CHAPTER 7

EXPERIMENT 3, COMPARISON OF THE SCA-METHODS, WHEN THE STRENGTH OF FACTORS WITH THE SAME INTERPRETATION DIFFERS ACROSS GROUPS

7.1 INTRODUCTION

In the third experiment, conducted in this study, the objective was to differentiate between the three SCA-methods SCA-W, SCA-P and SCA-S, when the factors differ in strength across groups, while maintaining the same interpretation. That is, in each group, the same variables pertain to a factor, and within each group all variables pertaining to a factor have the same true loading, but the sizes of the loadings for the variables pertaining to a certain factor differ across groups.

For the comparison, three conditions of the independent variable PatternComb were chosen for two groups (Experiment 3a), and one condition of PatternComb was chosen for three groups (Experiment 3b), to simulate, for instance, a longitudinal development, where some factors grow stronger and some factors grow weaker over time. In the analysis of the results, a new success criterion (described in Section 7.3.2) was used in addition to those used in the previous experiments.

7.2 MANIPULATIONS OF THE DATA

The following values were chosen for the independent variables. In the simulated data sets in Experiment 3a there were two samples (p=2), in Experiment 3b there were three samples (p=3). In both experiments there were 12 or 24 variables (=m), 2 or 4 factors (=q) and sample sizes of n=50, n=100 or n=150. The true pattern matrices ($\mathbf{P}_b$, $\mathbf{P}_d$, $\mathbf{P}_e$, $\mathbf{P}_m$, $\mathbf{P}_n$, $\mathbf{P}_p$ and $\mathbf{P}_q$) used in Experiments 3a and 3b are given in Appendix A. There were three conditions of PatternComb in Experiment 3a and one in Experiment 3b, and, in both experiments there were two different conditions of

PhiComb ($\mathbf{\Phi}_t$=**I** and $\mathbf{\Phi}_t$=$\mathbf{\Phi}_{.4}$; see Section 4.1.2). Hence, in total there were (3×2×2×2×3=) 72 different conditions in Experiment 3a and (1×2×2×2×3=) 24 different conditions in Experiment 3b to be analyzed. In each condition, ten replications of simulated data were used. The methods of analysis used were SCA–W, SCA–P, SCA–S.

The conditions of PatternComb were chosen such that the strength of the factors differs across groups. In the conditions using pattern combination '$\mathbf{P}_e\mathbf{P}_m$' there are both strong and weak factors present in each group. The strong factors in the first group are weak in the second, and the weak factors in the first group are strong in the second group. In condition '$\mathbf{P}_d\mathbf{P}_n$' the same goes for strong factors and intermediate factors, and in condition '$\mathbf{P}_p\mathbf{P}_q$' the same goes for intermediate factors and weak factors.

In Experiment 3b, there is one PatternComb condition: '$\mathbf{P}_e\mathbf{P}_b\mathbf{P}_m$'. The patterns reflect the longitudinal process simulated by this condition of PatternComb: In the first group ('$\mathbf{P}_e$'), the first half of the factors is strong and the second half of the factors is weak (see Appendix A); In the second group ('$\mathbf{P}_b$'), all factors are of intermediate strength; Finally, in the third group ('$\mathbf{P}_m$'), the first half of the factors is weak and the second half of the factors is strong. Looking at these patterns from a longitudinal (developmental) point of view, the first half of the factors has grown less important (less strong) over time, while the second half of the factors has gained importance.

SIFASP–ML was not included in this experiment. This was done, because (in factor analysis) the strength of factors is usually determined by looking at the size of the estimated loadings. For SIFASP–ML (as defined by Jöreskog, 1971), however, the loadings are kept equal across groups, making this approach pointless. Using the percentage of explained variance of each factor (as is done in component analysis) was considered an approach too far removed from what is usual in SIFASP. For these reasons, it was decided that SIFASP–ML was not a suitable method for the conditions used in the present experiment.

7.3 DEPENDENT VARIABLES

For each data set of two groups, the amount (percentage) of variance explained was calculated for 1, 2, 3, 4, 5 and 6 components drawn. As in Experiments 1 and 2, in the 4 factor condition, the explained variance for 7 components drawn was also calculated.

Because of the results of Experiment 1, only the measure QDA was used for dimension indication. For each solution with the correct dimension, the measures RR and DFC were calculated (see Section 4.4.1. and Section 4.4.2).

