Journal of Educational and Behavioral Statistics
Summer 2002, Vol. 27, No. 2, pp. 105–145

Constrained Principal Component Analysis: Various Applications

Michael A. Hunter
Yoshio Takane
University of Victoria

Constrained Principal Component Analysis (CPCA) is a method for structural analysis of multivariate data. It combines regression analysis and principal component analysis into a unified framework. This article provides example applications of CPCA that illustrate the method in a variety of contexts common to psychological research. We begin with a straightforward situation in which the structure of a set of criterion variables is explored using a set of predictor variables as row (subjects) constraints. We then illustrate the use of CPCA using constraints on the columns of a set of dependent variables. Two new analyses, decompositions into finer components and fitting higher order structures, are presented next, followed by an illustration of CPCA on contingency tables, and CPCA of residuals that includes assessing reliability using the bootstrap method.

Keywords: *correspondence analysis (CA), principal component analysis (PCA) projection, singular value decomposition (SVD)*

Accompanying the increased use of multivariate designs in behavioral research has been a corresponding growth in the development of new methods of multivariate data analysis. Of particular interest to researchers in clinical, social, and developmental psychology are methods that investigate the underlying structure (e.g., factor structure) of sets of variables and simultaneously the interrelationships among those underlying structures. Adding to traditional methods for addressing these issues (e.g., canonical correlation analysis) are techniques such as Structural Equation Modeling (SEM) (Joreskog, 1970), and if the data are categorical, Correspondence Analysis (CA) (Greenacre, 1984; Nishisato, 1980). A recently developed multivariate technique is constrained principal component analysis (CPCA) (Takane & Shibayama, 1991; Takane & Hunter, 2001). CPCA is a very general method that combines regression analysis and Principal Component Analysis (PCA) into a unified framework that can be used with data measured on

This research was supported by grant A6394 from the Natural Sciences and Engineering Research Council of Canada and by grant 410-89-1498 from the Social Sciences and Humanities Research Council of Canada to the second author. The authors thank Joe Rodgers and Forrest Young for providing data for the analysis.

TABLE 1
Mezzich's (1978) Data on Four Prototypical Psychiatric Patients by 11 Psychiatrists

		SC	A	EW	CD	GF	T	MP	GR	DM	H	S	HB	MR	U	UT	BA	E
1. Manic-depressive (Depressed)	1	4	3	3	0	4	3	0	0	6	3	2	0	5	2	2	2	1
	2	5	5	6	2	6	1	0	0	6	1	0	1	6	4	1	4	0
	3	6	5	6	5	6	3	2	0	6	0	5	3	6	5	5	0	0
	4	5	5	1	0	6	1	0	0	6	0	1	2	6	0	3	0	2
	5	6	6	5	0	6	0	0	0	6	0	4	3	5	3	2	0	0
	6	3	3	5	1	4	2	1	0	6	2	1	1	5	2	2	1	1
	7	5	5	5	2	5	4	1	1	6	2	3	0	6	3	5	2	3
	8	4	5	5	1	6	1	1	0	6	1	1	0	5	2	1	1	0
	9	5	3	5	1	6	3	1	0	6	2	1	1	6	2	5	5	0
	10	3	5	5	3	2	4	2	0	6	3	2	0	6	1	4	5	1
	11	5	6	6	4	6	3	1	0	6	2	0	0	6	4	4	6	0
2. Manic-depressive (Manic)	1	2	2	1	2	0	3	1	6	2	3	3	2	1	4	4	0	6
	2	0	0	0	4	1	5	0	6	0	5	4	4	0	5	5	0	6
	3	0	3	0	5	0	6	0	6	0	3	2	0	0	3	4	0	6
	4	0	0	0	3	0	6	0	6	1	3	1	1	0	2	3	0	6
	5	3	4	0	0	0	5	0	6	0	6	0	0	0	5	0	0	6
	6	2	4	0	3	1	5	1	6	2	5	3	0	0	5	3	0	6
	7	1	2	0	2	1	4	1	5	1	5	1	1	0	4	1	0	6
	8	0	2	0	2	1	5	1	5	0	2	1	1	0	3	1	0	6
	9	0	0	0	6	0	5	1	6	0	5	5	4	0	5	6	0	6
	10	5	5	1	4	0	5	5	6	0	4	4	3	0	5	5	0	6
	11	1	3	0	4	1	4	2	6	3	3	2	0	0	4	3	0	6
3. Simple Schizophrenic	1	3	2	5	2	0	2	2	1	2	1	2	0	1	2	2	4	0
	2	4	4	5	4	3	3	1	0	4	2	3	0	3	2	4	5	0
	3	2	0	6	3	0	0	5	0	0	3	3	2	3	5	3	6	0
	4	1	1	6	2	0	0	1	0	0	3	0	1	0	1	1	6	0
	5	3	3	5	6	3	2	5	0	3	0	2	5	3	3	5	6	2
	6	3	0	5	4	0	0	3	0	2	1	1	1	2	3	3	6	0
	7	3	3	5	4	2	4	2	1	3	1	1	1	4	2	2	5	2
	8	3	2	5	2	2	2	2	1	2	2	3	1	2	2	3	5	0
	9	3	3	6	6	1	3	5	1	3	2	2	5	3	3	6	6	1
	10	1	1	5	3	1	1	3	0	1	1	1	0	5	1	2	6	0
	11	2	3	5	4	2	3	0	0	3	2	2	0	0	2	4	5	0
4. Paranoid Schizophrenic	1	2	4	3	5	0	3	1	4	2	5	6	5	0	5	6	3	3
	2	2	4	1	1	0	3	1	6	0	6	6	4	0	6	5	0	4
	3	5	5	5	6	0	5	5	6	2	5	6	6	0	5	6	0	2
	4	1	4	2	1	1	1	0	5	1	5	6	5	0	6	6	0	1
	5	4	5	6	3	1	6	3	5	2	6	6	4	0	5	6	0	5
	6	4	5	4	6	2	4	2	4	1	5	6	5	1	5	6	2	4
	7	3	4	3	4	1	5	2	5	2	5	5	3	1	5	5	1	5
	8	2	5	4	3	1	4	3	4	2	5	5	4	0	5	4	1	4
	9	3	3	4	4	1	5	5	5	0	5	6	5	1	5	5	3	4
	10	4	4	2	6	1	4	1	5	3	5	6	5	1	5	6	2	4
	11	3	5	5	5	2	5	4	5	2	4	6	5	0	5	6	5	5

Note. SC = Somatic Concern, A = Anxiety, EW = Emotional Withdrawal, CD = Conceptual Disorganization, GF = Guilt Feelings, T = Tension, MP = Mannerisms and Posturing, GR = Grandiosity, DM = Depressive Mood, H = Hostility, S = Suspiciousness, HB = Hallucinatory Behavior, MR = Motor Retardation, U = Uncooperativeness, UT = Unusual Thought Content, BA = Blunted Affect, E = Excitement.

interval or categorical scales, or both, to address a wide variety of multivariate questions focusing on the interrelationships among the underlying structures of sets of variables. For example, suppose we have a multivariate data set in which the variables are considered to represent a smaller number of underlying constructs. Applying simple PCA to these data would be one way to explore these constructs. Suppose further that the data are considered to represent criterion variables and that they are accompanied by additional information on subjects considered as predictor variables. In this case a more elaborate analysis (than simple PCA) might be possible by incorporating the predictor variables into the analysis of the criterion variables.

In CPCA this is accomplished by first partitioning the total variability in the criterion data into variability that is related to the predictors and variability that is independent of them (we refer to this first step as the external analysis). Then, PCA is applied separately to each of these sources of variability to explore possible underlying structures that are specifically related (or unrelated) to the predictor variables. This step is called the internal analysis.

As an example, consider the data in Table 1. The data, originally reported in Mezzich (1978), are ratings of each of 11 psychiatrists on four archetypal psychiatric patients, 1. Manic-Depressive-Depressed (MDD), 2. Manic-Depressive-Manic (MDM), 3. Simple Schizophrenic (SS), and 4. Paranoid Schizophrenic (PS), on 17 psychopathological items from the Brief Psychiatric Rating Scale. The 17 items are: 1. Somatic Concern (SC), 2. Anxiety (A), 3. Emotional Withdrawal (EW), 4. Conceptual Disorganization (CD), 5. Guilt Feelings (GF), 6. Tension (T), 7. Mannerisms and Posturing (MP), 8. Grandiosity (GR), 9. Depressive Mood (DM), 10. Hostility (H), 11. Suspiciousness (S), 12. Hallucinatory Behavior (HB), 13. Motor Retardation (MR), 14. Uncooperativeness (U), 15. Unusual Thought Content (UT), 16. Blunted Affect (BA), and 17. Excitement (E).

Applying simple PCA to these data, we find that the first two components account for 66.1% of the total variation (see Table 2). The two-dimensional loading matrix is illustrated in Figure 1a where it can be seen that items characteristic of simple schizophrenic symptoms (items EW and BA) are located in the first quadrant, those related to paranoid schizophrenia (items H, S, HB, U, and UT) are in the second quadrant, those characteristic of manic state (items GR and E) are in the third quadrant, and those characteristic of depression (GF, DM, and MR) are in the fourth quadrant. Matrices such as the one illustrated in Figure 1a are variously called loading matrices, pattern matrices, and structure matrices. Pattern matrices contain regression-like weights that when applied to component scores provide the best approximation to the data. Structure matrices contain correlations between variables and components. When components are orthogonal (uncorrelated) as is true of all the example applications included in this article, pattern matrices and structure matrices are one and the same and are referred to as loading matrices. Thus, loading matrices contain correlations between variables and components, *and* these correlations can be considered as weights which when applied to component scores provide the best approximation to the data.

These results reflect the structure of the total variation in the data. We may also partition this total variation into the between-patient effect and the within-patient effect using regression analysis with dummy coding and then separately explore the structure of each part. CPCA performs just such an analysis. Applying it to the current data, we find that 74.7% of the total variation is accounted for by the between-patient effect (Table 2). The first two components of this between-patient variation account for 83.8% of the overall between-patient variation (and 62.5% of the total variation). The two-dimensional between-patients loading matrix is plotted in Figure 1b, where we see a striking similarity to the results presented in Figure 1a. In contrast, Table 2 and Figure 1c show that the two-dimensional structure of the within-patient effect accounts for only 25.3% of the total variation and is very different from the previous two configurations. Typically, configurations associated with minor percentages of the data are not interpreted. We include Figure 1c here

TABLE 2
SS and Percent SS for Mezzich's (1978) Data

Source	External Analysis	Internal Analysis 1	2	1 + 2
Total	748.0	364.0	133.5	497.5
	(100.0%)	(48.7%)	(17.4%)	(66.1%)
Between	558.5	356.1	111.5	467.1
Percent total	(74.7%)	(47.6%)	(14.9%)	(62.5%)
Percent between	(100.0%)	(63.8%)	(20.0%)	(83.8%)
Within	189.5	37.2	26.0	63.2
Percent total	(25.3%)	(5.0%)	(3.5%)	8.5
Percent within	(100.0%)	(19.6%)	(13.7%)	(33.3)

Note. Percent SS are shown in parentheses.

