (0.076–0.080)	17.64 (17.02–18.29)
1.13	τ, κ	1872	6386.3	0.079 (0.077–0.081)	17.48 (16.87–18.12)
1.14	Φ, κ	1893	6321.9	0.078 (0.075–0.080)	17.21 (16.60–17.84)
1.15	Φ, κ, $\boldsymbol{\Theta}$	1941	6962.3	0.082 (0.080–0.084)	18.61 (17.96–19.27)

Note: $N = 390$, $J = 9$, $K = 7$, $P = 12$, $Q = 4$, $R = 3$; Δ restriction n.a.; RMSEA and ECVI values in parentheses denote 90% confidence intervals.

the Λ restriction (equation (11)) imposed with

$$\Lambda_S = \begin{vmatrix} \lambda_{S11} & 0 & 0 & 0 \\ \lambda_{S21} & 0 & 0 & 0 \\ 0 & \lambda_{S32} & 0 & 0 \\ 0 & \lambda_{S42} & 0 & 0 \\ 0 & 0 & \lambda_{S53} & 0 \\ 0 & 0 & \lambda_{S63} & 0 \\ 0 & 0 & \lambda_{S73} & 0 \\ 0 & 0 & 0 & \lambda_{S84} \\ 0 & 0 & 0 & \lambda_{S94} \end{vmatrix}, \quad \Lambda_F = \begin{vmatrix} \lambda_{F11} & 0 & 0 \\ \lambda_{F21} & 0 & 0 \\ 0 & \lambda_{F32} & 0 \\ 0 & \lambda_{F42} & 0 \\ 0 & \lambda_{F52} & 0 \\ 0 & 0 & \lambda_{F63} \\ 0 & 0 & \lambda_{F73} \end{vmatrix} \tag{44}$$

Table 4. Fit results for three-mode models with seven factors

Model	Restrictions	df	CHISQ	RMSEA	ECVI
2.0	none	1818	5210.0	0.069 (0.067–0.071)	14.74 (14.19–15.30)
2.1	Λ	1911	5964.3	0.074 (0.072–0.076)	16.20 (15.61–16.80)
2.2	τ	1866	6240.3	0.078 (0.076–0.080)	17.14 (16.53–17.76)
2.3	$\boldsymbol{\Theta}$	1866	5979.2	0.075 (0.073–0.077)	16.47 (15.88–17.08)
2.4	Λ, τ	1959	7081.3	0.082 (0.080–0.084)	18.82 (18.17–19.49)
2.5	Λ, τ, $\boldsymbol{\Theta}$	2007	7756.4	0.086 (0.084–0.088)	20.31 (19.62–21.01)

Note: $N = 390$, $J = 9$, $K = 7$, $P = 7$, $Q = 4$, $R = 3$; Δ, Φ and κ restrictions n.a.; RMSEA and ECVI values in parenthesis denote 90% confidence intervals.

Table 5. Fit results for exploratory three-mode models with varying numbers of factors; after Bentler & Lee (1978; 1979)

Model	No. of factors	df	CHISQ	RMSEA	ECVI
Λ_S, Λ_F, Γ echelon					
3.1	2 + 2 = 4	1919	6648.0	0.080 (0.078–0.082)	17.91 (17.28–18.56)
3.2	3 + 2 = 5	1905	6226.7	0.076 (0.074–0.079)	16.90 (16.30–17.53)
3.3	3 + 3 = 6	1884	5659.9	0.072 (0.070–0.074)	15.55 (14.98–16.14)
3.4	4 + 3 = 7	1858	5521.7	0.071 (0.069–0.073)	15.33 (14.77–15.91)
3.5	4 + 4 = 8	1821	4924.9	0.066 (0.064–0.068)	13.99 (13.46–14.53)
3.6	5 + 4 = 9	1777	4639.4	0.064 (0.062–0.067)	13.48 (12.97–14.01)
Λ_S, Λ_F echelon, Γ diag.					
4.1	2 × 2 = 4	1924	6713.8	0.080 (0.078–0.082)	18.06 (17.42–18.71)
4.2	3 × 2 = 6	1917	6285.8	0.077 (0.074–0.079)	16.99 (16.38–17.62)
4.3	3 × 3 = 9	1912	5792.9	0.072 (0.070–0.074)	15.75 (15.17–16.35)
4.4	4 × 3 = 12	1906	5496.2	0.070 (0.067–0.072)	15.02 (14.46–15.60)
4.5	4 × 4 = 16	1902	4937.8	0.064 (0.062–0.066)	13.60 (13.08–14.15)
4.6	5 × 4 = 20	1897	4824.5	0.063 (0.061–0.065)	13.34 (12.82–13.87)

Note: $N = 390$; only the Λ restriction is applied.

and $\Gamma = \mathbf{I}_{QR\times QR}$. This model, with the identification restrictions of (42) replaced by

$$\lambda_{S11} = \lambda_{S32} = \lambda_{S53} = \lambda_{S84} = \lambda_{F11} = \lambda_{F32} = \lambda_{F63} = 1, \tag{45}$$

gives a fit of CHISQ(1866) = 5678.3 and RMSEA = 0.073.

