In H.G. Law, C.W. Snyder Jr., J.A. Hattie, & R.P. McDonald (Eds.), Research methods for multimode data analysis. (pp. 64–103). 1984

3

Three-Mode Principal Component Analysis: Illustrated with an Example from Attachment Theory

Pieter M. Kroonenberg

The three-mode principal component model—here referred to as the *Tucker3* model—was first formulated within the context of the behavioral sciences by Ledyard Tucker (1963). In subsequent articles, Tucker extended the mathematical description and its programming aspects (1964, 1966). In the context of multidimensional scaling, references to his model occur frequently (Carroll and Chang 1972; Takane, Young, and de Leeuw 1977; Jennrich 1972), as the Tucker3 is the general model comprising many other individual differences models. A discussion of the relation between multidimensional scaling and three-mode principal component analysis can be found in Tucker (1972), Carroll and Wish (1974), Takane, Young, and de Leeuw (1977), and Carroll and Arabie (1980). Other approaches to three-mode analysis include three-mode common factor analysis within the context of linear structural equation models (Bloxom 1968; Bentler and Lee 1978, 1979; Law and Snyder 1981). Sands and Young (1980) presented a restricted form of three-mode principal component analysis in the spirit of Harshman's PARAFAC2 model (1972), but they included an optimal scaling phase in their algorithm to accommodate data with lower measurement levels, missing data, and different data conditionalities (see also Young 1981).

In this chapter, we first present the three-mode principal component model. on a conceptual level by providing various informal ways of looking at it. Secondly, we provide an outline of some technical aspects connected with analyzing this type of model. Finally, an example treating data from attachment theory

---

I am particularly grateful to Frits Goossens for allowing me to use his data even before publication of his thesis. The conclusions with respect to the example should be seen as part of his work rather than mine. I would also like to thank Rien van IJzendoorn and other members of the "Attachment project" for their assistance in the preparation of the manuscript.

is used to illustrate some of the major aspects and possibilities of analyzing three-mode data with the three-mode principal component model.* (The method to solve the estimation of the model used and described here has been treated in full by Kroonenberg and de Leeuw [1980].)

THEORY

Informal Descriptions

In this section, we present three more or less different ways of looking at three-mode principal component analysis. First, we start with questions a researcher might ask about three-mode data and discuss the way in which these questions fit into the framework of a three-mode principal component model. Next, we take a structural point of view, postulating some structural relationships and investigating how real data might be described by a combination of structural parameters. Third, we will take a methodological point of view and demonstrate how three-mode principal component analysis is a generalization of standard principal component analysis and so-called singular value decomposition.

Research Questions Arising from Three-Mode Data

After collecting information from a number of subjects on a large number of variables, one often wants to know whether the observed scores could be described as combinations of a smaller number of more basic variables or so-called *latent variables*. As a first approximation, one generally looks for linear combinations of such underlying variables, which either account for the larger part of the variation—*principal components*—or reproduce the covariation matrix—*factors*.

As an example, one could imagine that the scores on a set of variables are largely determined by linear combinations of such latent variables as the arithmetic and verbal content. The latent variables—arithmetic and verbal content—can be found by a standard principal component analysis.

Suppose now, in the same example, that the researcher has administered the variables a number of times under various conditions of stress and time limitations. The data are now classified by three different types of quantities or *modes* of the data: subjects, variables, and conditions. First of all, the researcher is again interested in the components of the variables that explain a larger part of the variation in the data. Second, he wants to know if general characteristics can be defined for subjects as well. To put it differently, the researcher wants to know if it is possible to see the subjects as linear combinations of "idealized" subjects. In the example, we could suppose that the subjects are linear combinations of an exclusively mathematically gifted person

---

*This chapter aims to be comprehensible for the relatively uninitiated. A basic working knowledge of standard principal component analysis is, however, essential, as is an insight into eigenvalue-eigenvector problems.

and an exclusively verbally gifted person. Such persons are clearly "ideal" types. Finally, a similar question could arise with respect to conditions: Can the conditions be characterized by a set of "idealized" or "prototype" conditions?

Each of these questions can be answered by performing principal component analyses for each type of quantity or mode. In fact, the same variation present in the data is analyzed in three different ways; therefore, the components extracted must in some way be related. The question is, of course, how? In order to avoid confusion in answering this question, we will call the variable components *latent variables*, the subject components *idealized subjects*, and the condition components *prototype conditions*.

Considering the relationship between the components of the three modes, one could ask: Do idealized subject 1 and idealized subject 2 react differently to latent variable 2 in prototype condition 1? Or, is the relation between the idealized subjects and the latent variables different under the various prototype conditions?

By performing three separate component analyses, such questions are not immediately answerable, as one does not know how to relate the various components. The three-mode principal component model, however, specifies explicitly how the relations between the components can be determined. The three-mode matrix that embodies these relations is called the *core matrix*, as it is assumed to contain the essential characteristics of the data.

Structure: Raw Scores Derived from Idealized Quantities

It is often useful to look at three-mode principal components starting from the other end—the core matrix. For example, we pretend to know how an exclusively mathematically gifted person scores on a latent variable that has only mathematical content and on a latent variable that has only verbal content. Furthermore, we pretend to know these scores under a variety of prototype conditions. In other words, we pretend to know how idealized subjects react to latent variables under prototype conditions. However, in reality, we deal with real subjects, variables, and conditions. Thus, we have to find some way to construct the actual from the idealized world. A reasonable way to do this is to suppose that a real subject is a mixture of the idealized individuals and then make an analogous assumption for variables and conditions; the real scores can then be thought of as combinations of mixtures of idealized entities.

What is still lacking is some rule that indicates how the idealized quantities can be combined into real values. One of the simplest ways to do this is to weight each and then add the weighted contributions. (For instance, each latent variable could be weighted according to its average contribution over all subjects and conditions.) In more technical terms, each real variable is a linear combination of the latent variables.

We will show how to construct the score of an individual i on a test j under condition k from known idealized quantities. Suppose we have at our disposal 2 idealized persons (p_1, p_2), 2 latent variables (q_1, q_2), and 2 prototype conditions (r_1, r_2). We also know the scores of:

subject p_1 on variable q_1 under condition r_1:
$c_{p_1 q_1 r_1}$ or c_{111};

subject p_1 on variable q_1 under condition r_2:
$c_{p_1 q_1 r_2}$ or c_{112};

subject p_1 on variable q_2 under condition r_1:
$c_{p_1 q_2 r_1}$ or c_{121}; and

subject p_1 on variable q_2 under condition r_2:
$c_{p_1 q_2 r_2}$ or c_{122}.

Similarly we know the scores of subject p_2:
c_{211}, c_{212}, c_{221}, and c_{222}.

In other words, we know all the elements of the core matrix. As mentioned above, we want to construct the score of real subject i on a real variable j under a real condition k. We will do this sequentially and assemble all of the findings at the end.

We start with the observation that the score of a real subject i on the latent variable q_1 under a prototype condition r_1 is a linear combination of the scores of the idealized persons p_1 and p_2, using weights g_{ip_1} and g_{ip_2}:

$$s_{iq_1r_1} = g_{ip_1} c_{q_1 r_1} + g_{ip_2} c_{p_2 q_1 r_1} \qquad \text{or}$$

$$s_{i11} = g_{i1} c_{111} + g_{i2} c_{211} \,.$$

Similarly, for variable q_2 under condition r_1:

$$s_{i21} = g_{ip_1} c_{p_1 q_2 r_1} + g_{ip_2} c_{p_2 q_2 r_1}$$

$$= g_{i1} c_{121} + g_{i2} c_{221} \,,$$

and the other variable-condition combinations:

$$s_{i12} = g_{i1} c_{112} + g_{i2} c_{212} \,,$$

$$s_{i22} = g_{i1} c_{122} + g_{i2} c_{222} \,.$$

The weights g_{i1} and g_{i2} thus indicate to what extent the idealized subjects p_1 and p_2 determine the real subject i. The assumption in this approach is that these g_{i1} and g_{i2} are independent of the test and the conditions under which the subject is measured. All interrelationships between subjects, variables, and conditions are the consequence of interrelationships between the idealized entities, as reflected in the core matrix (see later section on interpretations of core matrices).

Our next step is to construct the scores for subject i on a real variable j instead of on the latent variables q_1 and q_2, as was done in the above procedure.

The score for subject i on real variable j under prototype condition r_1 is:

$$v_{ij1} = h_{jq_1} s_{iq_1 r_1} + h_{jq_2} s_{iq_2 r_1}$$
$$= h_{j1} s_{i11} + h_{j2} s_{i21} \; .$$

Similarly, on real variable j under prototype condition r_2, we have

$$v_{ij2} = h_{j1} s_{i12} + h_{j2} s_{i22} \; ,$$

where the weights h_{j1} and h_{j2} indicate to what extent the latent variables determine the real variable j.

Finally, we combine the idealized conditions. Subject i's score on test j under condition k may be written as

$$z_{ijk} = e_{k1} v_{ij1} + e_{k2} v_{ij2} \; ,$$

where the weights e_{k1} and e_{k2} indicate to what extent each idealized condition determines the real condition k.

Assembling the results from the three steps we get:

$$z_{ijk} = \sum_{r=1}^{2} e_{kr} v_{ijr} = \sum_{r=1}^{2} e_{kr} \{ \sum_{q=1}^{2} h_{jq} s_{iqr} \}$$

which can be compactly written as

$$z_{ijk} = \sum_{r=1}^{2} e_{kr} \{ \sum_{q=1}^{2} h_{jq} \, (\sum_{p=1}^{2} g_{ip} c_{pqr}) \} \; ,$$

where

$\sum_{p=1}^{2} g_{ip} c_{pqr}$ is the linear combination of subjects p_1 and p_2;

$\sum_{q=1}^{2} h_{jq} (\sum_{p=1}^{2} g_i c_{pqr})$ is the linear combination of variables q_1 and q_2; and

$\sum_{r=1}^{2} e_{kr} \{ \sum_{q=1}^{2} h_{jq} (\sum_{p=1}^{2} {}_{ip} c_{pqr}) \}$ is the linear combination of conditions r_1 and r_2

or

$$z_{ijk} = \sum_{p=1}^{2} \sum_{q=1}^{2} \sum_{r=1}^{2} g_{ip} h_{jq} e_{kr} c_{pqr} \; ,$$

as it is usually written.

As can be seen in the next section—Formal Descriptions—this is the definition of the three-mode principal component model. In Bloxom (chapter 4), the nested form of the three-mode model is described as well, but there, the model is developed as an example of a third-order factor analysis model, in which the s are the second order and the v the third order factors.

Methodology: Extending Standard Principal Component Analysis

From a methodological point of view, three-mode principal component analysis is a generalization of standard principal component analysis, or rather, of *singular value decomposition*. Figure 3–1 schematically shows the relationship between standard principal component analysis and singular value decomposition. In essence, singular value decomposition—(or *two-mode principal component analysis*)—is a simultaneous analysis of both the individuals and the variables, in which the relationship between the components of the variables and the subjects is represented by the core matrix **C**. In Figure 3–1, the core matrix is diagonal with s diagonal elements c_{pp} ($p = 1, \ldots, s$). These c_{pp} are equal to the square roots of the eigenvalues associated with the pth components of both the variables and the subjects. When **G** and **C** are combined to form **A**, as shown in Figure 3–1, we have the standard principal component solution; and when **H** and **C** are combined, we have what could be called (in Cattell's terms [1966]) "Q"-principal component analysis. Figure 3–2 shows the decomposition of a three-mode matrix according to the three-mode principal component model. Comparison of Figure 3–1 and Figure 3–2 shows the analogy between the singular value decomposition and three-mode principal component analysis. The core matrix now has three modes, and the relationships between the singular values or elements of the core matrix and the eigenvalues of the various modes are less simple than in the two-mode case (see later section on interpretations of core matrices).

Examples of Applications

In this section, we present some examples of the types of problems that can be handled successfully by three-mode principal component analysis.