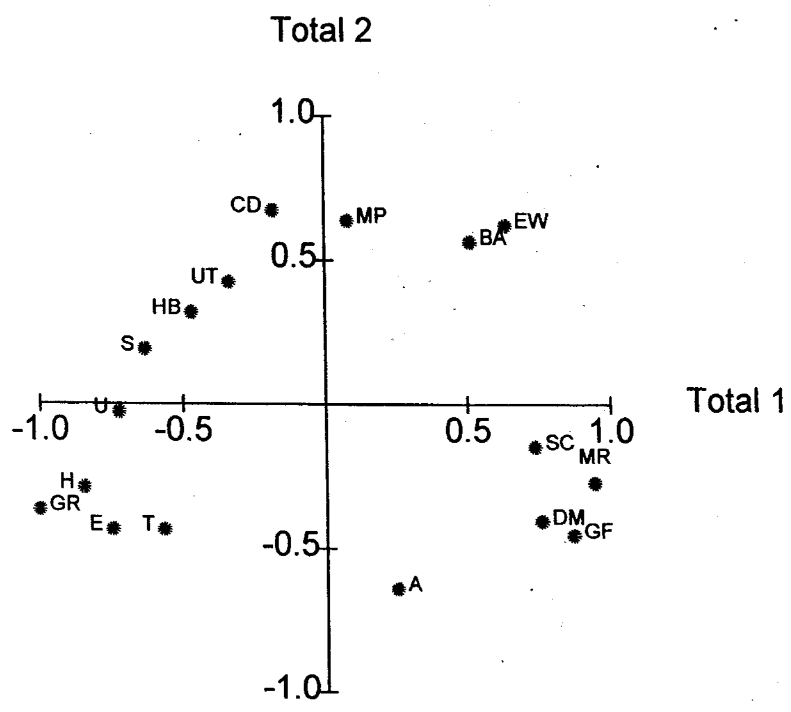

FIGURE 1a. *Component loadings from the Total of Mezzick's Rating Scale Data.*[1]

and similar low percentage configurations throughout the ensuing examples for illustrative purposes only. Combined, the current results suggest that most of the variation in the original data set, and the structure arising out of it, is due to differences among patient groups.

In addition to uncovering the structure of the predictable portion of the criterion data as shown in Figure 1b, the structure of the corresponding predictor variables is also of interest and is shown in Figure 1d (i.e., Figures 1b and 1d complement each other and are interpreted together). The sign and size of the predictor loadings indicate that the difference between the combined depressed and simple schizophrenic patients and the combined manic and paranoid schizophrenic patients accounts for most of the variation (47.6%) in the ratings. The second component, accounting for 14.9% of the total variation, contrasts depressed and manic patients with the two groups of schizophrenics.

The analysis described above may be summarized as follows. We have a matrix of criterion variables, **Z**, which in the current example is the 44 by 17 matrix given in Table 1. We also have a matrix of predictor variables, **G**, which is a 44 by 4 matrix of dummy-coded variables indicating the four patient cate-

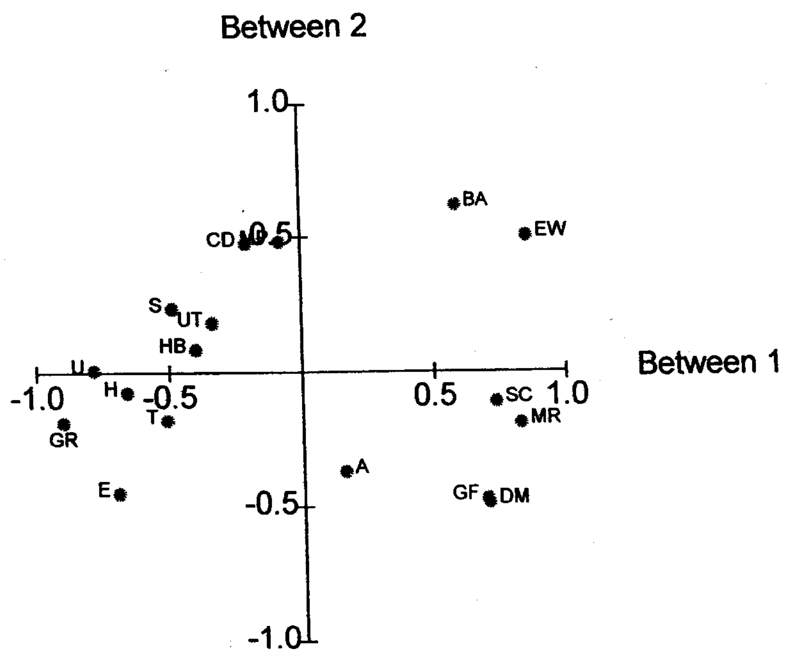

FIGURE 1b. *Component loadings from the between analyses of Mezzick's Rating Scale Data.*

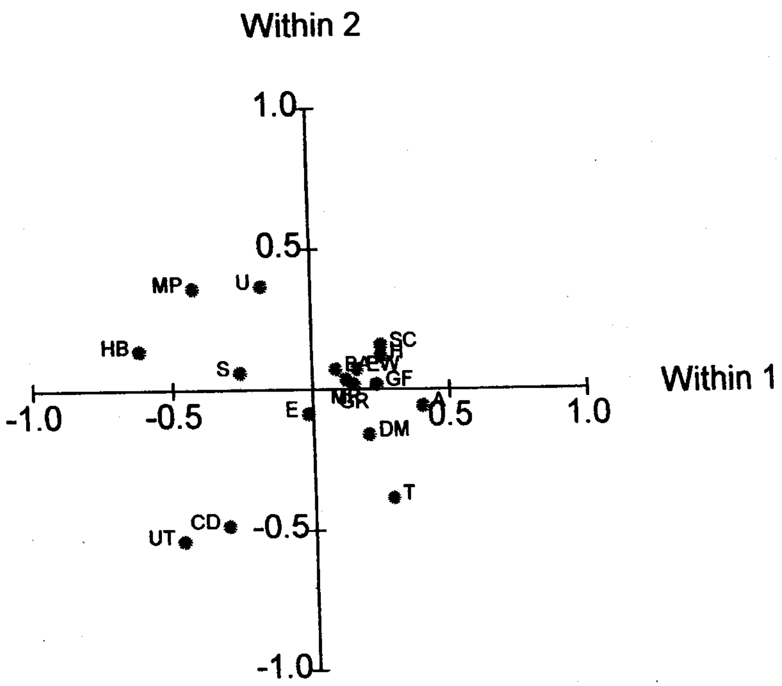

FIGURE 1c. *Component loadings from the within analyses of Mezzick's Rating Scale Data.*

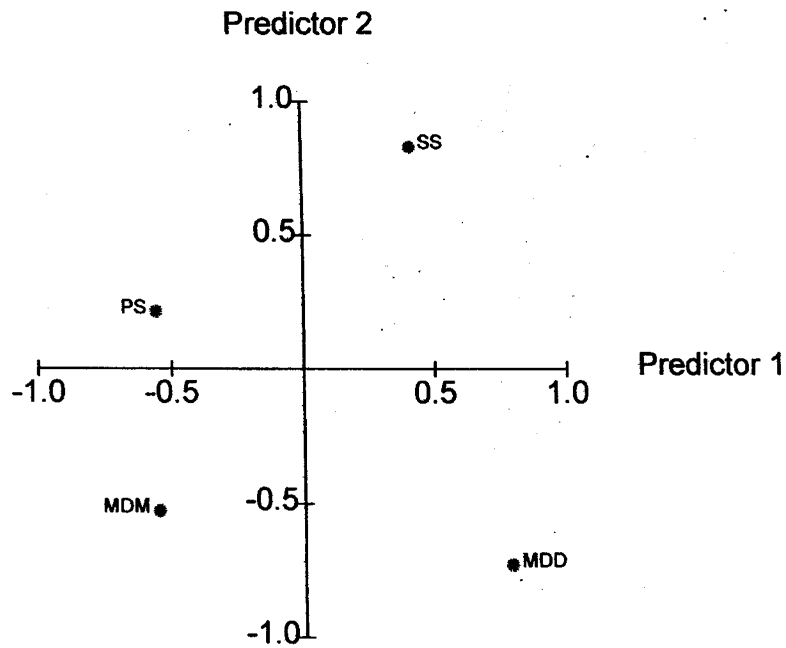

FIGURE 1d. *Corresponding predictor loadings on components from the constrained solution.*

gories. This particular choice of coding scheme produces predicted values equal to the group means. Alternative coding schemes such as effects coding and contrast coding that reflects specific a priori comparisons could be used without affecting the percentage of variation accounted for by the between-subjects effect, although the weights will change. The choice among coding alternatives is a substantive issue determined by investigators in exactly the same way as such decisions are made in standard ANOVA. Regardless of the coding scheme used, matrix **Z** is decomposed into two parts, what can be explained by **G** and what cannot be explained by **G**, using multivariate multiple regression (e.g., Reinsel & Velu, 1998). That is,

$$\mathbf{Z} = \mathbf{GC} + \mathbf{E}, \tag{1}$$

where $\mathbf{C} = (\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\mathbf{Z}$ is a matrix of regression coefficients which when applied to **G** produces a matrix of predicted scores, $\hat{\mathbf{Z}}$, and **E** is a matrix of error components equal to $\mathbf{Z}-\hat{\mathbf{Z}}$. This initial partitioning of **Z** is referred to as the external analysis.

We then apply PCA separately to $\mathbf{GC} = \hat{\mathbf{Z}}$ and $\mathbf{E} = \mathbf{Z}-\hat{\mathbf{Z}}$, referred to as the internal analysis. Technically, this is equivalent to the Singular Value Decomposition

(SVD) of a rectangular matrix. Let **X** be the rectangular matrix whose SVD is to be obtained. The SVD decomposes **X** into a product of three matrices:

$$\mathbf{X} = \mathbf{UDV}', \tag{2}$$

such that $\mathbf{U}'\mathbf{U} = \mathbf{I}$, $\mathbf{V}'\mathbf{V} = \mathbf{I}$ and **D** is diagonal with positive diagonal elements. Matrix **U** is called the matrix of left singular vectors, which when scaled as NU are equivalent to component scores. Matrix **V** is the matrix of right singular vectors, equivalent to component weights, and $\mathbf{VD}/\sqrt{\mathrm{N}}$ gives the matrix of component loadings. The best reduced-rank (lower dimensional) approximation of **X** can be obtained by discarding appropriate portions of **U**, **V**, and **D**, making PCA useful for extracting the few most important dimensions in the data. Figures 1a, 1b, and 1c show the loadings (truncated to two dimensions) obtained by the SVD of **Z**, $\hat{\mathbf{Z}}$, and $\mathbf{Z}-\hat{\mathbf{Z}}$ respectively. Figure 1d shows the predictor variable loadings. These result from correlating the predictor variables (**G**) and the component scores ($\sqrt{\mathrm{N}}\mathbf{U}$) obtained from the SVD of the predictable part of the criterion variables, $\hat{\mathbf{Z}}$. Thus, CPCA finds predictor variates that explain a maximum proportion of the variation in the criterion variables and uncovers the structure of the corresponding criterion variates. Additionally, CPCA provides a separate PCA of the variation in **Z** that remains after the effects of **G** have been removed by partialling (i.e., a separate PCA of $\mathbf{Z}-\hat{\mathbf{Z}}$).

When the data in **Z** are continuous, CPCA assumes that they are measured on an interval scale (although extending the method to accommodate ordinal data would be straightforward). Scale of measurement is determined by the type of functional relationship (e.g., one-to-one, monotonic, linear) between the data and the predictions from a model (e.g, between **Z** and $\hat{\mathbf{Z}}$). This relationship is often approximately linear, even with data that appear on the surface to be ordinal (such as those in the previous example), making the interval-scale assumption justifiable in most practical situations.

Like PCA, CPCA has rotational indeterminancy. Thus, in the previous analysis, **Z**, $\hat{\mathbf{Z}}$, and $\mathbf{Z}-\hat{\mathbf{Z}}$ could all have been rotated if desired. The choice of whether to rotate and which type of rotation to use is (as for standard PCA) at the discretion of the investigator, although if one opts to rotate, the same type of rotation should be applied to each matrix. In general, configurations that are similar prior to rotation will also be similar after rotation. However, it is *not* the case that either the predictable portion of the data or the unpredictable portion will necessarily produce components that resemble those from the total data, even when they account for a relatively high percentage of variation. The components of the total data might misrepresent both the predictable and unpredictable variation in much the same way as an overall correlation might misrepresent the relationship between two variables in different groups with mean differences on either or both variables. Indeed, sorting out the relationships among the structure of total, predictable, and unpredictable variation is a primary motivation for

CPCA. In the examples that follow rotations were not applied in order to simplify presentation.