Model 1.1 is more parsimonious than model 1.0. In model 1.0, Λ is a 63×12 matrix, containing 63 non-zero factor loadings, 12 of which are fixed to provide scales for the factors (42). This leaves 51 factor loadings free to be estimated. In model 1.1, where the Λ restriction is applied, we have only 16 non-zero factor loadings (equation (44)), 7 of which are fixed to provide scales (45), leaving only 9 parameters free to be estimated. Hence, we gain 42 degrees of freedom. Actually, it turns out that the higher CHISQ value of model 1.1 is almost made up for by the gain in degrees of freedom, so that the relative fit of models 1.0 and 1.1 is about the same, as indicated by the almost coincident 90% confidence intervals for the RMSEA.

Table 6 gives the parameter estimates for Λ_S and Λ_F in model 1.1. The parameters can be interpreted as a kind of overall factor loadings for the economic activities and perception scales on the slow and fast mode factors (see Table 2). That is, although, for instance the economic activity 'being employed as a teacher' was evaluated seven times, on seven different perception scales, there is just one loading for that item on the EMPD factor. From the parameter estimates in Table 6 it appears that most constituent factors are measured evenly by their different indicators. An exception is the RISK factor, where the factor loadings show ($\lambda_{F63} = 1.00$ and $\lambda_{F73} = 0.28$) that one perception scale ('Well known') is three times more indicative than the other ('Risky').

Model 1.2 includes the τ restriction (12). In this model the identification restrictions of (43) are replaced by

$$\tau_{S1} = \tau_{S3} = \tau_{S5} = \tau_{83} = \tau_{F1} = \tau_{F3} = \tau_{F6} = 1. \tag{46}$$

These intercepts have been fixed at unity rather than zero, because the use of zeros would restrict the composite vector τ more than is necessary for identification. Model 1.2 has the same number of parameters to be estimated as model 1.1, but its fit is worse (Table 3).

In model 1.0 separate intercepts have been estimated for each combination of economic activities and perception scales. In model 1.2 slow and fast mode contributions to these separate intercepts are estimated. The estimates of parameters in τ_S and τ_F are given in Table 6. They can be interpreted as the (reversed) difficulty of the abstract $\mathbf{x}_S$ and $\mathbf{x}_F$ variables of the $f_{\mathbf{x}}$ function (21). For example, looking at the τ_F parameter estimates, it appears that, regardless of the economic activity, it is easier to agree with something being 'Beneficial for society' ($\tau_{F2} = 2.06$) than with something being 'Moral' ($\tau_{F1} = 1.00$). Or, looking at τ_S, it appears that it is easier to agree with questions about gambling ($\tau_{S8} = 1.00$) than with questions about bribes ($\tau_{S9} = -0.31$), regardless of the perception scales.

Here all *QR* combinations of slow and fast mode components transform into separate factors, so we can check whether the variances, covariances and means of these factors conform to a multiplicative structure. Imposing the Φ restriction (equation (15)) with $\varphi_{S11} = 1$ gives CHISQ(1887) = 5824.8 (model 1.3). Imposing the κ restriction (equation (16)) instead, with $\kappa_{S1} = 1$, yields CHISQ(1830) = 5924.5 (model 1.4). The RMSEA values in Table 3 indicate that model 1.3 fits about as well as model 1.0. Yet the point estimate of the ECVI for model 1.3 only just falls within the 90% confidence interval estimate of the ECVI for model 1.0. Model 1.4 fits worse than model 1.0.

