Computational Statistics & Data Analysis 18 (1994) 39–72
North-Holland

PARAFAC: Parallel factor analysis

Richard A. Harshman and Margaret E. Lundy
Department of Psychology, University of Western Ontario, London, Ontario, Canada

Abstract: We review the method of Parallel Factor Analysis, which simultaneously fits multiple two-way arrays or 'slices' of a three-way array in terms of a common set of factors with differing relative weights in each 'slice'. Mathematically, it is a straightforward generalization of the bilinear model of factor (or component) analysis ($x_{ij} = \Sigma_{r=1}^{R} a_{ir} b_{jr}$) to a trilinear model ($x_{ijk} = \Sigma_{r=1}^{R} a_{ir} b_{jr} c_{kr}$). Despite this simplicity, it has an important property not possessed by the two-way model: if the latent factors show adequately distinct patterns of three-way variation, the model is fully identified; the orientation of factors is uniquely determined by minimizing residual error, eliminating the need for a separate 'rotation' phase of analysis. The model can be used several ways. It can be *directly* fit to a three-way array of observations with (possibly incomplete) factorial structure, or it can be *indirectly* fit to the original observations by fitting a set of covariance matrices computed from the observations, with each matrix corresponding to a two-way subset of the data. Even more generally, one can simultaneously analyze covariance matrices computed from different samples, perhaps corresponding to different treatment groups, different kinds of cases, data from different studies, etc. To demonstrate the method we analyze data from an experiment on right vs. left cerebral hemispheric control of the hands during various tasks. The factors found appear to correspond to the causal influences manipulated in the experiment, revealing their patterns of influence in all three ways of the data. Several generalizations of the parallel factor analysis model are currently under development, including ones that combine parallel factors with Tucker-like factor 'interactions'. Of key importance is the need to increase the method's robustness against nonstationary factor structures and qualitative (nonproportional) factor change.

Keywords: Three-way exploratory factor analysis; Unique axes; Parallel proportional profiles; Factor rotation problem; Three-way data preprocessing; Three mode principal components; Trilinear decomposition; Trilinear model; Multidimensional scaling; Longitudinal factor analysis; Factor analysis of spectra; Interpretation of factors; 'Real' or causal or explanatory factors; L.R. Tucker; R.B. Cattell.

1. Introduction

Parallel factor analysis extends the ideas and methods of standard two-way factor analysis to three-way data. A key motivation for this is the possibility that

Correspondence to: Dr. R.A. Harshman, Department of Psychology, University of Western Ontario, London, Ont., Canada, N6A 5C2.

SSDI 0167-9473(92)00004-N

'simultaneous factor analyses' of several matrices in parallel can sometimes lead to a unique set of factors for which any rotation of axes will cause a loss in total fit. This is similar to the way in which solving simultaneous equations allows the family of possible solutions for any single equation to be reduced to a unique solution for each unknown that fits all the equations in parallel.

The original inspiration for PARAFAC came from R.B. Cattell, who argued that the rotational indeterminacy of factor analysis can be overcome by seeking "Parallel Proportional [factor] Profiles" in two solutions at once (Cattell, 1944; Cattell and Cattell, 1955). He reasoned as follows: if the same factors are present in two different datasets, but change their relative proportions of variance-accounted-for by distinct amounts from one dataset to the next, then there is only one *unique* set of axis orientations in both spaces which will reveal this parallel proportional relationship; hence, by discovering that unique position, one can *empirically* determine the approximate orientation that the factors must have had when the data were generated (cf. Cattell, 1978, Section 6.3).

2. Some PP models of three-way variation

This special issue of CSDA describes many different methods of multi-way data analysis. To avoid repetition of points made by other authors, we omit a general literature review and concentrate directly on our family of Parallel Profiles multi-way models and analysis procedures (Harshman, 1970; Harshman and Lundy, 1984a); the simplest of these is PARAFAC1 (also known as PARAFAC).

PARAFAC1 is based on the trilinear model given in Equation 3 of Section 2.1. Harshman (1970) developed it by generalizing Cattell's (1944) idea of Parallel Profiles factor rotation. Carroll and Chang (1970) independently arrived at the same trilinear model by generalizing Horan's (1969) work on individual differences in multidimensional scaling. They called the trilinear model 'CANDECOMP,' for 'CANonical DECOMPosition,' and developed from it the widely used MDS (MultiDimensional Scaling) procedure INDSCAL. Other important early contributors include, for example, Bloxom (1968), and Sands and Young (1980).

Considered by itself, the trilinear model would seem to be somewhat restrictive: it requires ratio scale data, and, in particular, systematic proportional changes from one level to the next in the pattern of influences of each latent factor in each of *three* ways (or 'modes' [1]) of a multimode data set (Harshman and Lundy, 1984a, p. 130). Yet even in this strongest form, the model has found

---

[1] Tucker (1963,1964) coined the term "mode" to reflect the fact that each way of a multiway dataset can represent a different "mode of measurement" or "mode of classification" of the data. Carroll and Arabie (1980) introduced a distinction between 'n-way,' which was the number of subscripts for the datapoint, and 'n-mode,' which was the number of differently labeled subscripts (i.e., the number of *distinct* 'modes of classification' of the data). Under their interpretation, a symmetric correlation matrix would be two-way but have only one mode, repeated for both ways.

a growing number of interesting applications, particularly in the 'harder' sciences, such as chemometrics (Burdick et al., 1990; Geladi, 1989; Ross et al., 1991; Smilde and Doornbos, 1991). *Statistical Science* recently published a useful expository introduction and overview of applications to spectroscopy (Leurgans and Ross, 1992; see also the accompanying commentaries by Burdick et al., and Kroonenberg). In such applications, the existence of 'true' factors is taken for granted, and the ability of the trilinear (PARAFAC–CANDECOMP) model to uncover such factors is considered an important advance.

Most data are not initially ratio-scale; variables may have arbitrary zero points and/or standard deviations. But these can be removed by array preprocessing methods discussed in Section 3.2.1. Even greater flexibility can often be obtained by applying PARAFAC to covariances or cross-products rather than raw scores. Some examples of applications using array preprocessing and/or indirect fitting methods include Cole and Ray (1985), Field and Gaupe (1991), Haan (1981), Harshman and DeSarbo (1984), Harshman et al. (1977), Kettenring (1983), Meyer (1980), Ossenkopp and Mazmanian (1985), Snyder et al. (1983) and Trick and Katz (1986).

2.1. Derivation of the basic trilinear factor analysis model

The trilinear model underlying PARAFAC or CANDECOMP may be the simplest nontrivial way to extend the algebra of the factor analysis model from two-way to three-way data. To see this, it is helpful to start with the scalar form of the two-way model:

$$x_{ij} = a_{i1} f_{j1} + a_{i2} f_{j2} + \cdots + a_{ir} f_{jr} + \cdots + a_{iR} f_{jR} + e_{ij}, \tag{1}$$

where x_{ij} is the entry in row i and column j of the I by J data matrix $\boldsymbol{X}$, and the terms $a_{i1} f_{j1}, \ldots, a_{ir} f_{jr}, \ldots, a_{iR} f_{jR}$ represent the additive contributions of factors 1 through R to this observed value. The e_{ij} is an error term or residual; its hypothesized properties depend on whether one is fitting the principal components or the common factor model (see Section 2.3.4).

Recall that x_{ij} is often obtained by measuring attribute i on entity j; in this case the weights a_{ir} are traditionally called factor 'loadings' while the weights f_{jr} are called factor 'scores'. These two sets of weights traditionally have different standardization conventions and often different methods of estimation. Fundamentally, however, both sets of weights play the same role: to represent the amount of a factor's influence or importance at a given level of a given mode of the data. Furthermore, both sets of weights represent variations in a factor's influence multiplicatively – by proportionally stepping up or down the size of all of a factor's contributions from one level of a particular mode to the next.

Because the same basic role is played by both sets of weights, we consider the tradition of distinctive naming and standardizing conventions for each set to be overly specialized. In some modern applications (e.g., chemistry) it is unclear which set of weights should be called 'factor loadings' and which 'factor scores.'

Therefore, we adopt a more neutral and general set of mode naming, weight naming, and standardizing conventions. We say that X is a two-way matrix in which the rows represent levels of 'Mode A' and the columns represent levels of 'Mode B.' The two-way model is then written:

$$x_{ij} = a_{i1}b_{j1} + a_{i2}b_{j2} + \cdots + a_{ir}b_{jr} + \cdots + a_{iR}b_{jR} + e_{ij}. \tag{2}$$

Standardization (e.g., to column mean squares of 1) is applied to one set of weights, and the scale of the data then fixes the sizes in the other set. The choice of which mode to standardize depends on the particular data involved and the analysis perspective adopted. The term 'loadings' can be applied loosely to either set of weights. In the mode that reflects the scale of the data, stronger interpretations of the variations in size of loadings are often possible. (For detailed discussion, see Harshman and Lundy, 1984a, pp. 192–203).

The two-way factor model is *bilinear*, since it is linear in one set of weights (e.g., the Mode A loadings) if the other set of weights (e.g., the Mode B loadings) are considered fixed, and vice versa. To generalize (2) into the three-way PARAFAC–CANDECOMP model, one simply adds a third set of weights. Weights in this third set play the same sort of role in the model as those in the other two sets. That is, they proportionally step up or down the contribution of particular factors at each level of a particular way or 'mode of measurement' of the dataset (in this case, the third way or mode); hence the generalized model is *trilinear*. The third way is called Mode C, and so the generalized form of the model becomes, in scalar notation,

$$x_{ijk} = a_{i1}b_{j1}c_{k1} + a_{i2}b_{j2}c_{k2} + \cdots + a_{ir}b_{jr}c_{kr} + \cdots + a_{iR}b_{jR}c_{kR} + e_{ijk}. \tag{3}$$

Here, x_{ijk} is an entry in a three-way array $\underline{X}$ with Modes A, B and C. For example, it might be the value of variable i as measured on entity j on occasion k. The a_{ir} gives the weight or loading of factor r on level i of Mode A; b_{jr} and c_{kr} give the weight or loading of the same factor on level j of Mode B and level k of Mode C, respectively; e_{ijk} is the residual or error term.

Once we view the weights for all modes in a uniform manner, it becomes natural to consider further *multilinear* generalizations incorporating additional sets of weights:

$$\begin{aligned} x_{ijkl\cdots} &= (a_{i1}b_{j1}c_{k1}d_{l1}\cdots) + (a_{i2}b_{j2}c_{k2}d_{l2}\cdots) + \cdots + (a_{ir}b_{jr}c_{kr}d_{lr}\cdots) \\ &\quad + \cdots + (a_{iR}b_{jR}c_{kR}d_{lR}\cdots) + e_{ijkl\cdots}. \end{aligned} \tag{4}$$

Carroll and Chang (1970) and Harshman (1970) are among those originally proposing such generalizations; Carroll and Chang actually developed a computer procedure with 7-way generality, but the computational inefficiency of carrying the extra subscripts when they were not needed led to the development of a version specifically for analysis of symmetric, 3-way data (SINDSCAL, see Pruzansky, 1975 as cited in Arabie et al., 1987). In this article, we will limit ourselves to consideration of three-way arrays.