Also like PCA, CPCA is not scale invariant (see concluding remarks), meaning that different results can be obtained depending on whether the analysis is based on raw data, centered data (scores from which the means are subtracted), or standardized data. Data are most often standardized, and we follow that convention in all our examples except those involving contingency tables (Examples 4 and 5).

The idea of relating two sets of multivariate data is by no means new, and several methods have been developed in multivariate analysis to accomplish this task. An obvious issue, therefore, is how CPCA relates to these procedures. One procedure is to use multivariate multiple regression (Reinsel & Velu, 1998), which is equivalent to CPCA without the internal analysis. Such an analysis would be appropriate if each criterion variable were considered conceptually unique and therefore dimension reduction was not considered a priority.

If dimension reduction is considered a priority, one could apply PCA to **Z** first to determine the most important dimensions in the criterion data, and then incorporate the predictor information by regressing the component scores onto **G**. In other words, one could reverse the order of regression analysis and PCA relative to CPCA. This is often done to aid the interpretation of the nature of the components. A potential advantage of this procedure is that when **G** is smaller than **Z**, more components of **Z** are possible because the maximum number of dimensions obtainable by CPCA equals the smaller number of variables in **G** or **Z**. A possible disadvantage is that the variance shared between corresponding predictor and criterion variates will not be maximized as is the case using CPCA, and the solution might be clouded by variability that is not uniquely related to the predictors.

Perhaps the most commonly used method for investigating the reduced-rank relationship between two sets of variables is canonical correlation (and its special cases, MANOVA and canonical discriminant analysis). Canonical correlation analysis obtains linear combinations of each set of variables under the constraint that corresponding pairs of components are maximally correlated. It can be thought of as a process in which the predictor and criterion variables are separately component analyzed, then the results orthogonally transformed to maximize the between-set correlations. From this characterization, it is apparent that highly correlated canonical variates could arise out of components that account for very little variance in their own sets (they could even be components that would be discarded from interpretation in separate analyses). As a result, a canonical variate of the predictor variables might have a high correlation with a canonical variate of the criterion set, but nevertheless account for only a negligible proportion of the overall variance in the criterion set (and vice versa).

The overall proportion of variance in one set that is explained by a component of another set is called redundancy, and it is this quantity that CPCA maximizes, thus guaranteeing finding components of the criterion variables that are maximally explained by the predictor variables. Indeed, in the context of this simple example, CPCA is equivalent to redundancy analysis (Rao, 1964; Van den Wollenberg, 1977). It is, however, much more general and widely applicable. In the preceding example external information (**G**) was available only on rows (subjects) of the criterion matrix **Z**. If information is available on the columns (variables) it may also be incorporated into the analysis of **Z**. Using **H** to denote the column information matrix, the CPCA model now becomes:

$$\mathbf{Z} = \mathbf{GMH}' + \mathbf{BH}' + \mathbf{GC} + \mathbf{E}, \tag{3}$$

where the first term on the right-hand side of the equation represents the portion of **Z** that can be explained by both **G** and **H**, the second term represents the portion that can be explained by **H** independent of **G**, the third term represents the portion that can be explained by **G** independent of **H**, and the fourth term represents the portion that can be explained by neither **G** nor **H**. As before, PCA is then applied separately to $\mathbf{GMH}' = {}_G\mathbf{Z}_H$^, $\mathbf{BH}' = \mathbf{Z}_H$^, $\mathbf{GC} = {}_G\hat{\mathbf{Z}}$, and $\mathbf{E} = \mathbf{Z} - {}_G\hat{\mathbf{Z}} - \mathbf{Z}_H$^ $-$ ${}_G\mathbf{Z}_H$^, to find concise representations of specific portions of **Z**.

In this article we illustrate various uses and extensions of CPCA through examples. We begin with an example that highlights column (variable) constraints because in psychological research column constraints typically are less familiar than row constraints. We then proceed to illustrate several new extensions to CPCA recently discussed by Takane and Hunter (2001), including decompositions into finer components, incorporating higher order structures, and the analysis of categorical (nominal) data.

Example 1: Family Composition Preferences

We discuss an application of CPCA that focuses on column constraints using data on student preferences for various family compositions (Rodgers & Young, 1981). The data were obtained by asking 100 students to rank order their preferences for the 18 family types shown in Table 3. These 18 preferences could be component analyzed to explore whether they reflect fewer salient underlying dimensions. Alternatively, CPCA can be used to evaluate how well a few a priori characterizations of the items predict the 18 preferences and their structure. For example, items can be categorized into those having the same number of boys and girls (items 1, 6, 7, 16) versus those having unequal numbers of boys and girls (items 2, 3, 4, 5, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18). Similarly, in items 2, 4, 8, 13, 14, 15, and 17, the number of boys is larger than the number of girls, whereas girls outnumber boys in items 3, 5, 9, 10, 11, 12 and 18. Contrasting these items would represent an a priori dimension of gender preference.

Such a priori characterizations of items (or stimuli in general) can be quantified and used in CPCA in the same way as a priori knowledge about subjects. An example of

TABLE 3
The 18 Family Types used in the Analysis of Rodgers and Young's (1981) Data and Five A Priori Contrasts

			H Matrix				
Item	Label	Description	C1	C2	C3	C4	C5
1	0	No children	7	0	8	1	1
2	B	One boy	−2	1	8	1	1
3	G	One girl	−2	−1	8	1	1
4	BB	Two boys	−2	1	8	1	1
5	GG	Two girls	−2	−1	8	1	1
6	BG	A boy then a girl	7	0	8	1	−1
7	GB	A girl then a boy	7	0	8	1	−1
8	3B	Three boys	−2	1	−7	1	1
9	3G	Three girls	−2	−1	−7	1	1
10	BGG	A boy then two girls	−2	−1	−7	1	−1
11	GBG	A girl then a boy then a girl	−2	−1	−7	1	−1
12	GGB	Two girls then a boy	−2	−1	−7	1	−1
13	GBB	A girl then two boys	−2	1	−7	1	−1
14	BGB	A boy then a girl then a boy	−2	1	−7	1	−1
15	BBG	Two boys then a girl	−2	1	−7	1	−1
16	4+=	Four or more, equal boys and girls	7	0	0	−5	−1
17	4+B	Four or more, more boys than girls	−2	1	0	−5	1
18	4+G	Four or more, more girls than boys	−2	−1	0	−5	1

Note. C1 contrasts items in which the number of boys and girls is equal (scored 7) with items having more boys or more girls (scored −2).
C2 contrasts items in which there are more boys than girls (1) with items in which there are more girls than boys (−1).
C3 contrasts items in which the total number of children is less than or equal to two (8) with items in which the total number of children equals three (−7).
C4 contrasts items in which the total number of children is less than or equal to three (1) with items in which the total number of children is four or more (−5).
C5 contrasts items in which children are all the same gender (1) with mixed-gender items (−1).

such a quantification for the current example appears in Table 3. The columns of this **H** matrix refer to:

Column 1: A contrast between items in which the number of boys and girls is equal (scored 7) and those having either more boys or more girls (−2).
Column 2: A contrast between items in which there are more boys than girls (1) and items in which there are more girls than boys (−1).
Column 3: A contrast between items in which the total number of children is less than or equal to two (8) and items in which the total number of children equals three (−7).
Column 4: A contrast between items in which the total number of children is less than or equal to three (1) and items in which the total number of children is four or more (−5).

Column 5: A contrast between items in which children are all the same gender (1) and mixed-gender items (−1).

These contrasts reflect a priori notions that family composition preferences are governed primarily by number of children (Contrasts 3 and 4), gender bias (Contrast 2), and homogeneity (Contrast 1 and 5).[2]

CPCA using column constraints is equivalent to an external analysis in which the data matrix **Z** is regressed onto **H** to obtain $\mathbf{Z_H}$^ and $\mathbf{Z} - \mathbf{Z_H}$^, followed by an internal analysis in which PCA is applied to each of these matrices. That is,

$$\mathbf{Z} = \mathbf{BH}' + \mathbf{E},$$

where $\mathbf{B} = \mathbf{ZH}(\mathbf{H'H})^-$ is a matrix of regression coefficients which when applied to $\mathbf{H}'$ produces predicted scores $\mathbf{Z_H}$^, and **E** equals $\mathbf{Z} - \mathbf{Z_H}$^. This is equivalent to regressing each subject's data separately onto **H**, resulting in as many regression equations as there are subjects. The resulting matrix (**B**) will contain the regression weights relating criterion scores to the contrasts in **H** for each subject.

CPCA was applied to the family preference data using the **H** shown in Table 3. The results of the unconstrained and constrained analysis are shown in Table 4. The five between-item contrasts comprising **H** accounted for 74.1% of the total variation, this value being the total redundancy.

When PCA is applied to that part of the preference rankings that is predictable from **H**, that is, $\mathbf{Z_H}$^, a three-dimensional solution is obtained accounting for 93.8% of the between-item variation and 69.5 % of the total variation. These three dimensions account for 46.1%, 38.3%, and 9.3% of the between-item variation, respectively (with corresponding redundancies of 34.2%, 28.4%, and 6.9%). The normalized loadings for the 18 criterion items are illustrated in Figure 2b. Table 4 contains the loading matrix for the five contrasts composing **H**. From these results, it is apparent that the a priori dimensions represented in **H** received strong support. The first predictor variate is composed primarily of contrasts designed to capture family size (Contrasts 3 and 4), and the corresponding criterion variate indeed distinguishes preferences for small, medium, and large families. Similarly, Contrasts 1 and 5, which represent family homogeneity or heterogeneity, have the highest loadings on the second predictor variate, and the corresponding criterion variate differentiates preferences for mixed-gender families from preferences for single-gender families. Finally, the third predictor variate, on which Contrast 2 (gender bias) loads most heavily, is associated with a criterion variate that differentiates preferences for girls versus preferences for boys. The unconstrained PCA of these data appears in Figure 2a. Comparing it to Figure 2b, we see that the effect of the constraints is to bring the structure of the main data into sharper, more interpretable focus.

Analyses that incorporate column constraints (stimulus or item information) appear to be relatively rare in psychological research. However, this may be more illusory than real. For example, in repeated measures designs trend analysis is equivalent to placing orthogonal polynomial coefficients in **H**. In family research, **H** could contain codes for contrasting fathers, mothers, and children. In cognitive research when variables represent stimuli, a feature matrix or a matrix of stimulus descriptor variables can be taken as **H**. In general, any within-subjects (or dyads, or families) comparison can be conceptualized as a constraint on the columns of **Z**. Whether dimension reduction is also desired is a matter of substantive concern, although some interesting possibilities immediately arise. For example, when multivariate observations are collected on two or more occasions, CPCA incorporating equality constraints on identical variables over time would be equivalent to evaluating invariance in factor structure across occasions. A similar approach to evaluating structural invariance across family members is another possibility.

