MULTIWAY DATA ANALYSIS
R. Coppi and S. Bolasco (Editors)

A TWO-STAGE PROCEDURE INCORPORATING GOOD FEATURES OF BOTH TRILINEAR AND QUADRILINEAR MODELS

Margaret E. Lundy, Richard A. Harshman
Psychology Dept., University of Western Ontario, London, Canada N6A 5C2
Joseph B. Kruskal
Statistics Dept., AT&T Bell Laboratories, Murray Hill, New Jersey, USA

Abstract

Applications of trilinear (PARAFAC-CANDECOMP) factor/component analysis to data requiring the more complex quadrilinear (Tucker T3 or T2) model sometimes produces uninterpretable "degenerate" solutions, in which two or more factors are highly negatively correlated (see Harshman, Lundy, & Kruskal, to appear; Kruskal, Harshman, & Lundy, this volume). The more general model does not have this problem, but it is subject an axis indeterminacy that leaves some interesting questions unanswered. Described here is a two-stage procedure that combines the strengths of the trilinear (PARAFAC) and quadrilinear (Tucker T3) models to better deal with such problems. An application to real data illustrates how it provides unique meaningful axes along with a core matrix that can give substantive insights into the data complexities that caused the degeneracies. More general models are also discussed.

1. INTRODUCTION AND BACKGROUND

Two well-known methods of three-way component analysis are PARAFAC- CANDECOMP (hereafter referred to as PARAFAC), which uses a trilinear model (Carroll and Chang, 1970; Harshman, 1970; Harshman and Lundy, 1984b) and Tucker's model (Tucker, 1964, 1966)—we refer to it as the T3 model—which is a quadrilinear one. As discussed in Harshman, Lundy and Kruskal (to appear), each model has strengths that make it particularly appropriate for certain situations and also weaknesses that sometimes hinder its application. The occurrence of "degenerate" solutions is a problem that has arisen with some PARAFAC applications; Harshman et al. describe the characteristics of such solutions in some detail. Kruskal, Harshman and Lundy (this volume) mathematically show how the presence of T3 structure in the data can give rise to these degenerate solutions.

It has been possible to "block" the degeneracy by imposing orthogonality constraints on the PARAFAC factors in one mode. An interpretable solution can thus be obtained, but it sheds no light on exactly what caused the degeneracy (Harshman et al., to appear). The procedure that is proposed here—we call it PFCORE—goes a step further in coping with degeneracies. It combines advantages of both the T3 and PARAFAC models, giving more insight into the complex structure of the data and how it violates the PARAFAC model.

In this paper, we first present PARAFAC solutions for some real data which illustrate the need for PFCORE; we then describe the PFCORE procedure and its application to the real data example. Finally, we note how the PFCORE procedure can be extended to more general models.

2. PARAFAC APPLICATION TO "TV" DATA

The data set to which PARAFAC was applied consists of ratings of 15 American television shows on 16 scales, made by 40 subjects in 1981 (see Figure 1). The ratings were made on 13-point bipolar scales; the subjects were introductory psychology students at the University of Western Ontario who were familiar with the shows.

The Three-way Array X

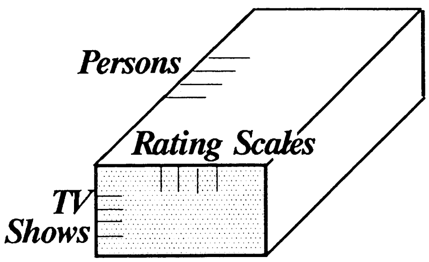

Figure 1. (above and right) Characteristics of the TV Dataset.