7.3.1 A problem with the method SCA-S in some of the conditions studied

In the present experiment, where the effect of different strengths of factors within and across groups is manipulated, it is crucial that the true factors have the same interpretation across groups, because otherwise factor strengths are not comparable. In the PatternComb conditions, used in the present experiment, the nonzero loadings differed both within and across groups, causing that for all samples in the condition where factors are correlated (PhiComb condition '$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$', used in half of the data sets studied), the true pattern matrices had perfectly congruent columns, while the columns of the true structure matrices were far from perfectly congruent. (In contrast, when PhiComb condition '**I I**' is used, the columns of the true structure matrices are perfectly congruent as well). This condition is similar to the conditions studied for two-or-more-populations data in Experiment 1. In that experiment, when factors in two-or-more-populations data were correlated, the columns of the true pattern matrices also had higher congruence values than the columns of the true structure matrices. However, in the present experiment, factors differ in strengths within and across groups as well, which causes a serious problem for SCA-S in PatternComb condition '$\mathbf{P}_e\mathbf{P}_m$' in Experiment 3a and, to a lesser degree in PatternComb condition '$\mathbf{P}_e\mathbf{P}_b\mathbf{P}_m$' in Experiment 3b. In PatternComb condition '$\mathbf{P}_e\mathbf{P}_m$', the factors in one group have, in terms of the true structure matrices, a congruence value with the factors in the other group, with which they should share the same interpretation, of .82 when there are 2 factors and of .77 when there are four factors present. However, the congruence value of the factors in one group with the *other* factors in the other group is .99 (!) when there are 2 factors and .89 when there are four factors present. In this condition we expect SCA–S to encounter serious difficulties in recovering the correct factors, in terms of the true structure matrices. In the other two PatternComb conditions in Experiment 3a, the factors, in terms of the true structure matrices, in one group do have the highest congruence value with the factors in the other group, with which they should share the same interpretation, so there should be no problem for SCA–S in those conditions. Therefore, SCA–S, in which the structure matrices are forced to be columnwise proportional, is not a proper method for describing data in the combination of PatternComb condition '$\mathbf{P}_e\mathbf{P}_m$' and PhiComb condition '$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$'. The same problem exists in PatternComb condition '$\mathbf{P}_e\mathbf{P}_b\mathbf{P}_m$', used in Experiment 3b, where the same two true pattern matrices are present. However, in PatternComb condition '$\mathbf{P}_e\mathbf{P}_b\mathbf{P}_m$', a third true pattern is present, because there are three groups in that experiment. In PhiComb condition '$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$' within this PatternComb condition, the factors, based on the true structure matrices for true pattern matrix '$\mathbf{P}_b$', have the highest congruence value with the factors in the two other groups (with true pattern matrices '$\mathbf{P}_e$' and '$\mathbf{P}_m$', respectively), with which they should share the same interpretation, so it is expected that SCA–S will give better results in this condition than it will in the combination of PatternComb condition '$\mathbf{P}_e\mathbf{P}_m$' and PhiComb condition '$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$' in Experiment 3a, but still will have trouble in retrieving the correct true structure for all three groups.

In practice, however, the true data structure is not known, and there is no a priori reason not to apply SCA–S. Therefore, in the present experiments (3a and 3b), we do apply SCA–S and inspect the recovery of the structure matrices. For comparative purposes, we also inspect the recovery of the pattern matrices for SCA–S in the problematic conditions mentioned above, because, for SCA–S, the pattern matrices are allowed to differ across groups.

7.3.2 A measure used for inspecting whether the relative strength of the recovered factors resembles the relative strength of the true factors: the Quotient of Component Strengths (QCS)

Besides the question whether the components, found with the SCA-methods, receive the same interpretation as the underlying factors, the question can be asked whether or not the relative strength of the components found is the same as the relative strength of the true factors. To answer this question, a new measure was devised as follows.

The true strength of a factor in the simulated population can be directly derived from the true structure matrix $\mathbf{S}_t$ by taking the sums of squares of the column of this matrix pertaining to that factor. From the description of the loadings in the patterns, used in this experiment (see Appendix A), it can be seen that in each PatternComb condition, there were always factors with two different (true) strengths. We take, for example, the condition where there are two factors present in each group, but of course the same applies in the condition where there are four factors in each group. The quotient, calculated as the strength of the weaker factor divided by the strength of the stronger factor, gives the strength ratio of the true factors in the population.

In the present experiments, for each sample, the strengths of the components, found with each method, were calculated from the structure matrices, found with each method. Next, the strengths of the components matching the weaker true factors are averaged and divided by the average strength of the components, matching the stronger true factors. The strength ratio found with each of the SCA-methods is compared to the strength ratio in the population. Because it is the strength ratio of the *components* found with the SCA-methods that is of interest, this measure is called the Quotient of Component Strengths (QCS). To avoid confusion, the strength ratio of the factors in the population is referred to as the "true QCS" (instead of quotient of factor strengths).

Firstly, it was checked whether or not the QCS, found with each method, was above or below the true QCS. Secondly, the absolute differences of the QCS and the true QCS were compared. Thirdly, as a combination of the RR measure and the QCS measure, the absolute differences of the QCS and the true QCS were compared when only those cases were included in which all factors were retrieved. That is, all data sets for which the RR is not 100%, were left out of the analysis, because it seems a useless course of action to inspect whether components that do not receive the correct interpretation do have the same relative strengths as the correct factors.

7.4 ANALYSIS

All the measures presented were analyzed with repeated measures analyses of variance. Sums of squares of main and interaction effects were calculated to give an impression of the size of the effects, and the associated averaged univariate F tests were employed for significance testing. In each of the separate analyses, only first and second order interactions among between subjects factors were included in the analysis. All precautions and adjustments, as presented in Section 5.4, were applied.