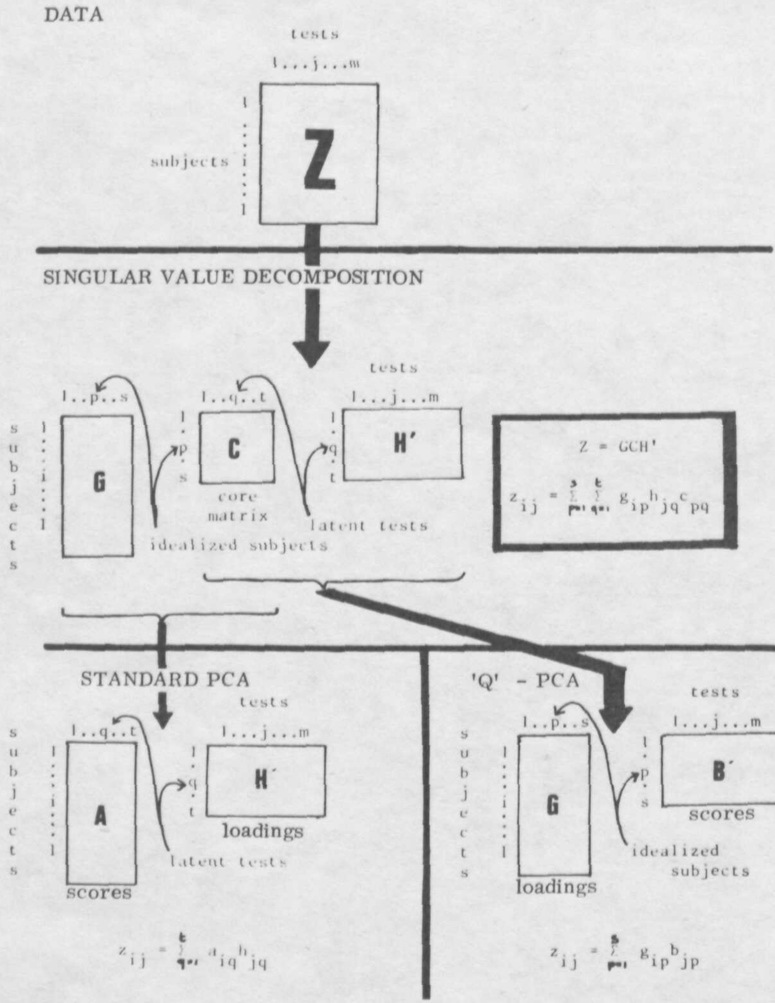
Semantic Differential Data

A classical example of three-way classified data can be found in the work of Osgood and associates (Osgood, Suci, and Tannenbaum 1957). In the development and application of semantic differential scaling, subjects have to judge various concepts using bipolar scales of adjectives. Such data used to be analyzed after averaging over subjects, but the advent of three-mode principal component analysis and similar techniques has made it possible to analyze the subject mode as well in order to detect individual differences with regard to semantic organization of the relations between the scales and the concepts. Examples of such studies can be found in Snyder and Wiggins (1970) and Kroonenberg (1983a).

Similarity Data

Three-way similarity data—consisting of stimuli by stimuli by subjects—are generally analyzed with individual differences scaling programs, such as INDSCAL (Carroll and Chang 1970) and ALSCAL (Takane, Young, and de Leeuw 1977). However, when

Figure 3-1. Singular Value Decomposition and Principal Component Analysis

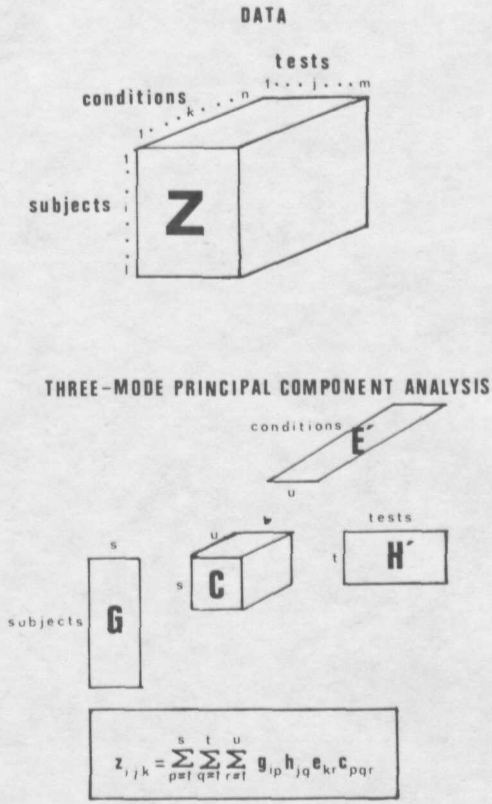


the data are asymmetric and/or a more general model is required, three-mode principal component analysis can provide useful insight. (See Carroll and Wish [1974] for details on individual differences scaling and its relation to three-mode component analysis.)

Asymmetric Similarity Data

Van der Kloot and Van den Boogaard (1978) collected data from 60 subjects who rated 31 stimulus persons on 11 personality trait scales. In the original report, the data—which can be considered asymmetric similarity data—were analyzed by canonical discriminant analysis using the stimulus persons as groups. This analy-

Figure 3–2. Three-Mode Principal Component Analysis



sis yielded a circular configuration for the scales and a similar configuration for the stimuli. Van der Kloot and Kroonenberg (1982) used three-mode principal component analysis on the original data, recovering essentially the same configurations for the scales and the stimuli. However, in their analysis, it was possible to show that the two spaces were in fact identical. In addition, it was possible to assess the individual differences between subjects. These differences manifested themselves primarily in the *size* rather than the *shape* of their configurations. These differences can be explained as differences in response style (extreme versus nonextreme), rather than differences in judgmental processes.

Multivariate Longitudinal Data

In the social sciences, multivariate longitudinal data pose problems for many standard techniques. There are often too few observations and/or too many points in time for the analysis of covariance approach (Jöreskog and Sörbom 1976), or too few points in time and/or too many variables for multivariate time series analysis by some kind of ARIMA model (see, for example, Glass, Willson, and Gottman 1975; Cook and Campbell 1979, ch. 6). In such situations, three-mode principal component analysis can be very useful, especially for exploratory purposes.

Lammers (1974) presented an example of longitudinal data with a relatively large number of variables and only a limited number of points in time. Data were available for 188 hospitals measured on 22 variables in 11 consecutive years. The aim of this study was to determine if various kinds of hospitals showed different patterns or rates of growth. The general results of this study revealed that, over the years, large hospitals stayed large in relation to the initially small ones and that all hospitals grew roughly in the same manner. There were, however, a small number of hospitals that showed a specific growth pattern in a special group of variables. A reanalysis of these data can be found in Kroonenberg (1983a). (A complete survey of applications of three-mode factor and principal component analysis can be found in Kroonenberg [1983b].)

Formal Descriptions

In this section, we present a rather superficial description of the Tucker3 and Tucker2 models since our purpose here is to provide just enough detail for understanding the main principles involved. A more detailed treatment can be found in Kroonenberg and de Leeuw (1980).

5

Tucker3 Model

The general three-mode principal component model, or Tucker3 model, can be formulated as the factorization of the three-mode data matrix $\mathbf{Z} = \{z_{ijk}\}$, such that:

$$z_{ijk} = \sum_{p=1}^{s} \sum_{q=1}^{t} \sum_{r=1}^{u} g_{ip} h_{jq} e_{kr} c_{pqr} \, ,$$

$$\text{for } i = 1, \ldots, l; \; j = 1, \ldots, m; \; k = 1, \ldots, n.$$

The coefficients g_{ip}, h_{jq}, and e_{kr} are the entries of the component matrices $\mathbf{G}(l \times s)$, $\mathbf{H}(m \times t)$, and $\mathbf{E}(n \times u)$; l, m, and n are the number of elements (= rows), and s, t, and u are the number of components of the first, second, and third mode, respectively. We will *always* assume that $\mathbf{G}$, $\mathbf{H}$, and $\mathbf{E}$ are *columnwise orthonormal* real matrices, with the number of rows larger than or equal to the number of columns. The c_{pqr} are the elements of the three-mode core matrix $\mathbf{C}(s \times t \times u)$.

In practice, the three-mode data matrix is not decomposed into all its components, since one is usually only interested in the first few. Therefore, one seeks an approximate decomposition $\hat{\mathbf{Z}}$ that is minimal according to a least-squared loss function. Specifically, one solves for a $\hat{\mathbf{Z}}$ such that:

$$\sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} (z_{ijk} - \hat{z}_{ijk})^2 \qquad \text{with}$$

$$\hat{z}_{ijk} = \sum_{p=1}^{s} \sum_{q=1}^{t} \sum_{r=1}^{u} g_{ip} h_{jq} e_{kr} c_{pqr}$$

attains a minimum. The algorithm to solve this minimization problem is implemented in the program TUCKALS3 (Kroonenberg 1981a). (Details about the existence and uniqueness of a minimum, the algorithm itself, and its implementation can be found in Kroonenberg and de Leeuw [1980].)

Tucker2 Model

An important restriction of the general Tucker3 model can be obtained by equating the component matrix **E** with the identity matrix. We will refer to this model as the Tucker2 model; it has also been called the generalized subjective metrics model. The Tucker2 model can be written as

$$z_{ijk} = \sum_{p=1}^{s} \sum_{q=1}^{t} g_{ip} h_{jq} c_{pqk} ,$$

or in matrix notation

$$\mathbf{Z}_k = \mathbf{G}\mathbf{C}_k\mathbf{H}' \quad (k = 1, \ldots, n) ,$$

where $\mathbf{Z}_k (l \times m)$ is the kth frontal plane or slice of the data matrix, and $\mathbf{C}_k(s \times t)$ is the extended core matrix, respectively.

The core matrix is called "extended" because the dimension of the third mode is equal to the number of conditions in the third mode rather than to the number of components, as is the case in the Tucker3 model. The Tucker2 model only specifies principal components for the l subjects and m variables but not for the n conditions. The relationships between the components of the subjects and the variables can be investigated for all conditions together, as well as for each condition separately.

The loss function for the Tucker2 model has the form

$$\sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} (z_{ijk} - \hat{z}_{ijk})^2 , \qquad \text{with}$$

$$\hat{z}_{ijk} = \sum_{p=1}^{s} \sum_{q=1}^{t} g_{ip} h_{jq} c_{pqk} .$$

The algorithm to solve this minimization problem is implemented in the program TUCKALS2 (Kroonenberg 1981b).

One important advantage of the methods discussed in this chapter over the standard procedures outlined by Tucker (1966, 297 ff) is that the estimates of the parameters are least-squares rather than estimates with ill-defined properties. Another advantage of the definition of loss functions is that it becomes possible to look at residuals (see later discussion on interpretation of residuals). A third advantage is that there exists a direct relationship between the eigenvalues of the configurations and the size of the elements in the core matrix (see introductory sections of this chapter).

Miscellaneous Topics

Various kinds of auxiliary information can be useful for the interpretation of results from a three-mode principal component analysis. Some of the most important ones will be presented here, including joint plots, component scores, use of residuals, scaling of input data, and rotations. Various ways to interpret core matrices will be discussed later in this chapter.

Joint Plots

After the components have been computed, the core matrix will provide the information about the relations between these components. It is very instructive to investigate the component loadings of the subjects jointly with the component loadings of, say, the variables, by projecting them together into one space, as it then becomes possible to specify what they have in common. The plot of the common space is called a *joint plot*.

Such a joint plot of every pair of component matrices for each of the components of the third mode—such as $\mathbf{E}$, in the TUCKALS3 case—and for the average core plane in the TUCKALS2 case, is constructed in such a way that $g_i (i = 1, \ldots, s)$ and $h_j (j = 1, \ldots, t)$ (the columns of $\mathbf{G}$ and $\mathbf{H}$, respectively) are close to each other. Closeness is measured as the sum of all $s \times t$ squared distances $d^2(g_i, h_j)$ over all i and j.