Rating Scales

1. Thrilling . . . . Boring
2. Intelligent . . . Idiotic
3. Erotic . . . . . Not Erotic
4. Sensitive . . . . Insensitive
5. Interesting . . Uninteresting
6. Fast . . . . . . Slow
7. Intellectually Stimulating . . Intellectually Dull
8. Violent . . . . Peaceful
9. Caring . . . . . Callous
10. Satirical . . . . Not Satirical
11. Informative . . Uninformative
12. Touching . . . "Leaves Me Cold"
13. Deep . . . . . . Shallow
14. Tasteful . . . . Crude
15. Real . . . . . . Fantasy
16. Funny . . . . . Not Funny

TV Shows

1. Mash
2. Charlie's Angels
3. All in the Family
4. 60 Minutes
5. The Tonight Show
6. Let's Make a Deal
7. The Waltons
8. Saturday Night Live
9. News (any channel; national edition)
10. Kojak
11. Mork and Mindy
12. Jacques Cousteau
13. Football
14. Little House on the Prairie
15. Wild Kingdom

Two- and three-dimensional analyses are discussed here, with the two-factor unconstrained solution, and factor 3 of the constrained 3-factor solution, presented in Figure 2. The solutions are scaled so that in Modes A and B, the mean squared factor loading on each factor is 1.0, and the Mode C loadings are adjusted in a compensatory way to reflect the scale of the data.

2.1 The 2-D PARAFAC Solution

The two-dimensional solution was well-behaved and interpretable. It accounted for 38.9% of the data variance. The correlations between columns of the factor loading matrices was low in all three modes: 0.03 in Mode A, -0.01 in Mode B and -0.35 in Mode C.

To aid in the interpretation of the solution, we construct one- way plots for each factor. Figures 2a and 2b show factor 1 and 2 of the 2D unconstrained solution. In each plot, Mode A and B loadings (i.e. for Scales and Shows) are plotted on separate vertical axes. Labels are positioned according to the size of their loading on the factor; those with loadings close to zero have been omitted. Note that the scales plot is unipolar; scales with a negative weight on the factor are plotted using the absolute value of the loading, and the label from the "negative" (opposite) end of the scale is written instead. A vertical line for the Subjects Mode is not included in any of the plots, because we had no additional information about the subjects that would help us interpret the factor. (Normally, the loadings in all three modes are used for factor interpretation.)

Figures 2a and 2b thus provide a visual summary of the relative strengths of the scales and shows that load substantially on the two factors. In Figure 2a, we see that "satirical" and "funny" (and "erotic") load highly in the scales mode; humorous shows such as Mork and Mindy and Saturday Night Live have relatively large positive weights in the shows mode while serious, informative shows like Jacques Cousteau, News and 60 Minutes have relatively large negative weights. Hence, we interpret factor 1 as a "Humor" dimension. Figure 2b indicates that the scales that load highly on factor 2 are "caring" and "sensitive" (and "touching"). The two family-life shows in the sample, The Waltons and Little House on the Prairie, load high in the positive direction while Football and News load negatively. Clearly, factor 2 is a "Sensitivity" dimension.

2.2 The 3-D PARAFAC Solution

The three-factor *unconstrained* PARAFAC solution was a typical "degenerate" solution (Harshman and Lundy, 1984a; Harshman et al., to appear). It had very high correlations between factors 1 and 2 in all three modes (0.97 in Mode A, -0.93 in Mode B and 0.82 in Mode C), the triple product of which was negative. This solution, which replicated from different starting positions, accounted for 44.4% of the data variance. Admittedly, the increase in variance-accounted-for, compared to the two-dimensional solution, is only 5 percent, but other checks lead us to believe that the third factor,