For the analysis of the results for the dimension indicator QDA, PCA–sep was included as a condition of the within subjects factor, so in that case the within subjects factor had 4 conditions. Outcomes concerning correct dimension indication were aggregated over the ten replications within each condition to obtain a reasonably continuous variable, in the same way as in Experiment 1.

7.5 QUESTIONS TO BE ANSWERED BY THIS EXPERIMENT

The present experiment was conducted to get answers as to what circumstances play a role in the success and failure of the SCA–methods in the recovery of factors and their relative strengths, in conditions where factors have considerably different strengths across and within groups. Answers will be given in separate subsections of Section 7.6 (Experiment 3a) and Section 7.7 (Experiment 3b), where the question numbers refer to the subsections in which answers to these questions will be given. Summarizing conclusions from the results of this experiment will be drawn in Section 7.8.

1a.) Which of the SCA-methods has the highest success rate on the dimension indicator QDA when factors differ in strength over samples?
1b.) Are there interesting interaction effects between independent variables and method of analysis for the success rate of the dimension indicator QDA?
2a.) Which of the SCA-methods can be judged as best capable of recovering the correct interpretation of factors when factors differ in strength over samples?
2b.) Are there interesting interaction effects between independent variables and method of analysis for the recovery of factors?
3a.) Which of the SCA-methods can be judged as best capable of recovering the relative strength of the factors when factors differ in strength over samples?
3b.) Are there interesting interaction effects between independent variables and method of analysis for the recovery of the relative strength of factors?
4a.) Which of the SCA-methods can be judged as best capable of recovering the correlations between factors when factors differ in strength over samples?
4b.) Are there interesting interaction effects between independent variables and method of analysis for the recovery of correlations between factors?

7.6 RESULTS EXPERIMENT 3A

In the present experiment, a significance level α=.001 was used.

Whenever differences between dependent variables or methods are reported, or when interaction effects are said to occur, effects were significant at $\alpha=.001$. In the following sections, each of the four main questions, asked at the end of Section 7.5, will be answered separately. At the start of each section, the questions will be repeated.

7.6.1 Finding the SCA-method with the best dimension indication

Question 1a: Which of the SCA-methods has the highest success rate on the dimension indicator QDA when factors differ in strength over samples?
Question 1b: Are there interesting interaction effects between independent variables and method of analysis for the success rate of the dimension indicator QDA?

As mentioned above, the dimension indicator QDA, found to be the dimension indicator with the highest percentage of correct dimension indication in Experiment 1, was used here. A brief comparison of the present results with all dimension indicators used in Experiment 1 (KA1, PA, QA and QDA), led to the same conclusions as in Experiment 1, thus justifying the decision to use only the dimension indicator QDA.

In Figure 7.1, the percentages of correct dimension indication by the dimension indicator QDA for the three SCA-methods are presented. Overall, SCA-P had the highest percentage of correct dimension indication (84.0%), followed by SCA-S (82.4%) and SCA-W (70.7%). The difference between SCA-P and SCA-S was not significant ($p=.149$), but the difference between either of these methods and SCA-W was. From this it may be clear that there was a large effect of Method. In fact, it explained 20% of the total within subjects variation. In contrast, in Experiment 1, SCA-P also had the highest percentage of correct dimension indication, but SCA-W followed very closely and SCA-S fell behind.

There were substantial and significant interactions of PatternComb $\times$ Method, PhiComb $\times$ Method, and NFac $\times$ Method, together explaining a further 34% of the total within subjects variation. The form of these interactions can be read from Figure 7.1. Specifically, it was found that

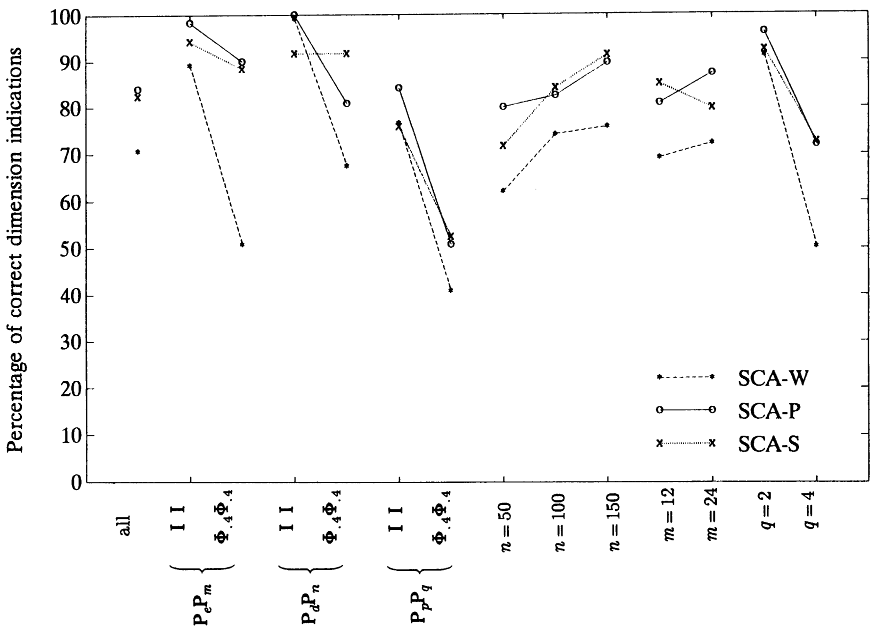

Figure 7.1 *Percentages of correct dimension indications for the three SCA–methods, averaged over (combinations of) the levels of the independent variables, using the dimension indicator QDA*

