Appendix A True pattern matrices (order 12×2 and 12×4) used in this study

$$P_a=\begin{pmatrix} .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \end{pmatrix},\ P_b=\begin{pmatrix} .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \end{pmatrix},\ P_c=\begin{pmatrix} .37 & 0 \\ .37 & 0 \\ .37 & 0 \\ .37 & 0 \\ .37 & 0 \\ .37 & 0 \\ 0 & .37 \\ 0 & .37 \\ 0 & .37 \\ 0 & .37 \\ 0 & .37 \\ 0 & .37 \end{pmatrix},\ P_d=\begin{pmatrix} .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \end{pmatrix},$$

$$P_e=\begin{pmatrix} .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ 0 & .37 \\ 0 & .37 \\ 0 & .37 \\ 0 & .37 \\ 0 & .37 \\ 0 & .37 \end{pmatrix},\ P_f=\begin{pmatrix} .37 & 0 \\ .55 & 0 \\ .89 & 0 \\ .37 & 0 \\ .55 & 0 \\ .89 & 0 \\ 0 & .37 \\ 0 & .55 \\ 0 & .89 \\ 0 & .37 \\ 0 & .55 \\ 0 & .89 \end{pmatrix},\ P_g=\begin{pmatrix} .89 & 0 \\ .55 & 0 \\ .37 & 0 \\ .89 & 0 \\ .55 & 0 \\ .37 & 0 \\ 0 & .89 \\ 0 & .55 \\ 0 & .37 \\ 0 & .89 \\ 0 & .55 \\ 0 & .37 \end{pmatrix},\ P_h=\begin{pmatrix} .55 & 0 \\ .37 & 0 \\ .89 & 0 \\ .55 & 0 \\ .37 & 0 \\ .89 & 0 \\ 0 & .55 \\ 0 & .37 \\ 0 & .89 \\ 0 & .55 \\ 0 & .37 \\ 0 & .89 \end{pmatrix},$$

$$P_l=\begin{pmatrix} .37 & 0 \\ .89 & 0 \\ .55 & 0 \\ .37 & 0 \\ .89 & 0 \\ .55 & 0 \\ 0 & .37 \\ 0 & .89 \\ 0 & .55 \\ 0 & .37 \\ 0 & .89 \\ 0 & .55 \end{pmatrix},\ P_m=\begin{pmatrix} .37 & 0 \\ .37 & 0 \\ .37 & 0 \\ .37 & 0 \\ .37 & 0 \\ .37 & 0 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \end{pmatrix},\ P_n=\begin{pmatrix} .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \end{pmatrix},\ P_p=\begin{pmatrix} .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ 0 & .37 \\ 0 & .37 \\ 0 & .37 \\ 0 & .37 \\ 0 & .37 \\ 0 & .37 \end{pmatrix},$$

$$P_q=\begin{pmatrix} .37 & 0 \\ .37 & 0 \\ .37 & 0 \\ .37 & 0 \\ .37 & 0 \\ .37 & 0 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \end{pmatrix},\ P_r=\begin{pmatrix} .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ .89 & 0 \\ .63 & .63 \\ .63 & .63 \\ .63 & .63 \\ .63 & .63 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \\ 0 & .89 \end{pmatrix},\ P_s=\begin{pmatrix} .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ .55 & 0 \\ .39 & .39 \\ .39 & .39 \\ .39 & .39 \\ .39 & .39 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \\ 0 & .55 \end{pmatrix}.$$

$$P_a=\begin{pmatrix} .89 & 0 & 0 & 0 \\ .89 & 0 & 0 & 0 \\ .89 & 0 & 0 & 0 \\ 0 & .89 & 0 & 0 \\ 0 & .89 & 0 & 0 \\ 0 & .89 & 0 & 0 \\ 0 & 0 & .89 & 0 \\ 0 & 0 & .89 & 0 \\ 0 & 0 & .89 & 0 \\ 0 & 0 & 0 & .89 \\ 0 & 0 & 0 & .89 \\ 0 & 0 & 0 & .89 \end{pmatrix},\ P_b=\begin{pmatrix} .55 & 0 & 0 & 0 \\ .55 & 0 & 0 & 0 \\ .55 & 0 & 0 & 0 \\ 0 & .55 & 0 & 0 \\ 0 & .55 & 0 & 0 \\ 0 & .55 & 0 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & 0 & .55 \\ 0 & 0 & 0 & .55 \\ 0 & 0 & 0 & .55 \end{pmatrix},\ P_c=\begin{pmatrix} .37 & 0 & 0 & 0 \\ .37 & 0 & 0 & 0 \\ .37 & 0 & 0 & 0 \\ 0 & .37 & 0 & 0 \\ 0 & .37 & 0 & 0 \\ 0 & .37 & 0 & 0 \\ 0 & 0 & .37 & 0 \\ 0 & 0 & .37 & 0 \\ 0 & 0 & .37 & 0 \\ 0 & 0 & 0 & .37 \\ 0 & 0 & 0 & .37 \\ 0 & 0 & 0 & .37 \end{pmatrix},$$

