PSYCHOMETRIKA—VOL. 68, NO. 1, 105–121
MARCH 2003

FOUR SIMULTANEOUS COMPONENT MODELS FOR THE ANALYSIS OF MULTIVARIATE TIME SERIES FROM MORE THAN ONE SUBJECT TO MODEL INTRAINDIVIDUAL AND INTERINDIVIDUAL DIFFERENCES

MARIEKE E. TIMMERMAN AND HENK A.L. KIERS

UNIVERSITY OF GRONINGEN

A class of four simultaneous component models for the exploratory analysis of multivariate time series collected from more than one subject simultaneously is discussed. In each of the models, the multivariate time series of each subject is decomposed into a few series of component scores and a loading matrix. The component scores series reveal the latent data structure in the course of time. The interpretation of the components is based on the loading matrix. The simultaneous component models model not only intraindividual variability, but interindividual variability as well. The four models can be ordered hierarchically from weakly to severely constrained, thus allowing for big to small interindividual differences in the model. The use of the models is illustrated by an empirical example.

Key words: three-way analysis, exploratory longitudinal analysis.

1. Introduction

A time series is a collection of usually quantitative observations made sequentially in time on the same observation unit. An example is the scores on a depression scale collected daily from the same subject. The observations can be made from more than one variable at each measurement occasion, in which case multivariate time series are obtained. In this paper, we deal with multivariate time series data from more than one subject, where the measurement occasions are not necessarily at equal time points for the subjects, but the scores are collected on the same variables for all subjects. It is assumed that, as far as scores are collected at unequal time points, the resulting "missing values" are missing completely at random (Little & Rubin, 1987). Furthermore, it is assumed that within a measument occasion of a particular subject the scores of all variables are obtained. If the latter is not the case, one should resort to data imputation or related techniques, taking the (presumable) mechanisms behind the missing data into account.

Various models have been proposed to model multivariate time series collected from one observation unit, in which the relationships between the observed variables are modeled using latent variables. In those models, it is assumed that the multivariate time series are generated by a latent uni- or multivariate time series, where the latent time series are of lower order than the observed multivariate time series. One such model is the one underlying the "P-technique", proposed by Cattell (1952, 1963), where a conventional cross-sectional factor analysis is applied to the multivariate time series of one subject. Anderson (1963) objected to this method since only simultaneous relations between variables are taken into account: possible relations between factor series at different times are not modeled in a P-technique analysis. Anderson (1963) proposed an alternative procedure, which has been elaborated by several authors. The elaboration is known under the name "dynamic factor analysis" (Engle & Watson, 1981; Immink, 1986; Molenaar, 1985). The various dynamic factor models differ in the way the latent time series are related to observed time series, the model of the latent time series and the estimation procedure of the parameters of the model.

This research has been made possible by funding from the Netherlands Organization of Scientific Research (NWO) to the first author. The authors are obliged to Tom A.B. Snijders, Jos M.F. ten Berge and three anonymous reviewers for comments on an earlier version of this paper, and to Kim Shifren for providing us with her data set, which was collected at Syracuse University.

Requests for reprints should be sent to Marieke E. Timmerman, Heymans Institute of Psychology, DPMG, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, THE NETHERLANDS. E-Mail: m.e.timmerman@ppsw.rug.nl

0033-3123/2003-1/2000-0796-A $00.75/0

A different approach is the use of *component* models, which are directed at fitting the data themselves, rather than their covariances. Bijleveld (1989) and van Buuren (1990) offered dynamic component models. The linear dynamic system model (Bijleveld, 1989) is a model in which the autocorrelation between successive component scores is explicitly modeled. van Buuren (1990) offered a very general dynamic component model, which is denoted as the canonical class model.

The models mentioned above aim at modeling multivariate time series of one subject, thus allowing to study intraindividual variability. Researchers are often interested in studying interindividual differences and similarities in intraindividual variability. One of the methods used so far is to compare the various fitted time series models obtained for all subjects separately. This approach is used quite often in P-technique factor analysis, as indicated in Jones' and Nesselroade's review (1990) of studies where multivariate time series have been analyzed using P-technique, but also in case of dynamic factor models (e.g., see Shifren, Hooker, Wood & Nesselroade, 1997). Another approach is to model the multivariate time series of more than one subject *simultaneously*. However, the multisubject extensions of the models for multivariate time series of one subject proposed so far leave little or no room for interindividual differences. The dynamic factor model for time series collected from more than one subject (Nesselroade & Molenaar, 1999) assumes that there are no interindividual differences in intraindividual variability in the data. The extension of the linear dynamic system model (Bijleveld & Bijleveld, 1997) offers only limited possibilities to model interindividual differences. The extension of the canonical class model (van Buuren, 1990) appears particularly useful to model data of a number of observers on the same subject, rather than modeling data of a number of subjects, and hence is not appropriate for modeling interindividual differences.

In this paper, we propose a class of four simultaneous component analysis (SCA) models for the exploratory analysis of multivariate time series for more than one subject, which explicitly model interindividual differences. Two of the models are new. In each of the four models, the multivariate time series of each subject is decomposed into a few series of component scores and a loading matrix. The loading matrix is assumed common for all subjects. The four SCA models differ with respect to the constraints imposed on the cross-products (covariances) of the component scores. Apart from the cross-product constraints, the component scores themselves are freely estimated, and the approach deviates from the usual time series analysis. The models can be ordered hierarchically from weakly to severely constrained, thereby allowing for big to small interindividual differences in the model. Which model is most appropriate depends on the degree of variability between subjects. After descriptions of the models themselves, the SCA models and alternating least squares algorithms to fit the models to data will be described. Finally, the methods will be illustrated by means of an empirical example.

2. The Four Simultaneous Component Analysis (SCA) Models

In the next sections, preprocessing steps to be taken on the raw data before performing a simultaneous component analysis (SCA; section 2.1), and the four SCA models will be discussed (sections 2.2 through 2.6). The transformational freedom in the four SCA models is treated in section 2.7, and issues in model selection are elaborated in section 2.8.

In the sequel, bold capitals refer to matrices, and italic characters to scalars. The scores on the J variables ($j = 1, \ldots, J$) for subject $i (i = 1, \ldots, I)$ are collected at K_i occasions ($k_i = 1, \ldots, K_i$), and represented by data matrix $\mathbf{X}_i (K_i \times J)$.

2.1. Preprocessing of Raw Data Before Fitting the SCA Models to Data

Before fitting one of the four SCA models, one has to decide whether the raw data should be analyzed or some preprocessed version thereof. In most cases in practice, the data can be considered to have interval level. Then, it is advised to center the raw scores across occasions per variable and per subject. Using this approach, constants are eliminated from the data without introducing artificial variation (Bro & Smilde, 2003; Harshman & Lundy, 1984b). In addition, then, the components are defined on the basis of the intra-individual covariances instead of the interindividual covariances between variables. Besides, the average component scores (over occasions) per subject can be shown to be zero. The latter is useful because then the different restrictions on the component score matrices in each of the four models can directly be interpreted in terms of different restrictions on the covariances between components, as will become clear later.

Scaling aims at eliminating artificial scale differences. In the case of SCA, scaling to equalize the importance of the several variables to the final solution appears to be most reasonable. We advise to normalize the (centered) scores "within variables" (i.e., per variable over occasions and subjects jointly), such that the sum of squares per variable over occasions *and* subjects is equal to the sum of the number of measurements of all subjects ($\sum_{i=1}^{I} K_i$, where K_i denotes the number of measurements of subject i, $i = 1, \ldots, I$). As a consequence, differences in intraindividual variability are preserved. Furthermore, this type of scaling does not affect the form of the structural model (refer to Bro & Smilde, 2003; Harshman & Lundy, 1984b).

2.2. *SCA with Invariant Pattern (SCA-P)*

The model for SCA with invariant Pattern (SCA-P; Kiers & ten Berge, 1994) was originally proposed for modeling multivariate data of a number of subjects drawn from more than one population. SCA-P can be used for modeling multivariate time series of a number of subjects as follows: Let $\mathbf{X}_i (K_i \times J)$ denote the matrix of (usually preprocessed) scores of the i-th subject ($i = 1, \ldots, I$) on J variables measured at K_i occasions ($k_i = 1, \ldots, K_i$). The SCA-P model is given by

$$\mathbf{X}_i = \mathbf{F}_i \mathbf{B}' + \mathbf{E}_i, \tag{1}$$

where $\mathbf{F}_i (K_i \times Q)$ denotes the Q component scores of subject i at time points $1, \ldots, K_i$, $\mathbf{B}$ $(J \times Q)$ denotes the loading matrix, and $\mathbf{E}_i (K_i \times J)$ denotes the matrix of residuals. The component scores matrix $\mathbf{F}_i$, $i = 1, \ldots, I$, is unconstrained. Thus, it is assumed that the true variable scores (i.e., without error) at occasion k_i are a linear combination of the component scores at occasion k_i.