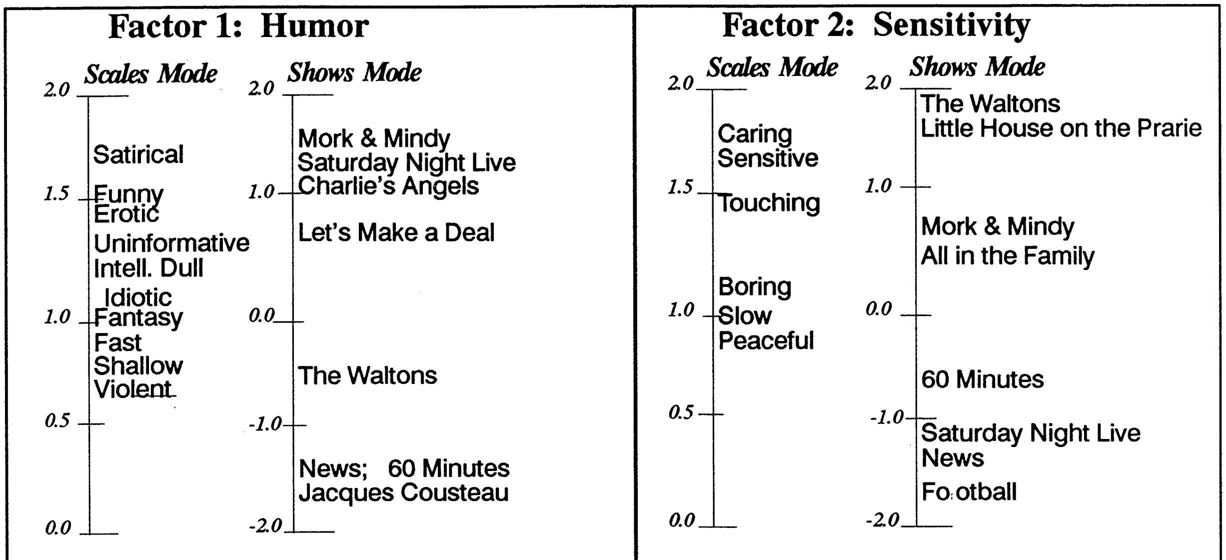

Figure 2: Two-factor unconstrained solution, TV Data

although small, is "real."

As is typical of degenerate solutions, the highly correlated dimensions were uninterpretable. They both appeared to be a combination of the funny and satirical items seen in the humor factor of the two-dimensional solution, mixed with violent items. As is also typical of many degenerate solutions, the third factor was easily interpretable. In this case it was the same "Sensitivity" factor seen as factor 2 of the two-factor solution.

2.3 The 3-D Constrained PARAFAC Solution

The three-factor analysis was repeated, but with the factors in Mode A constrained to be orthogonal. The R^2 fit value dropped only slightly, from 0.444 to 0.438. Again, factor 1 was interpreted as "Humor" and factor 2 "Sensitivity" (minor changes in weights of some scales and shows did not change the interpretation). The new information was contained in factor 3. As is apparent from Figure 3, factor 3 is clearly a "Violence" dimension, with "violent" loading extremely high in the scales mode and Football, Charlie's Angels and Kojak loading in a high positive direction in the shows mode. In contrast, nonviolent family shows Mork and Mindy, The Waltons, Little House on the Prairie and All in the Family have large negative weights.

This constrained solution is well-behaved, replicable from different starting positions and interpretable. It does not, however, explain what caused the degeneracy in the unconstrained solution. And, some information has presumably been lost, as indicated by the (small) decrease in fit.

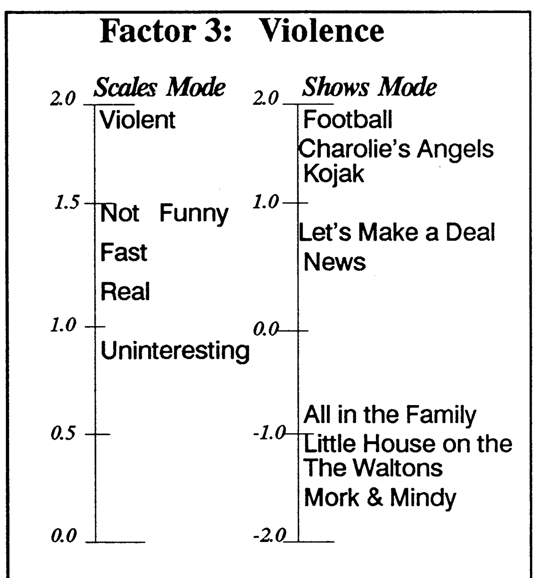

Figure 3: Third dimension of the 3D constrained Parafac analysis, TV Data.

3. PFCORE PROCEDURE

As discussed elsewhere, the degenerate PARAFAC solution is a result of some complexity in the data structure that violates the PARAFAC model. Empirical studies (Harshman and Lundy, 1984a; Harshman et al., to appear) and mathematical work (Kruskal et al., this volume) show that in fact T3 structure in the data will cause degenerate PARAFAC solutions; this suggests the need to reanalyze with a more general model like Tucker's. However, the rotational indeterminacy of T3 solutions and possible interpretational difficulties mean that the T3 model is not a perfect alternative. The PFCORE procedure that we propose combines advantages of both models: the intrinsic axis property of PARAFAC (hence no rotation is necessary) and the T3 core array (which provides information about factor interactions, not allowed for by PARAFAC).