$$P_d=\begin{pmatrix} .89 & 0 & 0 & 0 \\ .89 & 0 & 0 & 0 \\ .89 & 0 & 0 & 0 \\ 0 & .89 & 0 & 0 \\ 0 & .89 & 0 & 0 \\ 0 & .89 & 0 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & 0 & .55 \\ 0 & 0 & 0 & .55 \\ 0 & 0 & 0 & .55 \end{pmatrix},\ P_e=\begin{pmatrix} .89 & 0 & 0 & 0 \\ .89 & 0 & 0 & 0 \\ .89 & 0 & 0 & 0 \\ 0 & .89 & 0 & 0 \\ 0 & .89 & 0 & 0 \\ 0 & .89 & 0 & 0 \\ 0 & 0 & .37 & 0 \\ 0 & 0 & .37 & 0 \\ 0 & 0 & .37 & 0 \\ 0 & 0 & 0 & .37 \\ 0 & 0 & 0 & .37 \\ 0 & 0 & 0 & .37 \end{pmatrix},\ P_f=\begin{pmatrix} .37 & 0 & 0 & 0 \\ .55 & 0 & 0 & 0 \\ .89 & 0 & 0 & 0 \\ 0 & .37 & 0 & 0 \\ 0 & .55 & 0 & 0 \\ 0 & .89 & 0 & 0 \\ 0 & 0 & .37 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & .89 & 0 \\ 0 & 0 & 0 & .37 \\ 0 & 0 & 0 & .55 \\ 0 & 0 & 0 & .89 \end{pmatrix},$$

$$P_g=\begin{pmatrix} .89 & 0 & 0 & 0 \\ .55 & 0 & 0 & 0 \\ .37 & 0 & 0 & 0 \\ 0 & .89 & 0 & 0 \\ 0 & .55 & 0 & 0 \\ 0 & .37 & 0 & 0 \\ 0 & 0 & .89 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & .37 & 0 \\ 0 & 0 & 0 & .89 \\ 0 & 0 & 0 & .55 \\ 0 & 0 & 0 & .37 \end{pmatrix}, \; P_h=\begin{pmatrix} .55 & 0 & 0 & 0 \\ .37 & 0 & 0 & 0 \\ .89 & 0 & 0 & 0 \\ 0 & .55 & 0 & 0 \\ 0 & .37 & 0 & 0 \\ 0 & .89 & 0 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & .37 & 0 \\ 0 & 0 & .89 & 0 \\ 0 & 0 & 0 & .55 \\ 0 & 0 & 0 & .37 \\ 0 & 0 & 0 & .89 \end{pmatrix}, \; P_l=\begin{pmatrix} .37 & 0 & 0 & 0 \\ .89 & 0 & 0 & 0 \\ .55 & 0 & 0 & 0 \\ 0 & .37 & 0 & 0 \\ 0 & .89 & 0 & 0 \\ 0 & .55 & 0 & 0 \\ 0 & 0 & .37 & 0 \\ 0 & 0 & .89 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & 0 & .37 \\ 0 & 0 & 0 & .89 \\ 0 & 0 & 0 & .55 \end{pmatrix},$$