2.2. *Dimensional uniqueness and the 'intrinsic-axis property'*

Our strongest motivation for developing and using three-way Parallel Profiles models is the potential for overcoming the 'rotation problem' of traditional factor analysis, at least in some circumstances. It is well known that the two-way model is underdetermined by the two-way data that it fits. Instead of calling upon additional criteria (such as 'simple structure'), to select among the multiplicity of solutions, one can collect stronger data – with systematic three-way variation – and fit the trilinear model. Given appropriate three-way system variation [2], this model is *not* undetermined by the data.

2.2.1. *Proofs*

Several mathematical proofs of the uniqueness of PARAFAC–CANDECOMP axes have been derived, each based on a different set of assumptions concerning the data (DeLeeuw and Pruzansky, 1978, p. 483; Harshman, 1972; Jennrich, in Harshman, 1970). The most general proof to date of the uniqueness of the (trilinear) model is due to Kruskal (1976, 1977). (PARAFAC uniqueness also follows as a special case of the uniqueness of the more general PT2 model, see Section 5.2.)

The implications of these uniqueness theorems can be briefly summarized as follows: PARAFAC factor axes will be uniquely oriented if, in each of at least three modes or ways of an *n*-way dataset, each factor has a pattern of variation in 'influence' (i.e., in size of weights) distinct from all other factors. 'Distinct' means that, for any two factors, we must be able to find two slices of the array in which the percentage increase or decrease between slices in the one factor's loadings is different from the percentage increase or decrease in the other's.

Mathematical study of the uniqueness of trilinear decompositions is a continuing area of research interest (e.g., Kruskal, 1989; Leurgans and Ross, 1992; Ten Berge, 1991; Ten Berge et al., 1988). An informal international symposium on rank, uniqueness, and related properties of three-way arrays and their trilinear (PARAFAC–CANDECOMP type) decompositions (focusing, in particular, on Kruskal's recent results) was held in Groningen, The Netherlands, in August, 1991; as of this writing, however, the proofs and discussion are only available by contacting the participants.

2.2.2. *Interpretation of uniqueness*

Uniqueness *per se* is not necessarily empirically meaningful or informative. For example, the uniqueness of the Singular Value Decomposition and (unrotated) Principal Components Analysis is obtained by imposing restrictions (of compo-

---

[2] 'System Variation' is coordinated proportional changes in the contributions of a given factor from one level to the next in at least three 'ways' of an *n*-way array. It is contrasted to 'Object Variation', in which the factors change proportionally in two modes (such as Variables and Persons) but not in the third mode (e.g., Occasions). In Object Variation, each object has its own idiosyncratic pattern of variations in the importance of factors across levels of the third mode (see Harshman, 1970, pp. 20–22). See also Sections 2.3.2.1 and 4.4.1.1.

nent orthogonality and/or successive variance maximization) that are mathematically convenient but empirically implausible, or, at best, arbitrary. Thus, there is no reason to expect the unique axes so determined to be oriented in the factor space so as to best correspond to empirical processes that generated the data. There are several conditions that, if fulfilled, make it more likely that uniqueness implies probable identification of the most scientifically meaningful orientation of factor axes in a given factor space. According to these principles the uniqueness should be:

1. *Well Motivated.* The uniqueness determining conditions should be non-arbitrary – they should be empirically motivated by known or anticipated traces or characteristic effects of the causal processes or 'organic unities' that underlie the systematic patterns in the data. Mathematically, the assumptions, restrictions, or characteristics of the formal model that determine uniqueness should do so in a way that is highly plausible empirically/scientifically for the data under study.

2. *Testable.* It should be possible to determine whether the data display enough of the uniqueness determining characteristics (the particular systematic characteristics that the formal model depends on to determine location of the factor axes) to in fact stably determine a meaningful set of axes. It should be possible for some datasets to fail this requirement and so for some solutions to be non-unique. Further, is should be possible to test a given solution/dataset and determine whether it is reliably unique or not.

3. *Univocal.* If a dataset fulfills the conditions allowing a particular model to define a set of unique axes, this should be subject to only one plausible interpretation – that this is the location of the axes most likely to correspond to the causal or organic unities underlying the data. There should be no plausible alternative interpretation of the result. In short, any unique axes obtained should be quite hard to explain scientifically *other than* by postulating that the factors have identified influences acting in the data and that the influences' direction of action and patterns of relative weights are roughly as represented in the factor analysis results.

When we use these principles to evaluate previous criteria for orienting factor axes (details omitted here due to space limitations), we find: (a) Thurstone's (1947, chapter 14; see also Cattell, 1978, Sections 6.4–6.7) rationale for the 'simple structure' search for hyperplanes is empirically well motivated, testable (with effort), and relatively univocal. When hyperplanes are well defined, the rotation they determine (e.g., graphically) is likely to place axes approximately in the direction of action of 'real' empirical influences. (b) VARIMAX, OBLIMIN, and related analytic rotation methods are more problematic, however, because their 'simplicity' criterion is flawed: it can be optimized by finding clusters as well as by hyperplanes. In studies with careful sampling of variables, analytic rotation will generally (though not always) tend to find hyperplanes, but in less well designed studies, these methods will often simply reflect the researcher's preconceptions back to him by placing factors through those clusters of similar or highly related variables that were built in at the start of the study.

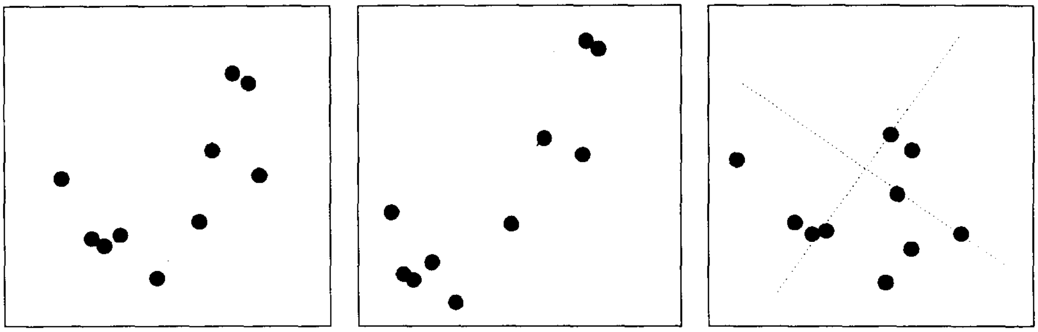

Fig. 1.

2.2.2.1. Evaluation of PARAFAC uniqueness PARAFAC axis orientation is not influenced by clusters, hyperplanes, or any other within-space configurations. It relies instead on finding systematic relationships between several parallel factor spaces.

1. *Motivation*. PARAFAC determines unique axis orientation by seeking out systematic proportional stretching or contraction of distances (from one two-way subset of the array to the next) along a consistent set of directions in the factor space (Figure 1). Such systematic proportional variations in inter-point distances along particular directions should only be present if there are, in fact, common influences each acting on a particular set of several variables in the space. These influences can be detected as distinct factors if the underlying influences change their importance relative to one another from one two-way slice to the next.

2. *Testability*. A standard part of recommended PARAFAC analysis procedure is to test the reliability of the axis orientations determined by PARAFAC using split-half or other methods (e.g., see Section 4.4.3). If the reliability test fails, PARAFAC factors cannot be trusted. If it succeeds, the axis agreement across split halves would seem to demonstrate that systematic expansions or contractions of interpoint distances are *reliably* found in particular directions in the factor space.

Other tests are also possible. For example, to test for required qualitative consistency of the factor structure across groups, ages, conditions, etc. one can often compare separate analyses of systematically different subsets of the data (e.g. young vs. older, condition 1 vs. 2, etc.).

3. *Univocality*. If proportional stretches and contractions of interpoint distances in the factor space are demonstrated by consistent split-half PARAFAC factors, it is hard to see how this three-way regularity could arise *except* through a common influence oriented along each direction of consistent stretching and contraction. However, is not such a common influence on several variables more or less what we mean by a 'common factor'?

Of course, it is always possible that someone will, in the future, come up with a persuasive alternate explanation of the systematic stretches and contractions, and/or of the replicating PARAFAC unique axes in a particular dataset. However, this is true of any scientific explanation. In the meantime, such a vague speculative 'possibility' does little to diminish the present fact that there is only one apparent explanation for the replicable unique axes in the dataset, and that explanation seems scientifically sensible and well motivated by what we know of the data.

2.2.2.2. Intrinsic axis property For PARAFAC–CANDECOMP and related parallel-profile factor models, finding the optimal orientation of the factor axes is an *intrinsic part* of optimizing the fit of the trilinear factor model to three-way data; if the factors have shown distinct patterns of change, then any other axis orientations will reduce the fit of the model to the data. Thus, such models are said to have an *intrinsic axis property*. (For further discussion, see e.g., Harshman and Lundy, 1984a, pp. 147–169, particularly 163–167.) This name distinguishes the kind of uniqueness of PP models from the more arbitrary uniqueness of certain other multidimensional models, such as principal components analysis. It also distinguishes them from models which are themselves intrinsically nonunique, but which determine unique solutions by adopting added side conditions, such as maximizing a simple structure criterion in the factor loading matrix.

2.3. Direct and indirect fitting

In the three-way case it becomes particularly important to distinguish two different approaches to fitting the factor model to data: (a) to directly fit a raw score or profile data matrix, or (b) to fit a matrix of covariances or cross-products (or pseudo scalar products) derived from the raw data. Following Kruskal (1978), we refer to the first as *direct fitting* and to the second as *indirect fitting* of the structure in a data matrix (see also Harshman, 1972; Harshman and Lundy, 1984a, pp. 133–140; Snyder et al., 1984, pp. 9–14).

Although these approaches yield equivalent factors in the two-way case (unless diagonal estimation is performed), they are not, in general, equivalent in the three-way case. (One special case in which they are equivalent arises when the dataset is a fully crossed three-way array and the columns of factor weights are orthogonal in at least two ways of the array.) Direct and indirect fitting each have their own distinct advantages and disadvantages.

2.3.1. Direct fitting

As described in Section 2.1, PARAFAC can directly fit the trilinear model to three-way 'raw data' (or, equivalently, centered and standardized raw score data). The scalar form of the model is given in (3); a matrix formulation is often written by giving a general expression for an arbitrary two-way slice:

$$\boldsymbol{X}_k = \boldsymbol{A}\boldsymbol{D}_k\boldsymbol{B}' + \boldsymbol{E}_k, \tag{5}$$

where $\boldsymbol{X}_k$ is I by J, the kth slice of an I by J by K array $\underline{\boldsymbol{X}}$; $\boldsymbol{A}$ is an I by R factor loading matrix for Mode A; $\boldsymbol{B}$ is a J by R loading matrix for Mode B; $\boldsymbol{D}_k$ is an R by R diagonal matrix with diagonal elements taken from the kth row of $\boldsymbol{C}$, a K by R loading matrix for Mode C; and $\boldsymbol{E}_k$ is the kth slice of the residual array $\underline{E}$. A Kronecker or tensor-product form is given in the article by R. Coppi elsewhere in this issue.