Suppose we have an m by n by p three-way data array **X**, and x_{ijk} is an element of **X**. The PFCORE procedure involves two steps:

1. Fit the constrained PARAFAC model to the data (the factors in one mode are constrained to be orthogonal) using

$$x_{ijk} \approx \sum_{r=1}^{q} a_{ir}\, b_{jr}\, c_{kr} \tag{1}$$

where a_{ir} represents the loading or weight of factor r on the i^{th} level of Mode A, b_{jr} the loading of factor r on the j^{th} level of Mode B, and c_{kr} the factor r loading on the k^{th} level of Mode C. a_{ir} is an element in **A**, the m by q factor loading matrix for Mode A. Similarly, b_{jr} is an element of **B** and c_{kr} is an element of **C**, the n by q and p by q factor loading matrices for Modes B and C respectively. Here and elsewhere in this paper, the symbol "$\approx$" is used to represent equality except for error terms, which are not specified.

2. Using the constrained solution obtained in step 1, estimate the corresponding T3 core array via

$$x_{ijk} \approx \sum_{r=1}^{q_a} \sum_{s=1}^{q_b} \sum_{t=1}^{q_c} a_{ir}\, b_{js}\, c_{kt}\, g_{rst} \quad . \tag{2}$$

The coefficients a_{ir}, b_{js} and c_{kt} are the PARAFAC estimates of the Mode A, B and C factor loadings respectively; hence all three modes have the same number of factors and so $q_a = q_b = q_c = q$; x_{ijk} is as defined for (1). g_{rst} is the entry from the q by q by q "core array" **G**, which gives the size of the interaction between factor r in Mode A, factor s in Mode B and factor t in Mode C.

Rewriting (2) as

$$x_{ijk} \approx \sum_{r=1}^{q} a_{ir} \Big(\sum_{s=1}^{q} b_{js} \Big(\sum_{t=1}^{q} c_{kt}\, g_{rst} \Big)\Big) \tag{3}$$

and premultiplying by the appropriate elements from the generalized inverses $\mathbf{A}^{+}$, $\mathbf{B}^{+}$ and $\mathbf{C}^{+}$ of the factor loading matrices **A**, **B** and **C** respectively, we obtain the estimate of the T3 core element

$$\sum_{k} c_{kt}^{+} \Big(\sum_{j} b_{js}^{+} \Big(\sum_{i} a_{ir}^{+}\, x_{ijk} \Big)\Big) \approx g_{rst} \quad . \tag{4}$$

(Note that the order shown for the summations in (3) and (4) is arbitrary.)

In effect, the PFCORE program computes the generalized inverses of the constrained PARAFAC factor loading matrices **A**, **B** and **C**, and then multiplies the data array **X** by these inverse matrices to estimate the q by q by q T3 core array **G**. We use the loading matrices from the constrained solution rather than the degenerate one, because the constrained factors are interpretable and hence the core array will describe relationships among meaningful components.

It is important to note that the variance extracted by the constrained PARAFAC analysis is not in general the same as that extracted, for the same dimensionality, by the T3 model. Sometimes the T3 extracts a substantial variance component due to interaction of different factors before extracting a smaller third or fourth factor. Hence it is not generally possible to rotate the q by q by q dimensional T3 analysis to obtain the same dimensions and core as a q dimensional PFCORE analysis. With the TV data, for example, only some of the PFCORE factors were recoverable in the rotated T3 solution. The rotated T3 solution may more clearly bring out the cause of the degeneracy, while less clearly displaying all the distinct dimensions and their core-matrix interactions. The relationships between these two kinds of solutions should be an interesting area for further research.

4. PFCORE APPLICATION

The PFCORE procedure was applied to the TV data set, using the three-factor constrained PARAFAC solution described above to estimate the T3 core array that is shown in Table 1. It is presented as three 3 by 3 matrices of scale factors by show factors, one matrix for each subject factor or "idealized person", as we prefer to call it. The (1,1), (2,2) and (3,3) entry for the first, second and third idealized person, respectively, are the "superdiagonal" elements of the core array (i.e., the (1,1,1), (2,2,2) and (3,3,3) elements).