$$P_m=\begin{pmatrix} .37 & 0 & 0 & 0 \\ .37 & 0 & 0 & 0 \\ .37 & 0 & 0 & 0 \\ 0 & .37 & 0 & 0 \\ 0 & .37 & 0 & 0 \\ 0 & .37 & 0 & 0 \\ 0 & 0 & .89 & 0 \\ 0 & 0 & .89 & 0 \\ 0 & 0 & .89 & 0 \\ 0 & 0 & 0 & .89 \\ 0 & 0 & 0 & .89 \\ 0 & 0 & 0 & .89 \end{pmatrix}, \; P_n=\begin{pmatrix} .55 & 0 & 0 & 0 \\ .55 & 0 & 0 & 0 \\ .55 & 0 & 0 & 0 \\ 0 & .55 & 0 & 0 \\ 0 & .55 & 0 & 0 \\ 0 & .55 & 0 & 0 \\ 0 & 0 & .89 & 0 \\ 0 & 0 & .89 & 0 \\ 0 & 0 & .89 & 0 \\ 0 & 0 & 0 & .89 \\ 0 & 0 & 0 & .89 \\ 0 & 0 & 0 & .89 \end{pmatrix}, \; P_p=\begin{pmatrix} .55 & 0 & 0 & 0 \\ .55 & 0 & 0 & 0 \\ .55 & 0 & 0 & 0 \\ 0 & .55 & 0 & 0 \\ 0 & .55 & 0 & 0 \\ 0 & .55 & 0 & 0 \\ 0 & 0 & .37 & 0 \\ 0 & 0 & .37 & 0 \\ 0 & 0 & .37 & 0 \\ 0 & 0 & 0 & .37 \\ 0 & 0 & 0 & .37 \\ 0 & 0 & 0 & .37 \end{pmatrix},$$

$$P_q=\begin{pmatrix} .37 & 0 & 0 & 0 \\ .37 & 0 & 0 & 0 \\ .37 & 0 & 0 & 0 \\ 0 & .37 & 0 & 0 \\ 0 & .37 & 0 & 0 \\ 0 & .37 & 0 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & .55 & 0 \\ 0 & 0 & 0 & .55 \\ 0 & 0 & 0 & .55 \\ 0 & 0 & 0 & .55 \end{pmatrix}.$$

Appendix B Identification constraints used for the LISREL program

Constrained values for the matrix Λ, used in the conditions where there were 12 variables for the six different models (with one to six factors), as specified for the method SIFASP–ML for the LISREL8 program. All zeros and ones stand for fixed loadings (fixed at zero or one). These values could not change during the iterations of the program. All nonfixed elements of Λ were left completely free. For Experiment 4c, the constrained values were somewhat adjusted (see below).

$$\Lambda_1=\begin{pmatrix} 1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix},\ \Lambda_2=\begin{pmatrix} 1 & 0 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 0 & 1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{pmatrix},\ \Lambda_3=\begin{pmatrix} 1 & 0 & 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0 & 1 & 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0 & 0 & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix},\ \Lambda_4=\begin{pmatrix} 1 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 1 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix},$$

$$\Lambda_5=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 1 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 1 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix},\ \Lambda_6=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}.$$

In the condition of 24 variables, the first 12 rows of the matrix Λ with starting values looked the same as above. The matrix was completed by stacking a matrix with 12 rows in which no values were fixed below this matrix.

For Experiment 4c, the constrained values, as presented below, were used.

$$\boldsymbol{\Lambda}_1=\begin{pmatrix}1\\ .\\ .\\ .\\ .\\ .\\ .\\ .\\ .\\ .\\ .\\ .\end{pmatrix},\ \boldsymbol{\Lambda}_2=\begin{pmatrix}1&0\\ .&.\\ .&.\\ .&.\\ .&.\\ .&.\\ .&.\\ .&.\\ .&.\\ .&.\\ .&.\\ 0&1\end{pmatrix},\ \boldsymbol{\Lambda}_3=\begin{pmatrix}1&0&0\\ .&.&.\\ .&.&.\\ .&.&.\\ 0&1&0\\ .&.&.\\ .&.&.\\ .&.&.\\ 0&0&1\\ .&.&.\\ .&.&.\\ .&.&.\end{pmatrix},\ \boldsymbol{\Lambda}_4=\begin{pmatrix}1&0&0&0\\ .&.&.&.\\ .&.&.&.\\ .&.&.&.\\ 0&1&0&0\\ .&.&.&.\\ .&.&.&.\\ 0&0&1&0\\ .&.&.&.\\ .&.&.&.\\ .&.&.&.\\ 0&0&0&1\end{pmatrix},$$

$$\boldsymbol{\Lambda}_5=\begin{pmatrix}1&0&0&0&0\\ .&.&.&.&.\\ .&.&.&.&.\\ 0&1&0&0&0\\ .&.&.&.&.\\ .&.&.&.&.\\ 0&0&1&0&0\\ .&.&.&.&.\\ 0&0&0&1&0\\ .&.&.&.&.\\ 0&0&0&0&1\\ .&.&.&.&.\end{pmatrix},\ \boldsymbol{\Lambda}_6=\begin{pmatrix}1&0&0&0&0&0\\ .&.&.&.&.&.\\ 0&1&0&0&0&0\\ .&.&.&.&.&.\\ 0&0&1&0&0&0\\ .&.&.&.&.&.\\ 0&0&0&1&0&0\\ .&.&.&.&.&.\\ 0&0&0&0&1&0\\ .&.&.&.&.&.\\ 0&0&0&0&0&1\\ .&.&.&.&.&.\end{pmatrix}.$$