The fact that the component scores matrices are unconstrained implies that the inner products between the components may vary across subjects. If the component scores matrix contains centered scores, the inner products of the component scores of subject i divided by K_i are covariances between the components of the subject concerned (as is proved in the Appendix for the four SCA models). If the analyzed data matrices have been normalized "within variables" (see section 2.1), the differences in intraindividual variability are preserved. In the sequel, it will be assumed that the raw scores are preprocessed as discussed in section 2.1.

The time series of each subject are decomposed into a number of series of component scores and a loading matrix, which is common for all subjects, and to all occasions. Thus, the loading matrix is assumed to be subject and time invariant. The component score at a certain occasion can be interpreted as the degree of the particular property as indicated by the particular component. The interpretation of the components is, as usual, based on the loadings. To investigate each individual's behavior on the various components, the series of component scores for each component, and for each subject, can be plotted against the time axis. Possible trends and deviating scores at certain occasions can be seen at once.

Differences between subjects in variance of a certain component can be interpreted as differences in intraindividual variability in terms of that particular component. Differences in covariances between components (within subjects) are easiest to interpret in terms of differences in correlations. Thus it is possible that certain components correlate highly for one subject, and almost zero for another.

2.3. *Constrained Versions of SCA-P*

In the SCA-P model the component scores matrices are unconstrained, implying that the variances of component scores and the covariances between component scores within subjects may vary across subjects. If no interindividual differences in covariation and/or variability are present in the data, it is warranted to use a more parsimonious model than SCA-P. By imposing proper constraints on the variances and covariances of the individual component scores, three models are defined that are restricted variants of SCA-P. The restrictions on the component scores of the four models are summarized in Table 1. The three restricted variants of SCA-P are described in sections 2.4 to 2.6.

TABLE 1.
The restrictions of the four SCA models on the covariances between and the variances of the component scores

	Covariances	Variances
SCA-P	free	free
SCA-PF2	equal across subjects	free
SCA-IND	equal to 0	free
SCA-ECP	equal across subjects	equal across subjects

Depending on the degree of interindividual differences, a weakly or strongly restricted SCA model can be chosen. The strength of the approach is that one can explicitly choose the most parsimonious model possible for the particular data set without ignoring important aspects of the data. The choice for the most parsimonious model is not only important in terms of the interpretation of the model. Fitting a less parsimonious model than the one that is underlying the data usually leads to so-called error fitting, that is, a part of the error term is mistakenly fitted in the model. This usually leads to unstable parameter estimates.

2.4. *SCA with PARAFAC2 Constraints (SCA-PF2)*

The model for SCA with PARAFAC2 constraints (SCA-PF2; Kiers, ten Berge & Bro, 1999) is a constrained version of the SCA-P model. Kiers, ten Berge and Bro (1999) named this model "direct fitting PARAFAC2", but we choose to use the name SCA-PF2 to stay in line with the other SCA models. The SCA-PF2 model is given by (1) with $\mathbf{F}_i'\mathbf{F}_i$ constrained to $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{D}_i\mathbf{\Phi}\mathbf{D}_i$, with $\mathbf{D}_i$ a $(Q \times Q)$ diagonal matrix and $\mathbf{\Phi}$ a $(Q \times Q)$ positive definite matrix, and without loss of generality, we further require $\mathbf{\Phi}$ to have unit diagonal elements. Thus, in SCA-PF2, the congruence coefficients (Tucker, 1951) between columns of $\mathbf{F}_i$ are invariant over all subjects $i = 1, \ldots, I$. The coefficient of congruence between two columns $\mathbf{x}$ and $\mathbf{y}$, φ_{xy}, is defined as the normalized inner product between the columns $\mathbf{x}$ and $\mathbf{y}$, namely as

$$\varphi_{\mathbf{xy}} = \frac{\mathbf{x}'\mathbf{y}}{\sqrt{\mathbf{x}'\mathbf{x}}\sqrt{\mathbf{y}'\mathbf{y}}}. \tag{2}$$

If the component scores matrices contain centered scores, the restriction on $\mathbf{F}_i, i = 1, \ldots, I$, implies that the components have the same mutual correlations for all subjects, and that the variances of the components may vary across subjects. These variances are given by the diagonal elements of $\mathbf{D}_i^2$, $i = 1, \ldots, I$. Thus, the SCA-PF2 model is suitable if the variables indicate concepts that are equally correlated for different subjects, and if the degree of intra-individual variability of the properties indicated by the concepts varies between subjects.

2.5. *SCA with INDSCAL Constraints (SCA-IND)*

The model for SCA with INDSCAL constraints (SCA-IND) is a constrained version of both SCA-P and SCA-PF2. That is, the SCA-IND model is given by (1) with $\mathbf{F}_i'\mathbf{F}_i$ constrained to $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{D}_i^2$, with $\mathbf{D}_i$ a diagonal matrix. Thus, the inner products between the components are zero and the sums of squares of the components may differ across subjects in the SCA-IND model. To the best of our knowledge, the SCA-IND model is new. If the component scores matrices contain centered scores, the restriction on $\mathbf{F}_i$, $i = 1, \ldots, I$, implies that the components are constrained to be uncorrelated, but the variances of the several components may vary across subjects. Again, the elements of $\mathbf{D}_i^2$ contain the variances of the components of subject i.

The name "SCA-IND" is obtained from a counterpart model for cross-product matrices. That is, taking the cross-products of the (error free parts of) the left and right hand side of (1), and imposing $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{D}_i^2$, $i = 1, \ldots, I$, we get

$$\frac{1}{K_i}\mathbf{X}_i'\mathbf{X}_i = \mathbf{B}\mathbf{D}_i^2\mathbf{B}' \tag{3}$$

which equals the scalar products version of the INDSCAL (INdividual Differences SCALing) model (Carroll & Chang, 1970) applied to the matrices $\frac{1}{K_i}\mathbf{X}_i'\mathbf{X}_i$, $i = 1, \ldots, I$.

The SCA-IND model should be used if interindividual differences in intraindividual variability for the separate components exist, but the separate components within subjects are uncorrelated for all subjects.

2.6. *SCA with Equal Average Cross-Products Constraints (SCA-ECP)*

The model for SCA with Equal average Cross-Products constraints (SCA-ECP) is a constrained version of the SCA-P, SCA-PF2, and SCA-IND models. The name of this new method expresses the constraints on the component scores. That is, the SCA-ECP model is given by (1) with $\mathbf{F}_i'\mathbf{F}_i$ constrained such that $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{\Phi}$, $i = 1, \ldots, I$, where $\mathbf{\Phi}$ is a $(Q \times Q)$ positive definite matrix. The restriction on the component scores implies that the average cross-products between the components are equal for all subjects. If the component scores matrix contains centered scores, the restriction implies that the correlations between the components as well as the variances of the components are equal for all subjects. The estimates of the data matrices $\mathbf{X}_i$, $i = 1, \ldots, I$, are insensitive to an orthogonal or oblique transformation of the component score matrices $\mathbf{F}_i$, provided that such a transformation is compensated in the loading matrix $\mathbf{B}$. Therefore, the sum of squares explained by the model does not alter by requiring $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{I}$ (instead of $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{\Phi}$), and thus it is clear that SCA-ECP is a constrained version of SCA-IND, and hence of the others as well.

In the SCA-ECP model, the variances of components and covariances between components within subjects are equal for all subjects. Thus, the model is suitable if the variables indicate concepts that are equally correlated over time for the subjects, and no differences in intra-individual variability can be found.

2.7. *Transformational Freedom in the SCA-P, SCA-PF2, SCA-IND and SCA-ECP Models*

As will be discussed further in section 3, the SCA models described above can all be fitted to a particular data set by least squares minimization of the residuals. In the sequel it is assumed that parameter estimates of $\mathbf{F}_i$ for at least one of the subjects have full rank, and the parameter estimates of $\mathbf{B}$ have full rank. The estimates of the data matrices $\mathbf{X}_i$, $i = 1, \ldots, I$, by the parameter matrices of the SCA-P, and SCA-ECP models are insensitive to orthogonal and oblique transformations of the loading matrix $\mathbf{B}$, provided that such a transformation is compensated in the component scores matrices $\mathbf{F}_i$, $i = 1, \ldots, I$. Standard rotational procedures (e.g., Varimax; Kaiser, 1958) can be used to obtain solutions which are easier to interpret.

Kiers, ten Berge and Bro (1999) have shown that, under some assumptions, SCA-PF2 solutions are "essentially unique", which means that SCA-PF2 estimates are unique up to trivial permutation, reflection and/or rescaling. In the proof, it is required, among other things, that the number of subjects relative to the number of components is rather large ($I \geq Q(Q+1)(Q+2)(Q+3)/24$, where I denotes the number of subjects and Q the number of components). However, Kiers, ten Berge and Bro (1999) report on the basis of simulations, that the uniqueness properties of PARAFAC2 appear to hold generally for $I \geq 4$.