One advantage of direct fitting is that it does not require orthogonal factors in any mode of the data, since the loading matrices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ need not be columnwise orthogonal. Another advantage is that the analysis directly provides weights or loadings for all three modes at once. A disadvantage of direct fitting is that it makes stricter assumptions about the three-way proportionality of the pattern of factor variation in the data, requiring System Variation [2]. Another disadvantage is that (because of the characteristics just mentioned) direct fitting is subject to 'degenerate solutions' (see Section 5.5), whereas indirect fitting (using PARAFAC1) appears imune to this problem.

2.3.2. Indirect fitting

2.3.2.1. Single sample indirect fitting If the algebraically identical trilinear model is applied to raw data *indirectly*, by using its trilinear structure to represent parallel cross-product or covariance matrices derived from raw data, it is not necessary that the raw data show System Variation, but only that the *covariances* have such structure, a condition usually easier to fulfill. For covariance matrices, it is necessary simply that the average importance of each factor varies from one covariance matrix to the next in a distinct fashion.

To indirectly fit the raw data array $\underline{X}$, one might, for example, compute covariances across Mode B at each level of Mode C, to obtain an I by I by K three-way array $\underline{Y}$ in which the kth slice consists of covariances among the I variables on occasion k. In matrix form the model becomes

$$\boldsymbol{Y}_k = \boldsymbol{A}\boldsymbol{D}_k^2\boldsymbol{A}' + \boldsymbol{E}\boldsymbol{E}_k', \tag{6}$$

where $\boldsymbol{Y}_k$ is I by I, the kth covariance matrix in an I by I by K array $\underline{Y}$; $\boldsymbol{A}$ and $\boldsymbol{D}_k$ are as defined in Section 2.3.1 for the direct fitting model, and $\boldsymbol{E}\boldsymbol{E}_k'$ is the matrix of residual or error covariances. Here, the second table of factor loadings ($\boldsymbol{B}$ in the direct fitting model) is identical to $\boldsymbol{A}$ as a consequence of the symmetry of the covariance matrices. Further, $\boldsymbol{D}_k^2$ contains the squares of entries obtained in the direct fitting case, where $\underline{X}$ itself is analyzed. Thus the indirect-fitting equivalent of (3) is

$$y_{ijk} = a_{i1}a_{j1}c_{k1}^2 + a_{i2}a_{j2}c_{k2}^2 + \cdots + a_{ir}a_{jr}c_{kr}^2 + \cdots + a_{iR}a_{jR}c_{kR}^2 + e_{(y)ijk}. \tag{7}$$

Note that reference to the pattern of loadings in Mode B disappears from the indirect fitting model. Thus, it is no longer required that the changes in Mode C (e.g., occasions) be proportional across levels of Mode B (e.g., individuals) since Mode B has 'vanished.' This indirect approach was used in a longitudinal personality analysis (Haan, 1981) to analyze the 'object variation' of personality change. Although direct fitting could have been performed, it was deemed too restrictive to assume that the pattern of change in importance for a given personality factor across time points in the life-span was proportional from one person to the next.

2.3.2.2. Multiple sample indirect fitting Alternatively, one can compute each covariance matrix using a different sample of Mode B levels. The generalized model for the raw data (ignoring, for the moment, details regarding the stochastic or error part) would be:

$$\boldsymbol{X}_k \approx \boldsymbol{A}\boldsymbol{D}_k\boldsymbol{B}_k'. \tag{8}$$

To obtain the corresponding covariance model, let

$$\boldsymbol{Y}_k \equiv \frac{1}{J_k}\boldsymbol{X}_k\boldsymbol{X}_k' \approx \frac{1}{J_k}(\boldsymbol{A}\boldsymbol{D}_k\boldsymbol{B}_k')(\boldsymbol{B}_k\boldsymbol{D}_k\boldsymbol{A}') = \boldsymbol{A}\boldsymbol{D}_k\left(\frac{1}{J_k}\boldsymbol{B}_k'\boldsymbol{B}_k\right)\boldsymbol{D}_k\boldsymbol{A}', \tag{9}$$

and let

$$\frac{1}{J_k}\boldsymbol{B}_k'\boldsymbol{B}_k = \boldsymbol{\Phi}_k, \tag{10}$$

where $\boldsymbol{\Phi}_k$ is the correlation (or covariance, or mean-cross-product) matrix for factor weights in Mode B at level k of Mode C.

If one allows $\boldsymbol{\Phi}_k$ to change arbitrarily from one value of k to the next, the PP solution is not unique (Meredith, 1964). One can subject the common factor space to an arbitrary nonsingular linear transformation $\boldsymbol{T}$, and absorb the inverse transformation into the weight and angle matrices.

Let $\overset{*}{\boldsymbol{A}} = \boldsymbol{A}\boldsymbol{T}$, then

$$\begin{aligned}\boldsymbol{A}\boldsymbol{D}_k\boldsymbol{\Phi}_k\boldsymbol{D}_k\boldsymbol{A}' &= \boldsymbol{A}\boldsymbol{T}\boldsymbol{T}^{-1}\boldsymbol{D}_k\boldsymbol{\Phi}_k\boldsymbol{D}_k\boldsymbol{T}'^{-1}\boldsymbol{T}'\boldsymbol{A}' \\ &= \overset{*}{\boldsymbol{A}}\left(\boldsymbol{T}^{-1}\boldsymbol{D}_k\boldsymbol{\Phi}_k\boldsymbol{D}_k\boldsymbol{T}'^{-1}\right)\overset{*}{\boldsymbol{A}}' = \overset{*}{\boldsymbol{A}}\overset{*}{\boldsymbol{D}}_k\overset{*}{\boldsymbol{\Phi}}_k\overset{*}{\boldsymbol{D}}_k\overset{*}{\boldsymbol{A}}_j 7.\end{aligned} \tag{11}$$

However, it is often theoretically plausible that factor intercorrelations will stay more or less the same across the particular set of k in a dataset. If we can require that

$$\boldsymbol{\Phi}_k = \boldsymbol{\Phi} \tag{12}$$

for all k, the axes can be uniquely recovered (Harshman and Lundy, 1991). Note that only after this restriction are we looking for *common directions* of stretch and contraction of the configuration across levels of Mode C. And, if stretches in common directions are reliably detectable in the data, this systematicity of the stretches provides the empirical evidence for a particular nonarbitrary orientation of common factor axes.

Of course, it is often unrealistic to expect *exactly* the same directions of stretch across multiple levels of Mode C. Fortunately, approximate stability of axes is all that is needed for the model to work, and so the assumption of common directions of stretch is often a reasonable simplifying approximation. However, if the variation in axis orientations is too great, then degenerate and/or unstable solutions will result if a 'common direction' simplifying assumption is attempted (see Section 5.5). This stability of axis orientation is an implicit assumption in direct fitting as well, since variations in factor weights are assumed to stay proportional at different levels of each mode.

2.3.3. *Further indirect fitting models*

We conjecture that even more general models, ones that allow $\boldsymbol{\Phi}_k$ to vary in certain constrained ways across different k, will still have unique solutions. In such cases, the empirical evidence for particular orientations of factor axes is modified. The evidence that was previously provided by finding more consistency in the directions of stretch or contraction than we could explain as due to chance, will be replaced by a subtler but still convincing systematicity. We will find more consistent patterns of constraints in the shifting directions of stretch than we could explain as due to chance (e.g., we will demonstrate split-half cross-validations of the constraint patterns). We are exploring some models of this kind (called PT3).

A natural special case of (12) arises if we assume orthogonality of the factors in the mode over which the covariances are computed (in this example, Mode B). If $\boldsymbol{\Phi} = (1/J_k)\boldsymbol{B}_k'\boldsymbol{B}_k = \boldsymbol{I}$, then

$$\boldsymbol{D}_k\boldsymbol{\Phi}\boldsymbol{D}_k = \boldsymbol{D}_k^2 \tag{13}$$

and

$$\boldsymbol{Y}_k = \frac{1}{J_k}\boldsymbol{X}_k\boldsymbol{X}_k' \approx \boldsymbol{A}\boldsymbol{D}_k^2\boldsymbol{A}'. \tag{14}$$

This is the same PARAFAC1 model of covariances that was derived in Section 2.3.2.1 from consideration of the more restricted case of a three-mode, single sample, repeated measures dataset with orthogonal factors in the mode over which the covariances are computed. To avoid this restrictive orthogonality assumption implicit in the use of indirect fitting with the PARAFAC1 model, one can indirectly fit a more general model, such as PARAFAC2 (Section 5.2).

Indirect fitting is quite useful when one wants to investigate factor variations across levels of fixed-effect categories such as sex, clinical diagnosis, etc., or to analyze cross-sectional longitudinal data.

2.3.4. *Indirect fitting to estimate communalities and fit the common factor model*

By analyzing covariance matrices and choosing the PARAFAC option to ignore all diagonal entries, one can fit a three-way version of the 'common factor' model. (Set parameter IGDIAG = 1 to continuously reestimate diagonals during the iterative fitting; at convergence, each matrix contains its own communality estimates.) Replacing self-covariance values, which are inflated by error and specific-factor variance, with less inflated values (i.e., estimated communalities) would seem as appropriate for three- as for two-way analysis.

Nonetheless, we agree with Tucker, Horst, Kruskal, and others who consider an analysis 'factor analysis' in the broad sense even if it has a principal-components-like error model, so long as it is used to identify latent factors. In particular, when PARAFAC is used for direct fitting (e.g., see example application, below), or for other 'component-like' analyses, one need NOT restrict interpretation to mere descriptive summarization of data. Many analytic models that are fit by least-squares to the raw data are quite appropriately used to discover latent sources of variation and/or estimate population parameters.

3. Algorithm and program characteristics

3.1. Estimation procedure

3.1.1. Unconstrained estimation

The fitting method used for PARAFAC is Alternating Least Squares, also known as Iterative Least Squares (Wold, 1966). The trilinear model is broken up into three sets of parameters, such that it is linear in each set given fixed values for the other two sets.

Each set (each mode's loadings) is estimated by simple linear regression methods. We can 'string out' the slices of the three-way data array into a two-way partitioned matrix consisting of a sequence of slabs, one for each level of Mode C. Then, we can write the model as

$$\boldsymbol{A}[\boldsymbol{D}_1\boldsymbol{B}' \mid \boldsymbol{D}_2\boldsymbol{B}' \mid \cdots \mid \boldsymbol{D}_K\boldsymbol{B}'] \approx [\boldsymbol{X}_1 \mid \boldsymbol{X}_2 \mid \cdots \mid \boldsymbol{X}_K]. \tag{15}$$

This is solved by postmultiplying both sides of (14) by the generalized inverse of $[\boldsymbol{D}_1\boldsymbol{B}' \mid \cdots \mid \boldsymbol{D}_K\boldsymbol{B}']$. The revised loadings for Mode B are obtained by the equivalent procedure, holding Modes A and C fixed, and for Mode C by holding A and B fixed. Each iteration thus gradually improves all the parameter estimates. In the PARAFAC program, an elementary linear extrapolation procedure was adopted to accelerate convergence (Harshman, 1970, pp. 32–33); note, however, that much more sophisticated methods have subsequently been proposed by Ramsay (1975), and may be worth exploring. It is also noteworthy that Kiers and Krijnen (1991) have recently developed a special version of the PARAFAC algorithm that can provide considerably faster estimation when one of the three modes is much larger than the other two, e.g., 20 stimuli, 30 rating scales, 250 raters.