Appendix C Deriving the largest possible difference in explained variance between the methods SCA-W and SCA-P in case two rank q matrices are analyzed

Here, it will be shown that the largest possible difference in explained variance between the methods SCA-W and SCA-P, in the case where two rank q correlation matrices of order $m \times m$ are analyzed, is 50% (see Section 3.3.3). This happens when the true pattern matrices $\mathbf{P}_{t1}$ and $\mathbf{P}_{t2}$ are chosen orthogonal. In this situation the components found with SCA-W explain 100% of the variance and the components found with SCA-P explain 50% of the variance. This can be proven as follows:

Define an $m \times q$ pattern matrix $\mathbf{P}_{t1}$ ($q \leq \frac{1}{2}m$), in which each component pertains to the same number of m/q variables, the q components do not overlap, and all nonzero loadings are $\pm$ 1. Choose a second pattern matrix $\mathbf{P}_{t2}$ in the same way, with the additional requirement that the columns of the matrices $\mathbf{P}_{t1}$ and $\mathbf{P}_{t2}$ are orthogonal. Choose the matrix $\mathbf{\Phi}$ as $\mathbf{I}$. The error free population correlation matrices are now written as $\mathbf{R}_{t1}=\mathbf{P}_{t1}\mathbf{P}'_{t1}$ and $\mathbf{R}_{t2}=\mathbf{P}_{t2}\mathbf{P}'_{t2}$. The total amount of variance to be explained is $2m$. In this situation, the q-dimensional solution, found with the method SCA-W, explains 100% of variance, as shown in Section 3.3.3. The method SCA-P, however, only explains 50% of the total variance as will now be shown. That SCA-P never explains less than 50% follows from the fact that q components can always explain 100% of variance in one of the two groups. The q-dimensional solution of the method SCA-P is found by taking the first q eigenvectors of $\mathbf{R}_{sum}$ $(=\mathbf{R}_{t1}+\mathbf{R}_{t2}=\mathbf{P}_{t1}\mathbf{P}'_{t1}+\mathbf{P}_{t2}\mathbf{P}'_{t2})$. Because the columns of the matrices $\mathbf{P}_{t1}$ and $\mathbf{P}_{t2}$ are orthogonal, the matrix $\mathbf{R}_{sum}$ has rank $2q$. Furthermore, $\mathbf{R}_{sum}=(\mathbf{P}_{t1}|\mathbf{P}_{t2})(\mathbf{P}_{t1}|\mathbf{P}_{t2})'$, hence $(\mathbf{P}_{t1}|\mathbf{P}_{t2})$ contains all eigenvectors of $\mathbf{R}_{sum}$ that correspond to nonzero eigenvalues. In fact, these eigenvalues are all m/q. Hence, taking $\mathbf{K}$ as any set of q normalized columns of $(\mathbf{P}_{t1}|\mathbf{P}_{t2})$ maximizes $\mathrm{tr}(\mathbf{K}'\mathbf{R}_{sum}\mathbf{K})$, and any such set of columns gives the best SCA-P solution. Each such component corresponds to an eigenvalue of m/q, hence together these q components explain $\frac{m}{2m} * 100\% =$ 50% of the variance.

Appendix D Examining real life data

In this appendix, the suggestions for practical use of SCA–P and SCA–S, made at the end of Chapter 9, will be illustrated on two empirical data sets. These have been analyzed previously, but not with SCA–P and SCA–S.