Table 6. Selected parameter estimates

Model	Selected parameter estimates
1.1	$\Lambda_S = \begin{vmatrix} \mathit{1.00} & & & \\ 1.18 & & & \\ & \mathit{1.00} & & \\ & 1.02 & & \\ & & \mathit{1.00} & \\ & & 1.22 & \\ & & 1.25 & \\ & & & \mathit{1.00} \\ & & & 1.72 \end{vmatrix}, \quad \Lambda_F = \begin{vmatrix} \mathit{1.00} & & \\ 0.68 & & \\ & \mathit{1.00} & \\ & 1.00 & \\ & 0.63 & \\ & & \mathit{1.00} \\ & & 0.28 \end{vmatrix}$
1.2	$\tau_S = \begin{vmatrix} \mathit{1.00} & 0.39 & \mathit{1.00} & 0.92 & \mathit{1.00} & 0.50 & 0.87 & \mathit{1.00} & -0.31 \end{vmatrix}'$ $\tau_F = \begin{vmatrix} \mathit{1.00} & 2.06 & \mathit{1.00} & 1.41 & 3.57 & \mathit{1.00} & 3.16 \end{vmatrix}'$
1.3	$\Phi_S = \begin{vmatrix} \mathit{1.00} & & & \\ 0.69 & 1.24 & & \\ 0.43 & 0.42 & 0.68 & \\ 0.09 & 0.30 & 0.27 & 0.64 \end{vmatrix}, \quad \Phi_F = \begin{vmatrix} 0.95 & & \\ 0.36 & 0.81 & \\ 0.45 & 0.26 & 1.68 \end{vmatrix}$
1.4	$\kappa_S = \begin{vmatrix} \mathit{1.00} & 0.93 & 0.78 & 0.50 \end{vmatrix}', \quad \kappa_F = \begin{vmatrix} 4.77 & 4.73 & 2.65 \end{vmatrix}'$
1.5	$\text{diag}(\mathbf{\Theta}_S) = (\mathit{1.00}, 1.20, 1.01, 1.17, 1.36, 1.78, 1.37, 1.78, 1.48)$ $\text{diag}(\mathbf{\Theta}_F) = (0.95, 1.26, 1.35, 1.26, 1.61, 1.73, 2.23)$

Note: Coefficients in italics are fixed parameters; diag ( ) designates the diagonal elements of the matrix concerned.

Table 6 gives the parameter estimates for Φ_S and Φ_F in model 1.3 and κ_S and κ_F in model 1.4. One possible interpretation of these parameters is that they are the variances, covariances and means of the constituent factors ξ_S and ξ_F of composite function f_ξ (22). Φ_S and Φ_F are easier to interpret if the scales of ξ_S and ξ_F are set by fixing their variances to unity. However, the Φ restriction would then only involve correlations and not the variances. Still, we can infer from Φ_S, for example, that the EMDP and CFIN factors hardly co-vary. The κ estimates are not very interesting in a single-group single-occasion model like the present model. We could just as well have provided origins for the ξ factors by fixing their means at zero ($\kappa = \mathbf{0}$). But that would have prevented the illustration of the κ restriction.

As Δ is an identity matrix here, the Δ restriction (13) is not applicable. However, the $\mathbf{\Theta}$ restriction (14) can be tested. By imposing the Θ restriction, we test the hypothesis of a multiplicative structure of the residual variances. In model 1.5 the $\mathbf{\Theta}$ restriction is imposed with matrices $\mathbf{\Theta}_S$ and $\mathbf{\Theta}_F$ both diagonal, with $\theta_{S11} = 1$. The fit of model 1.5 is not good (Table 3). Parameter estimates for $\mathbf{\Theta}_S$ and $\mathbf{\Theta}_F$ in model 1.5 are given in Table 6. They can be interpreted as the variances of the ε_S and ε_F variables that feature in the composite function f_ε (23). The values of the residual variances cannot be interpreted relative to each other as the variances of the observed variables have not been standardized.

In models 1.1–1.5 we have imposed the Kronecker-product restrictions one at a time. In models 1.6–1.9 we have added restrictions to model 1.1. In model 1.1 only the Λ restriction is imposed. In model 1.6 the τ restriction is added, so that all measurement parameters are restricted (as in (35)–(37)). The fit of model 1.6 is poor (Table 3). Nevertheless, for the sole purpose of illustration, we have fitted three more models. In models 1.7, 1.8 and 1.9, the Kronecker-product restrictions on matrices Φ, κ and $\boldsymbol{\Theta}$ are added. The fit of model 1.9, where all applicable restrictions are imposed, serves as a test of the mean and covariance structures given by (17) and (18). Not surprisingly, the fit of models 1.7–1.9 is poor. The same goes for model 1.10, with Kronecker-product restrictions on all measurement parameter matrices (Λ and τ), and the matrix of residual variances ($\boldsymbol{\Theta}$).

Other models that were fitted are model 1.11, with the common covariance structure parameters restricted (Λ and $\boldsymbol{\Theta}$); model 1.12, with all covariance structure parameters restricted (Λ, Φ and $\boldsymbol{\Theta}$); model 1.13, with all mean structure parameters restricted (τ and κ); model 1.14, with all common mean and covariance structure parameters restricted (Φ and κ); and model 1.15, with all structural parameters restricted (Φ, κ and $\boldsymbol{\Theta}$). Model 1.15 can be seen as the counterpart of model 1.6 where all measurement parameters are restricted.