TABLE 4
SS and Percent SS for Rodgers and Young's (1981) Data

		Internal Analysis			
	External	Dimension			
Source	Analysis	1	2	3	1 + 2 + 3
Total	48387.5	18738.8	16055.4	4037.9	38832.2
	(100.0%)	(38.7%)	(33.2%)	(8.4%)	(80.2%)
Between Item	35860.4	16547.4	13744.1	3337.0	33628.6
Percent total	(74.1%)	(34.2%)	(28.4%)	(6.9%)	(69.5%)
Percent between	(100.0%)	(46.1%)	(38.3%)	(9.3%)	(93.8%)
					1 + 2
Residuals	12527.0	6420.0	1416.1		7836.1
Percent total	(25.9%)	(13.2%)	(2.9%)		(16.1%)
Loadings for the Predictor Variates					
	C1	0.199	−0.431	0.043	
	C2	−0.034	0.159	0.980	
	C3	0.601	0.084	0.165	
	C4	0.788	0.199	−0.072	
	C5	0.038	−0.963	0.155	

Note. There were only two dimensions remaining after between-item variation was removed. Therefore, there is no entry for Residuals under Dimension 3. Additionally, the entries listed under Dimension 1 + 2 + 3 for the Residuals are in fact totals for two (rather than three) dimensions. Percent SS are shown in parentheses.

Example 2: Decompositions into Finer Components

To this point **G** has been treated as a single set of variables. In many applications **G** is actually composed of two or more meaningful subsets. For example, in facto-

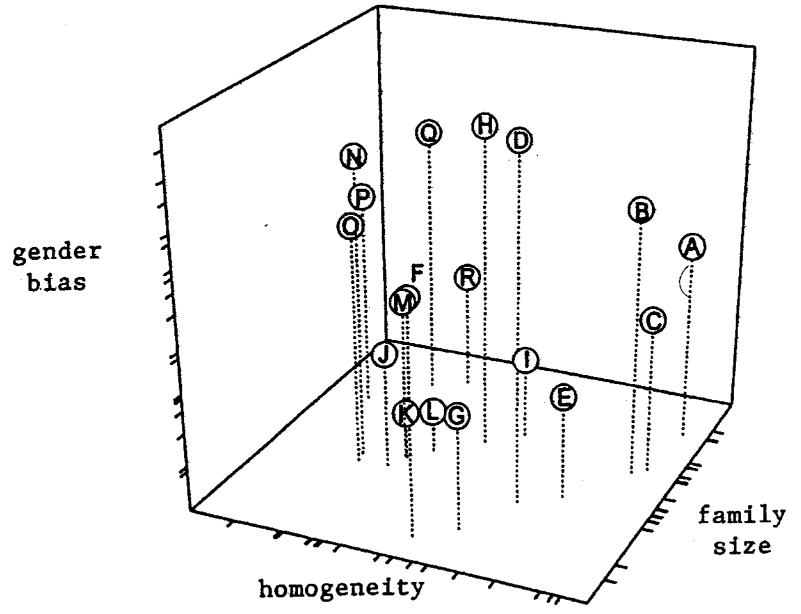

FIGURE 2a. *Total solutions from the analyses of the Rodgers and Young, (1981) data. Letters not enclosed in circles have exactly the same values as their adjacent circled letter. Their values are equal due to the imposed equality constraints. Labeling follows the same convention as in Figure 1.*

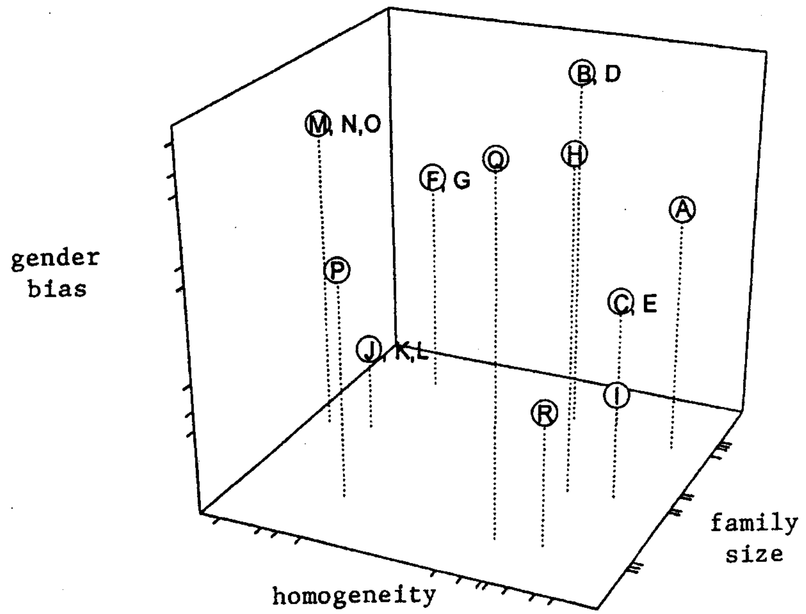

FIGURE 2b. *Between items solutions from the analyses of the Rodgers and Young Data. Letters not enclosed in circles have exactly the same values as their adjacent circled letter. Their values are equal due to the imposed equality constraints. Labeling follows the same convention as in Figure 1.*

rial designs, different sets of variables are used to represent main effects and interactions. In multivariate analysis of covariance designs, **G** can be subdivided into treatment or group effects and covariates. Even when **G** contains exclusively continuous variables, they often can be considered as forming distinguishable subsets (e.g., demographic variables versus personality scales, verbal versus performance subscales of intelligence tests, measures of auditory versus visual information processing, etc).

In all of these cases, statistical analysis usually focuses on the independent effects of each subset of **G** rather than on the overall effect of **G**. Takane and Hunter (2001) have recently shown that when subsets of constraints are available in **G** (or **H**, or both), it is possible to decompose **Z** into finer components associated uniquely with each subset. The external analysis is equivalent to regressing **Z** on that part of each **G** (or **H**) subset from which the remaining subsets have been partialled (i.e., one subset of **G** is regressed on the remaining subsets of **G** and the residuals from that analysis are used to predict **Z**). Alternatively, a hierarchical approach can be taken in which subsets are analyzed in a logical sequence with each subset having all previously entered subsets partialled prior to analysis with **Z**.

The external analysis in this case is equivalent to Cohen's (1982) multiple set correlation procedure. The internal analysis then proceeds as usual via PCA of the predictable and unpredictable parts of **Z** for each partialled subset in **G**.

We illustrate decompositions into finer components using selected data from a study investigating the effects of child sexual abuse and adult attachment style on psychological adjustment (Roche, Runtz, & Hunter, 1999). The subjects were 307 female university students, 85 of whom had been sexually abused during childhood, and 222 of whom had not been sexually abused (Non abuse: NA). Of those who had been abused, 54 had been abused outside the family (Extrafamilial Abuse: EA) and 31 had been abused by a family member (Intrafamilial Abuse: IA). All the women completed the Relationship Questionnaire (RQ) (Bartholomew & Horowitz, 1991) measure of adult attachment from which scores on four aspects of attachment were derived: Secure, Dismissing, Preoccupied, and Fearful. Psychological adjustment was measured using the anxiety (ANX), anger (ANG), depression (DEP), avoidance (AVD), sexual dysfunction (SEXD), intrusive experiences (INT), impaired self-reference (SRF), and sexual concerns (SEXC) scales from the Traumatic Symptom Inventory (Briere, 1996). Preliminary analyses indicated that the predictor sets, abuse and adult attachment, were related to each other and that each of them was separately related to psychological adjustment.

The results of an unconstrained PCA are shown in Table 5 and Figure 3a, where it can be seen that two dimensions accounting for 79.3% of the total variation serve to differentiate the scales characteristic of externalizing problems (anxiety, anger, avoidance, sexual concerns), from the scales characteristic of internalizing problems (depression, intrusive thoughts, impaired self reference, sexual dysfunction).

TABLE 5
Analyses of Roche et al., (1999) Psychological Adjustment Data

Source	External Analysis	Internal Analysis: Dimension 1	2	1+2
Unconstrained	2456.0	1379.9	570.2	1950.1
	(100.0%)	(56.1%)	(23.2)	(79.3%)
Attachment and abuse combined	1004.5	561.6	423.1	984.7
Percent unconstrained	(40.9%)	(22.9%)	(17.2%)	(40.1%)
Percent constrained	(100.0%)	(55.9%)	(42.1%)	(98.0%)
Residuals	1451.5	857.1	183.7	1040.8
Percent unconstrained	(59.1%)	(34.9%)	(7.5%)	(42.4%)
Abuse with attachment partialled	104.4	61.1	43.2	104.3
Percent unconstrained	(4.3%)	(2.5%)	(1.8%)	(4.3%)
Percent constrained	(100.%)	(57.6%)	(42.4%)	(100.0%)
Attachment with abuse partialled	717.8	387.7	322.6	710.3
Percent unconstrained	(29.2%)	(15.8%)	(13.1%)	(28.9%)
Percent constrained	(100.0%)	(54.0%)	(44.9%)	(98.9%)

Note. Both attachment and abuse are used as row constraints. Abuse is used as a constraint after attachment has been partialled. Attachment is used as a constraint after abuse has been partialled. Percent SS are shown in parentheses.

CPCA was used to find components of psychological adjustment that were maximally related to sexual abuse independent of attachment, and attachment independent of sexual abuse. For illustrative purposes we first show an analysis in which adult attachment and sexual abuse were treated as a single set of variables. The results of this analysis are shown in Table 5 and Figures 3b, 3c, and 3d.

Combined, childhood sexual abuse and adult attachment style account for 41% of the variance in psychological adjustment (this total redundancy corresponds to an average multiple correlation across the six adjustment scales of 0.64). Ninety-eight percent of the predictable variation is accounted for by two dimensions (56% and 42%, respectively, amounting to 22.9% and 17.2% of the total variation in psychological adjustment).[3] Figure 3b shows that these two dimensions differentiate the psychological adjustment scales in a pattern very similar to the unconstrained PCA, but with sharper focus; that is, they have simpler structure (and are reflected) compared to the unconstrained loadings. The predictor variable loadings shown in Figure 3d suggest that the first dimension (internalizing problems) is most related to the difference between the Secure-Dismissing and Preoccupied-Fearful attachment styles and between the non abuse and intrafamilial abuse groups. The difference between the

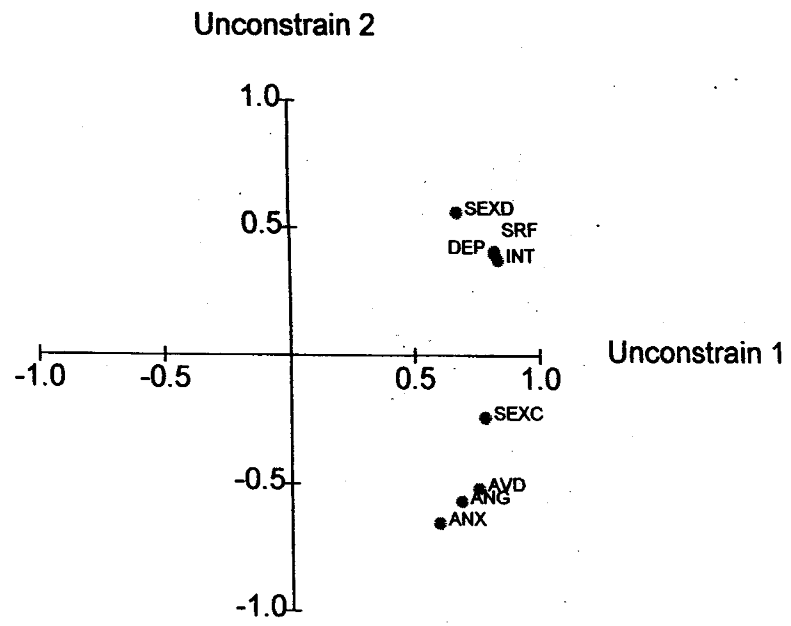

FIGURE 3a. *Component loadings from the unconstrained analyses of the Roche et. al., (1999) Psychological Adjustment Data. Figures 3a and 3d are complementary and should be interpreted together.*