If the structure in the data could be entirely fit by the PARAFAC model, all elements other than the superdiagonals would be zero. However, there are some substantial off-superdiagonal values for idealized persons 1 and 3, which indicate the presence of T3 structure. Off-superdiagonal values with a magnitude greater than 0.30 (an arbitrary value) are interpreted. Kroonenberg (1983, 1984) discusses various ways to interpret the elements of the core array, but here we concentrate on their meaning as interactions among the factors in the three modes of the data.

Table 1: Core Matrix for PARAFAC Factors

First Idealized Individual (Core Slice One)

C_{F1}	B_{F1} *Funny Shows*	B_{F2} *Sensitive Shows*	B_{F3} *Violent Shows*
A_{F1} *Humor Scales*	1.058	-.257	.301
A_{F2} *Sensitivity Scales*	-.084	.102	-.216
A_{F3} *Violence Scales*	-.095	.093	-.057

Second Idealized Individual (Core Slice Two)

C_{F2}	B_{F1} *Funny Shows*	B_{F2} *Sensitive Shows*	B_{F3} *Violent Shows*
A_{F1} *Humor Scales*	.014	-.013	.089
A_{F2} *Sensitivity Scales*	-.071	.951	.113
A_{F3} *Violence Scales*	.016	-.049	-.025

Third Idealized Individual (Core Slice Three)

C_{F3}	B_{F1} *Funny Shows*	B_{F2} *Sensitive Shows*	B_{F3} *Violent Shows*
A_{F1} *Humor Scales*	-.125	.485	-.667
A_{F2} *Sensitivity Scales*	.244	-.099	.194
A_{F3} *Violence Scales*	.135	.087	1.061

Idealized person 1 rates violent shows as more humorous, whereas idealized person 3 rates sensitive shows as more humorous, but violent shows as less humorous.

The superdiagonal values indicate a large positive interaction between corresponding factors in all three modes, as expected for PARAFAC factors (factor 1 in the scales mode interacts with factor 1 in the shows mode for the first idealized person, etc.). Some of the off-superdiagonal values indicate substantial interactions amongst other combinations of factors, as allowed by the T3 model. In particular, there is a positive interaction of 0.301 between the "Humor" scale factor and the "Violence" show factor for idealized person 1. In other words, this person tends to rate violent shows as more humorous than average (or nonviolent shows as less humorous). Another way of interpreting this is to say that his/her humor ratings for violent shows are about 30 percent as great as his/her humor ratings for funny shows. (Furthermore, s/he tends to rate sensitive shows as somewhat less humorous than average, as indicated by the interaction value of -0.257 in the (1,2) cell.)

In contrast, idealized person 3 shows an opposite pattern of interactions between the "Humor" scale factor and the "Sensitivity" and "Violence" show factors. The positive interaction of 0.485 indicates that s/he sees more humor in sensitive shows than the average person does (or less humor in insensitive shows); the interaction of -0.667 indicates that s/he rates violent shows as substantially less humorous than average (or nonviolent shows as more humorous).

Thus we see that the complexities of this data set are related to differences in people's sense of humor. (The ratings of the actual subjects are weighted combinations of the idealized people.) They present problems for the simple PARAFAC model, and the large off-superdiagonal interactions for idealized persons 1 and 3 show why. They are geometrically equivalent to substantial differences in obliqueness between factors, the "Humor" and "Violence" factors especially. They cannot be represented by PARAFAC, which requires that the obliqueness between factors be invariant across individuals. PARAFAC's attempt to fit this structure was the degenerate three-dimensional solution, which revealed two highly correlated factors (negatively correlated in one mode) on which both humorous and violent scales and shows loaded highest.

5. MORE GENERAL MODELS

5.1 The T2, T1[A] and T1[B] models

Table 2 presents a list of models for fitting three-way data, starting with the most restrictive, PARAFAC, and then the more general T3 model, that were previously discussed. Next is Tucker's "extended core" model (Kroonenberg, 1983), which is called T2, followed by two even more general models that we refer to as T1(A) and T1(B). x_{ijk}, a_{ir} and b_{js} for the T2 and T1 models are as defined for T3 in (2), and the g's are elements in "core" arrays of various sizes.