SCA–W had a relatively low percentage of correct dimension indication in PatternComb condition '$\mathbf{P}_e\mathbf{P}_m$', where factors are strong and weak in the two groups; in PhiComb condition '$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$', where factors are strongly oblique; and when there are four factors present instead of two. The most interesting difference with the results in Experiment 1 is that here it is SCA–W that mainly causes the interactions, whereas in Experiment 1, SCA–S caused interactions.

Answer 1a: The methods SCA–P and SCA–S have comparable success rates on the dimension indicator QDA when factors differ in strength over samples. SCA–W clearly falls behind.

Answer 1b: SCA–W especially falls behind the other methods when there are both strong and weak factors present (where the strong and weak factors trade place over groups), when factors are strongly oblique, and when there are four factors present instead of two.

7.6.2 Finding the SCA-method with the best recovery of factors

Question 2a: Which of the SCA-methods can be judged as best capable of recovering the correct interpretation of factors when factors differ in strength over samples?
Question 2b: Are there interesting interaction effects between independent variables and method of analysis for the recovery of factors?

Overall, recovery is high for all methods, and comparable to results found for similar conditions in Experiment 1. In Figure 7.2, the RR's are presented for all the conditions of the independent variables. For SCA-S the RR, based on both the pattern matrices and the structure matrix, are presented, for reasons explained in Section 7.3.1. First we will look at these two RR's for SCA-S. Overall, the RR for SCA-S based on the pattern matrices is higher (95.6%) than the RR based on the structure matrices (89.8%). However, as can be seen from Figure 7.2, only in the combination of PatternComb condition '$\mathbf{P}_e\mathbf{P}_m$' and PhiComb condition '$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$', the RR, based on the structure matrix, is lower than the RR, based on the pattern matrix. This was the very condition where we expected SCA-S to perform poorly in terms of recovery of the structure matrix. In all other conditions the structure matrix is to be preferred, and will therefore be employed in the sequel. We will now compare the three SCA-methods. SCA-P has the highest RR (97.3%), followed by SCA-W (93.0%) and SCA-S, using the structure matrix (89.8%). The differences between the three SCA-methods were all significant. There was a significant effect of Method, and there were significant and substantial interactions of PatternComb × Method, PhiComb × Method, and a significant and substantial three-way interaction of PatternComb × PhiComb × Method, together explaining 64% of the total within subjects variation. The deviant behavior of the RR for SCA-S in PatternComb condition '$\mathbf{P}_e\mathbf{P}_m$'/PhiComb

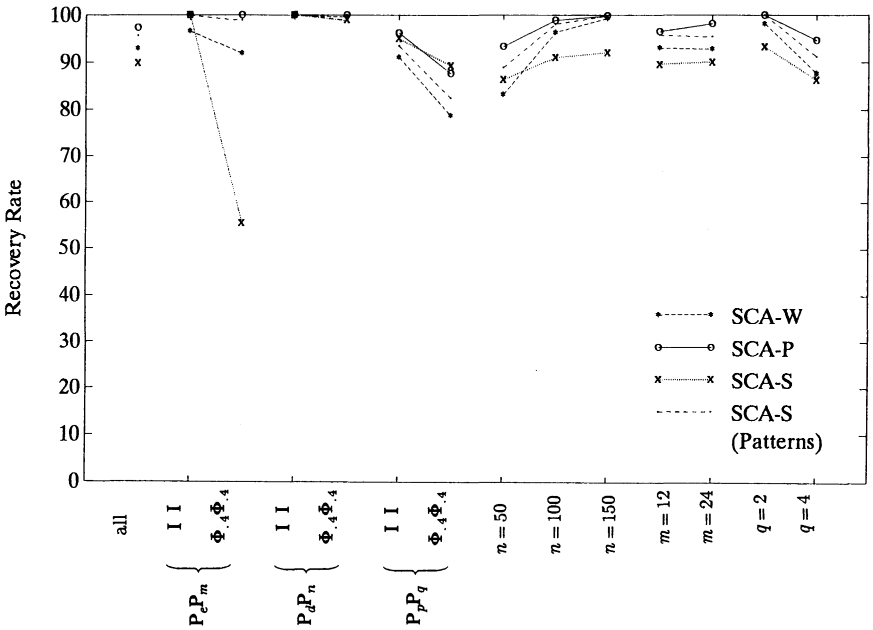

Figure 7.2 *Recovery Rates for the three SCA–methods (and a second time for the method SCA–S, when based on the pattern matrix), averaged over (combinations of) the levels of the independent variables*

condition '$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$' clearly accounts for the largest of the interactions. SCA–W especially has lower RR's than SCA–P in PatternComb conditions '$\mathbf{P}_e\mathbf{P}_m$' and '$\mathbf{P}_p\mathbf{P}_q$' (the conditions with weak factors).