Assuming that there is at least one pair $\mathbf{D}_i, \mathbf{D}_{i'} (i \neq i')$, such that $\mathbf{D}_i \mathbf{D}_{i'}^{-1}$ has no equal pair of diagonal elements, it can be proven that estimates of SCA-IND are essentially unique. The proof is largely based on the uniqueness proof for PARAFAC, as given by Harshman (1972).

The assumptions for essentially uniqueness of SCA-PF2 and SCA-IND estimates are not directly comparable. In practice, however, the uniqueness assumptions for SCA-PF2 are stronger than for SCA-IND. Hence, one could encounter data for which no unique SCA-PF2 estimate exists, whereas there is one for SCA-IND. In practice, the reverse is unlikely to occur.

2.8. *Model Selection*

Generally, fitting a model to observed data aims at obtaining an interpretable model with a small degree of overall error, which refers to the lack of fit of the current model as it is fitted to the current data set, to the population data (Browne & Cudeck, 1992). In an SCA, one aims at optimally decomposing observed data into a systematic part, described by an interpretable model that would fit the population data well, and a residual part. As one does not know which part of the data is systematic, the choice of a particular SCA-model and number of components is a fairly complicated matter. Besides, the call for an interpretable model introduces a certain subjective aspect. However, one can combine different criteria to come to a model decision, which hopefully point in the same direction.

One can perceive whether a model fits the observed data well, although one would be more interested in whether a model fits the systematic part of the data. However, the degree to which a model fits the observed data can be used for comparison of different models. One could try to find the model that covers the most important or most salient aspects of the data, thereby ignoring aspects of little importance (e.g., because they pertain to a small number of subjects, variables, or occasions). Thus, it is desirable to choose the most constrained model with a relatively small number of components that still fits the data well.

Possible ways for investigating the degree of overall error of a particular model are cross-validation and split-half analysis. In practice, we apply these methods to each of the four SCA-models for a reasonable range of numbers of components to select the model with a relatively low degree of overall error.

Cross-validation assesses the predictive validity of the estimated model parameters. A well known approach is K-fold cross-validation (Hastie, Tibshirani & Friedman, 2001, pp. 214–217), which can be summarized as follows: The observed data set is split into K roughly equal-sized parts. The idea of cross-validation now is to predict the scores in the k-th part, denoted as the test set, by using a model based on the other $K-1$ parts of the data, denoted as the training set. Then, the prediction error of the model when predicting the k-th test set is computed. The procedure is repeated for all K test sets. Finally, the K estimates of prediction error are combined to assess the predictive validity of the model at hand.

K-fold cross-validation can be used to assess the predictive power of a particular SCA model of multisubject multivariate time series. The two main decisions in applying this procedure pertain to how to split the data set, and how to predict the scores in the k-th test set by the model estimates based on the associated training set. In the current SCA context, various approaches for defining the k-th test set of the data are reasonable. For example, one could randomly select elements from the full data set, as we will do, or randomly select the scores on a particular measurement occasion (hence, rowwise selection from the data matrices). Subsequently, one has to choose the method for the prediction of the k-th test set by the model estimates based on the training set. This may follow directly from the estimation procedure, as is the case in our approach, but will be less apparent in other cases.

We propose to use a kind of K-fold cross-validation, which is based on the Expectation-Maximization cross-validation (EM-CV) method (Louwerse, Smilde, & Kiers, 1999). The Expectation Maximization cross-validation procedure for SCA (EM-SCA) can be summarized as follows. The full observed data set (consisting of the data of all I subjects) is split into K (about equal-sized) parts, by random assignment of the elements to each of those K parts. Then, each training set is constructed from the full data set by leaving out one such random part; the left-out part serves as the test set. In total, K different training sets are constructed by leaving out each of the different parts in turn. The left-out parts in the ensuing training sets are treated as missing data. The training sets are next preprocessed in the usual way, while disregarding the missing values (see sec. 2.1). A sensible starting value is imputed for the missing value in the preprocessed data array, and this data array is analyzed by the SCA-model at hand. The SCA estimation procedure is iterative, as will be discussed in section 3, and after each cycle the "missing element" is estimated on the basis of the current model parameters, and it is imputed. The cycles are repeated until convergence, and the estimated value of the missing element is retained. The procedure is repeated until all elements have been eliminated once. Finally, the predictive residual error sum of squares (PRESS) is calculated as the sum of squares between estimated values and observed values. A high PRESS value is indicative of a model with low predictive value, either because the model underfits or overfits the data. Underfitting may be caused by estimating too few components and/or a too much restricted model, whereas the reverse leads to an overfitted model. Thus, only models with small PRESS values have a high predictive value. The size of the PRESS value is partly dependent on the degree of noise in the data, and therefore it is not possible to give a generally valid bench-mark value for PRESS. We advise to consider only models that have relatively low PRESS values, compared to the PRESS values of the other models considered. A disadvantage of the EM-SCA cross-validation approach is that a large number of analyses is needed, depending on K, the number of sets defined.

Split-half analysis assesses the stability of a model. Split-half analysis is applied, for instance, in the PARAFAC model (Harshman & Lundy, 1984a), and is here used for the SCA models. The observed data are randomly split in two parts over the mode that can be viewed as replications, thus the subject or the occasion mode. Then, the SCA model at hand is fitted to each of the (preprocessed) data halves. We propose to compare the two estimated loading matrices, after rotation to simple structure in case of SCA-P and SCA-ECP (e.g., Varimax rotation; Kaiser, 1958). If the solution is stable, the respective columns of the two loading matrices should be (approximately) equal up to permutation and/or reflection. Hence, permutation and reflection should be taken into account in a stability measure. As a stability measure, we propose to use the average of the mean absolute difference of the columns $\mathbf{b}_{q,1}$ and $\mathbf{b}_{q,2}$, $q = 1, \ldots, Q$, where $\mathbf{b}_{q,1}$ and $\mathbf{b}_{q,2}$ denote the q-th column of the loading matrices $\mathbf{B}_1$ and $\mathbf{B}_2$, and where the columns of $\mathbf{B}_1$ are ordered and reflected so as to yield the lowest average mean of the absolute differences. This stability measure will be denoted as "split-half stability coefficient". We advocate to replicate the procedure a number of times (e.g., 50 times) and to use the average thus obtained split-half stability coefficients, in order to increase the reliability of the stability study. Just as for the PRESS value, no generally valid benchmark can be given for the SHS value, and one should consider only models that have a relatively low stability coefficient.

A third topic to consider is the interpretability of a model. A model should make sense, substantively, because our main aim is to give a useful description of the data. In the SCA-models, one should consider the interpretability of the loading matrix, the (lack of) interindividual differences in covariances between components and variances of components, and the development of the various component scores series in the course of time. The degree of interpretability can be a subject for debate, because the interpretation is partly dependent on the theoretical ideas and knowledge a researcher has.

Ideally, the three criteria point in the same direction. It is possible that a few competing models remain. Then, it is up to the researcher to choose the final model. It might also happen that all criteria indicate different models, which might be indicative of unsuitability of the SCA models to describe the present data.

3. Fitting the Four SCA Models to Data

To fit each of the four models for SCA to observed data, we propose to minimize the sum of squared residuals. Hence, we minimize

$$F(\mathbf{F}_i, \mathbf{B}) = \sum_{i=1}^{I} \left\| \mathbf{X}_i - \mathbf{F}_i\mathbf{B}' \right\|^2, \tag{4}$$

subject to the constraint imposed in the particular SCA. Thus, the total sum of squares that is explained by the model is maximized. To identify the solution partly, we require for the four SCA models that

$$\frac{1}{\sum_{i=1}^{I} K_i} \sum_{i=1}^{I} \operatorname{diag}(\mathbf{F}_i'\mathbf{F}_i) = \mathbf{I},$$

where K_i denotes the number of measurements of subject i, $i = 1, \ldots, I$, and $\operatorname{diag}(X)$ the diagonal of matrix $\mathbf{X}$. This identification constraint can always be invoked after a solution has been obtained, namely by a simple scaling transformation of $\mathbf{B}$ and the $\mathbf{F}_i$'s. If the component scores matrices $\mathbf{F}_i$, $i = 1, \ldots, I$, are centered, this identification constraint implies that the variance per component over all subjects is one. This constraint is satisfied automatically in the SCA-ECP model (since the component scores are restricted to $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{\Phi}$, which is equivalent to requiring that $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{I}$).

The degree to which the estimated model describes the data is expressed by the proportion of sum of squares explained by the model, which we call the "fit" in the sequel. The fit is defined as

$$1 - \sum_{i=1}^{I} \left\| \mathbf{X}_i - \mathbf{F}_i\mathbf{B}' \right\|^2 \Big/ \sum_{i=1}^{I} \|\mathbf{X}_i\|^2 . \tag{5}$$

The fit is often expressed as a percentage, upon multiplication by 100.

The SCA-P algorithm aims at minimizing (4) over arbitrary $\mathbf{F}_i$, $i = 1, \ldots, I$, and $\mathbf{B}$. Kiers and ten Berge (1994) give an explicit solution to this problem.