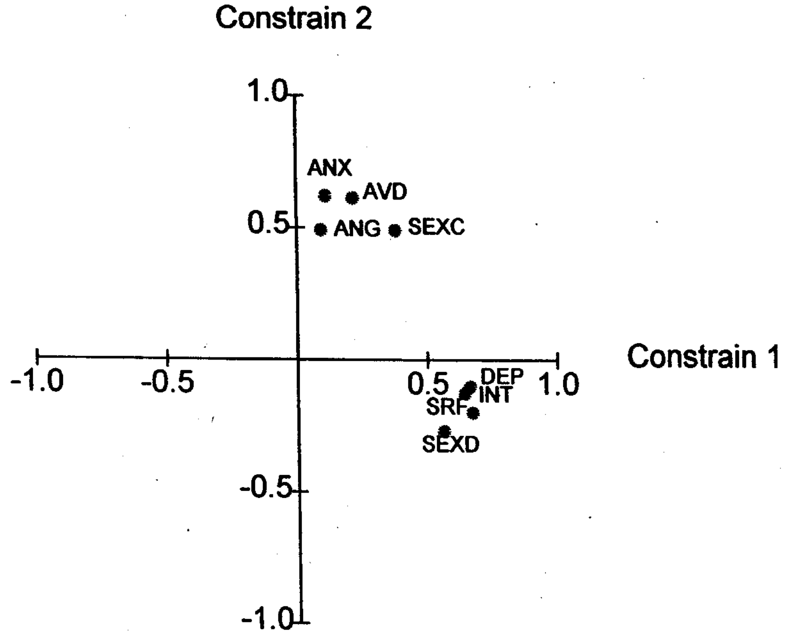

FIGURE 3b. *Component loadings from the constrained Analyses of the Roche et al. (1999) Psychological Adjustment Data.*

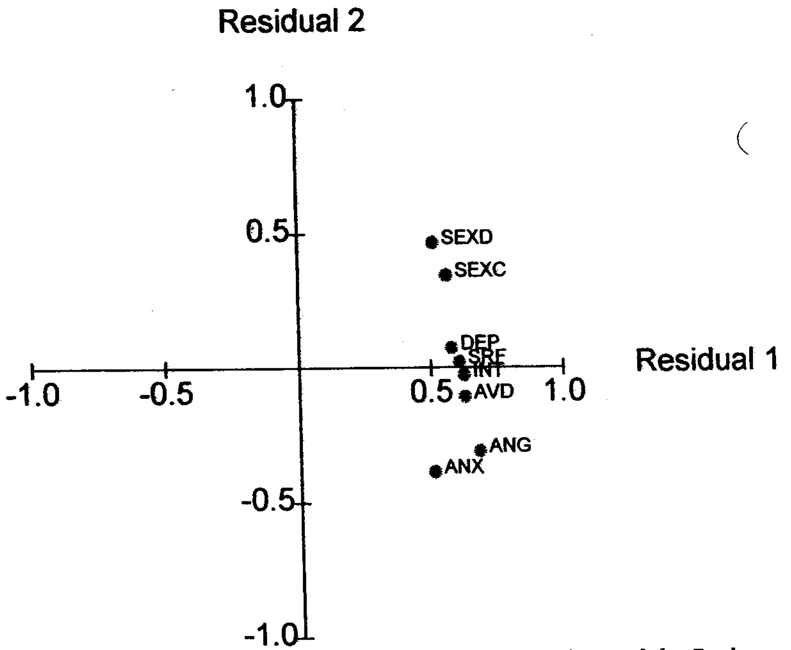

FIGURE 3c. *Component loadings from the residuals analyses of the Roche et. al., (1999) Psychological Adjustment Data.*

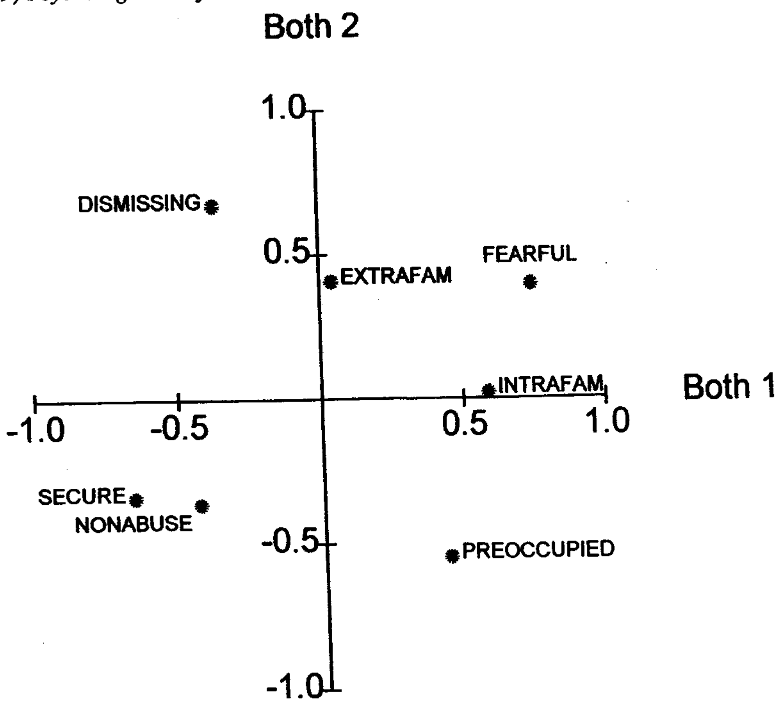

FIGURE 3d. *Corresponding predictor loading when the predictors included both Sexual Abuse and Attachment Style. Figures 3a and 3d are complementary and should be interpreted together.*

Secure-Preoccupied and Dismissing-Fearful attachment styles and between the non abuse and extrafamilial abuse groups is most related to the second dimension (externalizing problems).[4]

The results of analyses relating psychological adjustment to abuse independent of attachment style and to attachment style independent of abuse, appear in Table 5 and Figures 4a, 4b, 4c, and 4d. After partialling attachment style, abuse only accounts for 4.3 % of the total variation. This relatively low redundancy is reflected in Figure 4a, which shows that the length of the two-dimensional loading vectors is attenuated compared to previous analyses, although the direction of the vectors still suggests a solution that differentiates internalizing from externalizing scales of psychological adjustment. The predictor loadings (Figure 4b) indicate somewhat different results for the effects of abuse compared to the unpartialled analysis. When attachment is partialled from abuse, the difference between the non abused group and the combined abuse groups is related to externalizing problems, whereas the difference between extrafamilial abuse and intrafamilial abuse is related to internalizing problems.

After partialling abuse history, attachment accounted for 29.2% of the total variation in adult psychological adjustment. Ninety-nine percent of the predictable

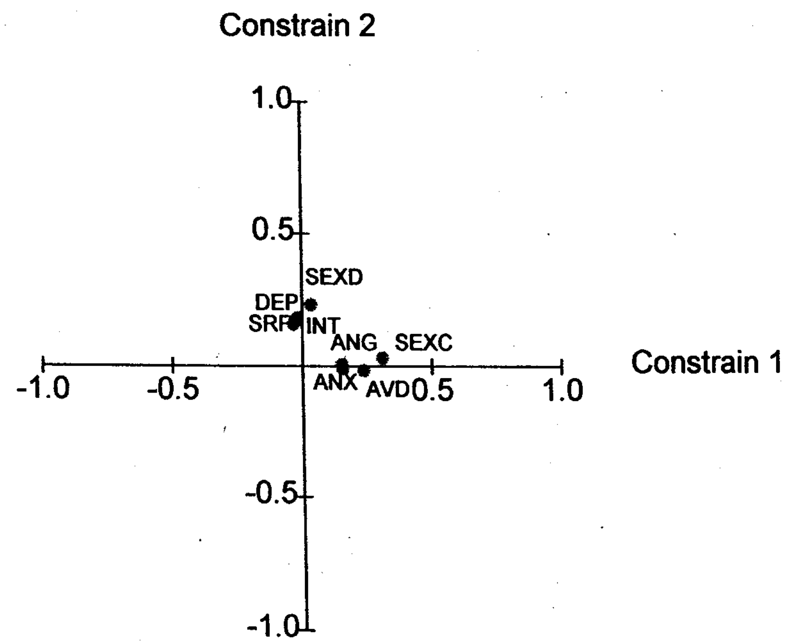

FIGURE 4a. *Component loadings from constrained analyses of the Roche et al. (1999) data isolating the effects of abuse independent of attachment.*

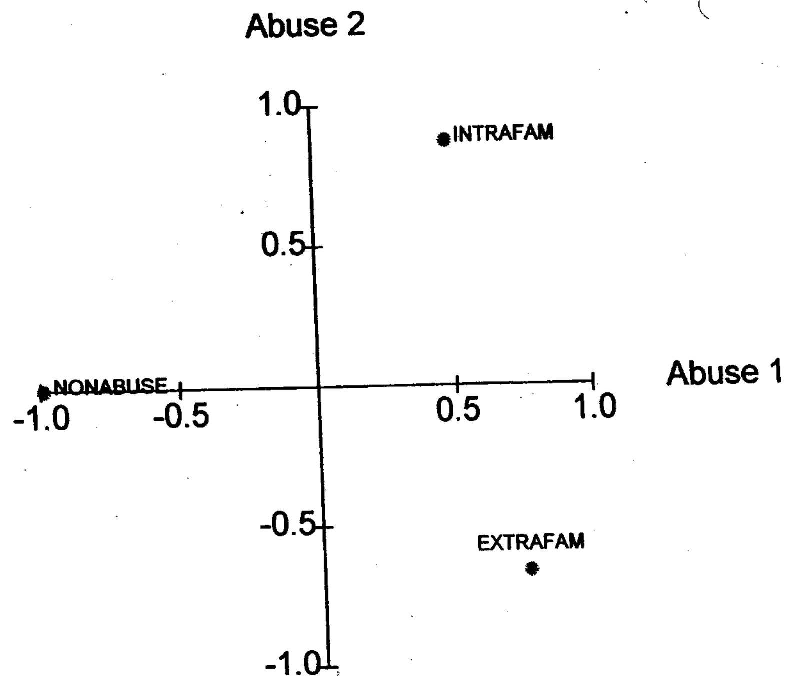

FIGURE 4b. *Component loading from the constrained analyses of the Roche et. al. (1999) data isolating the effects of abuse independent of attachment.*

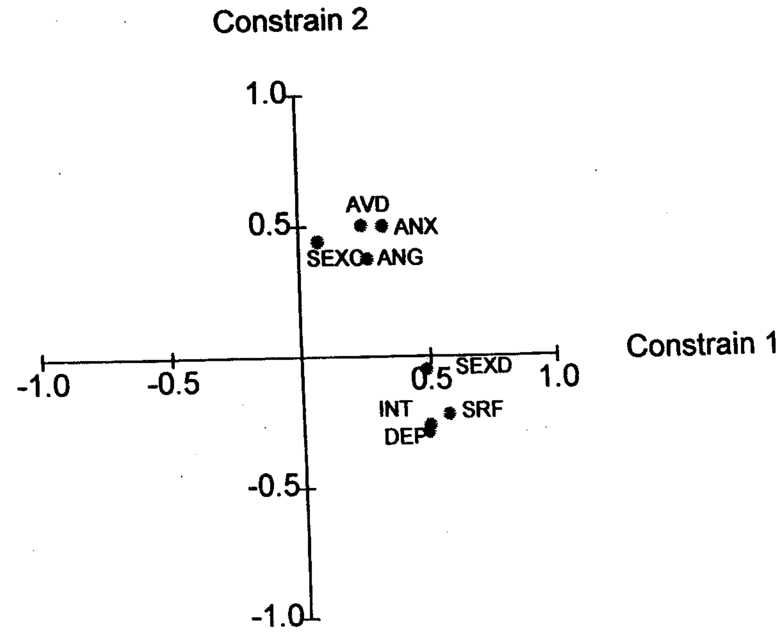

FIGURE 4c. *Component loadings from constrained analyses of the Roche et. al. (1999) data isolating the effects of attachment independent of abuse.*