Comparing the results for the three conditions of PatternComb, with those in similar PatternComb conditions in Experiment 1, we find that RR's are higher for all methods in PatternComb condition '$\mathbf{P}_e\mathbf{P}_m$' than in condition '$\mathbf{P}_e\mathbf{P}_e$', especially for SCA–P and SCA–S, with the exception of PatternComb condition '$\mathbf{P}_e\mathbf{P}_m$'/PhiComb condition '$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$' for SCA–S. RR's in PatternComb condition '$\mathbf{P}_d\mathbf{P}_n$' closely match those in condition '$\mathbf{P}_d\mathbf{P}_d$' (RR's close or equal to 100%). The RR's in PatternComb condition '$\mathbf{P}_p\mathbf{P}_q$', finally, take a position in between those in PatternComb conditions '$\mathbf{P}_b\mathbf{P}_b$' and '$\mathbf{P}_c\mathbf{P}_c$', as was expected, considering that the average loadings in PatternComb condition '$\mathbf{P}_p\mathbf{P}_q$' are also in between those in PatternComb conditions '$\mathbf{P}_b\mathbf{P}_b$' and '$\mathbf{P}_c\mathbf{P}_c$'.

Answer 2a: The methods SCA-P and SCA-S are best capable of recovering factors when factors differ in strength across samples, while maintaining the same interpretation, with the exception for SCA-S in the situation where there are strong and weak factors in both samples that change place over samples and are strongly oblique.
Answer 2b: SCA-W especially falls behind the other methods when there are weak factors present.

7.6.3 Finding the SCA-method with the best recovery of the relative strength of factors

Question 3a: Which of the SCA-methods can be judged as best capable of recovering the relative strength of the factors when factors differ in strength over samples?
Question 3b: Are there interesting interaction effects between independent variables and method of analysis for the recovery of the relative strength of factors?

In Table 7.1, the mean over and underestimation of the true QCS (see Section 7.3.2) is given for each method, together with the number of times over and underestimation occurred.

Table 7.1: *Over and underestimation and mean absolute difference from the true QCS by the SCA-methods*

	SCA-W(cases)	SCA-P(cases)	SCA-S(cases)
Mean overestimation	.16 (674)	.16 (676)	.13 (398)
Mean underestimation	.03 (46)	.03 (44)	.12 (322)
Mean abs. diff.	.15 (720)	.15 (720)	.13 (720)

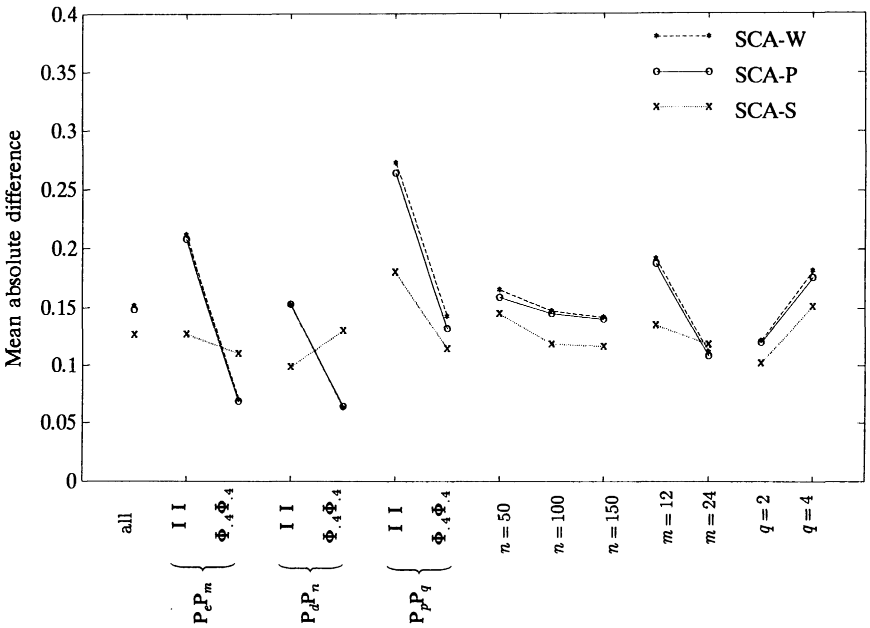

Figure 7.3 *Mean absolute differences of the QCS for each SCA–method and the true QCS, averaged over (combinations of) the levels of the independent variables*

It can be seen from Table 7.1 that SCA–W and SCA–P generally overestimate the QCS. That is, these methods tend to overestimate the relative strength of the weaker components found. SCA–S seems to display no such bias.