For SCA-PF2, SCA-IND and SCA-ECP we use alternating least squares (ALS) algorithms: the matrices over which the function has to be minimized are alternatingly updated until convergence. The algorithm is said to have converged if from one cycle (i.e., update of all parameters) to another, the residual sum of squares decreases less than a prespecified value. The fitting of the SCA-PF2, SCA-IND, and SCA-ECP models to data will be treated successively.

3.1. Fitting the SCA-PF2 Model

The SCA-PF2 algorithm aims at minimizing (4), subject to

$$\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{D}_i\mathbf{\Phi}\mathbf{D}_i, \qquad i = 1, \ldots, I,$$

where $\mathbf{D}_i$ is a $(Q \times Q)$ diagonal matrix and $\mathbf{\Phi}$ a $(Q \times Q)$ positive definite matrix with unit diagonal elements. Kiers, ten Berge and Bro (1999) proposed an ALS algorithm for the equivalent problem of minimizing (4) subject to $\mathbf{F}_i'\mathbf{F}_i = \tilde{\mathbf{D}}_i\tilde{\mathbf{\Phi}}\tilde{\mathbf{D}}_i$ with $\tilde{\mathbf{D}}_i$ a $(Q \times Q)$ diagonal matrix and $\tilde{\mathbf{\Phi}}$ an arbitrary ($Q \times Q$) positive definite matrix. Their algorithm is essentially based on the fact that every matrix $\mathbf{F}_i$ that meets the constraint $\mathbf{F}_i'\mathbf{F}_i = \tilde{\mathbf{D}}_i\tilde{\mathbf{\Phi}}\tilde{\mathbf{D}}_i$ can be written as $\mathbf{F}_i = \mathbf{P}_i\tilde{\mathbf{F}}\tilde{\mathbf{D}}_i$ provided that $\mathbf{P}_i'\mathbf{P}_i = \mathbf{I}_Q$, $\tilde{\mathbf{F}}$ an arbitrary ($Q \times Q$) matrix, and $\tilde{\mathbf{D}}_i$ a diagonal ($Q \times Q$) matrix, $i = 1, \ldots, I$. The SCA-PF2 algorithm as proposed by Kiers, ten Berge and Bro (1999) boils down to minimizing

$$f_1(\mathbf{P}_i, \tilde{\mathbf{F}}, \tilde{\mathbf{D}}_i, \mathbf{B}) = \sum_{i=1}^{I} \left\| \mathbf{X}_i - \mathbf{P}_i\tilde{\mathbf{F}}\tilde{\mathbf{D}}_i\mathbf{B}' \right\|^2 \tag{6}$$

subject to $\mathbf{P}_i'\mathbf{P}_i = \mathbf{I}_Q$, and $\tilde{\mathbf{D}}_i$ a diagonal matrix, $i = 1, \ldots, I$. The function in (6) is minimized by updating $\mathbf{B}$, $\mathbf{P}_i$, $\tilde{\mathbf{F}}$, and $\tilde{\mathbf{D}}_i$ alternatingly. We propose to use this algorithm to find solutions for $\mathbf{B}$, $\mathbf{F}_i$, $\tilde{\mathbf{\Phi}}$, and $\tilde{\mathbf{D}}_i$. Solutions for $\mathbf{D}_i$ and $\mathbf{\Phi}$ such that $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{D}_i\mathbf{\Phi}\mathbf{D}_i$ can then be obtained by taking

$$\mathbf{\Phi} = \operatorname{diag}(\tilde{\mathbf{\Phi}})^{-\frac{1}{2}}\tilde{\mathbf{\Phi}}\operatorname{diag}(\tilde{\mathbf{\Phi}})^{-\frac{1}{2}}, \quad \text{and} \quad \mathbf{D}_i = \frac{1}{\sqrt{K_i}}\tilde{\mathbf{D}}_i\operatorname{diag}(\tilde{\mathbf{\Phi}})^{\frac{1}{2}}.$$

3.2. Fitting the SCA-IND Model

The SCA-IND algorithm aims at minimizing (4) subject to $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{D}_i^2$, $i = 1, \ldots, I$, with $\mathbf{D}_i$ a ($Q \times Q$) diagonal matrix. The ALS algorithm to find estimates of the parameters of the SCA-PF2 model (Kiers, ten Berge and Bro, 1999) can be used to find estimates of $\mathbf{F}_i$, $i = 1, \ldots, I$ subject to $\mathbf{F}_i'\mathbf{F}_i = \tilde{\mathbf{D}}_i^2$ by keeping $\tilde{\mathbf{\Phi}} = \mathbf{I}$. In the SCA-PF2 algorithm this is obtained by keeping $\tilde{\mathbf{F}}$ fixed as $\tilde{\mathbf{F}} = \mathbf{I}$, and only updating $\mathbf{B}$, $\mathbf{P}_i$, and $\tilde{\mathbf{D}}_i$. With this algorithm we find solutions for $\mathbf{B}$, $\mathbf{F}_i$ and $\tilde{\mathbf{D}}_i$. Solutions for $\mathbf{D}_i$ such that $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{D}_i^2$ can be obtained by taking $\mathbf{D}_i = \frac{1}{\sqrt{K_i}}\tilde{\mathbf{D}}_i$.

3.3. Fitting the SCA-ECP Model

The SCA-ECP algorithm aims at minimizing (4) subject to $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{\Phi}$, $i = 1, \ldots, I$, which is equivalent to (i.e., without affecting the model fit) imposing that $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{I}_Q$. Updating $\mathbf{B}$, and $\mathbf{F}_i$, $i = 1, \ldots, I$, alternatingly can solve this problem. The problem to find an update for $\mathbf{B}$ is analogous to finding an update of $\mathbf{B}$ in the SCA-PF2 algorithm (Kiers, ten Berge & Bro, 1999). The next problem is to find, for every value of i, an update for $\mathbf{F}_i$, subject to $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{I}_Q$. An update for $\mathbf{F}_i$ can be obtained by minimizing

$$f_2(\mathbf{F}_i) = \left\| \mathbf{X}_i - \mathbf{F}_i\mathbf{B}' \right\|^2 \tag{7}$$

subject to $\frac{1}{K_i}\mathbf{F}_i'\mathbf{F}_i = \mathbf{I}_Q$. Upon substitution of $\tilde{\mathbf{F}}_i = \frac{1}{\sqrt{K_i}}\mathbf{F}_i$, this is equivalent to maximizing $\operatorname{tr}(\tilde{\mathbf{F}}_i'\mathbf{X}_i\mathbf{B})$, subject to $\tilde{\mathbf{F}}_i'\tilde{\mathbf{F}}_i = \mathbf{I}$. Consider the singular value decomposition $\mathbf{X}_i\mathbf{B} = \mathbf{U}_i\mathbf{\Delta}_i\mathbf{Q}_i'$, with

$$\mathbf{U}_i'\mathbf{U}_i = \mathbf{Q}_i'\mathbf{Q}_i = \mathbf{Q}_i\mathbf{Q}_i' = \mathbf{I}_Q,$$

and $\mathbf{\Delta}_i$ a diagonal matrix with nonnegative diagonal elements in weakly descending order. Then the maximum of $\operatorname{tr}(\tilde{\mathbf{F}}_i'\mathbf{X}_i\mathbf{B})$ is given by $\tilde{\mathbf{F}}_i = \mathbf{U}_i\mathbf{Q}_i'$ (Cliff, 1966), hence an update of $\mathbf{F}_i$ is given by $\mathbf{F}_i = \sqrt{K_i}\mathbf{U}_i\mathbf{Q}_i'$.

3.4. Starting Values of the Parameters

Each of the iterative algorithms (i.e., for SCA-PF2, SCA-IND and SCA-ECP) has to be initialized with certain starting values. The starting matrices can be drawn randomly from, for example, a normal distribution. As a so-called rational start, the (explicit) SCA-P solution for the matrix **B** can be used in all algorithms. In case of SCA-IND and SCA-PF2, the starting matrices for the diagonal matrices $\mathbf{D}_i$, $i = 1, \ldots, I$, are set at identity matrices ($Q \times Q$). In case of SCA-PF2, the identity matrix ($Q \times Q$) is also used as a starting matrix for $\mathbf{\Phi}$. These starting values suffice to start the iterative process (by updating $\mathbf{P}_i$ or $\mathbf{F}_i$). In practice, it is recommended to use several differently started runs in order to decrease the chance of missing the global minimum of the function.