D.1 MEREDITH (1964) DATA

The first data set used was the same as the one analyzed by Meredith (1964). Meredith used a selection of tests from a monograph by Holzinger and Swineford (1939). The purpose of Meredith's analysis was to illustrate the performance of two procedures for rotating any number of factor pattern matrices based on different populations to conform to a single "best fitting" factor pattern matrix. Here, on the other hand, this data set is used to illustrate the SCA–methods. The data consist of scores of seventh and eight grade students from two schools on nine tests. The two schools had a different socioeconomic character. One (the Pasteur school) enrolling children of factory workers and one (the Grant–White school) enrolling children in a middle–class suburban area. In the selection of nine tests used by Meredith three different types of tests were administered. They were chosen so that there were a "space" factor, a verbal factor, and a memory factor present, each factor represented by three tests. These tests were: A Visual Perception test involving logical sequences of abstract figures, a Cubes test with designs on the surfaces, and a Paper Form Board test for the "space" factor; a General Information test, a Sentence Completion test, and a Word Classification test for the verbal factor; and a Figure recognition test, an Object–Number paired associates test, and a Number–Figure paired associates test for the memory factor. A complete description of the samples and tests can be found in Holzinger and Swineford (1939).

Meredith (1964) split the two groups of children into two groups each, by splitting at the median (within each school) on one of the speeded tests (an addition test; in the selection of nine tests by

Meredith, speeded tests were not included). This yielded four groups, denoted by Pasteur–Low (speed), Pasteur–High, Grant–White–Low, and Grant–White–High. The four groups had sample sizes of 77, 79, 74 and 71, respectively.

Data from the four groups were first analyzed with SCA–P. That is, the eigenvalues of the sum of the weighted correlation matrices were calculated and the values of the QDA measure were derived from these (see Section 4.2.4) to determine the number of components to retain. Only for the one, two and three–dimensional solution, the QDA value was acceptable (only for those three solutions the additional amount of explained variance was larger than the amount of variance explained by one variable). These three QDA values for the one, two and three–dimensional solution were 8.42, .45 and 5.42, respectively, indicating that one or three components should be retained. On theoretical grounds, three components were retained.

For both SCA–P and SCA–S, the three–dimensional solution was calculated. The solution found with SCA–P was rotated using the HKIC rotation (see Section 2.4.1); the solution found with SCA–S was rotated using the simple structure rotation described in Section 2.4.3. The pattern matrix, found with SCA–P, was normalized so that $\mathbf{P'P}=\mathbf{I}$. Note that with the normalization the pattern equals the weights matrix. Also for the three–dimensional solution, the mean percentage of explained variance of four separate PCA's for the four groups was calculated. The percentage of explained variance for the three–dimensional solution was 64.3% for the separate PCA's, 62.9% for SCA–P and 60.9% for SCA–S. The difference between SCA–P and SCA–S is 2.0%, which is considered small, but not negligible, suggesting that there are small differences between the groups. The congruences of the columns of the pattern matrix found with SCA–P and the corresponding columns of the structure matrix found with SCA–S were .93, .89 and .89 for the three components, respectively. This suggests the samples were drawn from the same population. As group size for the discriminant function (5.1), the mean group size of the four groups 75.25 was chosen. The value of the discriminant function thus found was 1.32, also suggesting that the samples were drawn from the same population. Besides inspecting the pattern matrix, found with SCA–P, also the structure matrices, found with SCA–P will be inspected, to detect possible small differences in behavior between the components in the four groups.

The loadings of the pattern matrix found with SCA–P are presented in Table D.1. As can be deduced from the loadings, the three components can be labeled as a "space" component (the ability to picture an object and perform actions (rotations) upon it), a verbal component, and a memory component. The relatively high loading of the test 'Figure Recognition' on the space component, besides the memory component, is not surprising, because this test clearly involves visual perception in addition to memory.

The structure matrix for each group is presented in Table D.2. As can be seen from Table D.2, the variables (tests) correlate highest with the component to which they pertain, with the exception of the variable 'Figure Recognition' in groups 1 and 4, that has a higher correlation with the 'space' component in those groups.

The most striking difference between the two Pasteur groups and the two Grant–White groups is that for the latter the components show much more overlap than for the Pasteur groups. This can also be seen in the correlations between the components, which are .19 when averaged over the

Table D.1: *Pattern matrix for the three dimensional solution, found with the method SCA–P*

	space	verbal	memory
Visual Perception	.49	.10	.10
Cubes	.58	−.07	−.10
Paper Form Board	.54	.02	−.11
General Information	.01	.58	−.05
Sentence Completion	−.05	.61	−.01
Word Classification	.05	.52	.09
Figure Recognition	.34	−.06	.39
Object–Number	−.14	.02	.67
Number–Figure	.05	−.02	.60

two Pasteur groups and .38 when averaged over the two Grant–White groups. As a consequence, the six components in the two Grant–White groups are all stronger than the corresponding components in the Pasteur groups. When comparing the Low–speed groups, we see that the differences in component strengths are considerable for both the verbal component and the memory component. When comparing the High–speed groups, the differences in component strengths are considerable for both the space component and the memory component. These findings can all be attributed