In Figure 7.3, the mean absolute differences between the QCS, found with each of the SCA–methods, and the true QCS are presented, when averaged over each condition of each independent variable. There was a significant effect of Method, explaining only 4% of the total within subjects variation, the QCS for SCA–S being the closest to the true QCS. There were substantial and significant interactions of PatternComb × Method and NVar × Method (see Figure 7.3), together explaining 11% of the total within subjects variation. The largest effects were the interaction of PhiComb × Method and the three-way interaction of PhiComb × NVar × Method, both explaining 21% of the total within subjects variation. The interaction of PhiComb × Method comes from the fact that while the QCS for SCA-W and SCA-P is larger when factors are uncorrelated (.21 for both methods) than when they are correlated (.09 for both methods), for SCA-S there is little difference on the average QCS for uncorrelated and correlated factors (.14 and .12, respectively). The three-way interaction of PhiComb × NVar × Method is described as follows. In PhiComb condition '**I I**', the QCS for *all methods* is smaller when there are 24 variables than when there are 12 variables present (.26 for SCA-W and SCA-P, and .20 for SCA-S at 12 variables, against .16 for SCA-W and SCA-P, and .07 for SCA-S at 24 variables). In PhiComb condition '$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$', the QCS for SCA-W and SCA-P is smaller when there are 24 variables than when there are 12 variables present (.12 at 12 variables against .06 at 24 variables), but the QCS for SCA-S is larger when there are 24 variables than when there are 12 variables present (.07 at 12 variables against .17 at 24 variables).

All cases are included in Table 7.1, but not for all cases all factors were recovered. This might contaminate the results for the QCS measure. Therefore, the mean absolute difference of the value of the QCS measure for each of the SCA-methods and the true QCS value, for those cases for which each of the SCA-methods had a RR of 100%, was inspected. SCA-W recovered all factors 577 times out of 720 (80.1%), SCA-P 681 times (94.5%) and SCA-S 544 times (75.6%, in terms of the structure matrix). In 485 cases, all three SCA-methods recovered all factors. In the combination of PatternComb condition '$\mathbf{P}_e\mathbf{P}_m$' and PhiComb condition '$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$', SCA-S had a RR of 100% in only 2 cases out of 120, so this condition is almost never present in this selection. For all other conditions, results led to practically the same conclusions as when all cases were included in the analysis. Mean absolute differences for the QCS measure were .16 (SCA-W), .16 (SCA-P) and .12 (SCA-S).

Answer 3a: SCA-S is most successful in correctly recovering the absolute and relative strength of factors when factors differ in strength over samples, followed by both SCA–P and SCA–W. However, there are strong and interesting interactions between independent variables and method of analysis.

Answer 3b: When factors are uncorrelated, SCA–S is best at recovering absolute and relative strength of factors, but while higher correlations between the factors and the presence of a larger number of variables strongly enhances the recovery of relative factor strength for SCA–P and SCA–W, this has only a small positive ('$\mathbf{P}_p\mathbf{P}_q$') or negative ('$\mathbf{P}_e\mathbf{P}_m$' and '$\mathbf{P}_d\mathbf{P}_n$') effect for SCA–S.

7.6.4 Finding the SCA-method with the best recovery of correlations between factors

Question 4a: Which of the SCA–methods can be judged as best capable of recovering the correlations between factors when factors differ in strength over samples?

Question 4b: Are there interesting interaction effects between independent variables and method of analysis for the recovery of correlations between factors?

The mean DFC values for each condition of each independent variable are presented in Figure 7.4. Compared to Experiment 1 (condition 'samples from one population'), overall mean DFC values are a little higher for SCA–P and SCA–W and a little lower for SCA–S. As in Experiment 1, the strongest interaction was that of PhiComb × Method, explaining 32% of the total within subjects variation. When factors are uncorrelated, SCA–S has the smallest DFC value (.071 versus .084 for SCA–W and .090 for SCA–P). When factors are correlated, SCA–S has the largest DFC value (.21 versus .15 for SCA–W and .14 for SCA–P).

A substantial effect was found for Method, in contrast to what was found in Experiment 1 in the condition 'samples from one population'. SCA–S had a higher DFC value than SCA–P and SCA–W. Other substantial effects, that were also not found in Experiment 1 in the condition 'samples from one population', were the interaction of PatternComb ×