The plots are constructed as follows. For each component r of $\mathbf{E}$, the components $\mathbf{G}$ and $\mathbf{H}$ are scaled by dividing the core plane associated with that component, $\mathbf{C}_r$, between them (by using singular value decomposition), and then weighting the scaled $\mathbf{G}$ and $\mathbf{H}$ by the relative number of elements in the modes to make the distances comparable:

$$\mathbf{D}_r = \mathbf{G}\mathbf{C}_r\mathbf{H}' = \mathbf{G}(\mathbf{U}_r\Lambda_r\mathbf{V}_r')\mathbf{H}'$$

$$= \left(\frac{l}{m}\right)^{1/4}(\mathbf{G}\mathbf{U}_r\Lambda_r^{1/2}) \left(\frac{m}{l}\right)^{1/4}(\mathbf{H}\mathbf{V}_r\Lambda_r^{1/2})' = \tilde{\mathbf{G}}_r\tilde{\mathbf{H}}_r' \; ,$$

with

$$\tilde{\mathbf{G}}_r = \left(\frac{l}{m}\right)^{1/4}\mathbf{G}\mathbf{U}_r\Lambda_r^{1/2} \quad \text{and} \quad \tilde{\mathbf{H}}_r = \left(\frac{m}{l}\right)^{1/4}\mathbf{H}\mathbf{V}_r\Lambda_r^{1/2} \; ,$$

$$r = 1, \ldots, u \; .$$

When $\mathbf{C}_r$ is not square, only the first $\min(s,t)$ components can be used. The procedure can be interpreted as rotating the component matrices by an orthonormal matrix, followed by a stretching or shrinking of the rotated components. Similar procedures for plotting two sets of vectors into one figure have been developed by Schiffman and Falkenberg (1968). (See also Schiffman et al. 1981, ch. 14; Gabriel 1971, biplot; Carroll 1972, MDPREF; Benzecri 1973, correspondence analysis; Gifi 1981, ch. 4.)

Component Scores

In some applications, it is useful to inspect the scores of all combinations of the elements of two modes on the components of the third mode. For instance, for longitudinal data the scores of each subject-time combination (or *ik*-combination) on the variable (j) components can be used to inspect the development of an individual's score on the latent variable over time. In the example presented here, these component scores in fact turn out to be the most successful summary of the relationships involved.

The component scores on the rth component of the third mode have the form

$$d_{ijr} = \sum_{p=1}^{s} \sum_{q=1}^{t} g_{ip} h_{jq} c_{pqr} \qquad \text{or} \qquad \mathbf{D}_r = \mathbf{G}\,\mathbf{C}_r \mathbf{H}' \ .$$

But by using other combinations of component matrices, three different sets of scores can be calculated. In general, only a few of these will be useful in a particular application.

One of the interesting aspects of the component scores d_{ijr} is that they are at the same time the inner products,

$$\sum_{p=1}^{\min(s,t)} \tilde{g}_{ip}^{r} \tilde{h}_{jp}^{r} \ ,$$

thus expressing the closeness of the elements from different modes in the joint plot.

Residuals

Kroonenberg (1983a) shows that for both the Tucker3 and the Tucker2 models the following is true:

$$\sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} z_{ijk}^2 = \sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} \hat{z}_{ijk}^2 + \sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} (z_{ijk} - \hat{z}_{ijk})^2 \ ,$$

where the $\hat{z}_{ijk}$ values are the data "reconstructed" from the estimated parameters. This is, of course, a standard result in least-squares analyses. Less numerically, this may be written as

$$\text{SS(Data)} = \text{SS(Fit)} + \text{SS(Residual)}.$$

In addition, it is shown that for each element f of a mode,

$$SS(Data_f) = SS(Fit_f) + SS(Residual_f).$$

By comparing the fitted sum of squares and the residual sum of squares for the fth element, one can gauge the correspondence of the fth element's configuration with the overall configuration. Large residual sums of squares indicate that a particular element does not fit very well into the structure defined by the other quantities.

Clearly, the size of the SS(Residual) depends on the SS(Total). Therefore, one should focus on the *relative residual sum of squares* (or *relative residual,* for short), which is equal to SS(Residual)/SS(Total) when assessing the role of a particular element in the final solution. Similarly, one could look at the *relative fit* (= SS(Fit)/SS(Total)). Of course, these two quantities convey essentially the same information.

The SS(Res) and the SS(Fit), as well as their relationships, can be shown directly in a so-called *sums-of-squares-plot,* which is explained and illustrated in the section on fit of the scales, episodes, and children.

Scaling of Input Data

In standard principal component analysis, the input data are often transformed into standard scores without much thought about the consequences. In other words, correlation matrices are generally analyzed with principal component analysis rather than cross-product matrices or covariance matrices. In three-mode analysis, the question of scaling the input data must be approached with more care, as there are many ways to standardize or center the data.

Two basic rules can be formulated with regard to the scaling of input data: (a) those means should be eliminated (set equal to zero) that cannot be interpreted or that are incomparable within a mode; and (b) those variances should be eliminated (set equal to one) that are based on arbitrary units of measurement or that are incomparable within a mode. If all quantities are measured in the same (possibly arbitrary) units, it is not necessary to eliminate the variances and perhaps is not even desirable.

Common scaling procedures include: (a) centering or standardizing the variables over all subject-condition combinations (*j-centering*), so that the grand mean of a variable over all subjects and conditions is zero and/or its total variance over all subjects and conditions is one; (b) centering or standardizing the variables over all subjects for each condition separately (*jk-centering*); and (c) double-centering, or centering per condition over both variables and subjects (*jk-, ik-centering*). (As before, subjects, variables, and conditions here indicate first, second, and third mode quantities, respectively.)

The decision as to which centering or standardization method is appropriate in any particular data set depends on the researcher's assessment of the origin of the variability of his data. In other words, one must assess which means and variances can be meaningfully interpreted. Harshman (chapters 5 and 6) and Kruskal (chapter 2) discuss these issues in greater detail. (For a somewhat different perspective, see Kroonenberg [forthcoming].)

Rotation of Components and Core Matrix

In standard principal component analysis and factor analysis, it is customary to rotate the solution of the variables to some kind of "simple structure," mainly by Kaiser's (1958) varimax procedure. This and other rotational procedures have been extensively applied in three-mode principal component analysis (see Kroonenberg 1983b). Various authors have advocated some particular rotation for a specific type of data. Lohmoller (1981), for instance, recommends rotation of time components to orthogonal polynomials, a proposal also put forward by Van de Geer (1974). Subject modes tend to be transformed in such a way that the axes coincide with centroids of clusters of individuals. Tucker discusses several of the above possibilities and advocates that the "first priority for these transformations should be given to establishing meaningful dimensions for the object space (variables)" (1972, 10–12).

The emphasis in the literature on rotating the component matrices first is clearly a consequence of the familiarity with such procedures in standard principal component analysis. In three-mode analysis, the core matrix is the most difficult to interpret, due to its trivariate character (see later section on interpretation of core matrices). This leads to the recommendation to concentrate on the simplicity of the core matrix rather than that of the component matrices. Simplicity here means a large number of zeros or very small values in the core matrix, preferably in the off-diagonal elements. The most simple structure would be a core matrix (for the Tucker3 model), with only non-zero elements on the body diagonal ($c_{pqr} \neq 0$, if $p = q = r$). In such a case, each component of a mode is exclusively linked to one component of another mode, so that they can be equated or at least be given the same interpretation. This model is then the orthonormal version of the PARAFAC/CANDECOMP model discussed in Harshman and Lundy (chapter 5).

An interesting observation drawn from a large number of applications is that the principal component solution already seems to produce such simple structures if they are present in the data and if they are compatible with the model employed. At present, this is just an empirical finding, but it is conjectured that it can be shown to be true for at least a number of specific cases. On the basis of this conjecture, it appears that rotating the component matrices to some kind of structure in fact destroys the simplicity of the core matrix and thus introduces unnecessary complications in its interpretation.

AN EXAMPLE FROM ATTACHMENT THEORY

Design and Data Description

To familiarize the reader with some practical aspects of three-mode analysis and to illustrate the main points of the previous sections, we will analyze data collected by Goossens (forthcoming) on the reactions of two-year-old children to a stranger and to their mothers in an unfamiliar environment within the context of a standardized observation procedure called the "Strange Situation," from *Patterns of Attachment* (Ainsworth et al. 1978). The practical aspects and theoretical considerations that form the foundation of the strange situation are covered in many publications (including the above) as the measurement procedure has become a standard one in developmental psychology. Our main purpose here is to illustrate three-mode principal component analysis rather than to dwell in detail on the strange situation itself. We will, therefore, treat its aspects only insofar as it is necessary to understand the data and the analysis.

In the course of the strange situation, the child is subjected to increasingly stressful circumstances (such as the arrival of a stranger, leaving of the mother, and being left alone) in order to elicit "attachment behaviors." *Attachment* itself is defined as "the affectional bond or tie that an infant forms between himself and his mother figure—a bond that tends to be enduring and independent of specific situations." *Attachment behaviors* are defined as "the class of behaviors that share the usual or predictable outcome of maintaining a desired degree of proximity to the mother figure" (Ainsworth et al. 1978, 302).

As Ainsworth et al. point out, the sequence of episodes was very powerful both in eliciting the expected behaviors and in highlighting individual differences (1978, 33). The major purpose of the procedure is to assess the quality of the attachment relationship of a child to its mother-figure. (A summary of the procedure is given in Table 3–1.) The major types of attachment are *secure* attachment (*B*-children), *anxiously resistant* attachment (*C*-children), and *anxiously avoidant* attachment (*A*-children). Ainsworth et al. (1978, ch. 3) have developed a more detailed classification system, which is presented in Table 3–2. The classifications of the children are made by trained judges on the basis of the children's scores on so-called *interactive scales,* which range from 1 to 7 (see Table 3–3). The child's behavior corresponding to each of the 7 categories has been explicitly defined and can be summarized as going from 1 (virtually nonexistent) to 7 (very often, very intense). The scores are awarded by trained observers while viewing videotapes of the strange situation. In the present analysis, the following scales were used: *proximity seeking* (*PROX*), *contact maintaining* (*CM*), *resistance* (*RES*), *avoidance* (*AVOI*), and *distance interaction* (*DI*).

The data of the present study consisted of observations on 65 two-year-old children on the 5 interactive scales during 4 episodes (*S*4, *M*5, *S*7, *M*8), where *S* indicates the presence of the stranger and *M* that of the mother. Details on the data and the reasons for discarding the earlier episodes can be found in Goossens (forthcoming). One might argue that a three-mode analysis is not a proper technique for these data; for instance, proximity seeking toward the stranger might not be the same variable as proximity seeking toward the mother. Also, the relationships between the scales in the stranger episodes might be different from those in the mother episodes. However, since the basic purpose of the strange situation is to assess children on the

TABLE 3-1. *Description of Strange Situation*

Situation Number	Persons Involved	Duration of Situation	Brief Description of Action
1	mother, child, observer	30 secs.	Observer introduces mother and baby to experimental room, then leaves.
2	mother, child	3 min.	Mother is nonparticipant while child explores; if necessary, play is stimulated after two minutes
3	stranger, mother, child	3 min.	Stranger enters. First minute: stranger silent. Second minute: stranger converses with mother. Third minute: stranger approaches child. After three minutes mother leaves unobstrusively.
4	stranger, child (S4)	3 min. or less 1)	First separation episode. Stranger's behavior is geared to that of the child.
5	mother, child (M5)	3 min. or more 2)	First reunion episode. Mother greets and/or comforts child, then tries to settle it again in play. Stranger leaves unobtrusively in the meantime. Mother leaves saying "bye bye."
6	child alone	3 min. or less 1)	Second separation episode.
7	stranger, child (S7)	3 min. or less 1)	Continuation of second separation. Stranger enters and gears behavior to that of the child.
8	mother, child (M8)	3 min.	Second reunion episode. Mother enters, greets child, then picks it up. Meanwhle stranger leaves unobtrusively.