4. An Empirical Example of Simultaneous Component Analyses: Mood in Individuals with Parkinson's Disease

In a study by Shifren, Hooker, Wood and Nesselroade (1997), mood structure was examined among 12 individuals diagnosed with Parkinson's disease. Positive and negative affect was measured with the Positive and Negative Affect Schedule (PANAS; Watson, Clark & Tellegen, 1988). This measure contains 10 positive and 10 negative affect items. Subjects were asked to rate the 20 adjectives on a 5 point scale (ranging from 1 (not at all) to 5 (all the time)) to indicate to what degree the subject experienced the particular affect on that day. Subjects scored the PANAS daily on successive days, ranging from 53 to 71 days. Over all 12 subjects, the scores on the 20 variables were obtained on 817 days. One was interested in the intraindividual structure of mood, as well as the interindividual differences. Watson (1988) showed that interindividual differences in mood (of healthy subjects) can be well described by two relatively independent dimensions, namely positive and negative affect. It is questioned whether this structure can also be used in describing intraindividual differences in mood of subjects suffering from Parkinson's disease, and whether the degree of intraindividual variability differed across subjects.

The scores on the items at successive days were analyzed by Shifren et al. (1997) using a dynamic factor analysis. A dynamic factor model was estimated for each subject separately, where, based on issues regarding content, the maximal number of factors was two, and the maximal "lag" was one. Items showing responses that were too stable over time (over 90% of the responses in the same category) were eliminated from the analyses. Also, any linear trend over time per variable per subject was removed from the data. Further information concerning the research, method of analyses and results can be found in Shifren et al. (1997).

Before performing the simultaneous component analyses, the data of each subject per variable were centered over the time points, and normalized within variables (i.e., over occasions and subjects jointly), such that the sum of squares per variable over occasions *and* subjects was equal to the sum of the number of occasions of all subjects ($\sum_{i=1}^{I} K_i$, $i = 1, \ldots, I$) (see section 2.1). In contrast to Shifren et al. (1997), we did not remove trends from the data, nor did we exclude variables from the analyses to keep as much of the information in the data as possible.

First we assessed the stability and the fit of the four models (see section 2.8) with one through five components, thus of 20 models in total. For each of the 20 models, the PRESS value was computed following the procedure for EM-SCA cross-validation, as discussed in section 2.8. We split the data set into 150 about equal-sized parts, each consisting of a random selection (without replacement) of the 16340 elements of the datamatrix (= 817 (total number of measurement occasions over all subjects) $\times$ (number of variables)). Out of the 150 sets, 149 sets consisted of 109 elements, and one set of the remaining 99 elements. As starting value, a zero was imputed for each removed observation, which is the average score per subject per variable (as a result of the preprocessing procedure). PRESS values were thus obtained for each of the 20 models.

The split-half procedure was applied following the guidelines as discussed in section 2.8. We repeated the split-half procedure 50 times, resulting in an average split-half stability coefficient over 50 replications. The SCA-algorithms used for computing the PRESS values and the SHS were started by the rational start (see section 3.4).

TABLE 2.
PRESS values, split-halve stability coefficients (SHS) and fit (%) of the four SCA-models with one through five components

Model	SCA-ECP			SCA-IND			SCA-PF2			SCA-P		
Measure	PRESS ($\times 10^4$)	SHS	Fit	PRESS ($\times 10^4$)	SHS	Fit	PRESS ($\times 10^4$)	SHS	Fit	PRESS ($\times 10^4$)	SHS	Fit
$Q = 1$*	1.39	0.06	24.0	1.24	0.07	30.8	1.24	0.07	30.8	1.24	0.07	30.8
$Q = 2$	1.49	0.08	31.8	1.16	0.09	42.8	1.16	0.58	42.8	1.14	0.09	43.5
$Q = 3$	1.76	0.09	37.1	1.18	0.13	49.5	1.18	1.03	49.8	1.12	0.20	51.2
$Q = 4$	2.02	0.12	41.0	1.27	0.14	55.3	1.20	2.12	56.3	1.13	0.10	58.4
$Q = 5$	2.72	0.13	44.0	1.51	0.14	60.1	2.05	2.00	61.1	1.25	0.12	63.2

*Note that for $Q = 1$, the models SCA-IND, SCA-PF2 and SCA-P are equivalent.

The average split-half stability coefficients (SHS), the PRESS values and the fit percentages (see (5)) of the SCA-P, SCA-PF2, SCA-IND, and SCA-ECP models with one through five components are reported in Table 2.

Using the SCA-IND model instead of the more constrained SCA-ECP model increases the fit considerably, for all models with one through five components, as can be seen in Table 2. Almost no fit is gained in using the even less constrained SCA-PF2 or SCA-P model instead of the SCA-IND model. Thus, on the basis of the fit percentages the use of SCA-IND is indicated, but the preferred number of components is not clearly indicated.

Relatively low PRESS values and relatively low split-half stability coefficients (SHS) are indicative of stable models. On the basis of a comparison of the present PRESS and SHS values, we deem models sufficiently stable if they have maximal values of 0.10 for SHS and 1.20×10^4 for PRESS. Thus, SCA-IND and SCA-P with 2 components, and SCA-P with 4 components will be considered. We start the discussion of the models by the most restricted model of the three, SCA-IND with 2 components. Note that using either SHS or PRESS to investigate the stability would lead to a selection of different models.

The loading matrix **B** of the SCA-IND solution is presented in the left hand side of Table 3. The components can be interpreted as "Introversion", and "Emotional Instability".

The size of the component scores can be compared *between* subjects: extremely high (and low) component scores indicate a large degree of variability in scores over time compared to the other subjects. The degree of variability across persons can easily be compared on the basis of the variances of the component scores, that equal the diagonal elements of $\mathbf{D}_i^2$ here (see section 2.5), and these values are presented in the left hand panel of Table 4. As can be seen in Table 4, for example, Subject 5 shows most, and Subject 10 shows least variability on the Introversion component.

A comparison of the size of the component scores *within* subjects reveals the degree of "Introversion" and "Emotional Instability" of that person compared to the degree at the other days. To give some insight into variation of the component scores over time, the Introversion scores of subjects 5 and 10, and the Emotional Instability scores of Subjects 2 and 4 are plotted in Figure 1.

Figure 1 illustrates the differences between subjects in intraindividual variability, and it also offers the possibility to look for trends in component scores. For example, Subject 5 shows a remarkable shift towards extraversion at day 15 (and, as has been verified, such a change was not found in the Emotional Instability component scores of Subject 5). Unfortunately, additional information about the subjects is lacking to explain these changes in component scores.

The SCA-P model with 2 components explains 43.5% of the variance, which is only 0.7% more than the SCA-IND model with 2 components. To be able to compare the loading matrices of SCA-IND and SCA-P, we orthonormally rotated the loading matrix of SCA-P to the loading

TABLE 3.
Loading matrix of the SCA-IND model with two components (columns 2 and 3), and Varimax rotated loading matrix of the SCA-P model with 4 components (columns 4 through 7) of the mood data. Loadings > 0.40 or < −0.40 are printed bold

	SCA-IND		SCA-P			
	Introversion	Emotional Instability	Arousal	Nervousness	Emotional Instability	Component IV
jittery	**0.59**	**0.43**	**0.45**	**0.50**	0.37	−0.15
distressed	−0.02	**0.65**	0.02	0.17	**0.66**	0.27
upset	−0.01	**0.74**	0.09	0.11	**0.75**	0.15
afraid	**0.61**	0.02	0.19	**0.75**	−0.00	0.11
scared	**0.63**	−0.04	0.18	**0.78**	−0.06	0.08
hostile	**0.44**	**0.52**	**0.41**	0.35	**0.46**	−0.13
irritable	−0.07	**0.76**	0.09	−0.01	**0.78**	0.09
guilty	−0.10	**0.66**	0.12	−0.18	**0.75**	−0.22
ashamed	**0.45**	0.22	0.26	**0.40**	0.29	**−0.40**
nervous	**0.55**	**0.40**	0.37	**0.53**	0.37	−0.12
inspired	**−0.71**	0.03	**−0.59**	**−0.41**	0.21	−0.12
excited	0.01	−0.24	−0.14	0.02	−0.07	**−0.69**
determined	−0.25	**−0.45**	**−0.58**	0.25	−0.29	−0.01
interested	**−0.75**	−0.24	**−0.78**	−0.27	−0.04	0.05
enthusiastic	**−0.68**	−0.29	**−0.72**	−0.28	−0.06	−0.08
attentive	**−0.69**	−0.25	**−0.77**	−0.19	−0.04	−0.07
proud	0.06	−0.23	0.04	−0.14	−0.12	**−0.68**
strong	−0.04	**−0.59**	**−0.45**	0.30	−0.37	−0.33
active	**−0.58**	**−0.46**	**−0.73**	−0.10	−0.29	0.08
alert	**−0.69**	−0.32	**−0.75**	−0.25	−0.11	−0.03

TABLE 4.
Variances ($\mathbf{D}_i^2$) of the component scores per subject of the mood data. *In* denotes Introversion, *EI* denotes Emotional Instability