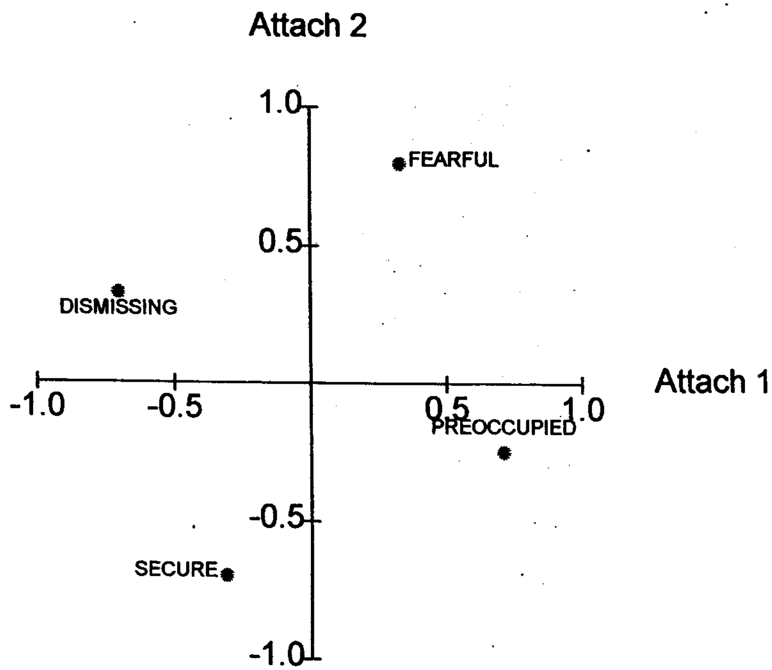

FIGURE 4d. *Component loadings from constrained analyses of the Roche et al. (1999) data isolating the effects of attachment independent of abuse.*

variation is accounted for by two dimensions (54.0% and 45.0%, respectively, amounting to 15.8% and 13.2% of the total variation in psychological adjustment). Again, these two dimensions differentiate externalizing problems (and sexual concerns) from internalizing problems (and sexual dysfunction; see Figure 4c). The corresponding predictor variates (Figure 4d) differentiate secure-dismissing from preoccupied-fearful attachment styles, and secure-preoccupied from dismissing-fearful attachment styles.

The predictor variates suggest that although four scores were used to measure adult attachment style, it might be captured in only two dimensions. Indeed, Bartholomew (1990) suggested that adult attachment can be organized in terms of Bowlby's (1982) notion of two internal working models—the model-of-self and the model-of-other. According to Bartholomew (1990), secure adult attachment reflects a positive model-of-self and a positive model-of-other; dismissing reflects a positive model-of-self and a negative model-of-other; preoccupied reflects a negative model-of-self and a positive model-of-other; and fearful reflects a negative model-of-self and a negative model-of-other (see Figure 5). The results illustrated in Figure 4d present some support for Bartholomew's hypothesis. In the next section, we give an example in which Bartholomew's hypothesis about the structure of adult attachment can be incorporated directly into CPCA.

	Positive Model of Self (Low Dependence)	Negative Model of Self (High Dependence)
Positive Model of Other (Low Avoidance)	**SECURE** Comfortable with intimacy & autonomy	**PREOCCUPIED** Preoccupied with relationships
Negative Model of Other (High Avoidance)	**DISMISSING** Dismissing of intimacy Counter-dependent	**FEARFUL** Fearful of intimacy Socially avoidant

FIGURE 5. *Relationship questionnaire two-dimensional model of attachment (Bartholomew, 1990).*

Example 3: Incorporating Higher Order Structures into CPCA

Takane and Hunter (2001) provide the detailed theory for fitting higher order structures into the CPCA model. In the context of the adult attachment and psychological adjustment example the model starts as

$$\mathbf{Z} = \mathbf{GC} + \mathbf{E},$$

where **Z** contains the eight psychological adjustment scores for all subjects, **G** contains the four attachment scores for all subjects, **C** is a matrix of parameters to be estimated, and **E** is a matrix of errors. Incorporating Bartholomew's (1990) hypothesis about the structure of attachment (Figure 5) is equivalent to estimating the model under constraints on the parameters in **C**. For example, if **C** contains parameters for weighting secure, dismissing, preoccupied, and fearful attachment styles, respectively, and constraints corresponding to Figure 5 are collected into a design matrix

$$\mathbf{A} = \begin{matrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1, \end{matrix}$$

then we can model $\mathbf{C} = \mathbf{AW} + \mathbf{E^*}$, where **W** is a matrix of weights applied to **A** in order to estimate **C** under the hypothesized higher order structure, and **E*** reflects variation that is predictable, but not by the hypothesized structure. The entire model then becomes

$$\begin{aligned} \mathbf{Z} &= \mathbf{G}(\mathbf{AW} + \mathbf{E^*}) + \mathbf{E} \\ &= \mathbf{GAW} + \mathbf{GE^*} + \mathbf{E}, \end{aligned}$$

which partitions **Z** into what can be explained by **GA** (in this context, the hypothesized structure of attachment), what can be explained by **G** independent of **GA**, and the residuals.

Applying this model to the current example (we ignore information on abuse in this example) produced the results shown in Table 6 and Figures 6a and 6b. The hypothesized structure of adult attachment fits the data well, accounting for 96% of that variation in psychological adjustment that is related to attachment. Figure 6a essentially replicates figure 4c, and the predictor loadings for the hypothesized attachment structure (Figure 6b) indicate that the model-of-self dimension is most related to internalizing problems, whereas the model-of-other is most related to externalizing problems.

The ability to incorporate nested higher order structures and decompositions into finer components into CPCA greatly expands the method. As shown in Takane and Hunter (2001) the original CPCA model and these new extensions can be combined and expressed in an equation very similar to that of COSAN for structural equation models (McDonald, 1978). The major difference between the two methods is that whereas structural equation models are fit to covariance matrices, CPCA models are fit to data matrices.

TABLE 6
Results of Fitting a Higher Order Structure on the Relationship of Attachment to Psychological Adjustment

		Internal Analysis		
	External	Dimension		
Source	Analysis	1	2	1 + 2
Attachment (Unstructured)	902.5	519.2	374.8	894.0
Percent unconstrained	(36.8%)	(21.1%)	(15.3%)	(36.4%)
Percent constrained	(100.0%)	(57.5%)	(41.5%)	(99.0%)
Residuals	1553.5	902.8	195.3	1098.1
Percent unconstrained	(63.3%)	(36.8%)	(8.0%)	(44.8%)
Attachment (Structured)	866.6	495.9	370.7	866.6
Percent unconstrained	(35.3%)	(20.2%)	(15.1%)	(35.3%)
Percent constrained	(96.0%)	(54.9%)	(41.1%)	(96.0%)

Note. Percent SS shown in parentheses.

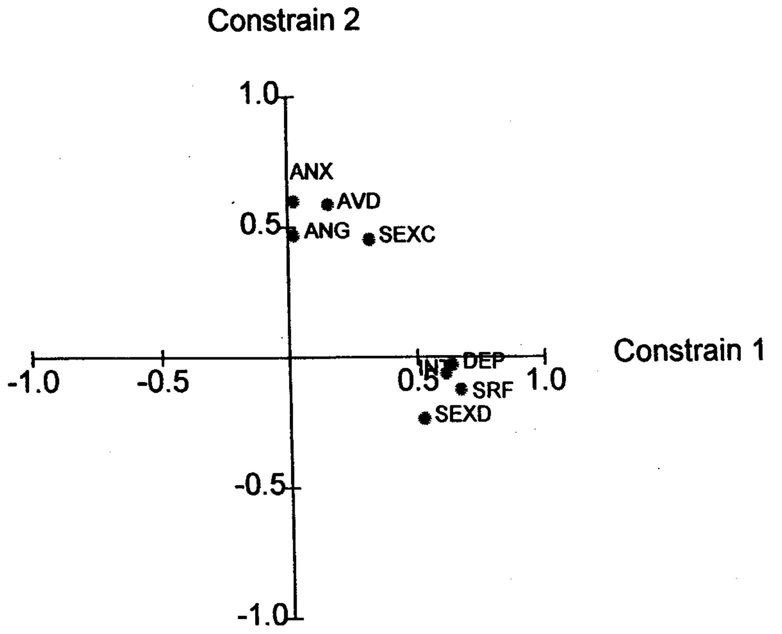

FIGURE 6a. *Component loadings from the higher order structure analysis of the Roche et al. (1999) data, showing the adjustment loadings.*

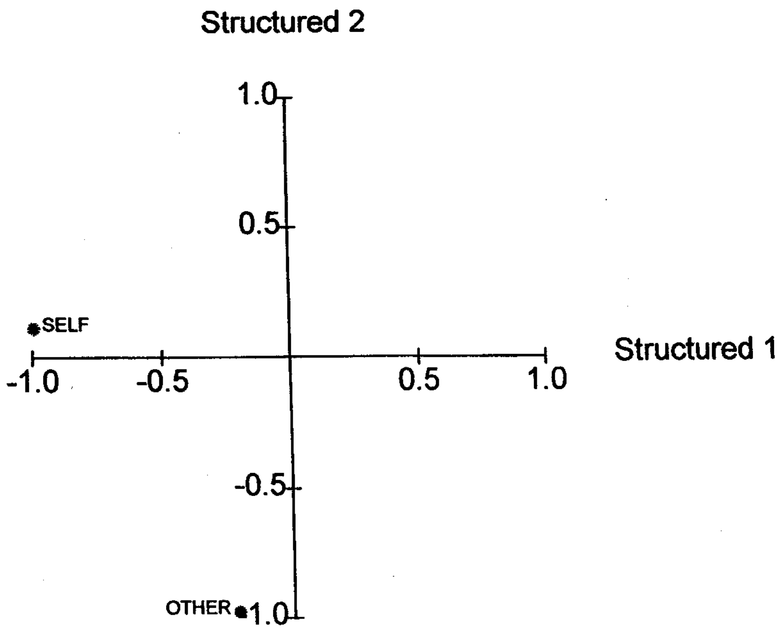

FIGURE 6b. *Component loadings from the higher order structure analysis of the Roche et al. (1999) data, showing the attachment loadings associated with incorporating Bartholomew's model of attachment into CPCA.*

Example 4: Analysis of Contingency Tables

Another important extension of CPCA is the inclusion of metric matrices into the procedure. In general, metric matrices are used to weight the columns and rows of a matrix to make them comparable. For example, standardizing a multivariate data set is equivalent to first subtracting from each score its column mean, then dividing the result by its standard deviations, thus transforming incomparable scales into comparable scales. In this case, the reciprocals (inverse) of the standard deviations form an example of a metric matrix. Although this is a simple example, the ability to incorporate metric matrices greatly enhances the scope of CPCA. One important class of analyses that emerges is reduced-rank contingency table analysis. This class of techniques seeks to represent the information in a contingency table in fewer dimensions, much as PCA seeks to represent the information in a set of variables in fewer dimensions, and includes such methods as Correspondence Analysis (CA) (Greenacre, 1984; Nishisato, 1980) and Latent Class Analysis (LCA) (Hagenaars, 1990; de Leeuw & Van der Heijden, 1991). These methods have become increasingly popular in sociology and anthropology and are starting to gain acceptance in psychology.

From the perspective of CPCA, an R-by-C frequency table (**F**) can be considered as an instance of **Z** once it is premultiplied by the square root of the inverse of the diagonal matrix of row totals ($\mathbf{D_R}^{1/2}$ = the row metric matrix), and postmultiplied by the square root of the inverse of the diagonal matrix of the column totals ($\mathbf{D_C}^{1/2}$ = the column metric matrix). This transformation divides the frequencies in each cell of the table by the product of the square root of their corresponding row and column totals, thus making comparable cell frequencies that are associated with row and column totals that differ in magnitude.