Comparing the fit results of all models, it appears that, in this example, there is perhaps a three-mode structure in the common covariance structure (Λ, Φ), but not in the mean structure (τ, κ) or in the residual covariance structure ($\boldsymbol{\Theta}$); see Table 3.

7.3. Models with seven factors

In the models discussed above all 12 combinations of slow and fast mode components were transformed into 12 separate factors. In the models with the Λ restriction this was done by choosing $\Gamma = \mathbf{I}_{QR\times QR}$. An alternative choice for Γ is

$$\Gamma = \begin{vmatrix} \gamma_{1,1} & 0 & 0 & 0 & \gamma_{1,5} & 0 & 0 \\ \gamma_{2,1} & 0 & 0 & 0 & 0 & \gamma_{2,6} & 0 \\ \gamma_{3,1} & 0 & 0 & 0 & 0 & 0 & \gamma_{3,7} \\ 0 & \gamma_{4,2} & 0 & 0 & \gamma_{4,5} & 0 & 0 \\ 0 & \gamma_{5,2} & 0 & 0 & 0 & \gamma_{5,6} & 0 \\ 0 & \gamma_{6,2} & 0 & 0 & 0 & 0 & \gamma_{6,7} \\ 0 & 0 & \gamma_{7,3} & 0 & \gamma_{7,5} & 0 & 0 \\ 0 & 0 & \gamma_{8,3} & 0 & 0 & \gamma_{8,6} & 0 \\ 0 & 0 & \gamma_{9,3} & 0 & 0 & 0 & \gamma_{9,7} \\ 0 & 0 & 0 & \gamma_{10,4} & \gamma_{10,5} & 0 & 0 \\ 0 & 0 & 0 & \gamma_{11,4} & 0 & \gamma_{11,6} & 0 \\ 0 & 0 & 0 & \gamma_{12,4} & 0 & 0 & \gamma_{12,7} \end{vmatrix}. \tag{47}$$

Retaining the Λ_S and Λ_F specifications of (44), this Γ transforms the 12 combinations of slow and fast mode components into seven factors. They are named after the components EMPD, EMPR, UFIN, CFIN, SOCV, ECOV, RISK (see Table 2). The factor ξ can be conceived as

$$\xi = \begin{vmatrix} \xi_S \\ \xi_F \end{vmatrix}. \tag{48}$$

We get a Λ matrix where each variable loads on two factors: one factor associated with a slow mode component and one associated with a fast mode component. Consequently, not all parameters representing covariations between these two types of factors are identified (as checked by evaluating the Jacobian's kernel). Therefore, part of the Φ matrix is fixed at zero:

$$\Phi = \begin{vmatrix} 1 & & & & & & \\ \varphi_{21} & 1 & & & & & \\ \varphi_{31} & \varphi_{32} & 1 & & & & \\ \varphi_{41} & \varphi_{42} & \varphi_{43} & 1 & & & \\ 0 & 0 & 0 & 0 & 1 & & \\ 0 & 0 & 0 & 0 & \varphi_{65} & 1 & \\ 0 & 0 & 0 & 0 & \varphi_{75} & \varphi_{76} & 1 \end{vmatrix}. \tag{49}$$

As indicated, the diagonal elements of Φ are fixed at unity, so that Φ can be interpreted as a correlation matrix. In addition, the scale-setting restrictions of (45) on Λ_S and Λ_F (44) are imposed as well, so that all Γ elements (47) can be left free to be estimated. Origins are imposed by fixing the factor means, by setting all κ elements at zero. This model is referred to as model 2.1. Note that, with the non-diagonal Γ matrix, the Kronecker-product restrictions are not applicable to matrix Φ and vector κ. So providing scales and origins through Φ and κ does not limit us in looking for multiplicative structures in Φ and κ, whereas it should give some leeway in investigating the Kronecker-product restrictions on Λ and τ.

Model 2.0 is the unrestricted version of model 2.1. It is similar to Jöreskog's (1971) factor model for multitrait-multimethod data (Wothke, 1996). The Λ matrices of models 2.0 and 2.1 have the same pattern of fixed zeros, but in model 2.0 there are no other restrictions on the factor loadings; all non-zero Λ elements are set free to be estimated.