	SCA-IND		SCA-P				SCA-IND compared to SCA-P	
	variances ($\mathbf{D}_i^2$)		variance		covariance	correlation	correlation	
Subject	*In*	*EI*	*In*	*EI*	*In/EI*	*In/EI*	*In/In*	*EI/EI*
1	0.73	1.07	0.61	0.87	0.39	0.54	0.95	0.96
2	0.18	5.30	0.21	5.71	−0.43	−0.39	0.92	1.00
3	0.38	0.53	0.38	0.55	−0.06	−0.13	0.99	1.00
4	2.04	1.20	2.69	1.64	−1.01	−0.48	0.99	0.93
5	6.42	0.97	5.99	0.89	0.42	0.18	1.00	0.99
6	0.31	1.27	0.31	0.97	0.31	0.56	0.90	0.99
7	0.63	0.30	0.53	0.23	0.11	0.31	0.99	0.96
8	0.16	0.38	0.16	0.33	0.04	0.19	0.99	1.00
9	0.65	1.50	0.61	1.45	0.12	0.13	0.99	1.00
10	0.08	0.17	0.10	0.19	−0.05	−0.35	0.98	0.99
11	0.20	0.06	0.19	0.05	0.02	0.24	0.99	0.97
12	0.17	0.28	0.17	0.26	0.01	0.06	1.00	1.00

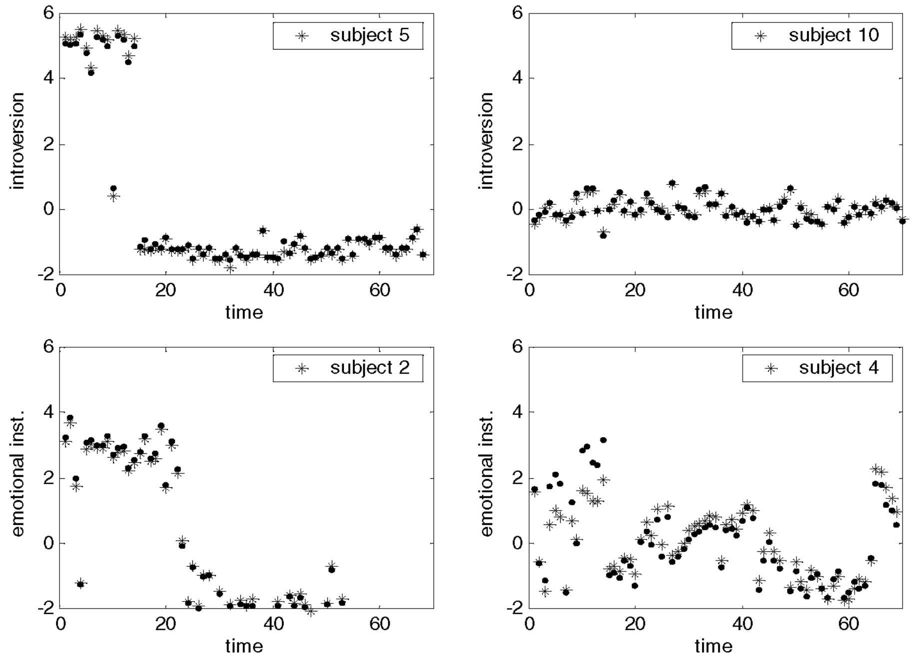

FIGURE 1.
Introversion component scores of Subjects 5 and 10, and Emotional Instability component scores of Subjects 2 and 4. SCA-IND estimates are denoted by "*", SCA-P estimates by "•".

matrix of SCA-IND, and this rotation is compensated in the component score matrices of all subjects. The thus obtained loading matrix resembles the loading matrix of the SCA-IND solution heavily, as is indicated by a maximal absolute difference of the loadings of 0.08, and an average absolute difference of 0.02. (Incidentally, even the normalized Varimax rotated loading matrix of SCA-P resembles the SCA-IND loading matrix considerably, with maximal and mean absolute difference of loadings 0.11, and 0.05, respectively). Therefore, the components for the transformed SCA-P solution are similarly interpreted as for the SCA-IND solution, namely as "Introversion", and "Emotional Instability".

The component scores of the transformed SCA-P solution can be compared to the ones of the SCA-IND solution in different ways. We computed per subject the average absolute difference between the component scores (both Introversion and Emotional Instability) of the SCA-P solution and the SCA-IND solution. On average, the SCA-P component scores deviate more than 0.10 from the ones of the SCA-IND solution only for Subjects 1, 2, 4, 6, and 7. Not surprisingly, those are the subjects with relatively high covariances between the two components in the transformed SCA-P solution, as reflected in correlations larger than 0.30 (see Table 4). The SCA-IND and SCA-P component scores can also be compared by investigating whether the score profiles (over time) of the two solutions are about equal. The correlation coefficients between the component scores of SCA-P and SCA-IND per component per subject are reported in Table 4. An impression of the differences in component scores between the two methods can be obtained from the plots in Figure 1. In Table 4, it can be seen that the correlation coefficients are rather high, thus the score profiles of Introversion and Emotional Instability as estimated by SCA-IND and SCA-P are about equal for all subjects. Even the estimated Emotional Instability scores of Subject 4, which show the lowest correlation among all subjects (0.93), do not differ so much that the interpretation of the development over time would be different. Thus, in comparing the SCA-P to the SCA-IND solution, the most striking is the presence of, to a certain extent, nonzero correlations between the components for a number of subjects.

The SCA-P model with 4 components explains 58.4% of the variance, which is as much as 14.9% more than the SCA-P model with 2 components. The Varimax rotated loading matrix is presented in the right hand side of Table 3. In comparison to the loading matrices of SCA-IND and SCA-P with 2 components, the Varimax rotated SCA-P loading matrix roughly shows a split of the introversion component: the positive items load high (< -0.40) on a different component than the negative items do. These 2 components can be labeled as "Arousal" and "Nervousness". Furthermore, the items "excited" and "proud", that were not assigned to a component in the SCA-IND solution, form together with "ashamed" one component (component IV in Table 3). It is interesting to note that one subject shows a large positive correlation between "Arousal" and "Nervousness". This is strange from an interpretational point of view, but it fully reflects the trend that is also perceptible in the raw data of this subject. The covariances between the components for the different subjects vary from large to small, but they are difficult to interpret.

The interpretation of the three models just discussed has much in common, but the models differ in the details. The less restricted the model is, the more details (whether of the variables or the subjects) of the data are captured in the model. The preference for a particular model depends on the degree of interest in those details. The SCA-IND model with 2 components is fairly simple to interpret, and it covers the main features of the data, as is indicated by a relatively high fit, at least compared to SCA-P with 2 components. In both SCA-P models (2 and 4 components) the components covary in a complicated manner for the different subjects. Besides, the magnitudes of the correlation between the sets of component scores in the two SCA-P models are not convincingly large to justify a lower dimensional interpretation, even for some subjects. The SCA-P model with 4 components offers a detailed insight into the variable structure, and the model fits the data much better than SCA-IND with two components. This model is preferred if a detailed description is desired.

On the basis of the estimated SCA-P models and SCA-IND model, one can conclude that the positive/negative dimensions, which are found using the PANAS scale in healthy subjects, is not found in subjects suffering from Parkinson's disease. Instead, the intraindividual differences in mood in Parkinson patients follow the "Emotional Instability" and "Introversion" dimensions mainly. Furthermore, the three models show that the degree of intraindividual variability differed much across subjects, indicating that the stability of state of mind varies across Parkinson patients.

5. Discussion

Four variants of models for Simultaneous Component Analysis, and their properties have been discussed in the preceding sections. The models for SCA are particularly useful for the exploratory analysis of multivariate time series collected from a number of subjects. The four models can be ordered hierarchically. The most restricted model, the SCA-ECP model, allows for least variation between subjects (in terms of average cross-products of component scores, or covariances, if the component scores are centered), whereas the least restricted model, the SCA-P model, allows for most variation between subjects. A prerequisite is that a chosen model has a small degree of overall error, which can be assessed via cross-validation and split-half analyses. The methods evaluate different aspects of stability, and therefore we advise to use both methods simultaneously. The final decision should be based on the interpretability of the model. The empirical example illustrated the use of the SCA models in practice. The example offered a nicely interpretable solution, in which the intraindividual as well as the interindividual structure is covered.

The four models are not only suitable for modeling multivariate time series of two or more subjects simultaneously. SCA-P has been used to analyze scores of two or more groups of individuals on the same variables (e.g., Kiers & ten Berge, 1994; Niesing, 1997). An application of SCA-PF2 to chemical data is given in Bro, Andersson and Kiers (1999). Those types of data can be modeled by the other three methods as well. Depending on the extent of differences between groups, samples, or individuals SCA-P, SCA-PF2, or the new SCA-IND and SCA-ECP models can be chosen.

In the SCA models, the components are defined on the basis of degree of linear independence between the variable scores over time, at least over all subjects simultaneously. Linear independence of the component scores over time certainly does not imply that the distinct series of component scores are unrelated. One could possibly extend the present models to models incorporating nonlinear relationships between component scores. In exploratory modeling, a simple approach to reveal possible nonlinear relationships between component scores is to visually inspect the various component scores series over time simultaneously per subject.