3.1.2. Constrained estimation

3.1.2.1. Theoretical constraints To test theoretical models, it is easy to hold loadings in a particular mode (or two modes) fixed at a preselected set of theoretical values and just update the loadings in other mode(s) plus overall scale for each of the factors. Less restrictive constraints can be enforced by applying projection matrices to the data before trilinear analysis (see Carroll et al., 1980); a potentially useful method of more flexible constraints has recently been suggested by Takane et al. (1991).

3.1.2.2. Orthogonality constraints It is sometimes useful to impose orthogonality or zero-correlation constraints when doing direct fitting. In this case, when the PARAFAC user requests orthogonality constraints for a given mode, a modified regression procedure is used to estimate the loadings for that mode.

Initially this was implemented via a Gram–Schmidt procedure inside the iterative loop. But this provided only an approximate least squares solution. A true conditional least squares regression solution was worked out in 1986 in collaboration with Martin Koschat at Bellcore. It is closely related to Cliff's (1966) method of orthogonal rotation to congruence. Again, taking Mode A as an example, the orthogonally constrained loadings are estimated (up to a rescaling of columns) by the modified regression formula

$$\mathring{A} \approx Y'F(F'YY'F)^{-1/2}, \tag{16}$$

where $\mathring{A}$ is the columnwise orthogonal estimate of Mode A loadings, F is the matrix of predictors ($[D_1B' \mid \ldots \mid D_KB']'$), and Y is the matrix of predicted variables ($[X_1 \mid \ldots \mid X_K]'$). Unfortunately, (16) also constrains the columns of $\mathring{A}$ to have sums of squares of one. To minimize the loss in fit caused by this scale restriction, a second stage is added to the constrained estimation procedure: F is rescaled (by R bivariate regressions) to maximize fit of $\mathring{A}F'$ to Y'. To obtain a least-squares solution unaffected by the scale restriction, this (two-part ALS) procedure must be iterated until convergence. In practice, however, this special scale adjustment can be omitted when (16) is inside the main ALS loop in PARAFAC (unless all three modes are constrained to be orthogonal), since estimation of nonconstrained mode(s) readjusts the column scale of F on every iteration.

It is only feasible to require orthogonality if, in the Mode to be constrained, the factors are 'bipolar' (i.e., have many positive and many negative loadings); it is not feasible if more than one of the factors are 'unipolar,' (i.e., have mostly loadings of a given sign and only a few small loadings – or none – of the opposite sign). Orthogonality constraints would force all but one of the unipolar factors to be bipolar, so that the between-factor loading-crossproducts sum to zero.

For this reason, PARAFAC also provides an option for zero-correlation constraints, a milder and more widely applicable method of insuring independence of factor variations. To impose this constraint, we use a modification of the orthogonality constraint procedure. Again, take Mode A as an example. On each PARAFAC iteration, we compute $\mathring{A}$ using (16), but with a modified data matrix $\overset{*}{Y}$ that has been centered across Mode A (i.e., rowwise); then we estimate column constants to add to $\mathring{A}$ in order to predict the means as well as deviations of Y. These column constants are given by $\bar{Y}'(F')^{+}$, where $\bar{Y} = Y - \overset{*}{Y}$, the matrix of row means removed from Y.

3.1.3. *Starting position for iteration*

The initial configuration or starting position for the iterative procedure can be random, theory-determined, or the result of some other analysis. We suggest using six random starts to confirm the within-sample uniqueness and optimality of an obtained solution. If all six solutions agree, then the probability is less than 0.05 that an equally or more likely alternative solution can be found using additional random starting positions from the same distribution. If alternative solutions do emerge, then the investigator can see if they are due to local minima or competing quasi-global solutions, caused perhaps by extraction of an insufficient number of factors (with different subsets of the total set of dimensions showing up in different solutions), or by inadequate system variation in the three way data to uniquely determine some or all of the axes. (For a discussion of the relevant diagnostics, see Harshman, 1984, pp. 580-583).

3.2. Input data characteristics for the program

Input data are usually organized as a series of matrices concatenated to form a 'strung out' 3-way array. Single matrices can be analyzed by treating them as an array with just one level of Mode C, but such an analysis would not have a unique solution; it would be equivalent to a regular two-way factor analysis. We will consider only the true three-way case here.

Input data may be covariance-like measures (with 2 identical ways) or raw score profile data (usually with 3 distinct ways). Correlational data are *not* strictly appropriate, since they impose different rescalings of the raw data in each correlation matrix (see Harshman and Lundy, 1984a, p. 141). All ways/modes of the data are treated similarly during the iterative process.

Missing data are allowed, and the program reestimates the values in missing data locations to make them consistent with the overall model. Starting estimates for missing values are either the input values, or the mean computed over all nonmissing values in the same location of the other matrices. Subsequent estimates are updated on each iteration, and are computed from the factor loadings at that iteration. To perform a three-way common factor analysis, communality estimates can be placed in the diagonals of each matrix, or the diagonals can be treated as missing (equivalent to iteration on the diagonals). Large blocks of systematically missing data and incomplete factorial designs are handled with the 'Incomplete Mode' variant of PARAFAC (see Section 5.).

3.2.1. Data manipulation and preprocessing

It is important to remove extraneous constants and/or two-way interactions before fitting the trilinear model. Failure to eliminate baseline offsets will throw off proportionality relationships and so may seriously interfere with the identification of parallel profile axes.

Preprocessing issues become considerably more subtle when three-way data are considered. Not all plausible transformations are in fact appropriate. Identifying the allowable preprocessing transformations (ones that do not distort the trilinear structure of the data) has been a subject of considerable study and some debate. (For mathematical analysis of 'appropriate' vs. 'inappropriate' preprocessing steps, the reader is referred to Harshman and Lundy, 1984b; and Kruskal, 1984).

Technically, the input data (after any necessary preprocessing) should be either interval scale or ratio scale (although it appears that rank data can sometimes be input and treated as interval scale data, e.g., see Weeks and Bentler, 1979). Before preprocessing, it can be fiber- or slab-conditional with respect to origin and slab-conditional with respect to scale.

3.2.1.1. Profile data PARAFAC has extensive capabilities for preprocessing raw score or profile data prior to direct fitting. It employs those types of centering and standardization combinations that do not alter the data structure. (Surprisingly, some common two-way methods distort the data structure when incorrectly generalized to three-way arrays; see, e.g., Harshman and Lundy, 1984b, pp. 225–253). Fiber-centering (removing row, column or 'tube' means) and slabwise size standardization or normalization (equalizing the mean square in data slices) are the basic types of preprocessing available. Centering and/or standardization can be done for any combination of one or more modes. Size standardization on more than one mode requires iterative preprocessing, as does centering and standardizing on the same mode. Usually the iterative procedure converges in 3 to 6 iterations.

Centering is used to remove constants and conditionality that would otherwise make the data inappropriate for the PARAFAC model. If such constants are not present, fiber-centering has no effect on the trilinear structure of the three-way array, except that the factor loadings for the centered mode(s) are themselves column-centered (Harshman and Lundy, 1984b, p. 234). Emphasis in the centered mode(s) is thus shifted from baseline values to differences across levels (i.e., patterns of change).

Standardization is used to remove differences in mean square values that might otherwise cause some variables to influence the solution too much, others not enough. The consequence of standardization for the solution is that each row of factor loadings (in a mode where the data were standardized) is itself 'restandardized,' i.e., the loadings for a given level of the output are multiplied by the same rescaling constant that was applied to that level of the input (Harshman and Lundy, 1984b, p. 246).

There are no definite rules for selecting among the appropriate data centering and/or standardizing options, just guidelines (see Harshman and Lundy, 1984b, pp. 257–259). Some experimentation with preprocessing applied to different combinations of modes may be necessary, in order to get the most interpretable solution.

We note in passing that preprocessing is also necessary in two-way factor analysis, since the two-way model also assumes ratio-scale data. However, well established conventions (e.g., conversion to correlations before factoring) allowed users to be unaware of these issues.

3.2.1.2. Covariance data For covariances, PARAFAC can do 'Equal–Average–Diagonal' scale standardization, which equates the variance associated with each variable in the total data array while preserving the proportional differences among slices, and hence among factors, from one two-way slab to the next. As mentioned in Section 3.2, correlational data are inappropriate for

PARAFAC analysis, because the computations involved impose different scaling on each matrix; thus, any proportionalities in relationships across the matrices are disturbed.

3.2.1.3. Dissimilarity / similarity or distance data PARAFAC can also be used to perform three-way multidimensional scaling by indirectly fitting the weighted Euclidean distance model to similarities or dissimilarities. This is equivalent to Carroll and Chang's (1970) INDSCAL procedure. The similarities/dissimilarities are first transformed to scalar products (e.g., by use of the DISTIN program that accompanies PARAFAC) and then the trilinear model is fit to them. The model to represent this is the same as the one given above for covariances, except now A is interpreted as the matrix of stimulus projections onto dimensions and the estimated D_k (often written D_k^2 as in equation (6)) gives the squares of the dimension saliences for the kth subject.

3.2.1.4. Elimination of preprocessing An alternative approach is now being developed which could often eliminate the need for preprocessing. It handles baseline offsets and other (either unwanted or interesting) constants by fitting additional parameters during data analysis (see Section 5.4). Multiplicative parameters are also included in this extended model, and can be used to equate or otherwise adjust 'corrected' variances that are present after extranious constants have been identified and removed.

3.2.2. Necessary input data characteristics for uniqueness

Where the assumptions of the PARAFAC model are more or less fulfilled (the three-way variation has a substantial part that can be considered 'system variation' [2], and the factors show consistent (qualitative) patterns of loadings, consistent orientations and distinct patterns of size variation across the levels of each mode) then the solutions are generally well-behaved – they converge and are unique and interpretable. If there are factors which do not show distinct variation in two of three modes (or in all three modes, of course) then these factors will be combined into a single composite factor by the analysis. If there are factors which do not show distinct variation in one mode, then the subspace spanned by those factors will be recovered, but will fail to have uniquely oriented axes. During data analysis, this might be revealed by the program converging on different 'rotations' for those dimensions from different starting positions, while the remaining dimensions are stable across starting positions. However, the presence of substantial error may produce a sample specific 'unique' solution (within-sample uniqueness). The safest test of generalizable unique axes (across-sample uniqueness) is obtained by comparing solutions based on random split halves of the data, as noted in the application example (Section 4) and illustrated in detail in Harshman and De Sarbo (1984).

Where the data have non-stationary patterns with more complex structure than provided for by the PARAFAC model (e.g., Tucker structure; see Kroonenberg, 1983; Lundy et al. 1989; Tucker, 1964, 1966, 1972) and as a result a large amount of systematic variance cannot be well approximated by the PARAFAC proportional model, nonconverging 'degenerate' solutions can occur. Degenerate solutions have two or more factors that are highly negatively correlated, are generally uninterpretible, and are reliably obtained from different starting positions; they are discussed in some detail in Kruskal et al. (1989) (see also Harshman and Lundy, 1984b, pp. 271–274). In most case, the degeneracy can be blocked and an interpretable solution obtained (with some loss of fit) by constraining the factors in one mode to be orthogonal or have zero-correlations, or by using indirect fitting (PARAFAC1) which is generally immune to degeneracies. Even better, a two-stage least-squares analysis can sometimes aid in recovery and interpretation of the structural complexities that caused the degeneracy (Lundy et al. 1989).