The fit of model 2.1, with only the Λ restriction imposed (with $\lambda_{S11} = 1$), may be considered reasonable (CHISQ(1911) = 5964.3, RMSEA = 0.074), but is much worse than the fit of model 2.0, since the point estimate of the RMSEA for model 2.1 does not fall within the 90% confidence interval for model 2.0's RMSEA (0.067–0.071; see Table 4). In models 2.2 and 2.3 the τ and $\boldsymbol{\Theta}$ restrictions are imposed instead of the Λ restriction, but these models do not give a better fit (Table 4). The fit deteriorates substantially if we combine restrictions, as in models 2.4 and 2.5 (Table 4).

It should be noted that the interpretation of the parameter estimates of the models with the Λ restriction is clearer than that of the models without the Λ restriction. For example, the composed factor loadings in the restricted Λ matrix of model 2.1 all have positive values, whereas the estimated factor loadings in model 2.0 show more variation, undermining the interpretation of some of the common factors. So, in spite of worse fit, model 2.1 may still be preferred to model 2.0 on account of better interpretability.

Comparison of the results in Tables 3 and 4 reveals that, although the unrestricted seven-factor model (model 2.0) fits better than the unrestricted 12-factor model (model 1.0), the S3MMs with 12 factors generally fit a little better than their counterparts with seven factors.

7.4. Models for exploratory analysis

If little is known about a data set, it is perhaps better to turn to real exploratory analysis, like the three-mode principal components analysis mentioned before (equation (1)). In fact, Veldscholte, Kroonenberg & Antonides (1998) have already carried out such analyses with the Antonides *et al.* (1996) data. However, Bentler & Lee (1978, 1979) show how stochastic three-mode models can be used in an exploratory way.

In exploratory research neither the number nor the content (meaning) of factors is known. Therefore, we investigate model fit when increasing numbers of slow and fast mode components are specified. In the first series Λ_S, Λ_F, and Γ are all of echelon form, with some necessary identification restrictions in Λ_S and Λ_F:

$$\Lambda_S = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \lambda_{S51} & \lambda_{S52} & \lambda_{S53} & \lambda_{S54} \\ \lambda_{S61} & \lambda_{S62} & \lambda_{S63} & \lambda_{S64} \\ \lambda_{S71} & \lambda_{S72} & \lambda_{S73} & \lambda_{S74} \\ \lambda_{S81} & \lambda_{S82} & \lambda_{S83} & \lambda_{S84} \\ \lambda_{S91} & \lambda_{S92} & \lambda_{S93} & \lambda_{S94} \end{vmatrix}, \Lambda_F = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \lambda_{F41} & \lambda_{F42} & \lambda_{F43} \\ \lambda_{F51} & \lambda_{F52} & \lambda_{F53} \\ \lambda_{F61} & \lambda_{F62} & \lambda_{F63} \\ \lambda_{F71} & \lambda_{F72} & \lambda_{F73} \end{vmatrix} \qquad (50)$$

and

$$\Gamma = \begin{vmatrix} \gamma_{1,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma_{2,1} & \gamma_{2,2} & 0 & 0 & 0 & 0 & 0 \\ \gamma_{3,1} & \gamma_{3,2} & \gamma_{3,3} & 0 & 0 & 0 & 0 \\ \gamma_{4,1} & \gamma_{4,2} & \gamma_{4,3} & \gamma_{4,4} & 0 & 0 & 0 \\ \gamma_{5,1} & \gamma_{5,2} & \gamma_{5,3} & \gamma_{5,4} & \gamma_{5,5} & 0 & 0 \\ \gamma_{6,1} & \gamma_{6,2} & \gamma_{6,3} & \gamma_{6,4} & \gamma_{6,5} & \gamma_{6,6} & 0 \\ \gamma_{7,1} & \gamma_{7,2} & \gamma_{7,3} & \gamma_{7,4} & \gamma_{7,5} & \gamma_{7,6} & \gamma_{7,7} \\ \gamma_{8,1} & \gamma_{8,2} & \gamma_{8,3} & \gamma_{8,4} & \gamma_{8,5} & \gamma_{8,6} & \gamma_{8,7} \\ \gamma_{9,1} & \gamma_{9,2} & \gamma_{9,3} & \gamma_{9,4} & \gamma_{9,5} & \gamma_{9,6} & \gamma_{9,7} \\ \gamma_{10,1} & \gamma_{10,2} & \gamma_{10,3} & \gamma_{10,4} & \gamma_{10,5} & \gamma_{10,6} & \gamma_{10,7} \\ \gamma_{11,1} & \gamma_{11,2} & \gamma_{11,3} & \gamma_{11,4} & \gamma_{11,5} & \gamma_{11,6} & \gamma_{11,7} \\ \gamma_{12,1} & \gamma_{12,2} & \gamma_{12,3} & \gamma_{12,4} & \gamma_{12,5} & \gamma_{12,6} & \gamma_{12,7} \end{vmatrix}. \qquad (51)$$