In the four SCA models, only the covariance structure of the component scores is constrained. No structure is imposed on the component scores themselves. This has the advantage that the model is flexible, and that possible trends over time can be revealed from the component scores. However, it might be attractive in certain cases to constrain the component scores themselves, for example by imposing the component scores to follow an auto-regressive moving average (ARMA) process. In imposing an ARMA process, the models are in fact three-way generalizations of a special case of the linear dynamic system model (Bijleveld, 1989). Alternatively, smoothness constraints or even a certain functional form could be imposed on the component scores. Approaches to the latter that can easily be generalized to the SCA models, could be found in Timmerman and Kiers (2002) and Timmerman (2001), respectively.

One of the factor models for modeling multivariate time series of one subject is dynamic factor analysis (Molenaar, 1985). In this model, multivariate time series are decomposed into latent factor score series, and the observed scores are related to the factor scores at the same time point as well as to the factor scores at U previous time points via $U+1$ loading matrices, and thus so-called lagged factor influences are modeled. We have implemented the idea to model not only simultaneous effects, but also lagged effects in the four SCA models as they are discussed above. Alternating least squares algorithms to fit those "lagged" SCA models to data, and rotational procedures for those models that are not essentially unique have been constructed. However, our first attempts to model empirical data reveal that the estimated models are extremely difficult to interpret (Timmerman, 2001).

A. Appendix

Let $\mathbf{X}_i(K_i \times J)$ denote the matrix of scores of the i-th subject ($i = 1, \ldots, I$) on J variables measured on K_i occasions ($k_i = 1, \ldots, K_i$). To fit the SCA models to observed data, the following function is minimized

$$F(\mathbf{F}_i, \mathbf{B}) = \sum_{i=1}^{I} \left\| \mathbf{X}_i - \mathbf{F}_i\mathbf{B}' \right\|^2, \tag{A1}$$

where $\mathbf{F}_i(K_i \times Q)$ contains the Q component scores of subject i at time points $1, \ldots, K_i$, $\mathbf{B}(J \times Q)$ denotes the loading matrix, and $\mathbf{E}_i(K_i \times J)$ denotes the matrix of residuals, and where $\mathbf{F}_i$ is subject to the constraint in the SCA at issue.

Assumption 1. $\text{rank}(\mathbf{X}_{\mathbf{sup}}) \geq Q$, where $\mathbf{X}_{\mathbf{sup}}$ contains the matrices $\mathbf{X}_1, \ldots, \mathbf{X}_I$ stacked below each other, $\mathbf{X}_i(K_i \times J)$ is the matrix with scores of the i-th subject ($i = 1, \ldots, I$), and Q is the number of components as estimated in the SCA-P model.

Theorem 1. If $\mathbf{X}_i(K_i \times J)$, the matrix with scores of the i-th subject ($i = 1, \ldots, I$), is centered column-wise, i.e., $\mathbf{1}'_{K_i}\mathbf{X}_i = \mathbf{0}'_J$, $i = 1, \ldots, I$, then, under Assumption 1, the component scores matrix $\mathbf{F}_i$, $i = 1, \ldots, I$, in the SCA-P model is centered column-wise as well, that is,

$$\mathbf{1}'_{K_i}\mathbf{F}_i = \mathbf{0}'_Q, i = 1, \ldots, I.$$

Proof. To fit the SCA-P model to observed data, the sum of squared residuals is minimized. The algorithm to find SCA-P estimates is essentially based on a singular value decomposition of the supermatrix $\mathbf{X_{sup}} = \mathbf{U}_r\boldsymbol{\Delta}_r\mathbf{Q}'_r$, where $\mathbf{X_{sup}}$ contains the observed score matrices $\mathbf{X}_1, \ldots, \mathbf{X}_I$ positioned below each other, and $\mathbf{U}'_r\mathbf{U}_r = \mathbf{Q}'_r\mathbf{Q}_r = \mathbf{Q}_r\mathbf{Q}'_r = \mathbf{I}_r$, $\boldsymbol{\Delta}_r$ a diagonal matrix with positive diagonal elements in weakly descending order, and r the rank of $\mathbf{X_{sup}}$. The data matrix $\mathbf{X}_i$ of subject $i, i = 1, \ldots, I$, can be written as $\mathbf{X}_i = \mathbf{U}_{ir}\boldsymbol{\Delta}_r\mathbf{Q}'_r$, where $\mathbf{U}_{ir}$ is obtained by selecting the rows of $\mathbf{U}_r$ that correspond to the rows with the observed scores for subject i of $\mathbf{X_{sup}}$. From $\mathbf{X}_i = \mathbf{U}_{ir}\boldsymbol{\Delta}_r\mathbf{Q}'_r$ it follows that $\mathbf{U}_{ir} = \mathbf{X}_i\mathbf{Q}_r\boldsymbol{\Delta}_r^{-1}$, which, given that $\mathbf{1}'\mathbf{X}_i = \mathbf{0}'$, implies that $\mathbf{1}'\mathbf{U}_{ir} = \mathbf{0}'$, $i = 1, \ldots, I$, as well. The matrices $\mathbf{F}_i$, $i = 1, \ldots, I$, are obtained as the first Q columns of $\mathbf{U}_{ir}$, where $Q \leq r$, as follows from Assumption 1, and hence $\mathbf{1}'_{K_i}\mathbf{F}_i = \mathbf{0}'_Q$. □

Theorem 2. If $\mathbf{X}_i(K_i \times J)$, the matrix with scores of the i-th subject ($i = 1, \ldots, I$), is centered column-wise, i.e., $\mathbf{1}'_{K_i}\mathbf{X}_i = \mathbf{0}'_J$, $i = 1, \ldots, I$, then, under a mild assumption specified in the proof, the component score matrix $\mathbf{F}_i$, $i = 1, \ldots, I$, in the SCA-PF2, SCA-IND, and SCA-ECP models is centered column-wise as well, that is, $\mathbf{1}'_{K_i}\mathbf{F}_i = \mathbf{0}'_Q$, $i = 1, \ldots, I$.

Proof. To fit the SCA-PF2, SCA-IND, or SCA-ECP models to observed data, the sum of squared residuals is minimized via an alternating least squares algorithm. The algorithm to fit the SCA-PF2 model, requiring that $\frac{1}{K_i}\mathbf{F}'_i\mathbf{F}_i = \mathbf{D}_i\boldsymbol{\Phi}\mathbf{D}_i$, with $\mathbf{D}_i$ a diagonal $Q \times Q$ matrix and $\boldsymbol{\Phi}$ a positive definite $Q \times Q$ matrix with unit diagonal elements, uses an ALS algorithm for the equivalent problem of minimizing (4), subject to the constraint $\mathbf{F}'_i\mathbf{F}_i = \tilde{\mathbf{D}}_i\tilde{\boldsymbol{\Phi}}\tilde{\mathbf{D}}_i$ with $\tilde{\boldsymbol{\Phi}}$ an arbitrary positive definite $Q \times Q$ matrix and $\tilde{\mathbf{D}}_i$ a diagonal $Q \times Q$ matrix. The algorithm is essentially based on the fact that every matrix $\mathbf{F}_i$ that meets the constraint $\mathbf{F}'_i\mathbf{F}_i = \tilde{\mathbf{D}}_i\tilde{\boldsymbol{\Phi}}\tilde{\mathbf{D}}_i$ can be written as $\mathbf{F}_i = \mathbf{P}_i\tilde{\mathbf{F}}\tilde{\mathbf{D}}_i$ provided that $\mathbf{P}'_i\mathbf{P}_i = \mathbf{I}_Q$, $i = 1, \ldots, I$, and $\tilde{\boldsymbol{\Phi}} = \tilde{\mathbf{F}}'\tilde{\mathbf{F}}$. The matrices $\mathbf{P}_i$, $\tilde{\mathbf{F}}$, $\tilde{\mathbf{D}}_i$, and $\mathbf{B}$ are alternatingly updated in the SCA-PF2 algorithm. The algorithm to fit the SCA-IND model uses the SCA-PF2 algorithm as follows: to find estimates of $\mathbf{F}_i$, $i = 1, \ldots, I$, in the SCA-IND algorithm, the SCA-PF2 algorithm is used, keeping $\tilde{\mathbf{F}}$ fixed at identity. An update of $\mathbf{P}_i$ in the SCA-PF2 algorithm, subject to $\mathbf{P}'_i\mathbf{P}_i = \mathbf{I}_Q$, can be found by first computing the singular value decomposition of

$$\mathbf{X}_i\mathbf{B}\tilde{\mathbf{D}}_i\tilde{\mathbf{F}}' = \mathbf{U}_{iQ}\boldsymbol{\Delta}_{iQ}\mathbf{Q}'_{iQ}, \mathbf{U}'_{iQ}\mathbf{U}_{iQ} = \mathbf{Q}'_{iQ}\mathbf{Q}_{iQ} = \mathbf{Q}_{iQ}\mathbf{Q}'_{iQ} = \mathbf{I}_Q,$$