Table D.2: *Variable – component correlations for the three dimensional solution, found by the method SCA – P, in the four groups, and the component strengths, in percentages of explained variance*

	Group 1: Pasteur–Low			Group 2: Pasteur–High		
	space	verbal	memory	space	verbal	memory
Visual Perception	.77	.38	.32	.70	.43	.15
Cubes	.73	.06	.04	.67	.12	–.15
Paper Form Board	.74	.08	.02	.64	.22	–.22
General Information	.17	.90	–.01	.29	.87	.13
Sentence Completion	.11	.91	.12	.35	.90	.09
Word Classification	.38	.82	.36	.30	.84	.16
Figure Recognition	.59	.19	.56	.33	.09	.53
Object–Number	.06	.06	.81	–.12	.06	.77
Number–Figure	.11	.10	.76	–.00	.11	.72
Component strength	24.6%	28.0%	20.0%	19.7%	28.3%	17.1%
	Group 3: Grant–W–Low			**Group 4: Grant–W–High**		
	space	verbal	memory	space	verbal	memory
Visual Perception	.75	.49	.38	.75	.34	.47
Cubes	.70	.20	.14	.70	.19	.16
Paper Form Board	.70	.35	.12	.67	.36	.29
General Information	.42	.86	.34	.39	.86	.11
Sentence Completion	.33	.91	.35	.29	.88	.18
Word Classification	.41	.83	.43	.39	.83	.26
Figure Recognition	.59	.23	.64	.66	.24	.65
Object–Number	.09	.36	.81	.14	.16	.84
Number–Figure	.37	.32	.78	.51	.13	.78
Component strength	27.8%	32.7%	25.0%	28.7%	28.6%	24.1%

to the presence of a stronger first unstated component (largest eigenvalue) in the Grant-White groups (explaining 36.6% and 39.4% of the total variance, respectively) than in the Pasteur groups (32.4% and 30.3%, respectively).

D.2 VAN SCHUUR (1984) DATA

The second data set used was taken from Van Schuur (1984), and was previously used by Ten Berge (1986c) to illustrate the crossvalidation of weights, and by Ten Berge and Kiers (1990) and Kiers and Ten Berge (1994a) to give an example of the application of SCA-W. The data set consists of scores of two samples of party-activists (an Italian sample of 718 persons and a Danish sample of 1565 persons) who rated 15 political issues on Likert scales. The scale ran from 1 (very much against) to 5 (very much in favor). The issues are Fight inflation (INFL), Accelerate European Integration (EURO), Reduce Public Control (PUBL), Women Decide About Abortion (ABOR), Fight unemployment (EMPL), Defense Against Superpowers (SUPP), Punish Terrorists (TERR), Develop Nuclear Energy (NUCL), Control Multinationals (MULT), Protect Environment (ENVM), Reduce Regional Differences (REGI), Equal Opportunity for Men and Women (EQUA), Increase Military Expenditure (MILT), Reduce Income Differences (INCO), and Concern Own Needs Versus needs of Third World (THRD).

The two groups were first analyzed by SCA-P. It must be noted that, because of the large group sizes, it was decided not to weight the matrices of the two groups according to the sample size, as is described in Sections 2.2.2 and 2.5.2. Both group sizes were considered large enough to produce fairly accurate estimates of population values. In the previous SCA analyses of this data set, weighting according to the group sizes was also not used. The eigenvalues of the sum of the correlation matrices were calculated to determine the number of components to retain, using the QDA measure. Only for the one, two and three-dimensional solution, the QDA value was acceptable (only for those three solutions the additional amount of explained variance was larger than the amount of variance explained by one variable). These three QDA values for the one, two and three-dimensional solution were 1.8, 13.0 and 14.4, respectively, indicating that two or three components should be retained. Although in the previous analyses of this data set, always two components were retained, in the present analysis, both the two and three-dimensional solutions were inspected.