Table 2: T3 & More General Models

$$x_{ijk} \approx \sum_{r=1}^{q_a} \sum_{s=1}^{q_b} \sum_{t=1}^{q_c} a_{ir}\, b_{js}\, c_{kt}\, g_{rst} \qquad \text{T3}$$

$$x_{ijk} \approx \sum_{r=1}^{q_a} \sum_{s=1}^{q_b} a_{ir}\, b_{js}\, g_{rsk} \qquad \text{T2}$$

$$x_{ijk} \approx \sum_{r=1}^{q_a} a_{ir}\, g_{rjk} \qquad \text{T1(A)}$$

$$x_{ijk} \approx \sum_{s=1}^{q_b} b_{js}\, g_{isk} \qquad \text{T1(B)}$$

The T2 core array has the same number of levels in the third mode as the data (i.e., it is q_a by q_b by p). In contrast to T3, each individual (assuming levels of Mode C are people) is allowed his/her own set of interactions between the factors in Modes A and B, rather than being represented by a weighted combination of idealized people's interactions. The T1 models reduce the data even less. For T1(A), the "core" array **G** is q_a by n by p; for T1(B), **G** is m by q_b by p. In other words, T1(A) estimates factors for Mode A only, and **G** gives the "interactions" between the Mode A factors and the levels of Mode B for each level of Mode C. With the TV data, for example, T1(A) would fit the common scales space and **G** would give the interactions between the scale factors and the individual TV shows for each subject (i.e., **G** would give each subject's "ratings" on the scale factors for each show). Similarly, T1(B) would fit the common shows space and **G** would give each subject's "ratings" for the show factors on each bipolar scale.

Besides estimating the T3 core array, the PFCORE program estimates the core arrays and computes R^2 fit values for the T2 and T1 models, using the PARAFAC estimates for the factor loading matrices **A**, **B** and **C**. Estimation of these "extended" core arrays is similar to the T3 core estimation except fewer generalized inverses are involved. For example, T2 can be written as

$$x_{ijk} \approx \sum_{r=1}^{q_a} a_{ir} \left(\sum_{s=1}^{q_b} b_{js}\, g_{rsk} \right) \quad .$$

We premultiply by the appropriate elements from the generalized inverses $\mathbf{A}^+$ and $\mathbf{B}^+$ of the PARAFAC factor loading matrices **A** and **B**, respectively to obtain the T2 core element

$$\sum_j b_{js}^+ \left(\sum_i a_{ir}^+ x_{ijk} \right) \approx g_{rsk} \quad .$$

(As in (3) and (4), the order of the summations is arbitrary.)

The T1(A) core is estimated by

$$\sum_i a_{ir}^+ x_{ijk} \approx g_{rjk} \quad ,$$

and the T1(B) core by

$$\sum_j b_{js}^+ x_{ijk} \approx g_{isk} \quad .$$

5.2. PFCORE Estimates of Fit

The fit values currently generated by PFCORE are the squared correlations between the real data and the data predicted by each model. When PFCORE is used to estimate fits for the family of models in Table 2, the pattern of values may sometimes suggest the presence of even more general structure in the data than T3. However, one needs a yardstick against which to compare the increases in fit, to see if they are larger than would be expected simply from the increase in the models' degrees of freedom. This can be obtained by using the PARAFAC data synthesis capability to generate synthetic data arrays with the same size and same systematic structure as in the constrained PARAFAC analysis of the real data, plus the appropriate amount of random noise added to produce fit values similar to those obtained with the real data. Such synthetic data arrays represent the null hypothesis, where there is no substantial structure more general than the trilinear structure of PARAFAC- CANDECOMP. In any real application, several such synthetic arrays are analyzed, to obtain an indication of stability.

The results for the TV data, where the three-factor constrained PARAFAC solution is used, are not particularly dramatic; they suggest modest T3 and perhaps T2 components, as well as some T1, but none of the comparisons are clear cut. Therefore we have used another dataset to demonstrate what can be found.