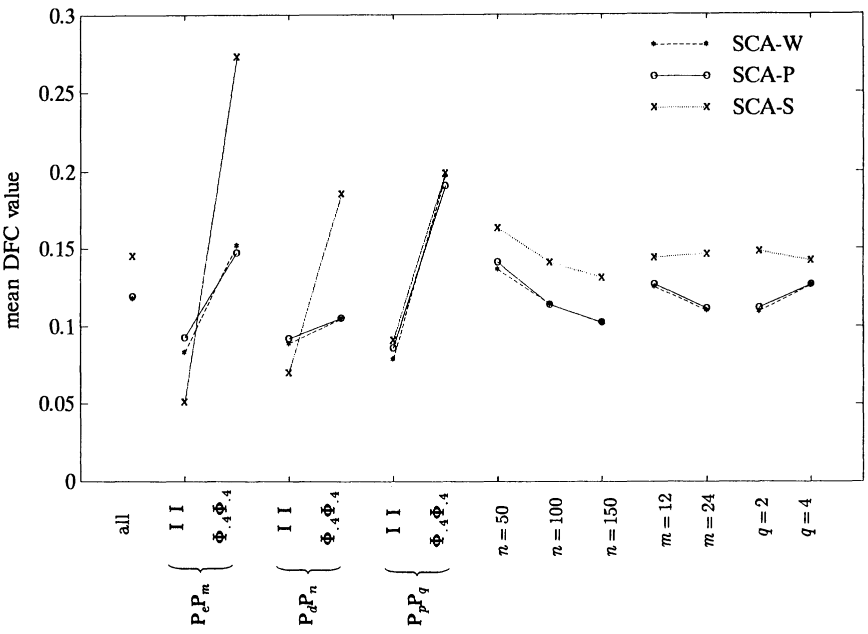

Figure 7.4 *Mean DFC values for the SCA-methods, averaged over (combinations of) the levels of the independent variables*

PhiComb × Method, and the interaction of PhiComb × NFac × Method, together with the Method effect explaining 43% of the total within subjects variation. Within PatternComb condition '$\mathbf{P}_p\mathbf{P}_q$', there was little difference between the mean DFC values of the three SCA-methods in PhiComb conditions '**I I**' and in PhiComb condition '$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$' (see Figure 7.4), while within the other PatternComb conditions, especially in '$\mathbf{P}_e\mathbf{P}_m$', SCA-S had a smaller mean DFC value than the other SCA-methods in PhiComb condition '**I I**' (.05 vs. .08 for SCA-W and .09 for SCA-P), but higher DFC values than the other SCA-methods in PhiComb condition '$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$' (.27 vs. .15 for both SCA-W and SCA-P). Within PhiComb condition '**I I**', SCA-W and SCA-P had larger DFC values when there were four factors present (average .09 and .10, respectively) than when there were two factors present (average .08 and .09) and SCA-S had smaller DFC values when there were two factors present (average .06) than when there were four factors present (average .08). Within PhiComb condition '$\Phi_{.4}\Phi_{.4}$', SCA–S gave high mean DFC values both when there were two factors (.24) and when there were four factors (.20), while SCA–W and SCA–P had smaller mean DFC values when there were two factors present (average .13 for both methods) than when there were four factors present (average .17 for both methods).

Answers 4a and 4b: The SCA–methods are about equally good in recovering the correlations between factors when there are only intermediate and weak factors present, and when the factors are uncorrelated. When there are strong factors present and when the factors are correlated, SCA–P and SCA–W are better able to recover the correlations between factors than SCA–S.

7.7 RESULTS EXPERIMENT 3B

In the present experiment, the same approach towards significance is taken as in previous experiments. In the following sections, the four main questions, asked in Section 7.5, will be answered separately.

7.7.1 Finding the SCA-method with the best dimension indication

SCA–P has the highest percentage of correct dimension indication with the measure QDA (89.6%), followed by SCA–S (85%) and SCA–W (79.2%). The results found lead to the same conclusions as those in Experiment 3a.

7.7.2 Finding the SCA-method with the best recovery of factors

In Table 7.2, the RR's are presented for the three SCA–methods, and a second time for SCA–S, in terms of the pattern matrices.

SCA–P has complete recovery, closely followed by SCA–W. The RR for SCA–S, in terms of the structure matrix, falls behind considerably. As in

Table 7.2: *Recovery rates for the three SCA-methods and for the method SCA-S, when based on the pattern matrices*

	SCA–W	SCA–P	SCA–S structure	SCA–S pattern	n
$\mathbf{P}_e\mathbf{P}_b\mathbf{P}_m$	97.8	100.0	85.9	99.1	240
I I	98.6	100.0	100.0	99.8	120
$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$	96.9	100.0	71.9	98.4	120
n=50	93.9	100.0	86.4	97.3	80
n=100	99.7	100.0	85.6	100.0	80
n=150	99.8	100.0	85.8	100.0	80
12 var	98.3	100.0	86.2	99.7	120
24 var	97.3	100.0	85.7	98.5	120
2 fact	99.7	100.0	88.2	99.7	120
4 fact	95.8	100.0	83.7	98.5	120

Experiment 3a, this only happens when the factors are correlated ('$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$'). A closer look at the RR's for each of the three groups ('$\mathbf{P}_e$', '$\mathbf{P}_b$' and '$\mathbf{P}_m$') separately, shows all factors in group '$\mathbf{P}_b$' are recovered, against an average of 54.4% in groups '$\mathbf{P}_e$' and '$\mathbf{P}_m$' (this is about the same percentage as found in Experiment 3a in PatternComb condition '$\mathbf{P}_e\mathbf{P}_m$'/PhiComb condition '$\mathbf{\Phi}_{.4}\mathbf{\Phi}_{.4}$'). The RR's for SCA–S, in terms of the pattern matrices, are very high, taking a position in between SCA–P and SCA–W. These results support the findings in Experiment 3a.