Note: The episode is curtailed if the child is unduly distressed, and the episode is prolonged if more time is required for the child to become reinvolved in play.

Source: Ainsworth et al. (1978, 37).

basis of their reactions to the entire strange situation—and not to specific parts of it—it seems justified to treat a scale as the same variable regardless of the adult to which the behavior is directed.

Before analysis, the overall scale means were removed; specifically, the scales were centered over all children-episode combinations (j-centering). (See previous section on Scaling of Input Data.) No equalization of variances was performed. This decision was based on the consideration that the individual differ-

TABLE 3-2. Ainsworth Classification System

	Behavior toward the mother					Most salient	Behavior toward
	PROX	CM	RES	AVOI	DI	feature	stranger
A1	-	-	-	++	-	disinterested	Treatment more or less
A2	+(+)	-	(+)	++	-	mixed feelings	like mother
B1	(+)	-	-	-	++	secure	Friendly toward stranger
B2	+(+)	(+)	-	(+)	+(+)	secure	but mother is clearly
B3	++	++	-	-	-/++	very secure	preferred and sought after
B4	++	++	(+)	-	-	secure	
C1	++	++	++	-	-	angry, ambivalent	Treatment more or less
C2	(+)	(+)	++	-	(+)	passive	like mother

- = low; (+) = low to moderate; + = moderate; +(+) = moderate to high; ++ = high.

Source: Ainsworth et al. (1978, 59-63); Sroufe and Waters (1977).

TABLE 3-3. Interactive Scales

Proximity (or contact) seeking	(PROX)	A measure for the degree of active initiative a child shows in seeking physical contact with or proximity to an adult.
Contact maintaining	(CM)	A measure for the degree of active initiative a child exerts in order to maintain physical contact with a person, once such contact is achieved.
Resistance	(RES)	A measure for the degree of angry and/or resistant behavior to an adult. It is shown by physically rejecting an adult who tries to come into contact or initiate interaction with the child.
Avoidance	(AVOI)	A measure for the degree of avoiding proximity and interaction with an adult, for instance by ignoring or looking away.
Distance interaction	(DI)	A measure for the degree in which a child interacts with an adult from a distance, for instance, by showing toys and talking.

ences between children were of more interest than the overall scoring levels of the children on the interactive scales. This centering ensures that the meaningful differences in scoring levels between episodes that carry important information are retained. However, a disadvantage of using the mean values for generalization is that they are sample-dependent. For more extensive studies, some standard norm for centering scales should be devised. (It should be mentioned that this centering is not recommended by Harshman [chapter 6] or Kruskal [chapter 2], but their starting point is different from ours [see Kroonenberg, forthcoming].)

Analyses and Fit

Analyses

The main analysis reported here is a Tucker3 (T3) analysis with two components each for the first mode (*episodes*), second mode (*interactive scales*), and third mode (*children*). It will be referred to as the *2 × 2 × 2-solution,* and it will be compared with a *3 × 3 × 3-solution* using the same data. We will also refer to a Tucker2 (T2) analysis with two components for the first two modes, or the *2 × 2-solution*. (It is, by the way, not necessary to have equal numbers of components, but it is often more convenient.)

Fit

Table 3-4 shows that the fit increases with an increasing number of components but that the increase in fit in going from the 2 × 2 × 2-solution (fit = .59) to the 3 × 3 × 3-solution (fit = .68) involves estimating an additional 93 parameters. At least 60% of the variation in the (*j*-centered) data is accounted for by the three-mode model. Considering the relative difficulty of reliably measuring children's behavior and the variability inherent in it, this seems quite satisfactory.

When using the Tucker2 model—computing only components for episodes and interactive scales—a better overall fit is possible than with the Tucker3 model using the same number of components (.67 for the 2 × 2-solution versus .59 for the 2 × 2 × 2-

TABLE 3-4. Characteristics of the Solution

	T3 2×2×2	T3 3×3×3	T2 2×2
Standardized total sum of squares--SS(Total)	1.00	1.00	1.00
Approximation of SS(Fit) from separate PCA			
on mode 1	.77	.91	.77
on mode 2	.83	.92	.83
on mode 3	.63	.71	--
Fitted sum of squares from simultaneous estimation--SS(Fit)	.59	.68	.67
Residual sum of squares from simultaneous estimation--SS(Res)	.41	.32	.33
Improvement in fit compared to initial configuration	.03	.01	.001
Parameters to be estimated	156	249	278

solution). But due to leaving the third mode uncondensed, there are more parameters in the former case (278 versus 156). Comparing the two T3-solutions, it is difficult to decide which is the "best" solution to look at in detail. No goodness-of-fit tests are available; furthermore, it seems largely a content-specific problem as to how much detail one wants to go into in describing the relations.

The "approximate fit" from the initial configuration for each of the modes, which are derived from the standard Tucker (1966) Method I solution, are upper bounds for the SS(Fit) of the simultaneous solution. Obviously, the smallest of the three is the least upper bound, which, in this case, is the one based on the third mode (.63). The initial configurations are used as starting points for the main TUCKALS algorithms. The improvement in fit indicates how much the iterative process improves the simultaneous solution over the starting solution. In this case, the improvement is not large—in other words, we might have settled for the Tucker method as far as fit is concerned. This does not mean, however, that the changes in the component matrices **G**, **H**, and **E** are also negligible.

Another point worth mentioning is the ratio of components to variables. In standard principal component analysis, it seems ill-advised to attempt to extract, for instance, three components for the four episodes. However, due to the presence of another mode, the order of the solutions in three-mode analysis may be larger than in the standard situation (see also Kruskal 1976, 1977).

Configurations of the Three Modes

One of the advantages of three-mode principal component analysis over separate analyses for each episode or each interactive scale is that one common space can be found for all episodes together, for instance, instead of one for each. The common component spaces for each mode are given in Tables 3–5, 3–6, and 3–7, respectively. In Figure 3–3, the components for scales and episodes are plotted and in Figure 3–4 those for the children are plotted. In Figure 3–3, but not Figure 3–4, the components have been multiplied by the square root of their component weights so that the plots reflect the relative importance of the axes.

The general remark can be made that the choice of a particular solution is not very crucial with respect to interactive scales and episodes. The first two components of both the scale space and the episode space are the same within reasonable bounds (roughly ± .05 and the order is preserved in all but two cases). The differences illustrate, by the way, that the solutions are *not* nested.

A point that should be made at the outset of the interpretation is that it is rather difficult to link the details of our results to those in Ainsworth et al. (1978); the latter refer mainly to one-year olds and Goossens' study deals with two-years olds. Previous research (summarized in Ainsworth et al. 1978) shows that the reaction of older children in the strange situation is different from that of one-year-olds for whom it has been validated (see also Goossens et al. 1982).

TABLE 3-5. Component Spaces--Episodes (Mode 1)

			T3: 2×2×2		T3: 3×3×3			T2: 2×2	
nr.	adult		E1	E2	E1	E2	E3	E1	E2
4	stranger	S4	.26	-.44	.25	-.37	.45	.26	-.45
5	*mother*	M5	.47	.25	.52	.28	.68	.48	.27
7	stranger	S7	.38	-.77	.41	-.80	-.23	.44	-.73
8	mother	M8	.75	.39	.71	.38	-.53	.71	.43
component weight (λ_p)			.37	.22	.41	.21	.07	.42	.25

Note: Labels for components: E1 = stress of situation; E2 = mother versus stranger; E3 = early versus late.

TABLE 3-6. Component Spaces--Interactive Scales (Mode 2)

		T3: 2×2×2		T3: 3×3×3			T2: 2×2	
Scales		S1	S2	S1	S2	S3	S1	S2
Proximity seeking	PROX	.32	.69	.37	.68	.04	.35	.67
Contact maintaining	CM	.26	.35	.26	.34	.14	.28	.34
Resistance	RES	.33	-.41	.30	-.39	.85	.30	-.39
Avoidance	AVOI	.27	-.48	.25	-.50	-.46	.25	-.53
Distance interaction	DI	-.81	.07	-.80	.12	.24	-.80	.10
Component weight (μ_q)		.37	.22	.43	.24	.02	.40	.27

Note: Labels for components: S1 = intensity of reaction; S2 = security seeking; S3 = interest in adult.

One of the aims of the present analysis, with regard to the content of the research, is to investigate how individual differences between the children can be traced back to their different behavior in the various episodes based solely on the interactive scales. These results will then be compared with the classification subcategories resulting from the scoring instructions in Ainsworth et al. (1978) (see Goossens, in preparation).

One qualification should be made in advance, as the research project from which these data have been derived is not yet finished. The results presented here should be seen as preliminary and not yet definitive; such final results will be published elsewhere at a later date.

TABLE 3-7. Component Spaces--Children (Mode 3)

Number	ACC	C1	C2	Number	ACC	C1	C2
55	B4	.34	.08	52	B3	-.03	.14
39	B4	.33	.12	32	B3	-.02	.14
38	B4	.30	.07	40	B3	-.05	.14
18	B4	.28	.09	19	B3	-.07	.14
62	B4	.27	-.03	7	B3	-.07	.14
20	B4	.25	-.03	33	B3	-.06	.13
48	B4	.22	.14	64	B3	-.06	.13
61	B4	.22	-.02	42	B3	-.05	.13
24	B4	.20	.05	59	B3	-.08	.12
3	B4	.19	.08	60	B3	-.08	.11
44	B3	.19	-.02	17	B2	-.04	.11
2	B4	.18	.06	47	B3	-.06	.10
41	C1	.18	-.01	30	B2	-.02	.09
11	B3/4	.15	.07	56	B3	-.04	.09
13	B3	.14	.14	16	B3	-.05	.09
34	B3	.08	.10	29	B2	-.09	.08
--------	--------	--------	--------	43	B3	-.09	.07
14	B3	.08	.26	26	B1	-.03	.07
57	B3	-.04	.21	6	B1	-.00	.06
4	B3	.01	.20	63	B2	-.03	.05
12	B3	-.01	.19	15	B2	-.05	.04
27	B3	.04	.19	21	B3	-.07	.04
22	B3	-.01	.18	10	B2	-.04	.00
50	B3	-.04	.18	31	B3	.04	-.01
65	B3	.01	.18	35	B1	-.02	-.02
28	B3	.02	.17	8	B1	-.02	-.04
9	B3	-.00	.17	49	?	.07	-.06
25	B3	-.09	.16	53	B2	-.04	-.08
5	B3	-.09	.16	51	B2	-.01	-.09
58	B3	-.07	.16	--------	--------	--------	--------
46	B3	-.03	.15	54	A1	-.03	-.17
1	B3	-.07	.15	37	A1	.08	-.21
36	B3	-.08	.14				
23	B3	-.10	.15	component weight		.50	.09
45	B3	-.00	.15	(ν_r)			

Note: ACC = Ainsworth's classification category; ? = unclassified; B3/4 = B3 or B4.