In the second series of models, only Λ_S and Λ_F are of echelon form, without further restrictions (apart from fixing one element of either Λ_S or Λ_F to remove the indeterminacy introduced by the Kronecker-product restriction), and Γ is an identity matrix, so that all *QR* combinations of slow mode components and fast mode components are transformed into separate factors. In this case it suffices to fix just one element of either Λ_S or Λ_F (e.g. $\lambda_{S11} = 1$). In both cases the factors are assumed orthogonal, and the factor variances are fixed at unity, so that Φ is an identity matrix. The diagonal matrix $\boldsymbol{\Theta}$ and vector τ are free, and κ is a zero matrix (i.e. the mean structure is not modelled).

The fit results are reported in Table 5. As expected, the fit improves with increasing numbers of components. In fact, it is not clear how many slow and fast mode components are required. It appears that with higher numbers of factors, the stricter models 4.4, 4.5 and 4.6 fit relatively better than their counterparts, models 3.4, 3.5 and 3.6.

Models 3.4 and 4.4 have confirmatory counterparts in models 2.1 and 1.1. It appears that the exploratory models, with their greater numbers of parameters, do fit better. However, it is impossible to interpret the parameter estimates of models like model 4.4 and especially model 3.4.

Concerning the interpretation of parameter estimates in an exploratory context, Bloxom (1968) proposes a simple structure rotation of Γ. Bentler & Lee (1979) deem interpretation of Γ unjustified, because of its arbitrary identification position. They go on to note that matrices Λ_S and Λ_F are also subject to arbitrary transformations. To ease interpretation, they suggest first rotating Λ into a simple interpretable form, and subsequently transforming Λ_S and Λ_F. However, we do not pursue that here.

7.5. Model evaluation

The preceding analyses were carried out for illustration only, to demonstrate various modelling aspects. It was not our goal to arrive at a 'best' model. Still, inspecting the fit results in Tables 3–5 combined, it appears that model 4.6 has by far the best fit, especially if judged by the ECVI index, which seems to be more sensitive than the RMSEA index. However, the parameter estimates for the exploratory S3MMs are difficult to interpret, and Model 4.6 is no exception. Following Browne (1984), we believe that model selection should be based not only on goodness of fit, but also on interpretability of the parameter estimates. Confirmatory S3MMs, with their simply structured matrices of factor loadings, are much easier to interpret.

Considering the fit of the confirmatory S3MMs, as reported in Tables 3 and 4, it appears that the unrestricted models 1.0 and 2.0 do fit better than models with one or more of the Kronecker-product restrictions imposed. Yet imposing such restrictions makes the models much more parsimonious. Generally, the fewer parameters a model has, the better the model will do in a cross-validation procedure. Cross-validation is probably the best method to evaluate the fit of competing models. However, we do not have a second sample here, neither have we split the present sample in two in view of the small sample size and large number of variables.

7.6. Practical considerations

As mentioned earlier, all analyses reported were carried out with the computer program Mx (Neale, 1997). Fitting a single S3MM with Mx to data sets like the one used in our illustration took between 2 and 12 hours on personal computers with Intel Pentium processors (166 MHz and 266 MHz), and sometimes even longer. Alternatively the same analyses could be done with standard SEM programs, which generally converge faster. However, in order to apply one of the Kronecker-product restrictions, the user must specify restrictions for almost all elements in the matrix concerned. For example, using LISREL to fit model 1.2, application of the τ restriction involves 48 restrictions of the form $\tau_9 = \tau_8 * \tau_2 * \tau_1^{-1}$. Even then, one often encounters non-convergence or improper solutions. The main advantage of Mx over other SEM programs is that restrictions such as the Kronecker-product restrictions can be applied by just typing the matrix equation involved on a single line. The ease of preparing Mx scripts compensates for the long running times.

8. Conclusion

In conclusion, we line up various models that are related to the stochastic three-mode models for mean and covariance structures described here.

Bloxom (1968) and Bentler & Lee (1978, 1979) have described the three-mode factor analysis model. Bentler *et al.* (1988) showed how this model can be extended with scale-absorbing parameters (after Lee & Fong, 1983), with more than three modes, and with structural relations between factors. These models are for the covariance structure only, which means that either deviation scores are analysed (i.e. $\mathbf{x} - \mu$ instead of $\mathbf{x}$), or the factor means are assumed zero ($\kappa = \mathbf{0}$) and the means of the observed variables disappear in the unrestricted intercepts ($\tau = \mu$). The three-mode factor analysis model can then be described in terms of our S3MMs as including only the Λ restriction.