Given $\mathbf{Z} = \mathbf{D_R^{-1/2} F D_C^{-1/2}}$ and either row constraints or column constraints or both, one can then proceed with CPCA essentially as usual by regressing **Z** onto the row (**G**) and column (**H**) information as in Equation 3, and component analyzing the results. The difference between decomposing the parts of Equation 3 with continuous information in **Z** versus with categorical information in **Z** is that, in keeping with the fact that metrics were applied to **Z**, the row metric must also be applied to **G** and the column metric to **H**. Technical details of the decomposition of $\mathbf{Z} = \mathbf{D_R^{-1/2} F D_C^{-1/2}}$ both with and without constraints appear in Takane and Hunter (2001).

Takane and Hunter (2001) and Takane, Yanai, and Mayekawa (1991) have shown that unconstrained PCA of $\mathbf{Z} = \mathbf{D_R^{-1/2} F D_C^{-1/2}}$ is equivalent to simple correspondence analysis (Greenacre, 1984; Nishisato, 1980), and that CPCA of $\mathbf{Z} = \mathbf{D_R^{-1/2} F D_C^{-1/2}}$ with row and column constraints is equivalent to Canonical Correspondence Analysis (CCA) (ter Braak, 1986). Thus, CPCA of contingency tables attempts to uncover in as few dimensions as possible the underlying structure of the association that exists between rows and columns of a contingency table, just as with continuous variables it attempts to uncover in as few dimensions as possible the underlying structure in the associations (e.g., correlations) that exist among those variables. When there are row or column constraints on a contingency table, CPCA evaluates how well these constraints (e.g., predictors) account for the association and underlying structure in a contingency table, as well as whether there is any structure remaining after what is known about the rows and columns has been taken into account.

We illustrate CPCA of contingency tables using data from a large scale survey on Japanese national characters conducted at the Institute of Statistical Mathematics in Tokyo in 1973. The data are a contingency table based on two multiple choice questions supposed to measure traditional versus modern views coexistent in Japanese culture. The questions are:

Q1. If you have no children, do you think it is necessary to adopt a child in order to continue the family line?
1. Would adopt
2. Would not adopt
3. Depends on circumstances

Q2. Some people say that if we get good political leaders, the best way to improve the country is for people to leave everything to them.
1. Agree
2. Disagree

For each question the first option (i.e., would adopt; agree) reflects traditional values, and the second option (would not adopt; disagree) reflects modern values. The third option of Question 1 (depends on circumstances) reflects indecisiveness. Responses to these questions were interactively coded to produce the columns of the contingency table shown in Table 7. That is, the three responses to Question 1 were factorially combined with the two responses to Question 2 such that the columns indicate the frequencies associated with Q11–Q21 (the frequency of responding with Option 1 of Question 1 as well as Option 1 of Question 2), Q12–Q21, Q13–Q21, Q11–Q22, Q12–Q22, and Q13–Q22, respectively.

A sample of over 2,000 Japanese citizens responded to the questions. They were cross-classified on three background variables *L*, *M*, *H*: 1. Level of education (*L*ow: elementary and junior high, *M*iddle: senior high, *H*igh: college); 2. Gender (Male, Female); and 3. Age (A1: 20–29, A2: 30–39, A3: 40–49, A4: 50–59, A5: 60 or older). The rows of Table 7 represent profiles on these three background variables. There are only 28 rows in the table because two combinations were unobserved.

Possible row and column constraints (matrices **G** and **H**) are presented in Table 8. Matrix **G** uses dummy codes to represent subject profiles as additive combinations of the main effects of the three background variables. Similarly, **H** represents item response profiles as additive combinations of the main effects of the two items. What is left out are the interactions among the subject profile variables, and the interaction between the two items.

The following CPCA models were evaluated:

(1.) Both rows and columns are unconstrained: This is equivalent to simple correspondence analysis in which PCA is used to find the most important dimensions of association between the rows and columns of the contingency table.

TABLE 7
The ISM Frequency Data on Traditional vs. Modern Views

No.	Education	Sex	Age Group	1-1	2-1	3-1	1-2	2-2	3-2	Row Total
1	L	M	20–30	6	7	1	15	13	4	46
2	L	M	30–40	15	13	2	22	22	9	83
3	L	M	40–50	16	10	0	31	23	9	89
4	L	M	50–60	17	11	1	34	17	11	91
5	L	M	60–	29	9	2	44	15	7	106
6	L	F	20–30	8	15	3	14	17	5	62
7	L	F	30–40	18	22	8	19	39	13	119
8	L	F	40–50	27	26	7	18	29	8	115
9	L	F	50–60	30	15	12	21	17	9	104
10	L	F	60–	63	16	9	19	9	4	120
11	M	M	20–30	7	7	2	37	58	32	143
12	M	M	30–40	8	13	4	21	41	16	103
13	M	M	40–50	9	6	4	26	25	13	83
14	M	M	50–60	4	3	1	14	11	5	38
15	M	M	60–	5	3	1	8	3	3	23
16	M	F	20–30	12	20	5	42	84	31	194
17	M	F	30–40	11	21	2	20	60	17	131
18	M	F	40–50	8	16	3	23	52	12	114
19	M	F	50–60	4	6	0	9	19	4	42
20	M	F	60–	3	3	0	5	5	2	18
21	H	M	20–30	1	5	1	15	40	15	77
22	H	M	30–40	4	6	0	11	20	3	44
23	H	M	40–50	1	4	0	5	10	4	24
24	H	M	50–60	0	1	0	3	4	4	12
25	H	M	60–	2	1	1	5	4	3	16
26	H	F	20–30	1	3	1	1	23	5	34
27	H	F	30–40	0	3	1	4	15	1	24
28	H	F	50–60	1	1	1	3	2	0	8
				310	266	72	489	677	249	2063

(2.) Column additivity: This is equivalent to CCA with column constraints, which in this case is equivalent to finding the structure of association in the table under the hypothesis that column profiles reflect the additive main effects of item types. That is, any differences among options for one item are constrained to be equal (parallel) across the response options of the other.

(3.) Row additivity: This is equivalent to CCA with row constraints. That is, the structure of association is sought under the hypothesis that it arises out of education, gender, and age effects, and that these effects do not interact.

(4.) Row and column additivity: This is equivalent to CCA with both row and column constraints. In the current example this model evaluates whether the

TABLE 8
Row and Column Design Matrices Under Additivity Hypotheses

	Education			Gender		Age Group (A)				
No.	L	M	H	m	f	1	2	3	4	5
1	1			1		1				
2	1			1			1			
3	1			1				1		
4	1			1					1	
5	1			1						1
6	1				1	1				
7	1				1		1			
8	1				1			1		
9	1				1				1	
10	1				1					1
11		1		1		1				
12		1		1			1			
13		1		1				1		
G = 14		1		1					1	
15		1		1						1
16		1			1	1				
17		1			1		1			
18		1			1			1		
19		1			1				1	
20		1			1					1
21			1	1		1				
22			1	1			1			
23			1	1				1		
24			1	1					1	
25			1	1						1
26			1		1	1				
27			1		1		1			
28			1		1			1		

	Q1			Q2	
	1	2	3	1	2
1	1			1	
2		1		1	
H = 3		1	1		
4	1				1
5		1			1
6			1		1

Note. Blank cells indicate zeros.

association between rows and columns and the structure that arises out of it can be accounted for by differential item effects for subjects who vary in education, gender, and age.

The sums of squares due to the terms analyzed in the previous four cases are given in Table 9. Whether the columns are constrained by additivity or not, over 70% of the sums of squares in the original data set can be explained under the row additivity hypothesis. Figure 7 displays two configurations of column points superimposed onto each other, one derived under the additivity, and the other without the additivity hypothesis on the columns (these points are analogous to loadings obtained when CPCA is applied to continuous data). In both

TABLE 9
Results (SS and Percent SS) for the ISM Data

Source		Internal Analysis			
			Dimension		
		External	1	2	1 + 2
Row	Column				
(1.) Unconstrained	Unconstrained	.244	.163	.045	.208
		(100.0%)	(66.9%)	(18.3%)	(85.3%)
(2.) Unconstrained	Additivity	.214	.160	.041	.200
	Percent out of (a)	(87.7%)	(65.6%)	(16.6%)	(82.2%)
	Percent out of (b)	(100.0%)	(74.6%)	(18.9%)	(93.5%)
(3.) Additivity	Unconstrained	.190	.145	.038	.183
	Percent out of (a)	(77.9%)	(59.5%)	(15.4%)	(74.9%)
	Percent out of (c)	(100.0%)	(76.5%)	(19.8%)	(96.3%)
(4.) Additivity	Additivity	.180	.144	.035	.178
	Percent out of (a)	(73.1%)	(58.9%)	(14.1%)	(73.1%)
	Percent out of (d)	(100.0%)	(80.0%)	(19.2%)	(99.2%)

Note. Percent SS are shown in parentheses.

cases, however, the row additivity constraint was imposed.[5] The two configurations are nearly equal in their accountability, and they account for most (96.3% and 99.2%, respectively) of the predictable variation associated with models 3 and 4.

The points representing item categories obtained without the column additivity constraint are indicated by stars in Figure 7. Clearly, if the stars representing categories 1-1, 2-1, and 3-1 (corresponding to the three response options in Q1 associated with Option 1 in Q2), and the vectors 1-2, 2-2, and 3-2 (corresponding to the three response options in Q1 associated with Option 2 in Q2) were connected, they would not exhibit the parallelism characteristic of strict additivity. Similarly, connecting 1-1 and 1-2 (corresponding to the two response options in Q2 associated with Option 1 in Q1), 2-1 and 2-2, and 3-1 and 3-2 would not produce parallel lines. In contrast, the points obtained under the additivity constraint (indicated by connected hollow circular dots) are perfectly parallel.

Weights attached to the subjects' background variables (analogous to predictor loadings as in Figure 1d) and those attached to response constraints (analogous to stimulus constraint loadings as in Table 4b) are displayed in Figure 8. In this Figure, 1-1, 1-2, and 1-3 denote the loadings attached to response options 1, 2, and 3, of Item 1, and 2-1 and 2-2 denote loadings attached to response options 1 and 2 of Item 2. For the response categories, it appears that the right side of the configuration represents more traditional views (response options 1-1 and 2-1, which, respectively, indicate favoring adoption and agreeing with leaving the