Episodes

With just four episodes, there is really no need to label the axes, but for further reference we will try to name them anyway. The first axis (*E*1) reflects the overall variability of the scores in the episodes, and it does not seem unreasonable to associate increasing variability with greater *stress* placed on the child. The second axis (*E*2) contrasts the behavior toward the *mother* with

Figure 3–3. Component Spaces (Scaled): Episodes and Interactive Scales

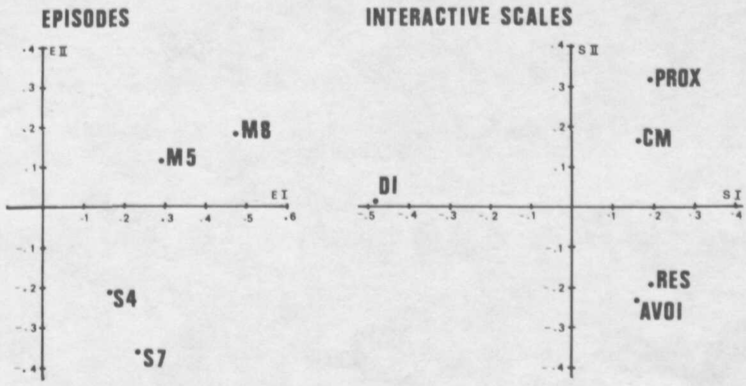
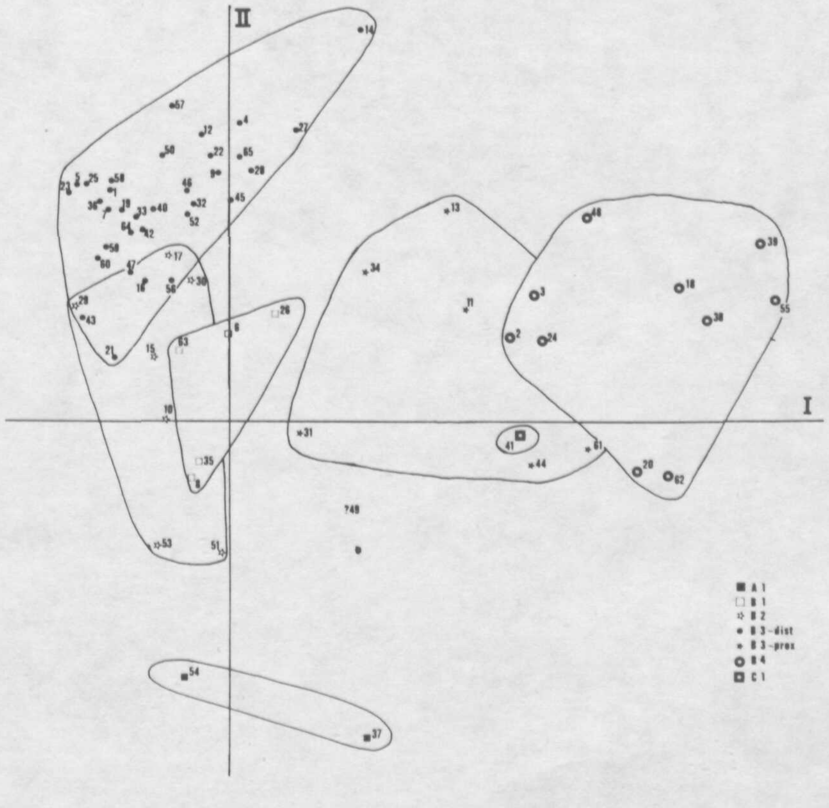


Figure 3–4. Child Space (Unscaled)



that toward a *stranger*. Finally, the third axis ($E3$) contrasts the *early* and *late* episodes, namely, those episodes before and after episode 6, in which the child has been left alone.

Interactive Scales

The first axis (*S*1) reflects the overall variability of the children-episode combinations around the overall scale mean. This variability is approximately equal for *PROX, CM, RES,* and *AVOI* and considerably larger for *DI*. It is clear that high scores on distance interaction reflect an opposite reaction compared to high scores on the other scales and that the same holds for low scores. This is, of course, to be expected as proximity seeking more or less precludes distance interaction and vice versa. The special position of distance interaction has been noted before; therefore, a number of researchers do not include it in their analyses (for example, see Waters 1977; Grossman et al, 1981). In Ainsworth et al. (1978), for instance, it is noted that for one-year-olds, distance interaction is a low-stress behavior of low intensity and that it differentiates less among the classification subcategories (Ainsworth et al. 1978, 246). Whether this is true for two-year-olds is still a matter for investigation; we will come back to this point later. An acceptable label for the first scale component therefore seems to be *intensity* of the reaction.

The second component (*S*2) distinguishes between attachment behaviors, proximity seeking and contact maintaining, and behaviors antithetical to attachment, including avoidance and resistance. It might be labeled as *security seeking*. We will not discuss the third axis (*S*3), due to the small amount of variation explained by it (2%), even though it shows a theoretically important contrast between resistance and avoidance.

Children

Table 3–7 and Figure 3–4 show the two-dimensional child space for the 2 × 2 × 2-solution. The children have been labeled both by a sequence number and their Ainsworth classification subcategory (see Table 3–2). These classifications are based on the same interactive scales as those in this analysis. However, it is primarily the behavior toward the mother that is taken into account, instead of that toward both the mother and the stranger, as is the case in our analysis. The classification instructions are contained in Ainsworth et al. (1978, 59–62; see also Swaan and Goossens 1982) and require extensive training. One of the aims of applying three-mode principal component analysis to these data is to assess the adequacy of the scoring instructions. For instance, it is known from psychological and medical research that people do not necessarily combine multivariate information in a very reliable way (see Sawyer 1966; Einhorn 1972).

With respect to these data, we will try to answer two questions: (a) whether the classification system is consistent; namely, whether the children who occupy the same region in the child space have the same Ainsworth classification; and (b) whether the same scales are responsible for the grouping of the children to the same extent, as is specified in the scoring instructions; it could be that the observed grouping in our analysis is the result of different combinations of scores. In other words, the present analysis is an attempt to validate the classification rules.

Ainsworth et al. (1978, ch. 6) applied discriminant analysis to check the adequacy of the classification system, but this involved the interactive scales twice: once to make the classification and then again to evaluate this classification by using the interactive scales as predictors in the discriminant functions. Here we use the interactive scales to group the children and to assess their contribution to this grouping simultaneously; only after that do we check the grouping against the classification. This provides a more adequate check of the appropriateness of the classification procedure.

The first general impression is that a reasonable separation is possible between the B-subcategories, although on the basis of our analysis alone the divisions could not have been made. In addition, the two $A1$-children are in their proper places, as their score pattern on the interactive scales should be the mirror-image of the $B3$-children (see Table 3–2). Furthermore, the one $C1$-child does not occupy a separate place. Finally, there are some $B3$-children seemingly belonging to the $B4$-children, and they have been labeled "$B3$-prox" for reasons to be discussed in the last section of this chapter, where we will also try to provide the answers to the above questions. In the meantime, we will use the Ainsworth classification to label the children, pretending we have already established its appropriateness.

Interpretations of Core Matrices

In this section, we will discuss the interpretational possibilities of the core matrices of the Tucker3 and Tucker2 models, both in general and within the context of the example. Three ways to interpret the values in the core matrix are given: (a) percentage of explained variation; (b) three-mode interactions; and (c) scores of idealized (latent) quantities.

Explained Variation

The core matrix indicates how the various components of the three modes relate to one another. For instance, the element c_{111} (= 19.9) of the T3 core matrix (Table 3–8) indicates the strength of the relation between the first components of the three modes, and c_{221} (= 13.5) indicates the strength of the relation between the second components of the first and second modes in combination with the first of the third mode. The interpretation of the elements of the core matrix is facilitated if one knows that the sum over all squared elements of the core matrix is equal to the unstandardized SS(Fit). In other words, c^2_{pqr} indicates how much the combination of the pth component of the first mode, the qth component of the second mode, and the rth component of the third mode contributes to the overall fit of the model, or how much of the total variation is accounted for by this particular combination of components. Thus, as seen in Table 3–8, 30% of the SS(Total) is accounted for by the combination of the first components of the three modes, another 14% by c^2_{221} and 3% each by c^2_{121} and c^2_{211}. Together the contributions of the elements of the first frontal plane add up to 50%, which is equal to the standardized weight of the first component of the third mode, as

13

TABLE 3-8. Core Matrices--TUCKALS3-Solution

2×2×2-solution (frontal planes)

Child component	Components of episodes		S1 Intensity of reaction	S2 Security seeking	Proportion variation explained			
			Components of interactive scales					
(C1): B4 versus REST	Stress of situation	E1	19.9	5.8	.30	.03	c_{111}	c_{121}
	Mother versus stranger	E2	-5.8	13.5	.03	.14	c_{211}	c_{221}
					$\nu_1 = .50$			
(C2): B3-dist versus A1	Stress of situation	E1	-6.7	3.0	.03	.01	c_{112}	c_{122}
	Mother versus stranger	E2	-2.1	7.7	.00	.05	c_{212}	c_{222}
					$\nu_2 = .09$			

3×3×3-solution (frontal planes)

C1	S1	S2	S3	C2	S1	S2	S3	C3	S1	S2	S3
E1	20.1	4.5	.6	E1	-6.5	2.8	2.6	E1	1.1	-6.7	-2.7
E2	-4.8	13.7	-2.2	E2	-2.0	6.8	0.1	E2	-1.7	-0.8	0.0
E3	-2.0	-2.3	-0.5	E3	-7.1	-0.3	0.6	E3	-4.9	-1.3	0.0
		$\nu_1 = .50$				$\nu_2 = .12$				$\nu_3 = .06$	

it should be. The core matrix thus breaks up the SS(Fit) into small parts, through which the complex relations between the components can be analyzed. It is in this way that we can interpret the core matrix as the generalization of eigenvalues or of the singular values of the singular value decomposition. It constitutes a further partitioning of the explained variation, as is indicated by the eigenvalues of standard principal component analysis.

In the present example, we see that the differences between the children on the first component ($C1$) explain half of the fitted variation. This 50% can be partitioned as follows:

a. due to c_{111} (30%): intensity of reaction ($S1$) due to the stress of situation ($E1$) for $B4$-children versus $REST$ ($C1$);
b. due to c_{221} (14%): security seeking ($S2$) with the mother versus stranger ($E2$) for $B4$-children versus $REST$ ($C1$);
c. due to c_{121} (3%): security seeking ($S2$) with stress of situation ($E1$) for $B4$-children versus $REST$ ($C1$); and
d. due to c_{211} (3%): intensity of reaction ($S1$) with mother versus stranger ($E2$) for $B4$-children versus $REST$ ($C1$).

The differences between the children on the second component ($C2$) contribute the remaining 9% explained variation, which can be divided as follows:

e. due to c_{112} (3%): intensity of reaction ($S1$) due to the stress of the situation ($E1$) for $B3$(dist)-children versus $A1$-children ($C2$);
f. due to c_{222} (5%): level of attachment ($S2$) with mother-stranger ($E2$) for $B3$(dist)-children versus $A1$-children ($C2$); and
g. due to c_{122} (1%): security seeking ($S2$) with stress of the situation ($E1$) for $B3$(dist)-children versus $A1$-children ($C2$).

Three-Mode Interactions

The percentages of explained variation only point to the important combinations but do not indicate the direction of the relationship. This information can be found in the original, not-squared core matrix. The problem is, however, what the c_{pqr} themselves represent. Their squares are variation explained; and the c_{pqr} themselves refer to what we call *three-mode interactions*.

To illustrate this three-mode interaction between loadings on components, we will look at c_{111} (= +19.9). The plus sign indicates that:

a. positive loadings on $C1$, $S1$, and $E1$ occur together the more $B4$-like children are, the more intensely they react (= the higher their scores are above average on all scales except DI) in more stressful situations (= $M5/S7$ and $M8$);
b. negative loadings on $C1$ and $S1$, occur together with positive loadings on $E1$: the more negative a child loads on $C1$ the less intensely it reacts (= scores below average on all scales except DI) in more stressful situations (= $M5/S7$ and $M8$);
c. positive loadings on $C1$ and negative loadings on $S1$ and $E1$

go together; and

d. negative loadings on *C*1 and *E*1 go together with positive scores on *S*1.

However, these combinations (*c* and *d*) do not occur in practice as all episodes load positively on *E*1.

Similarly, for each element c_{pqr} of the core matrix, such a set of statements can be made. Clearly, having four statements to explain each element of the core matrix is not particularly easy to comprehend. The situation can be simplified by omitting statements about negligible elements in the core matrix (here, c_{212} and c_{122}) and by making "conditional statements." For instance, *a, b, c,* and *d* can be simplified as:

a. for *B*4-children (with positive loadings on *C*1), intensity of the reaction (*S*1) and stress of the situation (*E*1) are positively related; and
b. for children with negative loadings on *C*1, intensity of the reaction and stress of the situation are negatively related.