3.3. *Program output characteristics*

Besides the factor loadings, PARAFAC supplies other information to help in assessing the solution. Fit values provided are R-squared, Stress and Mean Squared Error. These are useful for 'fit versus dimensionality' plots. Also, the Root Mean Squared (RMS) contribution of each factor is printed. When the data have been centered in at least one mode (as is usually the case) and the factors are mutually orthogonal in at least one mode, the squared RMS values are the variances accounted for by the factors. Interfactor correlations are cross-products are given for each mode. These indicate factor profile similarity, and in one sense orthogonality or lack thereof; thus they help to confirm whether a solution is 'degenerate.' Goodness of fit of the solution may be studied in more detail with the aid of an error analysis table, which lists the Mean Square Error and Stress values for each slice of each mode of the data array. This indicates which parts of the data are fit better or worse than others. Exceptionally large mean squares for particular levels may indicate problems in the data (e.g., data entry errors, out-of-range values).

We suggest split-half analyses as the best way to confirm the solution (Harshman, 1984, pp. 587–591). If essentially the same (or a similar) solution is replicated in both split halves, you have confirmed that the PARAFAC axes reflect genuine and reliable patterns of stretch or contraction of the factor space across the third mode (see Section 4.4.3).

The CMPARE program which accompanies PARAFAC can be used to compare solutions quantitatively; it computes interfactor correlations and cross-products for each mode, within and between solutions. The PFPLOT program which also accompanies PARAFAC can display the loadings projected onto each factor axis (useful when interpreting the factors) or projected onto the coordinate plane defined by a pair of factor axes (useful for showing clusters of points or shears in the space).

Additional output information is also optionally provided, such as data residuals, rate of change in parameters, current estimates of missing values, etc. To save space, these are not discussed here.

4. Example application: Hemisphere involvement in task-directed movements

4.1. *Theoretical rationale for the study*

In humans, each cerebral hemisphere tends to control muscles on the 'opposite' (contralateral) side of the body. This is particularly true for distal muscles such as those involved in fine movements of the fingers. However, little information is available on how the two sides of the brain interact during complex tasks that involve use of both hands. Does each hemisphere simultaneously control the contralateral hand in a semi-independent but cooperative fashion?

Likewise, we don't know whether motor control is influenced by the relative hemisphere involvement in the intellectual aspects of the task being performed. What would happen, for example, if the task required such sophisticated nonverbal cognitive processing that the right hemisphere could better determine which actions were appropriate; would the right hemisphere then transmit its conclusions to the dominant left hemisphere for execution, or would it issue motor commands directly to the hands? Finally, since the right hemisphere can better control fine movements of the left hand, would a greater reliance on left hand manipulations result?

We consider here data from an experiment that addressed such questions (Hampson and Kimura, 1984). In this experiment, the investigators studied the hand movements of normal people performing similar tasks in several different circumstances. By manipulating the cognitive demands of physically equivalent motor tasks, they sought to discover how changes in the degree of right vs. left hemisphere *cognitive involvement* would affect the pattern of right vs. left hemisphere *motor involvement*.

4.2. *Previous data collection and analysis*

4.2.1. *Data collection*

Twenty-four right-handed college students, 12 male and 12 female, were individually videotaped while they assembled blocks to accomplish each of eight tasks set by the experiments. The cognitive nature of the tasks differed: some were highly verbal (e.g., making particular kinds of words using alphabet blocks) and some were much more spatial (e.g., reconstructing designs or pictures using blocks painted with lines and shaded sections). The physical part of all the tasks was more or less the same: rearrange small wooden blocks on a tabletop.

4.2.2. *Data classification and coding*

An initial study of videotapes of several pilot subjects was used to develop descriptions of all the different kinds of hand movements observed under these circumstances. Then, by 'combining closely related descriptions' (p. 111), Hampson and Kimura developed a taxonomy of 19 different movement *types* (e.g., 'pick up block,' 'lean/rest,' etc., see Table 1, column 2).

Data coding then proceeded as follows: For each subject, for each task, Hampson examined the videotape record during a 1 minute period beginning 30 seconds after the start of the task. She counted the number of right hand, left hand, and bimanual movements [3] of each of the 19 types during this period. These counts were the 'raw data' used by Hampson and Kimura (1984), and are the data which we were given to reanalyze. Data from the seven experimental tasks were used in our reanalysis (see Table 2); we omitted the neutral 'control task' on Hampson's recommendation, as zero frequencies were obtained for almost all movement types, due to the simplicity of the task. On the other hand, we included a 'sequencing' task that Hampson and Kimura had omitted as problematic.

4.2.3. Hampson and Kimura's analysis

Rather than statistically assessing changes in each of the 19 movement types individually, Hampson and Kimura (1984) combined functionally similar movement types into broader categories for analysis. To reduce the number of variables involved, they classified each of the 19 individual movement *types* into one of 5 broader movement *categories*: Active, Auxiliary, Non-contact, Self-touching/rest, and Miscellaneous. Separate right and left hand composite scores for each movement category were then computed for each subject during each task period. To assess relative hand usage, the composite left and right hand counts were converted into {(R minus L)/(R plus L)} movement ratios. This resulted in 5 ratios per task per subject. These ratios were then analyzed by various methods including multivariate ANOVA. Their statistical results showed that shifts in $\{(R-L)/(R+L)\}$ movement ratios were indeed observed, but, as predicted, "only for movements that played a functional role in task performance" (Hampson and Kimura, 1984, p. 102).

4.3. Motivation for a three-way factor analysis

4.3.1. Experimenter's motivation

(a) To test and confirm (or revise) taxonomy of movements. The pattern of R–L shifts observed is of course affected by the movement classification scheme that is used to combine individual movement types into larger functional movement categories. Since this scheme is influenced by the theoretical framework in which the authors formulated the study, it would be interesting to see if an 'atheoretical' factor analysis of the original, unclassified 19 movement types would confirm the authors' five movement categories. In this regard, the rotational uniqueness or intrinsic-axis property of PARAFAC is particularly attractive because it provides a stronger kind of confirmation if its unique factors line up with Hampson and Kimura's a priori categories.

---

[3] The bimanual movement counts were included in some analyses but were too sparse to substantially affect the results. For simplicity we concentrate here on the right and left movements.

(b) To test for category specificity of cognitive task effects with factor-derived and hence more theoretically neutral categories.

(c) To obtain a more comprehensive and detailed picture of experimental results. PARAFAC can provide a single, integrated solution in which each factor has loadings for each of the original 19 movements, 7 experimental tasks, and 24 people. Instead of R–L ratios (which do not tell how much each component of the ratio is changing) the raw counts for each hand can be analyzed. Such detailed information on broad patterns (factors) arising from the experiment might lead to new insights into the movement mechanisms under study.

4.3.2. Methodologist's motivation

(a) This is a nontrivial problem. These movement data have an interesting mix of suitable and unsuitable three-way structure. The useful answers to be obtained are not obvious in advance, yet there is an objective character to the data and a considerable neuropsychological body of information to help decide whether the results are scientifically meaningful.

(b) Causal issues. The experimental nature of the Hampson and Kimura study makes consideration of causal issues more straightforward. Hence it provides a way to discuss questions of how factor axis orientations do or don't approximate 'real' *causal* factors.

4.4. Three-mode analysis procedure

4.4.1. Choosing direct vs. indirect fitting

Before performing a PARAFAC analysis, one needs to consider whether the patterns of data variation in the three-way array will be appropriate for direct fitting or whether it requires indirect fitting (i.e., whether system variation is present at the raw data level, or only at the level of covariances, see Section 2.3). For the Hampson movement data, we decided that while both kinds of variation would be present, it would be most useful to perform direct fitting, so as to concentrate on the experimental effects.

4.4.1.1. Inappropriateness and appropriateness of the direct-fit model The subjects' *spontaneous movement patterns* would probably *not* vary in parallel profiles form from one subject to the next. Each person could be expected to have nonrandom but idiosyncratic kinds of movements that would be used to achieve particular results. For example, some people might be habitual 'block sliders,' while others might be 'block placers,' etc. Consequently, one person would not generally use the various movement types in the same relative proportions as someone else. Three-way data like this, where a large part of the systematic variation follows patterns that are idiosyncratic for each data generating object, are said to show 'object variation'. These patterns can be explored by using indirect fitting (and we have, in fact, tried some indirect fitting analyses, but we have no space to discuss these here).

In contrast, the *experimentally induced variations* in subjects' movement usage might well obey the coherent three-way proportionality needed for direct fitting PARAFAC analysis. Such variation is called 'system variation' because it can be thought of as arising out of the three-mode system as a whole rather than separately for each object (for further discussion, see, e.g., Harshman and Lundy, 1984a, pp. 130–133 and 140; Harshman, 1970, pp. 20–24).

4.4.1.2. An argument for a system variation component to the variation There is a theoretical argument to support the idea that Hampson and Kimura's experimentally induced shifts in movement usage might show the desired system variation to allow them to be identified and extracted by PARAFAC. It is based on the notion that the influence of greater left or right cerebral hemisphere cognitive engagement (due to the nature of the task being solved) might act proportionally across levels of each of the three modes. In Mode A, the strength of the effect may vary proportionally from one movement type to the next, being greatest for those types most directly related to task solution, such as picking up blocks, placing blocks, and so on; it would also be proportionally increased in all the left hand and decreased in all the right hand categories, or vice versa. In Mode C, it should proportionally change from one task to the next. For example, a proportionally greater number of right hand movements would be expected in all affected movement categories during the verbal tasks. Finally, it is less certain, but still plausible, that the degree of asymmetry of cerebral engagement in problem solving movements would proportionally change from one subject to the next, depending on the general tendency of any particular subject to rely more on verbal or spatial task solution strategies. Subject differences might also arise if different subjects had different degrees of cerebral lateralization of the spatial or verbal cognitive functions (or motor functions) engaged by the tasks.

Thus direct fitting will not detect all the patterns of object variation that would be detected by indirect fitting, but it would concentrate most effectively on the experimental variation of interest. For the purpose of this analysis, therefore, we chose direct fitting.

4.4.2. *Preprocessing the three-way array*

Psychological and behavioral measurements usually have arbitrary baselines which obscure the proportionality of changes; hence one almost always need to 'center' (subtract out means) across at least one mode (Harshman and Lundy, 1984b, p. 217 ff). We initially centered the Hampson movement data across levels of Mode C (tasks) before performing the PARAFAC analysis (centered $x_{ijk} = x_{ijk} - (1/K)\Sigma_{k=1}^{K} x_{ijk}$). However, the resulting solutions showed that this centering over-emphasized object variation (i.e., idiosyncratic subject-task interactions) in the array. By using bootstrapping we could identify a meaningful reliable pattern related to tasks and movements in the third factor, but by this point the analysis had become overly complicated.