For both SCA-P and SCA-S, first the two dimensional solution was calculated. The solutions, found with SCA-P and SCA-S, were obtained by the same rotations as in the previous section. Also for the two-dimensional solution, the mean percentage of explained variance of two separate PCA's for the two groups was calculated. The percentage of explained variance for the two-dimensional solution was 44.5% for the separate PCA's, 43.5% for SCA-P and 40.3% for SCA-S. The difference between SCA-P and SCA-S is 3.2%, which is considered not negligible. The congruences of the columns of the pattern matrix found with SCA-P and the corresponding columns of the structure matrix found with SCA-S were .97 and .91 for the two components, respectively. Because the sample sizes were very large, the fact that SCA-P and SCA-S lead to virtually the same interpretations of the components suggests that the samples were drawn from the same population. The loadings of the pattern matrix found with SCA-P, and the structure matrices for the two groups are presented in Table D.3.

The loadings (which equal the SCA-P weights), reported in the first two columns of Table D.3, differ very little from the weights found by Kiers and Ten Berge (1994a) with SCA-W. In fact, all SCA-P weights differ no more than .02 from the corresponding SCA-W weights reported by Kiers and Ten Berge (1994a), with the exception of the loadings for the Abortion issue, which differ .05 and .04 from the corresponding weights from SCA-W, for the two components, respectively. Therefore, the two components can be labeled the same as was done by Kiers and Ten Berge (1994a). The first component is thus labeled as a measure for 'Conservatism' and the second component as a measure for 'Progressiveness'. The correlations between the components are equal (when rounded) to the ones found by Kiers and Ten Berge (1994a): respectively −.14 in the Italian sample and −.47 in the Danish sample. The highest difference in correlations between variables and components, compared to the solution found with SCA–W, is again found for the Abortion issue (still a negligible difference of .03). The main differences between the Danish and the Italian sample are that in the Danish sample the issues 'Control Multinationals' and 'Reduce Income Differences' are more strongly opposed by the conservatives than in the Italian sample and that the 'Women Decide About Abortion' issue is strongly opposed by the conservatives in the Italian sample, while this is not an issue for the conservatives in the Danish sample. For a more thorough inspection of the two dimensional solution, refer to Kiers and Ten Berge (1994a).

Table D.3: *Component pattern and structure matrices for the two dimensional solution, found by the method SCA–P, and the component strengths, in percentages of explained variance*

	Simultaneous Pattern		Structure in Danish sample		Structure in Italian sample	
component	1	2	1	2	1	2
INFL	.27	.27	.36	.18	.42	.42
EURO	.38	.06	.73	−.27	.69	.03
PUBL	.37	−.04	.70	−.43	.78	−.15
ABOR	−.17	.10	−.17	.22	−.58	.29
EMPL	.08	.36	−.03	.45	.00	.64
SUPP	−.08	.24	−.34	.46	−.20	.45
TERR	.41	.04	.72	−.25	.82	−.08
NUCL	.40	−.03	.78	−.40	.80	−.15
MULT	−.07	.38	−.52	.71	−.12	.63
ENVM	−.00	.41	−.38	.65	−.04	.71
REGI	.10	.43	−.07	.58	.04	.74
EQUA	−.08	.36	−.40	.69	−.27	.61
MILT	.38	−.05	.77	−.36	.73	−.23
INCO	−.15	.31	−.61	.64	−.26	.56
THRD	.30	.00	.62	−.33	.52	−.01
Percentage of explained variance			29.0	22.6	26.0	20.8

As a next step, for both SCA–P and SCA–S the three dimensional solution was calculated. The solutions, found with SCA–P and SCA–S, were again obtained in the same way as in the previous section. The percentage of explained variance for the three–dimensional solution was 52.5% for separate PCA's, 50.3% for SCA–P and 49.3% for SCA–S. The difference between SCA–P and SCA–S is now only 1.0% (compared to 3.2% for the two–dimensional solutions), which is considered negligible. The congruences of the columns of the pattern matrix found with SCA–P and the corresponding columns of the structure matrix found with SCA–S were .95, .84 and .67 for the three components, respectively. This suggested the samples were drawn from different populations. The loadings of the pattern matrix found with SCA–P, and the structure matrices for the two groups are presented in Table D.4.

From Table D.4 it can be seen that the loadings for the first component differ .06 or less from the loadings for the first component in the two–dimensional solution. The variable–component correlations in the two groups never differ more than .05 with the correlations found in the two–dimensional solution. Thus, the first component can again be labelled as a measure for 'Conservatism'.