7.7.3 Finding the SCA-method with the best recovery of the relative strength of factors

The major difference of the present experiment with Experiment 3a is that in the present experiment we have three groups instead of two. In each group there were factors of two different strengths, and over the three groups, factors with the same interpretation had three different strengths: weak, intermediate and strong. This meant that there were three measures available for the relative strength. We compared strengths of the weak and the strong factors, the intermediate and the strong factors, and the weak and the intermediate factors. Mean over- and underestimation, as well as the mean absolute differences for the factors, found with the SCA-methods, for each of the three QCS measures, are given in Table 7.3.

From Table 7.3, it can be seen that SCA-W and SCA-P behave the same for this measure. Both methods overestimated the relative strengths of the weak factors and the intermediate factors, compared to the strong factors in 93% and 95% of the cases, respectively, and overestimate the relative strength of the weak factors compared to the intermediate factors in 70% of the cases. SCA-S underestimates the relative strength of the weak factors about just as often as it overestimates the relative strength of the weak factors. The comparison of the weak and the strong factors (upper section of Table 7.3) leads to the same conclusion as in Experiment 3a.

All cases are included in Table 7.3, but not for all cases all factors were recovered. This might contaminate the results for the QCS measure. Therefore, in Table 7.4, the mean absolute differences for the

Table 7.3: *Mean over- and underestimation, and mean absolute differences from the true QCS for the SCA-methods*

	SCA-W(cases)	SCA-P(cases)	SCA-S(cases)
Weak vs. strong factors			
Mean overestimation	.15 (224)	.14 (224)	.13 (126)
Mean underestimation	.02 (16)	.02 (16)	.11 (114)
Mean abs. diff.	.14 (240)	.14 (240)	.12 (240)
Intermediate vs. strong factors			
Mean overestimation	.13 (229)	.13 (229)	.10 (149)
Mean underestimation	.03 (11)	.03 (11)	.08 (91)
Mean abs. diff.	.13 (240)	.13 (240)	.09 (240)
Weak vs. intermediate factors			
Mean overestimation	.21 (167)	.21 (168)	.21 (120)
Mean underestimation	.12 (73)	.12 (72)	.22 (120)
Mean abs. diff.	.18 (240)	.18 (240)	.21 (240)

Table 7.4: *Mean absolute differences from the true QCS for the SCA-methods, including only the* 109 *cases in which all three SCA-methods had a RR of* 100%

	SCA–W	SCA–P	SCA–S
Weak vs. strong factors			
Mean abs. diff.	.20	.19	.13
Intermediate vs. strong factors			
Mean abs. diff.	.15	.15	.09
Weak vs. intermediate factors			
Mean abs. diff.	.24	.24	.21

components, found with the SCA–methods, for each of the three QCS measures are given, when including only those cases for which all three SCA–methods had a RR of 100%. This was the case in a total of 109 out of 240 cases (45%).

For SCA–W and SCA–P, the mean absolute differences in the cases where all factors are recovered are higher than when averaged over all cases, while for SCA–S there is hardly any difference. So, surprisingly, the results for SCA–W and SCA–P are actually worse when only looking at cases where factors were correctly recovered than when looking at all cases. This was not found in Experiment 3a.

7.7.4 Finding the SCA-method with the best recovery of correlations between factors

The mean DFC values for SCA–W and SCA–P were about equal (.1158 and .1157, respectively). SCA–S had a higher mean DFC value of .1697. Results for the DFC measure led to the same conclusions as in Experiment 3b.

7.8 CONCLUSION

When factors differ in strength across groups, while maintaining the same interpretation, SCA–P and SCA–S have comparable success rates on the dimension indicator QDA. SCA–W falls behind in some conditions, mentioned in Section 7.6.1.

The methods SCA–P and SCA–S are best capable of recovering factors, except for SCA–S in the special situation where there are strong and weak factors in the (two or more) groups that switch places over groups and are strongly oblique. In the latter case, recovery for SCA–S is exceptionally low. SCA–W performs relatively badly when there are weak factors present.

SCA–S is most successful in correctly recovering the absolute and relative strength of factors, followed by SCA–P and SCA–W, which give about equal results. It must be noted, however, that, in the exceptional condition mentioned above, SCA–S almost never correctly recovers all factors, devaluating the correct recovery of relative factor strength. High correlations between the factors and the presence of a larger number of variables enhances the recovery of relative factor strength for SCA–P and SCA–W, while it has little effect for SCA–S. In the three group condition used in Experiment 3b, SCA–W and SCA–P had a worse recovery of factors strength when averaged over cases with only correctly recovered factors than when averaged over all cases. This result was not found in similar two group conditions in Experiment 3a.

When there are strong factors present and when the factors are correlated, SCA–P and SCA–W are better able than SCA–S to correctly recover the correlations between factors. In the other conditions the SCA–methods perform about equally well.