We decided to try, instead, an analysis of the data without any centering.

While this is usually ill-advised, it seemed possible that in these particular data the experimental effects were directly measured at ratio-scale, without spurious baselines [4].

We did apply one modest preprocessing adjustment to the data. The data were size-standardized to have unit mean-squares within each level of Mode C. This gives the data from each task potentially equal influence on the total solution.

4.4.3. Factor analyses

The data were analyzed at one through four dimensions. A diagram was constructed showing improvements in fit value (R-squared) vs. number of factors. This is interpreted in the same way as in standard two-mode factor analysis. The shape of the diagram suggests the presence of one major factor and a second moderate sized factor. Additional smaller factors might be present, but this is somewhat ambiguous. To save space, we omit the plot. The r-squared values were: 1D = 0.471, 2D = 0.556, 3D = 0.595, 4D = 0.625. Stress values were = 0.639, 0.585, 0.559, 0.537.

As noted earlier, our main test of dimensionality is how many factors are demonstrably replicable. We used the 'orthogonal split half' approach previously employed, for example, in Harshman and DeSarbo (1984). That is, subjects are randomly divided into four quarter-samples, which we called QA, QB, QC, and QD. This allows us to create split-half samples in two different 'orthogonal' ways: (a) Half1 = (QA + QB) vs. Half2 = (QC + QD); and (b) Half3 = (QA + QC) vs. Half4 = (QB + QD). We then perform PARAFAC analyses on each of these half samples, and compare the results. Replication of a factor between *either* Half1 and Half2 *or* between Half3 and Half4, demonstrates that it is generalizable, because such a replication across either pair of independent subsamples could not be explained simply on the basis of chance.

In the Hampson and Kimura data, the results of our split-half tests were as follows: the one-dimensional solution replicated across both sets of split halves. In the two-dimensional solution, factor 2 replicated across Half3 vs. Half4 but not Half1 vs. Half2. Factor 1 of the 2D solution replicated across both sets of splits. These results indicate that we can cross-replicate at least two stable, generalizable, factors. In contrast, the 3D and 4D solutions did not replicate

---

[4] Since the data were counts of movements, there is a superficial sense in which the data are clearly ratio-scale. Zero movements is a 'true zero' and 10 movements is clearly twice as much as 5 movements. However, the key question is not whether the data are ratio-scale in this superficial sense, but whether the measurement of the underlying psychological processes is truly ratio scale. Is the natural baseline for the behavior that we are measuring zero movements, or some spontaneous level of movements that would be present even in the absence of the tasks? The prior analysis by Hampson and Kimura (1984) indicated that the effects of the experimental manipulations were selectively modulating the generation of task directed or problem solving movements. Perhaps, in the absence of a problem, the natural level of such movements would indeed be zero. If so, then the raw counts of movements in this class might indeed provide ratio-scale measurement of the strength of the underlying movement generation processes of interest.

across either set of splits. Apparently, we are unable to recover more than two stable factors from these data. (For more details on method, including quantitative assessment of agreement across split halves, see, for example, Harshman and DeSarbo, 1984).

4.5. Interpretation of the solution

4.5.1. Descriptive level

Loadings from the two-factor PARAFAC analysis of the full Hampson movement data are shown in Tables 1 and 2. We interpret these loadings in ways similar to the way loadings are interpreted in two-way factor analysis. In this case, neither Mode A nor C are scaled to look like correlations, but we can think of them instead as regression weights.

In Mode A the factors are highly 'hand specific,' with factor 1 showing large positive loadings for Active movements generated by the right hand and factor 2 showing large positive loadings for Active Movements generated by the left hand. Some Auxiliary movement types show movements of the opposite hand, but this is consistent with the fact that this category of movements perform a secondary or 'assisting' role.

In Mode C, factor 1 seems to be a general factor active during all tasks, though slightly more active for the verbal tasks. In contrast, factor 2 has much higher loadings on the nonverbal tasks than on the verbal ones.

(a) Taxonomy and (b) selectivity of hand movement effects: The key movement distinction made by Hampson and Kimura was between activities involved in task solution, (mainly Active and Auxiliary categories) and those not so involved. Recall from Sections 4.4.1.1 and 4.4.1.2 that PARAFAC factors should have substantial loadings on those movement categories that are sensitive to the effects of task manipulation, but not on other categories. We find that PARAFAC loads highly on just Active and Auxiliary movement categories. This provides strong support for the basic experimental distinction proposed by Hampson and Kimura, and the direction of the selective effects on these categories also confirm the particular predictions of Hampson and Kimura as to the nature of verbal and nonverbal task involvement [5].

---

[5] Actually, to unambiguously test the taxonomy one would want either to standardize variance of the levels of Mode A on input, or (if this would amplify noisy variables too much) adjust the loadings on output for differences in input variance. And/or we might use bootstrapping to determine which loadings are statistically significant. The bootstrapping of our prior analysis of these data (when centered across Mode C – not reported here) did, in fact, confirm that only the Active and Auxiliary variables loaded significantly on the factors – with one or two interesting exceptions, particularly 'Hover Moving.' Space limitations preclude consideration of either a bootstrap of the uncentered solution or an additional PARAFAC analysis of the uncentered data but with Mode A standardized, to investigate this question more adequately. Thus for expository reasons we concentrate on the one solution deliberately ignoring input differences in levels of Mode A.

Table 1

Mode A: Movement Loadings

No.	Movement	Category	Hand	Factor 1	Factor 2
1	Slide Block	Actv	L	-0.142	4.009
			R	2.745	0.379
2	Pick Up Block	Actv	L	0.362	1.568
			R	3.404	-1.085
3	Place Block	Actv	L	-0.105	2.683
			R	3.745	-1.475
4	Slide & Reject Block	Actv	L	-0.196	1.458
			R	0.478	0.761
5	Pick up & Reject Block	Actv	L	0.043	0.112
			R	0.227	-0.028
6	Transfer from one hand to other-most active hand	–	L	0.268	0.178
			R	0.054	0.332
7	Adjust Block	Actv	L	0.338	0.527
			R	1.146	-0.290
8	Stabilize Block	Aux	L	0.722	0.488
			R	-0.233	1.485
9	Stationary Contact	Misc	L	0.307	0.406
			R	0.684	0.158
10	Rotate Block in Air	Misc	L	-0.025	0.150
			R	0.043	0.010
11	Rotate Block on Tabletop	Actv	L	-0.245	1.385
			R	0.384	0.424
12	Self Touching Movement	Self Tch	L	0.260	0.061
			R	0.158	0.159
13	Finger Tapping on Table	Misc	L	0.008	0.028
			R	0.124	0.054
14	Finger Tapping on Blocks	Misc	L	0.006	0.018
			R	0.040	0.018
15	Lean/Rest	Self Tch	L	0.806	0.163
			R	0.354	0.531
16	Point	Non Con	L	0.032	0.007
			R	0.162	-0.088
17	Hover Moving	Non Con	L	0.004	0.594
			R	0.536	-0.137
18	Hover Stationary	Non Con	L	0.447	0.204
			R	0.340	0.184
19	Rotate Hand/Limb without contact	Non Con	L	0 *	0 *
			R	0.043	-0.019

* No movements of this type made with this hand.

This factor analysis does not allow us to test the full taxonomy of movement categories, including those not related to task solution, since as noted in Sections 4.4.1.1 and 4.4.1.2, the use of direct fitting with these data concentrated the analysis on the task-specific effects. To examine the full set of movement

Table 2

Mode C: Task Loadings

No.	Task Name	Type	Factor 1	Factor 2
1	Palindromes	V	1.063	0.614
2	Crossword	V	1.079	0.599
3	Cued Crossword	V	1.044	0.596
4	Nonverbal Sequencing *	?	1.114	0.527
5	Modified Jigsaw	NV	0.945	1.170
6	Rotated Figure	NV	0.953	1.047
7	Figure Reconstruction	NV	0.756	1.779

* Not used in Hampson and Kimura (1984) analysis.

patterns it would be necessary to do indirect fitting analysis (see Section 2.3.2 for mathematical discussion of indirect fitting).

(c) New insights: Since the factor analysis was conducted on raw counts for each hand, rather than L–R ratios, it is possible to see where the locus of hand movement change lies. The loadings suggest that it is primarily in the frequency of left hand movements. Some other new results are discussed in terms of causal scientific interpretation of the factors, Section 4.5.2.1.

4.5.2. 'Explanatory' (scientific) level

It is striking how well the properties of the PARAFAC factors mesh with neuropsychological theory. From the Hampson and Kimura article it is clear that the key underlying causal constructs behind their experiment are the cognitive-motor activity of the right and left cerebral hemispheres. Elegantly, the PARAFAC factors are directly interpretable as reflecting the 'voluntary' action of left hemisphere (factor 1) and right hemisphere (factor 2) in the experimental situation. The right hand specificity of loadings on factor 1, and the left hand specificity of loadings of factor 2 suggest this interpretation. In addition, factor 2 is mainly sensitive to activation by non-verbal tasks, as would be appropriate for the right hemisphere. More detailed consideration of the factor loadings reveals further details consistent with a linking to these neuropsychological theoretical constructs.

4.5.2.1. New results The factor analysis reveals some details not easily apparent in the Hampson and Kimura ANOVA results. For example, from Mode C loadings it appears that the left hemisphere's influence is almost as strong on nonverbal as on verbal tasks, whereas the right hemisphere's is much stronger during nonverbal tasks. This might be interpreted as some kind of consequence of left hemisphere 'dominance' in right handed individuals, or it might indicate that a good deal of analytic-verbal mediation is involved in all the tasks used in the experiment. Further experimentation would seem necessary to resolve this. This pattern could also conceivably be produced by something peculiar to this experiment. So this should be checked in new data before strong scientific conclusions are drawn.

The need for two factors rather than one shows that the process controlling these hand movements is 'two-dimensional' (rank-2). Two factors are necessary to adequately capture the complexity of the hand shift patterns. This supports the idea that the two hemispheres exert their effects quasi-independently (at least in this situation).

Various other potentially useful quantitative findings emerging from the factor analysis include the following: (a) The experimentally manipulated shifts (as fit by the factor analysis) account for roughly half of the variation in hand movements in the data. (b) The relative size of left hemisphere (factor 1) and right hemisphere (factor 2) influence across the experiment as a whole is 0.77 vs. 0.40.

4.5.2.2. New questions Some questions arising from the factor analysis might have potential theoretical importance. For example, why is it that only factor 2 has substantial negative loadings on the 'opposite' hand within some movement types? Does this tell us something about reciprocal inhibition, or replacement of one movement by another? Likewise, we might ask what differentiates those subjects (Mode B, not shown) with equal factor 1 and factor 2 weights from those with large factor 1 and vanishingly small factor 2 weights?

4.6. General methodological conclusions

Perhaps the most interesting and controversial goal of exploratory factor analysis is the attempt to uncover 'real' functional unities, each of which reflect effects of a distinct cause acting in the situation (e.g., see Yates, 1987, pp. 1–6). Because the Hampson and Kimura study is a true experiment, issues of causality are central to a scientific understanding of these data. We see from the above Hampson and Kimura's theoretical constructs underlying the experiment and its interpretation appear to correspond well with the interpretation of latent constructs one would make based on the PARAFAC-determined orientation of factor axes. In this sense, then, the factor analysis may indeed have uncovered statistical approximations of 'real' functional unities.