The composite direct product models of Browne (1984) are also for covariance structures only. They are special cases of S3MMs with Λ diagonal, with the Φ restriction and optionally the $\boldsymbol{\Theta}$ restriction imposed. Browne (1984) shows that the parallel batteries model is a special case of a direct product model, and he discusses the relationships with three-mode factor analysis models. Wothke & Browne (1990) demonstrate that composite direct product models can be parameterized as second-order factor models. Browne (1989) also shows that direct product models are difficult to distinguish from the covariance component analysis models of Bock (1960), as they give approximately the same fit—although Wothke (1996) mentions counterexamples.

The longitudinal three-mode model is an S3MM for multivariate longitudinal data. It has the Λ and τ restrictions imposed, with $\Lambda_S = \mathbf{I}_{J \times J}$ and $\tau_S = \mathbf{u}_{j \times 1}$. Oort (1999) describes how further modelling of the Φ, κ and $\boldsymbol{\Theta}$ matrices leads to special cases such as second-order factor analysis models, McDonald's (1984) invariant factors model, latent growth-curve models, and autoregressive models for mean and covariance structures.

Other special cases of S3MMs like the one given by (36) and (37) originate from further modelling of Φ, κ and $\boldsymbol{\Theta}$, with or without the Kronecker-product restrictions.

Tucker's (1966) three-mode principal components model of (1) is not a special case of the S3MMs. In Tucker's model the levels of all three modes are fixed. Bloxom's (1968) model of equation (2), which *is* a special case of the S3MMs, is obtained by taking the levels of the first mode as random. That is why Bloxom's three-mode factor analysis model is called a stochastic version of Tucker's three-mode principal components model. Kroonenberg (1983) and Kiers (1991) describe several special cases of the principal components model, such as the parallel factor analysis model of Harshman (1970) and Carroll & Chang (1970). Of course, stochastic versions of special cases of three-mode principal components models are themselves special cases of the S3MMs discussed here.

Tucker's three-mode principal components model is a model for the observed scores. However, it can be applied to cross-product (or variance-covariance) matrices as well (Tucker, 1972). It is then known as the three-mode scaling model (Kroonenberg, 1983).

Kroonenberg (1983) and Kiers (1991) also mention several special cases of the three-mode scaling model, such as the INDSCAL model of Carroll & Chang (1970). Again, stochastic versions of the three-mode scaling model and its special cases are special cases of the S3MMs presented here. Apart from that, it should be noted that S3MMs can also be used in a non-stochastic way, yielding results that are comparable with the results of a three-mode scaling analysis of the variance-covariance matrix. This can be achieved by using the S3MM with the Λ restriction, setting the residual variances and covariances equal to zero ($\boldsymbol{\Theta} = \mathbf{0}_{JK \times JK}$), analysing the covariance structure only, thereby imposing a minimal number of constraints, for identification only, and estimating the model parameters through the unweighted least squares method.

Two obvious extensions of the S3MMs should be mentioned. Firstly, by using the fit function of (20), S3MMs can be used in multiple group analysis. If the measurement parameters (in Λ and τ) are constrained to be equal across groups, then the other structural parameters can be used to investigate differences in means, variances and covariances across groups. Secondly, S3MMs are easily extended to multi-mode models. For example, the Λ restriction in a stochastic n-mode model, where the levels of the first mode are considered random, becomes $\Lambda = (\Lambda_2 \otimes \Lambda_3 \otimes \ldots \otimes \Lambda_n)\Gamma$, where Λ_2, Λ_3, and Λ_n contain the component loadings for the second mode, third mode and nth mode, respectively (see Bentler *et al.*, 1988, p. 111, eq. 7). The other Kronecker-product restrictions ((12)–(16)) can be modified similarly. Extending the model given by (29) in the same way yields the multi-mode direct product model discussed by Verhees & Wansbeek (1990).

Finally, note that the meaning of the parameters in the special cases of S3MM is always the same as the meaning of the same parameters in more general S3MMs. Thus, irrespective of whether the S3MM parameters in special models are fixed at zero, fixed at unity, fixed at another value, or free to be estimated, their meaning is determined by their function in the model. Therefore our description and interpretation of S3MMs facilitates the understanding of special cases of S3MMs.

Acknowledgements

This research was conducted at Leiden University, as part of a research project, 'Theory and practice of multivariate longitudinal analysis', funded by the Netherlands Organization of Scientific Research (NWO grant 575-30-005). The author thanks P. M. Kroonenberg, M. W. Browne, H. A. L. Kiers, C. V. Dolan, and an anonymous referee for their helpful comments, and G. Antonides for making his data available for secondary analysis.

References

Antonides, G., Farago, K., Ranyard, R. & Tyszka, T. (1996). 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.