Table 3 gives some fit values for a dataset based on 40+ subjects who used bipolar rating scales to evaluate a set of metaphors. The two-dimensional unconstrained PARAFAC solution was degenerate; the constrained two-factor solution is used here. The left column shows fit values for the "real" metaphor data, and the right two columns show fit values for two synthetic data sets, these have 2D PARAFAC structure like the real metaphors, plus random noise.

As in Table 2, the models are ordered from most restrictive to most general (T1(A) and T1(B) are equally general). It is interesting that the table suggests some T3 and T2 structure, but mainly indicates a large amount of T1(A) structure in the data. This is consistent with other knowledge about these data: the raters could easily understand and agree on the dimensions underlying the rating scales, but they had very idiosyncratic responses to the metaphors, many of which were "bad" or bizarre.

	Data set being fit		
Model	*Metaphor Data (2D Sol)*	*Synthetic Data Set 1*	*Synthetic Data Set 2*
Parafac-Cc	.247	.255	.253
T3	.250	.261	.257
T2	.306	.268	.262
T1(A)	**.640**	.356	.357
T1(B)	.370	.324	.317

Table 3: PFCORE fit values

5.3 Further Development

Further research will include Monte Carlo studies, which are needed to answer two types of questions. The first is how big must the off- superdiagonal core elements be to be "real", since small values may be due to error? If we knew this, we would not use some arbitrary value, such as 0.30 with the

TV data, when deciding which core elements to interpret. The second question is how much must the fit increase to confirm the need for a more general model, since small increases may be due only to more degrees of freedom? A separate Monte Carlo study for each data set can be conducted as described above to answer this question, but a general rule would perhaps eliminate this time-consuming procedure in most cases.

6. SUMMARY AND CONCLUSION

To summarize, PFCORE incorporates good feature of two factor analysis models for three-way data—the intrinsic axis property of PARAFAC and the core array of the T3 model—in a two-step procedure. There is thus no rotation problem and the core array provides additional information about the data structure. PFCORE also provides fit values for successively more general models: Tucker's T2 model, and models that we call T1(A) and T1(B). The motivation for using PFCORE is usually a "degenerate" unconstrained PARAFAC solution, although you might sometimes input a nondegenerate unconstrained solution to see what the associated T3 core looks like (e.g., you might want to look at the T3 core corresponding to the two-factor solution for the TV data).

The application of PFCORE to the TV data set provides a good example of the extra understanding of the data that can be gained by using PFCORE. The three-factor constrained PARAFAC solution is comprehensible enough, but it is the T3 core array that sheds light on the interesting and important complexities in the data that cannot be represented by PARAFAC. A TV network executive who is planning programming based on this data, for example, would need to know not only that "Humor", "Sensitivity" and "Violence" aspects are salient to the viewer, but also how these aspects interact differently for individuals.

REFERENCES

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

Harshman, R. A. (1970). Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-mode factor analysis. UCLA Working Papers in Phonetics, 16, 1-84.

Harshman, R. A., and Lundy, M. E. (1984a). 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 (pp. 216-284). New York: Praeger.

————. (1984b). The PARAFAC model. In H. G. Law, C. W. Snyder, Jr., J. A. Hattie, and R. P. McDonald (Eds.), Research methods for multimode data analysis (pp. 122-215). New York: Praeger.

Harshman, R. A., Lundy, M. E., and Kruskal, J. B. (1985, July). Comparison of trilinear and quadrilinear methods: Strengths, weaknesses, and degeneracies. Paper presented at the annual meeting of the Classification Society, St. John's, Newfoundland.

Kroonenberg, P. M. (1983). Three-mode principal component analysis: Theory and applications. Leiden, The Netherlands: DSWO Press.

————. (1984). Three-mode principal component analysis: Illustrated with an example from attachment theory. In H. G. Law, C. W. Snyder, Jr., J. A. Hattie, and R. P. McDonald (Eds.), Research methods for multimode data analysis (pp. 64-103). New York: Praeger.

Kruskal, J. B., Harshman, R. A., and Lundy, M. E. (1985, July). Several mathematical relationships between PARAFAC-CANDECOMP and three-mode factor analysis. Paper presented at the annual meeting of the Classification Society, St. John's, Newfoundland.

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

————. (1966). Some mathematical notes on three-mode factor analysis. Psychometrika, 31, 279-311.