Although the proportional profiles factor orientation is apparently quite interpretable and useful, can we say that it is the empirically 'best' orientation of factors in the space spanned by the solution? Suppose we ignore (or override) the uniqueness property and force-rotate the factors to alternative positions, in order to explore alternative possible theoretical constructs and interpretations that they would suggest. Because there are only two dimensions in this particular solution, we can examine the range of alternative rotations by hand. When we do this, we discover that in this case other rotations lead to less intelligible constructs and interpretations. The most distinct alternative in the two dimensional space spanned by the PARAFAC factors corresponds to a 45 degree rotation of axes. This gives us a set of two constructs in which one factor

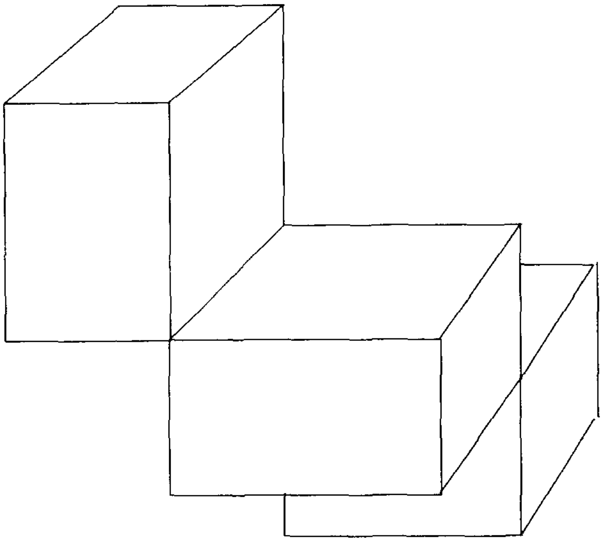

Fig. 2.

controls overall activity of both hands, while the other factor controls the distribution of this activity between left and right hands. This implies a considerably different neuropsychological mechanism of motor control than the PARAFAC factors, and in fact seems inconsistent with much current knowledge in the area.

5. New developments in parallel factor analysis

5.1. 'Linked mode' and 'incomplete mode' PARAFAC analysis

5.1.1. Linked mode PARAFAC

This analysis can be though of as 'multi-Multimode' fitting, where several multi-way arrays are solved simultaneously, subject to restrictions that a particular mode of one array is 'Linked' to some specific mode of another, so that they are both represented by the same latent factors (Harshman and Lundy, 1984a, p. 281). More than two arrays may be linked; three linked arrays – linked on different modes – are shown diagrammatically in Figure 2.

Naturally, Linked-Mode PARAFAC models could have many different specific forms. Here is one example (ignoring error for simplicity):

$$\begin{aligned}
&\text{(i)}\quad \boldsymbol{X1}_{k,} \approx \boldsymbol{A1D1}_{k,}\boldsymbol{B1}\\
&\text{(ii)}\quad \boldsymbol{X2}_{k,} = \boldsymbol{A2D1}_{k,}\boldsymbol{B2} \qquad (17)\\
&\text{(iii)}\quad \boldsymbol{X3}_{k.} = \boldsymbol{A2D3}_{k.}\boldsymbol{B3}.
\end{aligned}$$

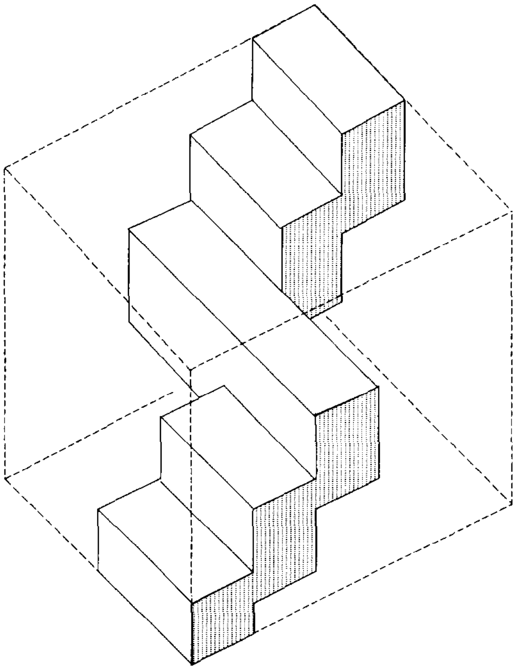

Fig. 3.

However, an easier way to write such linked models is simply to specify the different arrays separately, and then list the linkage conditions.

Linked Mode analysis provides a useful way of integrating several different three-way studies that involve the same measurement scales, or the same stimuli, etc. Another interesting case arises when a two-mode array is linked on one of its modes to one or more three mode array(s).

5.1.2. Incomplete mode PARAFAC

The special handling of data in computer memory in the Linked Mode program also allows analysis of an array where large blocks of the data are 'missing' or undefined, as in Figure 3. Consider, for example, a study where there are too many stimuli to be rated by a single individual. One might have individuals 1 through 10 rate stimuli 1 through 15, subjects 10 through 20 rate stimuli 10 through 25, etc. with all subjects using the same set of rating scales. The Incomplete-Mode (Linked Mode) fitting procedure could solve for a solution that unites all the partially overlapping sets of stimuli into one consistent least squares estimate of a comprehensive space in which all stimuli reside. In an even more economical design, subsets of the subjects could use partially overlapping subsets of the rating scales (should there be too many scales, as well as too many stimuli, for one person to rate). The dataset shown in Figure 3 has incomplete overlapping subsets in two of the three modes. (Note, however, that this figure has been simplified for pedagogical reasons; a more complex 'balanced' structure would be needed in many situations)

5.2. PARAFAC2

When PARAFAC is used for indirect fitting, the factors are necessarily orthogonal (see, e.g., Harshman and Lundy, 1984a, p. 139–140). A new program, PARAFAC2 allows the factors to be oblique [6] (and still uniquely resolved, Harshman and Lundy, 1991). Because it deals with multiple matrices, PARAFAC2 requires a consistency assumption: the angle between any particular pair of factors should remain the same (or at least roughly similar) in all the covariance matrices in the three-way array.

The PARAFAC2 mathematical model has the following form:

$$\mathbf{Cov}_k = \boldsymbol{A}\boldsymbol{D}_k\boldsymbol{\Phi}\boldsymbol{D}_k\boldsymbol{A} \tag{18}$$

We have a working program that fits this model, but it is in 'bare bones' research/testing form. We have not yet converted it to distribution level status.

PARAFAC2 can also be used as an oblique-axis generalization of INDSCAL multidimensional scaling (Carroll and Chang, 1970), if applied to pseudo-scalar product data derived from similarity data. For a discussion of oblique axis MDS see Dunn and Harshman (1982).

5.3. PARATUCK (PARAFAC–TUCKER) models

We have recently been exploring a set of quite general parallel factor models that combine the intrinsic axis capabilities of PARAFAC with some of the greater structural generality of Tucker's Three-mode Factor Analysis (Kroonenberg, 1983; Tucker, 1964). Of these 'PARATUCK' models, the currently best studied and furthest developed is the PARATUCK2 family (so called because it combines aspects of PARAFAC and with aspects of Tucker's T2; for information on T2 see Kroonenberg, 1983). The PARATUCK2 (also known as PT2) model has the form

$$\boldsymbol{X}_k = \boldsymbol{A}^{\mathrm{L}}\boldsymbol{D}_k\boldsymbol{H}^{\mathrm{R}}\boldsymbol{D}_k\boldsymbol{B}'. \tag{19}$$

This model has a unique intrinsic axis solution (given 'adequate' data and appropriate output standardization conventions concerning column order, size of left vs. right D matrix, etc.) Models such as PARAFAC, and PARAFAC2 can be considered special cases of PT2. We are currently using it to explore four versions of three-way DEDICOM (Harshman and Lundy, 1992).

---

[6] 'Orthogonal' and 'oblique' are used here in exactly the same sense as in two-way factor analysis. For example, in the orthogonal case, the implicit factor scores for different factors are assumed to be orthogonal across people, or – more generally – across the levels of the mode over which covariances are computed; that is, across levels the mode that 'disappears' when one converts to covariances.

Although working programs for PARATUCK2 (and several special cases) do exist, these are minimal programs, set up for our own research and testing of the model; they are not yet suitable for general distribution.

Some practical information on PARAFAC software

Source Language and Program Installation. PARAFAC and programs that accompany it are written in Fortran, and run in batch mode. All except PFPLOT conform to PFORT specifications (a portable subset of American National Standard Fortran X3.9-1966). Thus they should compile successfully on any modern Fortran compiler (requiring only that the compiler include the Fortran 66 extensions – which most do). The installation is simple, in part because of the batch-mode design. PARAFAC has run successfully on a wide variety of computers from Crays to DECs to PCs.

Sample Memory Requirements. On a Cyber 170 computer, the core memory required for the compiled PARAFAC program is approximately 27K words (not bytes), not including space required by I/O buffers and arrays. The array space required for a ten-factor analysis of 18 by 18 by 35 data is an additional 14K words on the Cyber. On an IBM AT computer, using a Ryan–McFarland compiler, PARAFAC requires 220K bytes, plus an additional 310K bytes for arrays when a 45 by 45 by 20 data arrays is to be analyzed. (Of course, less space is required for smaller data sets.)

Documentation. Documentation for PARAFAC is provided by an extensive user manual (Lundy and Harshman, 1985; includes update notes) with over 100 pages plus index. For those who like to study the source, there are many comment statements in the source code.

Availability and Distribution Format. PARAFAC and related programs have been placed in the public domain. You may get copies from others, and give away copies, as you please. [7]

On request, we will provide copies of the latest 'standard' version free of charge (except for postage and handling, see below). Our software is currently distributed on high density DOS (IBM) PC floppy diskettes. Specify your preference for high density 5.25 or high density 3.5 (DOS format), or other (e.g., low density 5.25).

Charges. To cover the costs of media, file copying and checking, mailers, postage and handling (and, in the case of the Manual, printing) please include

---

[7] However, if you change the code, please put a line at the very beginning of the source as follows: "'Lines-,-,-,- modified by ..., on data ..., to produce the effect that ...'." This is an important courtesy for others who may later get copies of the modified program.)

$10.00 for each disk requested, and $30 for each Manual requested [8]. Items are sent by Air Parcel Post. If faster mail is desired (first class airmail, express mail, or courier service), please specify and include enough extra fees to cover the extra postage.

Address requests to: Richard Harshman, Department of Psychology, Social Science Center, University of Western Ontario. London, Ontario, Canada N6A 5C2.

References

Arabie, P., J.D. Carroll and W.S. DeSarbo, *Three-way Scaling and Clustering*, Quantitative Applications in the Social Sciences, No. 65, series no. 07-001, (Sage, Beverly Hills, CA, 1987).

Bloxom, B., A note on invariance in three-mode factor analysis, *Psychometrika*, **33** (1968) 347–350.