and $\boldsymbol{\Delta}_{iQ}$ a diagonal matrix with nonnegative diagonal elements in weakly descending order, and then taking $\mathbf{P}_i = \mathbf{U}_{iQ}\mathbf{Q}'_{iQ}$, where Q is the number of components. Now assuming that $\text{rank}(\mathbf{X}_i\mathbf{B}\tilde{\mathbf{D}}_i\tilde{\mathbf{F}}') = Q$, which can be expected to be satisfied in practice, the equality $\mathbf{X}_i\mathbf{B}\tilde{\mathbf{D}}_i\tilde{\mathbf{F}}' = \mathbf{U}_{iQ}\boldsymbol{\Delta}_{iQ}\mathbf{Q}'_{iQ}$ implies that $\mathbf{U}_{iQ} = \mathbf{X}_i\mathbf{B}\tilde{\mathbf{D}}_i\tilde{\mathbf{F}}'\mathbf{Q}_{iQ}\boldsymbol{\Delta}_{iQ}^{-1}$, which, given that $\mathbf{1}'\mathbf{X}_i = \mathbf{0}'$, implies that $\mathbf{1}'\mathbf{U}_{iQ} = \mathbf{0}'$, $i = 1, \ldots, I$, as well. After convergence, for SCA-PF2, $\mathbf{F}_i = \mathbf{P}_i\tilde{\mathbf{F}}\tilde{\mathbf{D}}_i$, whereas for SCA-IND $\mathbf{F}_i = \mathbf{P}_i\tilde{\mathbf{F}}\tilde{\mathbf{D}}_i$ with $\tilde{\mathbf{F}}$ fixed at identity. Given that $\mathbf{F}_i = \mathbf{P}_i\tilde{\mathbf{F}}\tilde{\mathbf{D}}_i$ and that every update of $\mathbf{P}_i$ is taken as $\mathbf{P}_i = \mathbf{U}_{iQ}\mathbf{Q}'_{iQ}$, it follows that if $\mathbf{1}'\mathbf{X}_i = \mathbf{0}'$, then $\mathbf{1} = \mathbf{F}_i = \mathbf{1}'\mathbf{U}_{iQ}\mathbf{Q}'_{iQ}\tilde{\mathbf{F}}\tilde{\mathbf{D}}_i = \mathbf{0}'\mathbf{Q}_{iQ}\tilde{\mathbf{F}}\tilde{\mathbf{D}}_i = \mathbf{0}'$.

An update of $\mathbf{F}_i$ in the SCA-ECP algorithm, subject to $\frac{1}{K_i}\mathbf{F}'_i\mathbf{F}_i = \mathbf{I}_Q$, can be found by first computing the singular value decomposition of

$$\mathbf{X}_i\mathbf{B} = \mathbf{U}_{iQ}\boldsymbol{\Delta}_{iQ}\mathbf{Q}'_{iQ}, \mathbf{U}'_{iQ}\mathbf{U}_{iQ} = \mathbf{Q}'_{iQ}\mathbf{Q}_{iQ} = \mathbf{Q}_{iQ}\mathbf{Q}'_{iQ} = \mathbf{I}_Q,$$

and $\boldsymbol{\Delta}_{iQ}$ a diagonal matrix with nonnegative diagonal elements in weakly descending order, and then taking $\mathbf{F}_i = \sqrt{K_i}\mathbf{U}_{iQ}\mathbf{Q}'_{iQ}$, where Q is the number of components. Assuming that $\text{rank}(\mathbf{X}_i\mathbf{B}) = Q$, which can be expected to be satisfied in practice, the equality $\mathbf{X}_i\mathbf{B} = \mathbf{U}_{iQ}\boldsymbol{\Delta}_{iQ}\mathbf{Q}'_{iQ}$ implies that $\mathbf{U}_{iQ} = \mathbf{X}_i\mathbf{B}\mathbf{Q}_{iQ}\boldsymbol{\Delta}_{iQ}^{-1}$, which, given that $\mathbf{1}'\mathbf{X}_i = \mathbf{0}'$, implies that $\mathbf{1}'\mathbf{U}_{iQ} = \mathbf{0}'$, $i = 1, \ldots, I$, as well. Given that $\mathbf{F}_i$ is taken as $\mathbf{F}_i = \sqrt{K_i}\mathbf{U}_{iQ}\mathbf{Q}'_{iQ}$, it follows that if $\mathbf{1}'\mathbf{X}_i = \mathbf{0}'$, then $\mathbf{1}'\mathbf{F}_i = \mathbf{1}'\mathbf{U}_{iQ}\mathbf{Q}'_{iQ}\sqrt{K_i} = \mathbf{0}'\mathbf{Q}'_{iQ}\sqrt{K_i} = \mathbf{0}'$. □

References

Anderson, T.W. (1963). The use of factor analysis in the statistical analysis of multiple time series. *Psychometrika, 28*, 1–25.

Bijleveld, C.C.J.H. (1989). *Exploratory linear dynamic systems analysis*. Leiden, Netherlands: DSWO Press.

Bijleveld, C.C.J.H., & Bijleveld, F.D. (1997). A comparison of the applicability of two approaches to linear dynamic systems analyze for N subjects. *Kwantitatieve Methoden, 55*, 37–57.

Bro, R., Andersson, C.A., & Kiers, H.A.L. (1999). PARAFAC2-Part II. Modeling chromatographic data with retention time shifts. *Journal of Chemometrics, 13*, 295–309.

Bro, R., & Smilde, A.K. (2003). Centering and scaling in component analysis. *Journal of Chemometrics, 17*, 16–33.

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

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.

Cattell, R.B. (1952). *Factor analysis*. New York, NY: Harper.

Cattell, R.B. (1963). The structuring of change by P-technique and incremental R-technique. In C.W. Harris (Ed.), *Problems in measuring change*. Madison, WI: The University of Wisconsin Press.

Cliff, N. (1966). Orthogonal rotation to congruence. *Psychometrika, 31*, 33–42.

Engle, R., & Watson, M. (1981). A one-factor multivariate time series model of metropolitan wage rates. *Journal of the American Statistical Association, 76*, 774–781.

Harshman, R.A. (1972). Determination and proof of minimum uniqueness conditions for PARAFAC1. *UCLA Working Papers in Phonetics, 22*, 111–117.

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

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

Hastie, T, Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning: Data mining, inference, and prediction*. New York, NY: Springer.

Immink, W. (1986). *Parameter estimation in Markov models and dynamic factor analysis*. Unpublished doctoral thesis, University of Utrecht, Utrecht, Netherlands.

Jones, C.J., & Nesselroade, J.R. (1990). Multivariate, replicated, single-subject, repeated measures designs and P-technique factor analysis: A review of intraindividual change studies. *Experimental Aging Research, 16*, 171–183.

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

Kiers, H.A.L., & ten Berge, J.M.F. (1994). Hierarchical relations between methods for simultaneous component analysis and a technique for rotation to a simple simultaneous structure. *British Journal of Mathematical and Statistical Psychology, 47*, 109-126.

Kiers, H.A.L., ten Berge, J.M.F., & Bro, R. (1999). PARAFAC2-Part 1. A direct fitting algorithm for the PARAFAC2 model. *Journal of Chemometrics, 13*, 275–294.

Little, R.J.A., & Rubin, D.B. (1987). *Statistical analysis with missing data*. New York, NY: Wiley.

Louwerse, D.J., Smilde, A.K., & Kiers, H.A.L. (1999). Cross-validation of multiway component models. *Journal of Chemometrics, 13*, 491–510.

Molenaar, P.M. (1985). A dynamic factor model for the analysis of multivariate time series. Psychometrika, *50*, 181–202.

Nesselroade, J.R., & Molenaar, P.C.M. (1999). Pooling lagged covariances structures based on short, multivariate time series for dynamic factor analysis. In R.H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 223–250). Thousand Oaks, CA: Sage Publications.

Niesing, J. (1997). *Simultaneous component and factor analysis methods for two or more groups: A comparative study*. Leiden, Netherlands: DSWO Press.

Shifren, K., Hooker, K., Wood, P., & Nesselroade, J.R. (1997). Structure and variation of mood in individuals with Parkinson's disease: A dynamic factor analysis. *Psychology and aging, 12*, 328–339.

Timmerman, M.E. (2001). *Component analysis of multisubject multivariate longitudinal data*. Unpublished doctoral thesis, University of Groningen, Groningen, Netherlands.

Timmerman, M.E., & Kiers, H.A.L. (2002). Three-way component analysis with smoothness constraints. *Computational Statistics and Data Analysis, 3*, 447–470.

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

van Buuren, S. (1990). *Optimal scaling of time series*. Leiden, Netherlands: DSWO Press.

Watson, D. (1988). Intraindividual and interindividual analyses of positive and negative affect: Their relation to health complaints, perceived stress, and daily activities. *Journal of personality and social psychology, 54*, 1020–1030.

Watson, D., Clark, L.A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070.

Manuscript received 1 MAR 2000
Final version received 3 JUN 2002