Bekker, P. A., Merckens, A. & Wansbeek, T. J. (1994). *Identification, equivalent models, and computer algebra*. Boston: Academic Press.

Bentler, P. M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.

Bentler, P. M. & Dudgeon, P. (1996). Covariance structure analysis: statistical practice, theory, and directions. *Annual Review of Psychology*, **47**, 563–592.

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 analyses. *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 and Data Analysis*, **6**, 107–118.

Bentler, P. M. & Weeks, D. G. (1980). Linear structural equations with latent variables. *Psychometrika*, **45**, 289–308.

Bloxom, B. (1968). A note on invariance in three-mode factor analysis. *Psychometrika*, **33**, 347–350.

Bock, R. D. (1960). Components of variance analysis as a structural and discriminal analysis of psychological tests. *British Journal of Statistical Psychology*, **13**, 151–163.

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. (1987). Robustness of statistical inference in factor analysis and related models. *Biometrika*, **74**, 375–384.

Browne, M. W. (1989). Relationships between an additive model and a multiplicative model for multitrait-multimethod matrices. In R. Coppi & S. Bolasco (Eds), *Analysis of multiway data matrices*. Amsterdam: Elsevier.

Browne, M. W. & Arminger, G. (1995). Specification and estimation of mean- and covariance-structure models. In G. Arminger, C. C. Clogg & M. E. Sobel (Eds), *Handbook of statistical modeling for the social and behavioral sciences*. New York: Plenum Press.

Browne, M. W. & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, **21**, 230–258.

Browne, M. W. & Shapiro, A. (1988). Robustness of normal theory methods in the analysis of linear latent variate models. *British Journal of Mathematical and Statistical Psychology*, **41**, 193–208.

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.

Char, B. W., Geddes, K. O., Gonnet, G. H., Leong, B., Monagan, M. B. & Watt, S. M. (1991). *Maple V language reference manual*. New York: Springer-Verlag.

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.

Graybill, F. A. (1983). *Matrices with applications in statistics*, 2nd ed. Belmont, CA: Wadsworth.

Harshman, R. A. (1970). Foundations of the PARAFAC procedure: Models and conditions for an 'explanatory' multi-mode factor analysis. *UCLA Working Papers in Phonetics*, **16**, 1–84.

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, **31**, 109–133.

Jöreskog, K. G. & Sörbom, D. (1996). *LISREL 8 user's guide*, 2nd ed. Chicago: Scientific Software International, Inc.

Jöreskog, K. G., Sörbom, D., Du Toit, S. H. C. & Du Toit, M. (1999). *LISREL 8: New statistical features*. Chicago: Scientific Software International, Inc.

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

Kroonenberg, P. M. (1983). *Three-mode principal component analysis*. Leiden: DSWO Press.

Kroonenberg, P. M. (1994). The TUCKALS line: A suite of programs for three-way analysis. *Computational Statistics and Data Analysis*, **18**, 73–96.

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.

McDonald, R. P. (1984). The invariant factors model for multimode data. In H. G. Law, Snyder, C. W. Jr, J. A. Hattie & R. P. McDonald (Eds), *Research methods for multimode data analysis*. New York: Praeger.

Neale, M. C. (1997). *Mx: Statistical modeling*, 4th ed. Richmond, VA: Department of Psychiatry, Virginia Commonwealth University (http://views.vcu.edu/mx).

Oort, F. J. (1996). Using restricted factor analysis in test construction. Doctoral dissertation, Faculty of Psychology, University of Amsterdam.

Oort, F. J. (1999). Three-mode models for multivariate longitudinal data. Manuscript submitted for publication. Department of Medical Psychology, Academic Medical Centre, University of Amsterdam.

Shapiro, A. (1986). Asymptotic theory of overparameterized structural models. *Journal of the American Statistical Association*, **81**, 142–149.

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

Tucker, L. R. (1972). Relations between multidimensional scaling and three-mode factor analysis. *Psychometrika*, **37**, 3–27.

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.

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.

Wolfram, S. (1991). *Mathematica: A system for doing mathematics by computer*, 2nd ed. Redwood City, CA: Addison-Wesley.

Wothke, W. (1996). Models for multitrait-multimethod matrix analysis. In G. A. Marcoulides & R. E. Schumacker (Eds), *Advanced structural equation modeling: Issues and techniques*. Mahwah, NJ: Lawrence Erlbaum.

Wothke, W. & Browne, M. W. (1990). The direct product model for the MTMM matrix parameterised as a second order factor analysis model. *Psychometrika*, **55**, 255–262.

Received 6 October 1998; revised version received 12 April 1999