Burdick, D.S., X.M. Tu, L.B. McGown and D.W. Millican, Resolution of multicomponent flourescent mixtures by analysis of the excitation-emission-frequency array, *J. Chemometrics* (1990) 15–28.

Carroll, J.D. and P. Arabie, Multidimensional scaling, *Annual Review of Psychology*, **31** (1980) 438–457.

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

Carroll, J.D., S. Pruzansky and J.B. Kruskal, CANDELING: A general approach to multidimensional analysis of many-way arrays with linear constraints on parameters, *Psychometrika*, **45** (1980) 3–24.

Cattell, R.B., 'Parallel Proportional Profiles' and other principles for determining the choice of factors by rotation, *Psychometrika4l, 9 (1944) 267–283.*

Cattell, R.B. and A.K.S. Cattell, Factor rotation for proportional profiles: Analytical solution and an example, British Journal of Statistical Psychology, **8** (1955) 83–92.

Cole, H.W. and W.J. Ray, EEG correlates of emotional tasks related to attentional demands, *International Journal of Psychophysiology*, **3** (1985) 33–41.

DeLeeuw, J. and S. Pruzansky, A new computational method to fit the weighted Euclidean distance model, *Psychometrika*, **43** (1978) 479–490.

Diaconis, P., and B. Efron, Computer-intensive methods in statistics, *Scientific American*, May (1983) 116–126, 128, 130.

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

Efron, B., The jacknife, the bootstrap and other resampling plans, *The Society for Industrial and Applied Mathematics Monographs*, **38** (1982).

Field, A.S. and D. Gaupe, Topographic component (parallel factor) analysis of multichannel evoked potentials: Practical issues in trilinear spatiotemporal decomposition, *Brain Topography*, **3** (4) (1991) 407– [illegible].

Geladi, P., Analysis of multi-way (multi-mode) data, *Chemometrics and intelligent laboratory systems*, **7** (1989) 11–30.

---

[8] For addresses in Canada, we accept $10 and/or $30 Canadian. For international addresses please send US $10.00 and/or US$30.00

Haan, N., Common dimensions of personality development: Early adolescence to middle life, in: D.H. Eichorn, J.A. Clausen, N. Haan, M.P. Honzik and P.H. Mussen (Eds.), *Present and Past in Middle Life* (Academic Press, New York, 1981) pp. 117–151.

Hampson, E. and D. Kimura, Hand movement asymmetries during verbal and nonverbal tasks, *Canadian Journal of Psychology*, **38** (1984) 102–125.

Harshman, R.A., Foundations of the PARAFAC procedure, *UCLA Working Papers in Phonetics*, **16**, (1970) 1–84. (University Microfilms International No. 10,085).

Harshman, R.A., Determination and proof of minimum uniqueness conditions for PARAFAC1, *UCLA Working Papers in Phonetics*, **22** (1972) 111–117. (University Microfilms International No. 10,085).

Harshman, R.A. and W.S. DeSarbo, An application of PARAFAC to a small sample problem, demonstrating preprocessing, orthogonality constraints, and split-half diagnostic techniques, in: H.G. Law, C.W. Snyder, Jr., J.A. Hattie and R.P. McDonald (Eds.), *Research Methods for Multimode Data Analysis* (Praeger, New York, 1984) pp. 602–642.

Harshman, R., P. Ladefoged and L. Goldstein, Factor analysis of tongue shapes, *Journal of the Acoustical Society of America*, **62** (1977) 693–707.

Harshman, R.A. and M.E. Lundy, The PARAFAC model for three-way factor analysis and multidimensional scaling, in: H.G. Law, C.W. Snyder, Jr., J.A. Hattie and R.P. McDonald (Eds.), *Research Methods for Multimode Data Analysis* (Praeger, New York, 1984a) pp. 122–215.

Harshman, R.A. and M.E. Lundy, Data preprocessing and the extended PARAFAC model, in: H.G. Law, C.W. Snyder, Jr., J.A. Hattie and R.P. McDonald (Eds.), *Research Methods for Multimode Data Analysis* (Praeger, New York, 1984b) pp. 216–284.

Harshman, R.A. and M.E. Lundy, *'Oblique and still unique?' Applying parallel profiles in more general models for factor analysis and multidimensional scaling*, Paper presented at the joint meeting of Psychometric and Classification Societies, New Brunswick, New Jersey, June, 1991.

Harshman, R.A. and M.E. Lundy, *Three-way DEDICOM: Analyzing multiple matrices of asymmetric relationships*, Paper presented at the meetings of the Psychometric Society, Columbus, Ohio, July, 1992.

Horan, C.B., Multidimensional scaling: Combining observations when individuals have different perceptual structures, *Psychometrika*, **34** (1969) 139–65.

Kettenring, J.R., A case study in data analysis, in: R. Gnanadesikan (Ed.), *Statistical Data Analysis* (Proceedings of Symposia in Applied Mathematics, Vol. 28) Providence, Rhode Island: American Mathematical Society (1983) pp. 105–139.

Kiers, H.A.K. and W.P. Krijnen, An efficient algorithm for PARAFAC of three-way data with large numbers of observation units, *Psychometrika*, **56** (1991) 147–152.

Kroonenberg, P.M., *Three-mode Principal Components Analysis: Theory and Applications* (DSWO Press, Leiden, The Netherlands, 1983).

Kruskal, J.B., More factors than subjects, tests than treatments: An indeterminacy theorem for canonical decomposition and individual differences scaling, *Psychometrika*, **41** (1976) 281–293.

Kruskal, J.B., Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics, *Linear Algebra and Its Applications*, **18** (1977) 95–138.

Kruskal, J.B., Factor analysis and principal components: Bilinear methods, in: W.H. Kruskal and J.M. Tanur (Eds.), *International Encyclopedia of Statistics* (Free Press, New York, 1978).

Kruskal, J.B., Multilinear Methods, in: H.G. Law, C.W. Snyder, Jr., J.A. Hattie and R.P. McDonald (Eds.), *Research Methods for Multimode Data Analysis* (Praeger, New York, 1984) pp. 36–62.

Kruskal, J.B., Rank, decomposition, and uniqueness for 3-way and N-way arrays, in: R. Coppi and S. Bolasco (Eds.), *Multiway Data Analysis* (North-Holland, New York, 1989) pp. 7–18.

Kruskal, J.B., R.A. Harshman and M.E. Lundy, How 3-MFA data can cause degenerate PARAFAC solutions, among other relationships, in: R. Coppi and S. Bolasco (Eds.), *Multiway Data Analysis* (North-Holland, New York, 1989) pp. 115–122.

Leurgans, S.E. and R.T. Ross, *Multilinear models: Application in Spectroscopy*, Tech. rep. No. 459, (Department of Statistics and Department of Biochemistry, Ohio State University, Columbus, 1991).

Lundy, M.E. and R.A. Harshman, *Reference Manual for the PARAFAC Analysis Package* (Scientific Software Associates, London, Canada, 1985).

Lundy, M.E., R.A. Harshman and J.B. Kruskal, A two-stage procedure incorporating good features of both trilinear and quadrilinear models, in: R. Coppi and S. Bolasco (Eds.), *Multiway Data Analysis* (North-Holland, New York, 1989) pp. 123–130.

Meredith, W., Notes on factorial invariance. *Psychometrika*, **29** (1964) 177–185.

Meyer, J.P., Causal attribution for success and failure: A multivariate investigation of Dimensionality, formation and consequences, *Journal of Personality and Social Psychology*, **38** (1980) 704–718.

Ossenkopp, K-P. and D.S. Mazmanian, The measurement and integration of behavioral variables: Aggregation and complexity as important issues, *Neurobehavioral Toxicology and Teratology*, **7** (1985) 95–100.

Ramsay, J.O., Solving implicit equations in psychometric data analysis, *Psychometrika*, **40** (1975) 337–360.

Ross, R.T., C-H. Lee, C.M. Davis, B.M. Ezzeddine, E.A. Fayyad and S.E. Leurgans, Resolution of the fluorescence spectra of plant pigment-complexes using trilinear models, *Biochimica et Biophysica Acta*, **1056** (1991) 317–320.

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

Smilde, A.K. and D.A. Doornbos, Three-way methods for the calibration of chromatographic systems: Comparing PARAFAC and three-way PLS, *Journal of Chemometrics*, **5** (1991) 345–360.

Smilde, A.K. and D.A. Doornbos, Simple validatory tools for judging the predictive performance of PARAFAC and three-way PLS, *Journal of Chemometrics*, **6** (1992) 11–28.

Snyder, C.W., Jr., W.D. Walsh and P.R. Pamment, Three-mode PARAFAC factor analysis in applied research, *Journal of Applied Psychology*, **68** (1983) 572–583.

Snyder, C.W., Jr., H.G. Law and J.A. Hattie, Overview of multimode analytic methods, in: H.G. Law, C.W. Snyder, Jr., J.A. Hattie and R.P. McDonald (Eds.), *Research Methods for Multimode Data Analysis* (Praeger, New York, 1984) pp. 2–35.

Takane, Y., H.A.L. Kiers and J. de Leeuw, *Principal component analysis with different sets of constraints on different dimensions*, Manuscript and Paper Presented at the 1991 Joint Meetings of the Psychometric and Classification Societies, New Jersey, June, 1991.

Ten Berge, J.M.F. Kruskal's Polynomial for $2 \times 2 \times 2$ Arrays and a generalization to $2 \times n \times n$ arrays. *Psychometrika*, **56** (1991) 631–636.

Ten Berge, J.M.F., H.A.L. Kiers and J. de Leeuw, Explicit Candecomp/Parafac solutions for a contrived $2 \times 2 \times 2$ array of rank three, *Psychometrika*, **53** (1988) 579–584.

Trick, L. and A.N. Katz, The domain interaction approach to metaphor processing: Relating individual differences and metaphor characteristics, *Metaphor and Symbolic Activity*, **1** (1986) 185–213.

Tucker, L.R., Implications of factor analysis of three-way matrices for measurement of change, in: C.W. Harris (Ed.) *Problems in Measuring Change*, (University of Wisconsin Press, Madison, WI, 1963).

Tucker, L.R., The extension of factor analysis to three-dimensional matrices, in: N. Frederiksen and H. Gulliksen (Eds.), *Contributions to Mathematical Psychology* (Holt, Rinehart, and Winston, New York, 1964).

Tucker, L.R., Some mathematical notes on three-mode factor analysis, *Psychometrika*, **31** (1966) 279–311.

Tucker, L.R., Relations between multidimensional scaling and three-mode factor analysis, *Psychometrika*, **37** (1972) 3–27.

Weeks, D.G. and P.M. Bentler, A comparison of linear and monotone multidimensional scaling models, *Psychological Bulletin*, **86** (1979) 349–354.

Weinberg, S.L., J.D. Carroll and H.S. Cohen, Confidence regions for INDSCAL using the jacknife and bootstrap techniques, *Psychometrika*, **49** (1984) 475–491.

Wold, H., Estimation of principal components and related models by iterative least-squares, in: P.R. Krishnaiah (Ed.), *Multivariate Analysis*, (1966).

Yates, A., (1987). *Multivariate Exploratory Data Analysis: A Perspective on Exploratory Factor Analysis* (State University of New York Press, Albany, N.Y., 1987) pp. 354.