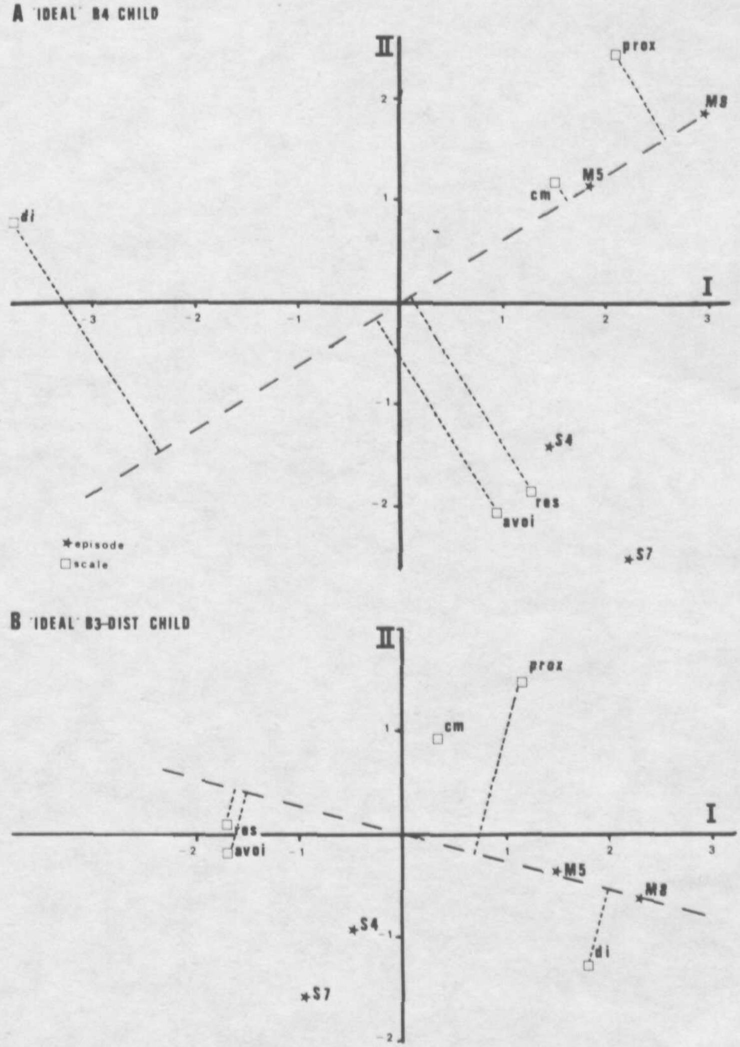
Another way to gain insight in the three-mode interactions is to try and produce plots of these relationships. Making plots "conditional" upon one of the modes turns out to be singularly effective. Both the *joint plots* (Figure 3-5, Parts A and B) and the plots of the *component scores* (Figure 3-6, Parts A and B) are examples of this approach. Here these plots are made conditional on the child components. In general, the subject matter and the way the data have been generated will determine which mode can be best used for conditioning.

Scores of Idealized Quantities

This interpretation was the basis for the earlier explanation of the model at the beginning of this chapter. Each element of the core matrix represents the score of a "pure" or "ideal" child on a latent interactive scale in a prototype episode. For Goossens' data, this means that an ideal *B*4-child reacts intensely in stressful situations (c_{111} = 19.9), seeks much security with its mother-figure (c_{221} = 13.5), seeks moderate security in stressful situations (c_{121} = 5.8), and reacts with moderately low intensity to the mother-figure (c_{211} = -5.8).

The difference with the interpretation in the previous section is that there the interpretations were based on relationships between loadings on components, and here we construct interpretations in terms of the components themselves. In some applications, the former method will be easier to handle and in other applications, the latter. In the present example, using very few elements in the episode and scale modes, the naming of components is somewhat uncertain and the former approach seems more helpful. In other cases, especially when the labeling of the components as continuous variables is more adequately defined, the latter approach will be easier to use.

Figure 3–5, Parts A and B. Joint Plots of Episodes and Interactive Scales

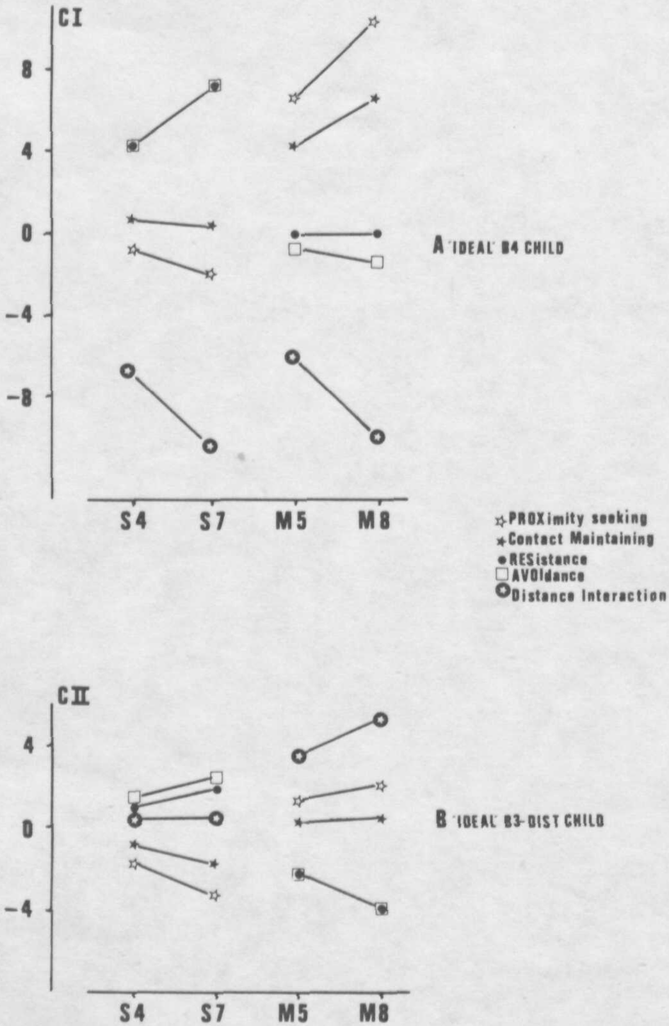


Extended Core Matrix of Tucker2 Model

So far we have only looked at the interpretation of the core matrix of the Tucker3 model. The extended core matrix can be interpreted in essentially the same way as the T3 core matrix, in terms of the amount of explained variation. Again the sum of the squared elements equals the fitted sum of squares, but now the sum of the squared elements of a frontal plane, $\mathbf{C}_k$, equals the contribution of the kth element (child) to this fitted sum of squares.

We already noted the near equality of the components for the interactive scales and the episodes in the 2×2-solution and

Figure 3–6, Parts A and B. Component Scores for Episode-Scale Combinations



15

2 × 2 × 2-solution in connection with Table 3–5 and Table 3–6, and thus interpretations of those spaces are the same as before. The relationships between these components, as embodied in the frontal planes of the T2 core matrix, are given for a few selected children in Table 3–9. Four of the children were chosen because they are relatively close to one of the axes in the child space (namely, 38, 57, 29, 37) and can thus be considered "ideal individuals" (Tucker and Messick 1963).

Thus, the frontal planes indicate how the axes of the common space are related for each child, as was the case in the Tucker3 model for "ideal" children. For instance, for child 38 (a *B*4-child), intensity of reaction (*S*1) and stress of the situation (*E*1) are positively related (see Table 3–9), as are security seeking ($S2$) and the mother versus stranger distinction ($E2$); the other combinations are immaterial. In comparison, for child 35 (a $B1$-child), none of the relationships seem very relevant (see the last section for a discussion of this phenomenon). Note also that the two $A1$-children (37 and 54) have very different patterns of relationships, despite their similar position in the child space (Figure 3–4).

Basically, one can conclude that children on the first child dimension ($C1$) weight the intensity-stress ($E1,S1$) combination and the mother versus stranger-security seeking ($E2,S2$) combination with a ratio similar to that of c_{111} to c_{221} in the T3 analysis. The overall size of the elements determines their position on the $C1$ component. High, positive numbers on the diagonal of the T2 core plane (for child 41 and child 38) lead to highly positive loadings on $C1$ and moderately negative numbers (for child 29) lead to moderately negative loadings. On the negative side of the second child component ($C2$), there are children who emphasize the ($E1,S1$) combination but not the ($E2,S2$) combination (37). On the positive side of $C2$ (57), the situation is reversed: ($E2, S2$) is high and ($E1,S1$) is low. This distinction corresponds with the opposite signs in the second frontal plane of the T3 analysis.

Direction Cosines

In those cases in which two modes are equal or the components define the same space, an additional interpretation of the core matrix is possible. For instance, within the context of multidimensional scaling of individual differences, the input similarity

TABLE 3-9. Core Matrices--TUCKALS2 Core Planes for Selected Children

	B4 (38)		B3 (57)		B2 (29)		B1 (35)	
	S1	S2	S1	S2	S1	S2	S1	S2
E1	5.7	0.9	-2.1	-0.7	-2.4	-0.2	-0.2	-1.0
E2	-0.5	5.1	-0.7	1.4	0.4	-0.6	1.1	0.1
*)	.30	.07	-.04	.21	-.09	.08	-.02	-.02
	A1 (54)		A1 (37)		C1 (41)		B2 (51)	
	S1	S2	S1	S2	S1	S2	S1	S2
E1	0.6	-2.2	3.0	-3.2	3.7	-.07	0.5	-0.0
E2	2.3	-0.8	-0.1	0.4	-1.0	2.9	0.2	-0.9
*)	-0.3	-.17	.08	-.21	.18	-.01	-.01	-.09

Note: B4 (38): child number 38--Ainsworth classification category B4; S1 (S2): first (second) scale component; E1 (E2): first (second) episode component; *): T3 component loadings (see Table 3-7).

matrices satisfy these conditions. Within this field, an interpretation has been developed in terms of correlations and direction cosines of the axes of the spaces common to two (generally the first and second) modes (see Tucker 1972, 7; Carroll and Wish 1974, 91).

In such situations, it makes sense to speak about the angle between the first and second component of the common space. This angle can be derived from the off-diagonal elements of the core planes, as they can be looked upon as a direction cosine or correlation between component p and component q, provided c_{pqr} is scaled by dividing it by $c^{1/2}_{ppk}$ and $c^{1/2}_{qqk}$, and that the components are standardized. The direction cosine indicates the angle under which the kth condition "sees" the axes or components of the common space. In the present example, the approach is not applicable, but in Kroonenberg (1981a) and Van der Kloot and Kroonenberg (1982), the method has been successfully used.

Joint Plots

The approach in the previous section toward the core matrices was in the spirit of Tucker's three-mode scaling (1972) and Harshman's PARAFAC2 (1972), as indicated in Carroll and Wish (1974) and Dunn and Harshman (1982). The joint plots, on the other hand, are more similar to Carroll and Chang's (1970, 1972) approach to treating the core matrix, in which the extended core matrix is decomposed by either eigenvalue-eigenvector or singular value decompositions. As pointed out earlier, the purpose of the joint plots is that elements of two modes can be plotted in one figure in order to express the relationships between the components in terms of the original variables, such as scales and episodes in our example. An advantage is that the components are now automatically scaled in accordance with their relative importance via the core plane.

