CHAPTER 6

CONCLUSIONS

A first purpose of this chapter is to give some recommendations for the analysis of three-way data. In particular, an overview will be given of the methods to analyze three-way data that have been the subject of this study. It will be assumed that the data consist of scores of persons on variables at various occasions.
The second purpose of this chapter is to draw some conclusions. Specifically, degenerate components and uniqueness will be revisited.

6.1 Exploratory analysis of three-way data

After a preprocessing method is chosen, a three-way array can be analyzed. The hierarchical relations between (in order of increasing fit) SUMPCA, Weighted PCA, PFORTA, PARAFAC and TUCKALS3, can be used as follows. In case it is found that the fit of a more constrained variant is considerably lower than a less constrained variant, the more constrained variant is not appropriate to fit the data satisfactorily. For instance, in case the PARAFAC fit to the data is considerably lower than the TUCKALS3 fit, it can be concluded that PARAFAC is overly restrictive, and that certain interactions between the components, represented in the core array, are necessary to fit the data satisfactorily. For such a case one is referred to Kroonenberg (1983). In case the discrepancy in fit between PARAFAC and TUCKALS3 is negligible and the Weighted PCA fit is considerably less than the PARAFAC fit, one may represent the data by PARAFAC or by one of the constrained PARAFAC methods that have been the object of this study. On the other hand, in case Weighted PCA, PFORTA, PARAFAC and TUCKALS3 fit the data almost equally well, one may represent the data by Weighted PCA.

To illustrate the use of the above hierarchical relations, the Affective response data were analyzed by SUMPCA, Weighted PCA, PFORTA, PARAFAC and TUCKALS3 with dimensionality 2. The percentages of variance explained are: 8.0, 42.9, 44.3, 46.5, 46.5, respectively. In Chapter 2 and 3 it has been reported that PARAFAC yields a degenerate solution for these data. Two clear conclusions can be drawn. First, the SUMPCA method is overly restrictive, and second, there is no need to consider certain interactions between the components revealed in the TUCKALS3 core array. These fit values suggest to represent the Affective response data by Weighted PCA or by PFORTA, see section 3.6.

As noted above, in case the discrepancy in fit between PARAFAC and TUCKALS3 is negligible and the Weighted PCA fit is considerably lower than the PARAFAC fit, one may want to represent the data by a constrained PARAFAC method. A number of constrained PARAFAC methods, including those discussed in the present study, are listed in Table 6.1. The rows correspond to the methods, and the columns correspond to the parameter matrices. In each cell the constraint that is imposed on a parameter matrix is depicted. An empty cell means that the corresponding method is not constrained in terms of the corresponding parameter matrix.

Table 6.1 *A schematic representation of the methods discussed in this study.*

Methods	A	B	C
PFORTA	Orthonormal		
PFORTB		Orthonormal	
PFORTAB	Orthonormal	Orthonormal	
PARAFAC prop. col.			Two Proportional columns
SUMPCA			Equal rows
Weighted PCA			All columns proportional
PFNC	$X_k'A$ non-neg.	Non-negative	Non-negative
PFCV		$W*B=B$, W fixed	
PFOCV		$W*B=B$, W free	

Table 6.1 summarizes the constrained methods that have been mentioned in the previous chapters. It can be seen that four types of constraints were used: orthonormality (see PFORTA), lower rank (see Weighted PCA), non-negativity (see PFNC) and constraining elements to zero (see PFOCV). In some methods, more than one parameter matrix is subjected to one (type of) constraint, see, for example, PFORTAB or PFNC. By simply putting other combinations of constraints on the parameter matrices, still other PARAFAC methods with different properties can be developed. An example of this is PARAFAC subject to the constraints that A, B, and C are all non-negative, see Carroll, De Soete and Pruzansky (1989). One may also want to combine certain constraints on one parameter matrix. For example, in PFCV one may

subject B to the additional constraint that it has no negative elements. Another example of two different constraints, resting on one of the parameter matrices, is to subject A in PFNC to the additional constraint that it has orthonormal columns. Such constrained PARAFAC methods may turn out useful in the future.

6.2 Various approaches to cope with degenerate solutions

In Chapter 1 the degeneracy problem was described and Harshman and Lundy's (1984b, p. 274) solution was mentioned: Imposing a column-wise orthonormality constraint on one of the parameter matrices. In addition to these solutions, it has been demonstrated that Weighted PCA, PFNC, PFCV and PFOCV yield non-degenerate solutions as well.

6.3 Examining the degree of uniqueness by constrained PARAFAC

In chapter 1 the uniqueness of the PARAFAC components was introduced as the most salient property of the PARAFAC method. In chapter 2 it was explained how the degree of PARAFAC uniqueness can be examined by constrained PARAFAC. Specifically, it appeared that by analyzing samples from populations having increasing degrees of uniqueness, increasing amounts of discrepancy were encountered between the PARAFAC fit and the Weighted PCA fit. In chapters 3, 4, and 5 it has been found for various empirical data sets that the fit of constrained PARAFAC is close to that of unconstrained PARAFAC, and that the constrained components allow for an easier (and thus a different) interpretation. It is clear that the existence of such an alternative PARAFAC representation indicates that the uniqueness of the unconstrained PARAFAC components is weak. So each of the constrained PARAFAC methods can detect weak uniqueness of the unconstrained PARAFAC components.

6.4 Representing three-way data by unconstrained and constrained PARAFAC

In case of uniqueness in the population, PARAFAC retrieved the population parameters in the sample, and its components were stable in the splithalf sense, even with relatively small sample sizes. Therefore, it seems that, for empirical data where PARAFAC shows strong uniqueness, PARAFAC is a useful method for the exploratory analysis of three-way data.
On the other hand, in case of weak uniqueness in the sample, the PARAFAC components were unstable, and PARAFAC failed to retrieve the population parameters, and its components contained contrast, whereas the population showed positive manifold. Therefore, it seems that, for empirical data where PARAFAC shows weak uniqueness, PARAFAC is less useful. In case of weak uniqueness, PFNC did retrieve the population parameters and its components were stable in the splithalf sense. By the results of analyzing various empirical three-way arrays it was demonstrated that, in case of weak uniqueness, alternative (constrained) PARAFAC components exist, which fit the data almost equally well as PARAFAC, allow for an easier

interpretation, and are less unstable or even stable in the splithalf sense. It seems that for empirical data, where PARAFAC shows weak uniqueness, constrained PARAFAC is useful for the exploratory analysis of three–way data.