The second component, however, is not equal to the second component from the two–dimensional solution. The component, labeled as a measure for 'Progressiveness' in the two–dimensional solution, is split into two components measuring different aspects of 'Progressiveness' in the three–dimensional solution. The first 'Progressiveness' component can be attached the label 'striving for equality of all individuals', as can be deduced from the large loadings (>.30) for the issues 'Control Multinationals', 'Reduce Regional Differences', 'Equal Opportunity for Men and Women' and 'Reduce Income Differences'. One could also label this component as a measure for 'Socialism'. Apparently, agreeing with these issues seems to coincide with support for the 'Protect the Environment' issue. The second 'Progressiveness' component is harder to interpret, because the large loadings (>.30) for the issues 'Women Decide About Abortion', 'Fight unemployment', 'Defense Against Superpowers' and 'Concern for Own Needs Versus needs of Third World' are not easily put under a common denominator.

The two 'Progressiveness' components overlap considerably, as can be seen from their positive correlation (.30 in the Danish sample and .50 in the Italian sample) against negative correlations with the 'Conservatism' component (−.48 and −.21 in the Danish sample and −.05 and −.33 in the Italian sample for the first and second 'Progressiveness' components, respectively).

The main differences with respect to the structure matrices between the Danish and the Italian sample (besides the same differences on the 'Conservatism' component as in the two-dimensional solution) are that, for the 'Progressiveness and striving for equality of all individuals' component, the issues 'Reduce Public Control', 'Develop Nuclear Energy', 'Increase Military Expenditure' and 'Concern for Own Needs Versus needs

Table D.4: *Component pattern and structure matrices for the three dimensional solution, found by the method SCA−P, and the component strengths, in percentages of explained variance*

	Simultaneous Pattern			Structure in Danish sample			Structure in Italian sample		
component	1	2	3	1	2	3	1	2	3
INFL	.30	.17	.25	.39	.07	.29	.45	.40	.20
EURO	.37	.10	−.12	.72	−.26	−.29	.69	.07	−.27
PUBL	.38	−.09	.05	.71	−.50	−.10	.77	−.13	−.30
ABOR	−.11	−.14	.57	−.13	.05	.61	−.55	.16	.69
EMPL	.13	.19	.43	.01	.32	.53	.05	.56	.61
SUPP	−.02	.05	.50	−.30	.32	.64	−.16	.33	.64
TERR	.41	.01	.00	.73	−.30	−.11	.82	−.06	−.29
NUCL	.39	.01	−.15	.77	−.39	−.33	.79	−.10	−.42
MULT	−.06	.41	.03	−.51	.74	.27	−.08	.63	.38
ENVM	−.00	.46	−.04	−.36	.69	.19	−.01	.74	.31
REGI	.10	.50	−.07	−.07	.64	.02	.07	.77	.34
EQUA	−.06	.34	.13	−.38	.68	.37	−.24	.60	.44
MILT	.36	−.00	−.16	.76	−.33	−.38	.71	−.20	−.40
INCO	−.14	.36	−.02	−.60	.68	.24	−.24	.58	.33
THRD	.33	−.15	.30	.63	−.45	.10	.54	−.08	.08
Percentage of explained variance				28.5	23.0	12.1	25.6	19.6	16.9

of Third World' are more strongly opposed by the progressive party activists in the Danish sample than in the Italian sample and that the 'Fight unemployment' and the 'Reduce Income Differences' issues are more strongly favored by the progressive party activists in the Italian sample than by the progressive party activists in the Danish sample. The only interesting difference on the third component between the Danish and the Italian sample is that the issue 'Reduce Regional Differences' is of more importance for the progressive party activists in the Italian sample than in the Danish sample, probably because there *are* larger regional differences in Italy than in Denmark.

The two examples presented above demonstrate the application of the simultaneous components analysis method SCA–P in practice. Compared to separate analyses of the groups in a data set, the simultaneous analysis has several advantages. The common components found are easily interpreted using the pattern matrix (which after HKIC rotation is equal to the weights matrix), while possible (small) differences are discovered by looking at the structure matrices found for each group in the simultaneous analysis separately. Thus, the method SCA–P offers more insight into what is going on in a data set than performing separate component or factor analyses on the different groups. In the latter case, it is, for instance, possible that some of the components obtained from most groups are present in the other groups as well, but are too weak to emerge among the first principal components drawn from those groups (see Section 1.4). In SCA–P this problem is circumvented in an elegant way.