With the joint plots, we can examine in some detail the relationships between the interactive scales and the episodes for each ideal-type child or child component. In Figure 3–5, Parts A and B, we present the joint plots for the two child components. The following characterization for the children loading on the positive side of the first component $C1$ can now be made.

a. They have high scores on proximity seeking and contact maintaining toward the mother (in episodes $M5$, $M8$), and they score about twice as high in $M8$ as in $M5$. With a high score, we mean relatively to the overall scale means, as we have removed these means for all interactive scales.
b. They have high scores on resistance and avoidance toward the stranger (in $S4$ and $S7$), nearly twice as high in $S7$ as in $S4$.
c. They show roughly average resistant and avoidant behavior toward the mother in $M5$ and $M8$ and even somewhat below average behavior on avoidance. Similarly, proximity seeking and contact maintaining toward the stranger have average values.
d. The scores on distance interaction do not discriminate between

the mother and the stranger and they are below average. There is less distance interaction in the later episodes.

These interpretations are derived from the fact that the scales can be seen as points and the episodes as vectors or directions in the common space, and vice versa. In this case, the former approach is to be preferred because the episodes are fixed; specifically, they are elements of the design. The relative importance of the various scales at any episode can then be assessed from their perpendicular projections on the vectors as is shown for *M*5 and *M*8 combined. The values of the projections are contained in the matrix of the scaled inner products $\mathbf{D}_r = \tilde{\mathbf{G}}_r \tilde{\mathbf{H}}_r'$ $= \mathbf{G}\mathbf{C}_r\mathbf{H}'$ (see earlier discussion of component scores).

For the positive scores on the second child component—the *B*3(dist)-children—the characterization is (see Figure 3–5, Part B):

a. low scores on resistance and avoidance toward the mother, coupled with average contact maintaining and proximity seeking; high distance interaction increasing further in *M*8;
b. low scores on proximity seeking and contact maintaining toward the stranger, with lower scores on proximity seeking; average resistance, avoidance, and distance interaction with a slight increase in the avoidance measures in *S*7.

For 37, an *A*1-child, the mirror image of the above observations is true as he lies on the negative side of the second child component (*C*2).

Component Scores

As remarked earlier in this chapter, the values of the inner products on which the above observations were made are at the same time the component scores on the child component in question; thus, they can serve as an intermediate level of condensation between the raw data and the three-mode model.

As we are looking here for characterizations of "ideal" children, it is not very useful to display the component scores in the two-dimensional child space. It is far more useful to plot the component scores of the interactive scales for each episode, as is done in Figure 3–6. In fact, for the present data—together with that of the children's loadings—these plots are the best summary of the results. Certainly they are easier to read than the joint plots with their projections on vectors. It has been noted that in longitudinal data, these plots have a similar appeal when the points in time are placed on the horizontal axis (see Kroonenberg 1983a).

Fit of the Scales, Episodes, and Children

In essence, the analysis could stop with the above interpretations. All that the technique has to offer toward breaking down complex relationships into small intelligible pieces is contained in the analysis so far. However, it is beneficial to have some additional information available to assess if there are no irregularities in the data, such as outliers, unduly influential points, and points that are not sufficiently accounted for. A useful way to investigate such questions is to inspect the residual sums of squares in conjunction with the fitted sums of squares (see earlier discussion of residuals). Whereas the core matrix informs us about the contributions of the components and their interrelationships, the sums of squares broken down by the elements or variables of the modes inform us about the contributions of these elements to the solutions.

In Table 3–10 and Table 3–11, the sums of squares for the scales and episodes are shown, respectively. From the SS(Total)s for episodes, we see that the variability as expressed by the sums of squares increases with the later episodes, as children deviate more from the scale means or perhaps show more variation among themselves. Which of the two is more important cannot be unequivocally determined from the present analysis and should be assessed separately. With respect to the scales, we see that contact maintaining has relatively little variability, while distance interaction has considerably more variation. From the residual sums of squares we note that the scales fit more or less equally well, irrespective of their total sum of squares, but that the configurations derived and discussed above are for a large part determined by the last two episodes. The structure described is therefore more representative of the later behaviors than the early ones. This explains, for instance, why an added third episode component shows an early versus late character; primarily the earlier episodes will then be fitted better.

Figure 3–7 is a so-called *sums-of-squares-plot,* which shows the residual sums of squares versus the fitted sums of squares for the children from the $2 \times 2 \times 2$-solution. By plotting the sums of squares directly, rather than taking the relative sums of squares, the total sums of squares are also contained in the plot, and unusually large elements can be spotted directly. Moreover, it can be seen if the larger SS(Fit)s resulted only from larger total sums of squares, as is to be expected from least-squares procedures. Furthermore, whether the variations of the elements

TABLE 3-10. Sums of Squares for Episodes (Mode 1)

		2×2×2-solution				3×3×3-solution	
	SS(Total)	SS(Fit)		SS(Res)		SS(Res)	
Episode	STD	STD	REL	STD	REL	STD	REL
S4	.16	.07	.40	.10	.60	.10	.59
M5	.21	.09	.44	.12	.56	.06	.27
S7	.29	.18	.63	.11	.37	.09	.30
M8	.33	.24	.74	.09	.26	.08	.23
Overall	1.00	.59		.41		.32	

Note: STD = standardized or divided by the overall SS(Total); REL = relative sum of squares, which is defined as:

$$\text{relative SS(Res) of episode S4} = \frac{\text{SS(Residual) of episode S4}}{\text{SS(Total) of episode S4}}$$

TABLE 3-11. Sums of Squares for Interactive Scales (Mode 2)

		2×2×2-solution				3×3×3-solution	
	SS(Total)	SS(Fit)		SS(Res)		SS(Res)	
Scale	STD	STD	REL	STD	REL	STD	REL
PROX	.23	.14	.61	.09	.39	.06	.27
CM	.10	.05	.54	.05	.46	.04	.41
RES	.15	.08	.52	.07	.48	.06	.41
AVOI	.17	.08	.44	.09	.56	.08	.46
DI	.35	.24	.68	.11	.32	.07	.21
Overall	1.00	.59		.41		.32	

Note: STD = standardized or divided by the overall SS(Total); REL = relative sum of squares, which is defined as:

$$\text{relative SS(Res) of PROX} = \frac{\text{SS(Residual) of PROX}}{\text{SS(Total) of PROX}}$$

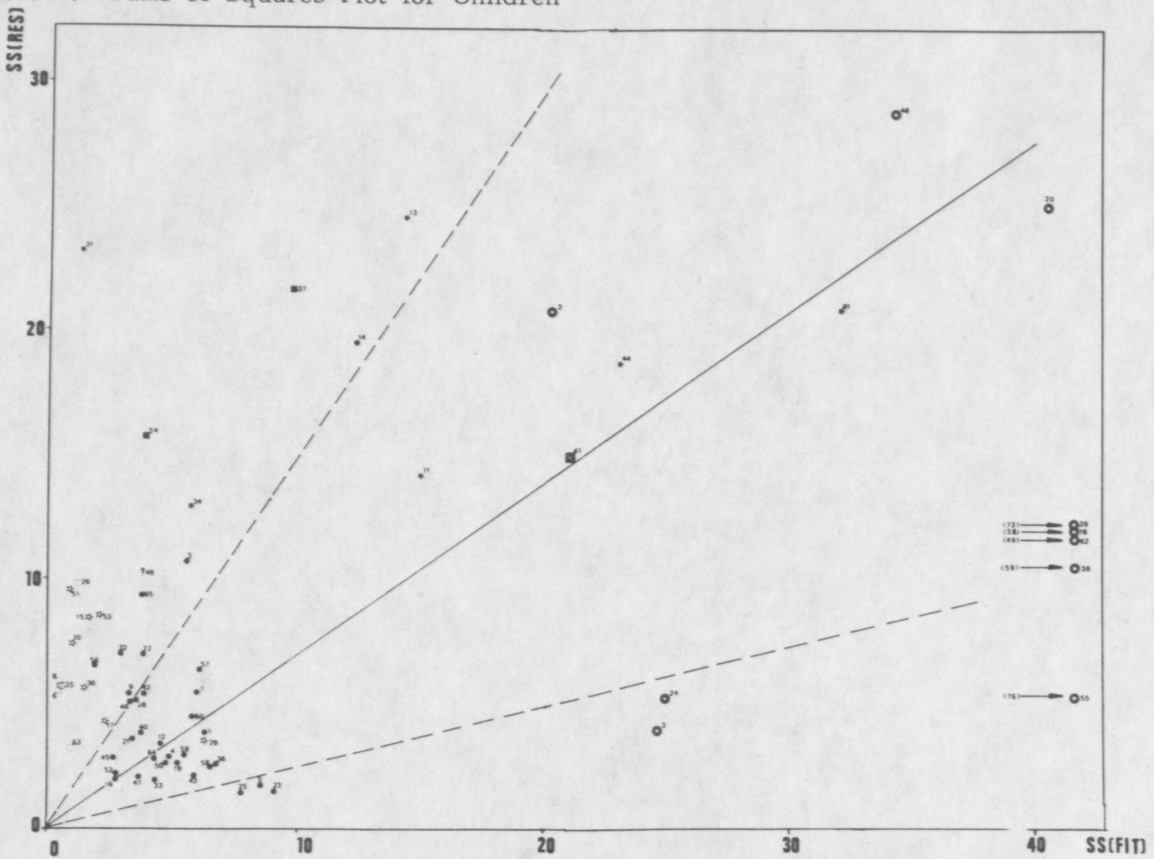
have been equalized is evident from the arrangement of the elements on a line at an angle of -45° to the positive x-axis. Note that because the axes represent sums of squares, the total sums of squares are obtained by directly adding the x-value and the y-value (according to the city-block metric).

Another interesting feature of these plots is that they show which elements have equal residual sums of squares with different total sums of squares. In other words, it becomes possible to separate points that have large residual sums of squares (because they do not fit well) from the points that have a large SS(Res) (because they have a large total sum of squares). Without a residual analysis, it is uncertain whether a point in the middle of a configuration on the first principal components is an ill-fitting point or just a point with little overall variation.

Finally, by drawing the line through (0,0) and (av.SS(Fit), av.SS(Res)), and appropriate similar lines above and below it, something similar to confidence bands can be constructed around the former line to assess the extremity of certain elements. The lines are the loci of points with equal relative residual sums of squares. A guide line for what is "appropriate" in this case—for instance, how much the individual element may deviate in relative residual sum of squares from the overall sum of squares—has not been developed yet.

A number of features should be noted for the present data. The $B4$-children fit well, have large sums of squares, and dominate the solution. (For a more detailed discussion of these $B4$-children and the relation to the other B-groups, see Van IJzendoorn et al. 1983.) Furthermore, there is a large group of $B3$-children that have small total sums of squares; thus they score about average on all scales and most of their variation is fitted well. Conversely, none of the $B1$- and $B2$-children fit very well into the overall pattern, but we have to remember that there are only few of them. Their total sums of squares are not very large, but their relative residual sums of squares are.

Figure 3–7. Sums-of-Squares Plot for Children



Finally, there is a number of peculiar children, and one should hesitate to draw definitive conclusions about them without further analysis. They couple considerable sums of squares with little fit, indicative of either another organization of the scale and episode relationships or of large amounts of random variation. In fact, the two *A*1-children (37 and 54) belong to this group.

Discussion

Keeping in mind the preliminary character of the data, there are some conclusions that can be drawn with respect to the example. In the first place, we note that three-mode principal component analysis succeeded in showing individual differences between the children and characterizing the kind and degree of these differences. Furthermore, the analysis presented here supports to a large degree the *consistency* of the classification procedures, as described by Ainsworth et al. (1978), especially for the *B*-children. The consistency follows from the grouping of children belonging to the same category. The presence of only two *A*-children and a single *C*-child precludes any serious statements about these classification categories, apart from the observation that their position in the child space (Figure 3–4) agrees with what one would expect, but this might be accidental.

We noted earlier the presence of two groups of *B*3-children. In Figure 3–4, they were labeled *B3-prox* and *B3-dist*. The classification instructions in Ainsworth et al. (1978, 61) for *B*3-children (see also Swaan and Goossens 1982) also suggest that there are two types of *B*3-children: those who actively seek physical contact with their mothers (*B*3-prox) and those who seem especially "secure" in their relationship with their mother and are thus content with mere interaction from a distance without seeking to be held (*B*3-dist). It is possibly due to the greater capabilities of communicating at a distance with the mother by two-year-olds that there are more children in the *B*3-dist than in the *B*3-prox group in Goossens' sample. For one-year-olds, the reverse seems to be true. (See Goossens, forthcoming, for further details.)

In Table 3–12, the characterizations of the children (derived from Figure 3–6) occupying the extremes of the axes in Figure 3–4 (child space) are presented. A comparison of this table with Table 3–2 shows global agreement and disagreement in detail. The most conspicuous differences are related to resistance and distance interaction. The comparison for resistance is probably biased by the absence of extremely resistant *C*-children; "high resistance" in Goossens' sample might be average when compared to the resistant behavior of *C*-children. The differences between distal behaviors are, of course, related to the age differences.

A number of problems remain. One is the low number of *A*-children compared to the number found in samples of one-year-old children. One possible explanation for this might be the less avoidant behavior of two-year-old children (see Goossens, forthcoming).

Another potential problem is the ill-fitting *B*1- and *B*2-children. Two reasons may explain this situation. First of all, these children have approximately average scores on all scales,

TABLE 3-12. Comparison of Ainsworth and TUCKALS Classifications (behavior toward the mother)

	Ainsworth*					TUCKALS**				
	PROX	CM	RES	AVOI	DI	PROX	CM	RES	AVOI	DI
A1	-	-	-	++	-	o	o	H	H	L
B3-dist						o	o	L	L	H
B3-prox	++	++	-	-	-	H	H	o	o	L
B4	++	++	(+)	-	-	HH	HH	o	o	LL

*For the Ainsworth classification: - = low; (+) = low to moderate; + = moderate; +(+) = moderate to high; and ++ = high.

**For the TUCKALS classification: LL = low; L = low to average; o = average; H = average to high; and HH = high.

so we are trying to fit their individual error rather than any meaningful variation. Another reason may be that their way of reacting to the Strange situation cannot be fitted very well together with the other children. The small number of these children may preclude finding a separate dimension especially for them. Clearly these conjectures are topics for further investigation.

19

In discussing three-mode principal component analysis, we have attempted to present as full and detailed an account as possible within the limited context of one chapter. In Kroonenberg (1983a), a more detailed account has been given of many of the points raised here. Although we have concentrated on the programs TUCKALS2 and TUCKALS3, which were developed by Kroonenberg and de Leeuw (1980) to illustrate the technique, many of the issues raised have wider relevance. Crucial to the whole approach, however, is the least-squares formulation of the method to solve the estimation of the parameters in the model. It is this formulation that allows the explanation of the core matrix in terms of amount of explained variation and makes the separation of the fitted and residual sums of squares possible.

REFERENCES

Ainsworth, M. D. S., M. C. Blehar, E. Waters, and S. Wall. 1979. *Patterns of attachment. A psychological study of the strange situation.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bentler, P. M., and S-Y. Lee. 1978. Statistical aspects of a three-mode factor analysis model. *Psychometrika* 43:343–52.

———. 1979. A statistical development of three-mode factor analysis. *British Journal of Mathematical and Statistical Psychology* 32:87–104.

Benzecri, J. P., et al. 1973. *Analyse des donnees. II Analyse des correspondences.* Paris: Dounod.

Bloxom, B. 1968. *Individual differences in multidimensional scaling (ETS Research Bulletin 68–45).* Princeton, N.J.: Educational Testing Service.

Carroll, J. D. 1972. Individual differences and multidimensional scaling. In *Multidimensional scaling: Theory and applications in the behavioral sciences*, vol. 1, ed. R. N. Shepard, A. K. Romney, and S. Nerlove. New York: Seminar Press.

Carroll, J. D., and P. Arabie. 1980. Multidimensional scaling. *Annual Review of Psychology* 31:607–49.

Carroll, J. D., and J. J. Chang. 1970. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika* 35:283–319.

————. 1972. IDIOSCAL (individual differences in orientation scaling): A generalization of INDSCAL allowing idiosyncratic reference systems as well as an analytic approximation to INDSCAL. Paper presented at the spring meeting of the Psychometric Society, Princeton, N.J., March, 1972.

Carroll, J. D., and M. Wish. 1974. Models and methods for three-way multidimensional scaling. In *Contemporary developments in mathematical psychology*, vol. 2, ed. D. H. Krantz, R. C. Atkinson, R. D. Luce, and P. Suppes. San Francisco: W. H. Freeman.

Cattell, R. B. 1966. The data box: Its ordering of total resources in terms of possible relational systems. In *Handbook of multivariate experimental psychology*, ed. R. B. Cattell. Chicago: Rand McNally.

Cook, T. D., and D. T. Campbell. 1979. *Quasi-experimentation. Design and analysis issues for field settings*. Chicago: Rand McNally.

Dunn, T. R., and R. A. Harshman. 1982. A multidimensional scaling model for the size-weight illusion. *Psychometrika* 47:25–45.

Einhorn, H. J. 1972. Expert measurement and mechanical combination. *Organizational Behavior and Human Performance* 7:86–106.

Gabriel, K. R. 1971. The biplot graphical display of matrices with applications to principal components. *Biometrika* 58:453–67.

Gifi, A. 1981. *Non-linear multivariate analysis*, preliminary edition. Leiden, The Netherlands: University of Leiden, Department of Data Theory.

Glass, G. V., V. L. Willson, and J. M. Gottman. 1975. *Design and analysis of time-series experiments*. Boulder, Colo.: Colorado Associated University Press.

Goossens, F. A. Forthcoming. The quality of the attachment relationship for two-year-old children of working and nonworking mothers and some associated factors. Dissertation in preparation, University of Leiden.

Goossens, F. A., J. Swaan, L. W. C. Tavecchio, M. M. Vergeer, and M. H. Van IJzendoorn. 1982. *The quality of attachment assessed (WEP-Reeks, WR 82-23-LE)*. Leiden, The Netherlands: Rijksuniversiteit Leiden.

Grossman, K. E., K. Grossman, F. Huber, and U. Wartner. 1981. German children's behavior towards their mother at 12 months and their father at 18 months in Ainsworth's strange situation. *International Journal of Behavioral Development* 4:157–81.

Harshman, R. A. 1972. PARAFAC2: Mathematical and technical notes. *UCLA Working Papers in Phonetics* 22:30–44 (University Microfilms No. 10,085).

Jennrich, R. A. 1972. A generalization of the multidimensional scaling model of Carroll & Chang. *UCLA Working Papers in Phonetics* 22:45–47.

Jöreskog, K. G., and D. Sörbom. 1976. Some models and estimation methods for analysis of longitudinal data. In *Latent variables in socioeconomic models,* ed. D. J. Aigner and A. S. Goldberger. Amsterdam, The Netherlands: North Holland.

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

Kroonenberg, P. M. Forthcoming. Scaling of input data for three-mode principal component analysis. Manuscript in preparation.

———. 1981a. *User's guide to TUCKALS3. A program for three-mode principal component analysis (WEP-Reeks, WR-81-6-RP).* Leiden, the Netherlands: Rijksuniversiteit Leiden.

———. 1981b. *User's guide to TUCKALS2. A program for three-mode principal component analysis with extended core matrix (WEP-Reeks, WR-81-35-RP).* Leiden, the Netherlands: Rijksuniversiteit Leiden.

———. 1983a. *Three-mode principal component analysis: Theory and applications.* Leiden: DSWO Press.

20

———. 1983b. Annotated bibliography of three-mode factor analysis. *British Journal of Mathematical and Statistical Psychology* 36:81–113.

Kroonenberg, P. M., and J. de Leeuw. 1980. Principal component analysis of three-mode data by means of alternating least-squares algorithms. *Psychometrika* 45:69–97.

Kruskal, J. B. 1976. More factors than subjects, tests and treatments: An indeterminacy theorem for canonical decomposition and individual differences scaling. *Psychometrika* 41:281–93.

———. 1977. Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra and its Applications* 18:95–138.

Lammers, C. J. 1974. *Groei en ontwikkeling van de ziekenhuisorganisaties in Nederland.* Leiden, The Netherlands: Rijksuniversiteit Leiden, Sociologisch Instituut.

Law, H. G., and C. W. Snyder, Jr. 1981. An introduction to the analysis of covariance structures: A general model for data analysis. In *Proceedings of a seminar on measuring social behavior in road research,* ed. J. M. Morris. Vermont South, Australia: Australian Road Research Board.

Lohmöller, J. B. 1981. Stabilität und Kontinuität in Längsschnittdaten, analysiert durch T- und trimodale Faktorenanalyse. Technical Report. Fachbereich Pädagogik, Hochschule der Bundegwehr München, Neubiberg, FRG.

Osgood, C. E., G. J. Suci, and P. H. Tannenbaum. 1957. *The measurement of meaning.* Urbana: University of Illinois Press.

Sands, R., and F. W. Young. 1980. Component models for three-way data: An alternating least-squares algorithm with optimal scaling features. *Psychometrika* 45:39–67.

Sawyer, J. 1966. Measurement *and* prediction, clinical *and* statistical. *Psychological Bulletin* 66:178–200.

Schiffman, H., and P. Falkenberg. 1968. The organization of stimuli and sensory neurons. *Physiology and Behavior* 3:197–201.

Schiffman, S. S., M. L. Reynolds, and F. W. Young. 1981. *Introduction to multidimensional scaling: Theory, methods, and applications*. New York: Academic Press.

Snyder, F. W., and N. Wiggins. 1970. Affective meaning systems: A multivariate approach. *Multivariate Behavioral Research* 5:453–68.

Sroufe, L. A., and E. Waters. 1977. Attachment as an organizational construct. *Child Development* 48:1184–99.

Swaan, J., and F. A. Goossens. 1982. *Handleiding bij de strange situation (WEP-Reeks, WR 82-24-EX)*. Leiden, The Netherlands: Rijksuniversiteit Leiden.

Takane, Y., F. W. Young, and J. de Leeuw. 1977. Non-metric individual differences multidimensional scaling: An alternating least-squares method with optimal scaling features. *Psychometrika* 42:7–67.

Tucker, L. R. 1963. Implications of factor analysis of three-way matrices for measurement of change. In *Problems in measuring change,* ed. C. W. Harris. Madison, Wis.: University of Wisconsin Press.

———. 1964. The extension of factor analysis to three-dimensional matrices. In *Contributions to mathematical psychology,* ed. N. Frederiksen and H. Gulliksen. New York: Holt, Rinehart and Winston.

———. 1966. Some mathematical notes on three-mode factor analysis. *Psychometrika* 31:279–311.

———. 1972. Relations between multidimensional scaling and three-mode factor analysis. *Psychometrika* 37:3–27.

Tucker, L. R., and S. Messick. 1963. An individual differences model for multidimensional scaling. *Psychometrika* 28:333–67.

Van de Geer, J. P. 1974. Toepassing van drieweg-analyse voor de analyse van multiple tijdreeksen. In *Groei en ontwikkeling van de ziekenhuis organisaties in Nederland,* ed. C. J. Lammers. Leiden, The Netherlands: Rijksuniversiteit Leiden, Sociologisch Instituut.

Van der Kloot, W. A., and P. M. Kroonenberg. 1982. Group and individual implicit theories of personality. An application of three-mode principal component analysis. *Multivariate Behavioral Research*. 17:471–92.

Van der Kloot, W. A., and T. Van den Boogaard. 1978. *Weights of information in impression formation*. (Report MT 001–78): Leiden, The Netherlands.

Van IJzendoorn, M. H., L. W. C. Tavecchio, F. A. Goossens, M. M. Vergeer, and J. Swaan. 1983. How B is B4? Attachment and security of Dutch children in Ainsworth's strange situation and at home. *Psychological Reports* 52:683–91.

Waters, E. 1978. The reliability and stability of individual differences in infant-mother attachment. *Child Development* 49:483–94.

Young, F. W. 1981. Quantitative analysis of qualitative data. *Psychometrika* 46:357–88.