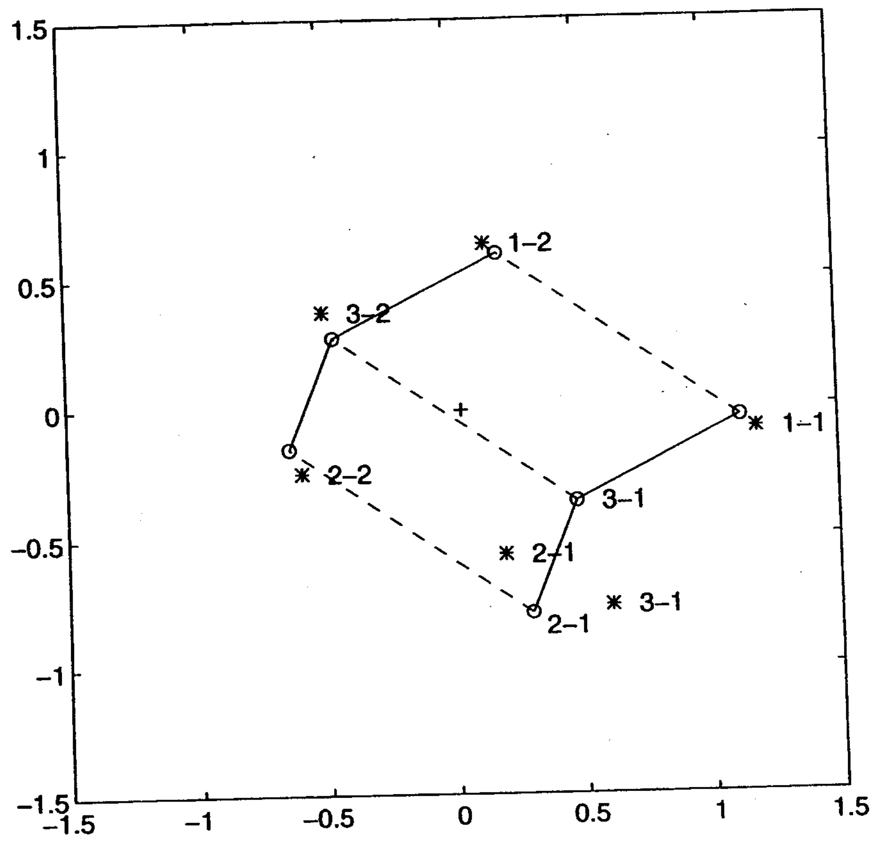

FIGURE 7. *Configuration of column points both under the additivity constraint and without the additivity constraint on the columns of the ISM Data. These points are analogous to loadings obtained when CPCA is applied to continuous data. 1-1, 2-1, and 3-1 denote the three response options in Q1 associated with option 1 in Q2, and the vectors 1-2, 2-2, and 3-2 denote the three response options in Q1 associated with option 2 in Q2.*

country in the hands of politicians), the left region represents more modern views (response options 1-2 and 2-2, which, respectively, indicate not favoring adoption and not agreeing with leaving the country in the hands of politicians), and the middle region represents "indecisiveness" (response option 1-3, "depends on circumstances"). The weights for background variables show that level of education follows the pattern of traditional, indecisive and modern. That is, people with Low education tend to be more traditional, those with Middle education more indecisive, and those with High education level more modern. There appears to be a sex difference, suggesting that Females have more modern views on family but are politically traditional, whereas the reverse is true of Males. The weights for different age groups are labeled A1 through A5. Younger generations are more modern, whereas older people tend to be more traditional. A point of caution is necessary in interpreting a joint representation of rows and columns such as Figure 8 (Carroll, Green, & Schaffer, 1989; Nishisato, 1988). Only directions from the origin can be meaningfully interpreted, not inter set distances (i.e., those between rows and columns).

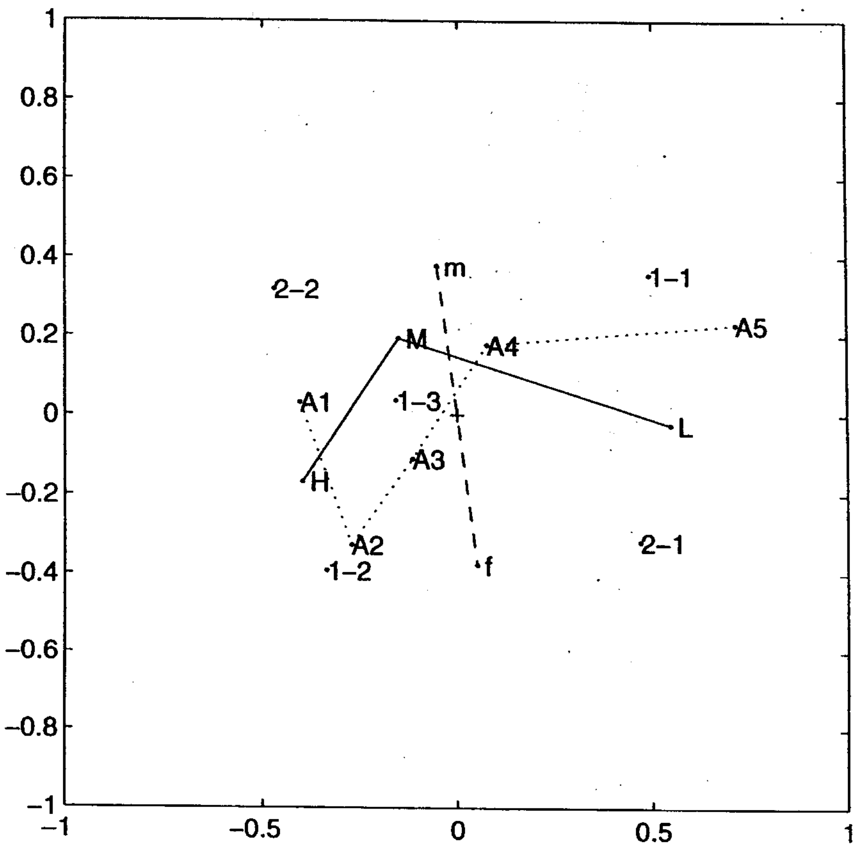

FIGURE 8. *Row and column weights obtained under the additivity constraints on both rows and columns of the ISM Data. 1-1, 1-2, and 1-3 denote the weights attached to response options 1, 2, and 3 of item 1, and 2-1 and 2-2 denote weights attached to response options 1 and 2 of item 2.* L = *low,* M = *middle,* H = *high,* M = *male,* F = *female.*

Example 5: Residual Analysis and Assessment of Reliability

To this point we have focused on analyzing the predictable part of a data matrix while simply mentioning the analysis of the unpredictable part, that is, the residuals, or that portion of the variability in a data set that is not accounted for by subject or stimulus information. In this section, we consider residuals analysis in more detail and give an example. We also illustrate the use of confidence regions to assess reliability in CPCA using the bootstrap method.

A CPCA of residuals can be viewed as a way to investigate the structure of a set of criterion variables after relations with predictor variables have been partialled, or statistically controlled. If the predictors are continuous variables such as age, socioeconomic status (SES), and IQ, CPCA of residuals finds structure in the criterion variables after (statistically) eliminating the effects of age, SES, and IQ. If the predictors represent groups indicated by a design matrix of dummy variables, differences in group means are statistically controlled and CPCA of residuals finds the average of within-group structure.

When the focus of CPCA is on residuals, the typical regression analysis view of residuals as a combination of measurement error, prediction error, and, specification error is relaxed. Instead, the residual variance becomes the focus of attention and the predictor variables are considered as confounding variables that when omitted or left uncontrolled produce spurious relationships (and structure) among the criterion variables.

Residuals do not always represent totally unsystematic variation, but can in some instances represent systematic variation and thus be directly interpretable. In the analysis of Mezzich's (1978) data discussed earlier, the total variation (T) was decomposed into between-patient (B) and within-patient (W) effects. The data were, however, collected from eleven psychiatrists, each rating four archetypal psychiatric patients. It would have been possible to incorporate the psychiatrist effect into the model, and further decompose W into two parts; the psychiatrist effect (P) and a residual (R). In this case, however, because there are no replicated observations within patients or psychiatrists, the residual effect (R) does not represent unexplained variation but instead the interaction between the patient effect (B) and the psychiatrist effect (P). In general, non replicated designs will have interpretable residuals.

Contingency tables can be viewed as an important class of non-replicated designs. In two-way tables the residuals from the row and column effects represent the row-by-column interaction effect, after which no variation remains (combined, the main effects and their interaction form a saturated model). Similarly, in three-way tables residuals from main effects and two-way interactions are three-way interactions, after which no further decomposition is possible. The idea that residuals from contingency tables can have specific meaning is reinforced by Takane, Yanai, and Mayekawa (1991) who showed that Canonical Correspondence Analysis (CCA) (ter Braak, 1986) and Canonical Analysis with Linear Constraints (CALC) (Böckenholt & Böckenholt, 1990), which are both constrained versions of correspondence analysis, analyze complementary parts of a contingency table, each essentially analyzing the residuals from the other. Takane and Hunter (2001) subsequently showed that both CCA and CALC are special cases of CPCA.

We demonstrate a CPCA of residuals using the data given in Table 10. These data were originally reported in Israels (1987, p. 76), and were reanalyzed by Van

TABLE 10
People Suspected of Shop-Lifting by Sex, Age, and Kind of Stolen Goods

Sex	Age (Year)	1 Clothing	2 Clothing Accessories	3 Provisions Tobacco	4 Writing Materials	5 Books	6 Records	7 Household Goods	8 Sweets	9 Toys	10 Jewelry	11 Perfume	12 Hobbies Tools	13 Other	Total
Male															
M1	<12	81	66	150	667	67	24	47	430	743	132	32	197	209	2845
M2	12–14	138	204	340	1409	259	272	117	637	684	408	57	547	550	5622
M3	15–17	304	193	229	527	258	368	98	246	116	298	61	402	454	3554
M4	18–20	384	149	151	84	146	141	61	40	13	71	52	138	252	1682
M5	21–29	942	297	313	92	251	167	193	30	16	130	111	280	624	3446
M6	30–39	359	109	136	36	96	67	75	11	16	31	54	200	195	1385
M7	40–49	178	53	121	36	48	29	50	5	6	14	41	152	88	821
M8	50–64	137	68	171	37	56	27	55	17	3	11	50	211	90	933
M9	65+	45	28	145	17	41	7	29	28	8	10	28	111	34	531
Female															
F1	<12	71	19	59	224	19	7	22	137	113	162	70	15	24	942
F2	12–14	241	98	111	346	60	32	29	240	98	548	178	29	58	2068
F3	15–17	477	114	58	91	50	27	41	80	14	303	141	9	72	1477
F4	18–20	436	108	76	18	32	12	32	12	10	74	70	14	67	961
F5	21–29	1180	207	132	30	61	21	65	16	12	100	104	30	157	2115
F6	30–39	1009	165	121	27	43	9	74	14	31	48	81	36	107	1765
F7	40–49	517	102	93	23	31	7	51	10	8	22	46	24	66	1000
F8	50–64	488	127	214	27	57	13	79	23	17	26	69	35	64	1239
F9	65+	173	64	215	13	44	0	39	42	6	12	41	11	55	715
Total		7160	2171	2835	3704	1619	1230	1157	2018	1914	2400	1286	2441	3166	33101

der Heijden, de Falguerolles, and de Leeuw (1989). The data are a contingency table consisting of frequencies of people of specific sex (*S*) and age (*A*) caught for shoplifting specific goods. Variables *S* and *A* are interactively coded to form rows of the table. Rows 1–9 represent males of different ages (1. 12 and younger, 2. 12–14, 3. 15–17, 4. 18–20, 5. 21–29, 6. 30–39, 7. 40–49, 8. 50–64, 9. 65 and older), and rows 10–18 represent the same age groups for females. The goods are: 1. clothing; 2. clothing accessories; 3. provisions, tobacco; 4. writing materials; 5. books; 6. records; 7. household goods; 8. sweets; 9. toys; 10. jewelry; 11. perfume; 12. hobbies, tools; and 13. other.

CPCA was applied to the data with various model specifications. Sums of squares accounted for each analysis are reported in Table 11. In Table 11:

(1.) "Unconstrained" means $G = I$, which is equivalent to a simple unconstrained CA.
(2.) *Age* and *Sex* Additivity. The rows are represented as additive functions of variables *A* and *S*, which means that appropriately coded sex and age variables are joined to form the row constraint matrix. This analysis is equivalent to CCA with the main effects of *A* and *S* incorporated as row constraints.
(3.) Residuals from analyses. These residuals are the *A*-by-*S* interaction. This analysis is equivalent to CCA with $G = AS$ where *AS* is the age-by-sex interaction effect, or to CALC with G equal to *A* and *S* main effects.

Figures 9, 10, and 11 display two-dimensional configurations of row and column points resulting from analyses 1–3. Again, these points are analogous to component loadings; and also, because these configurations are joint plots of row and column points, the same cautionary remark given previously applies here as well. To recapitulate, only the relative directions of row and column points can be interpreted, not the distances between them. In each figure, item points are indicated by boldfaced numbers from 1 to 13, while subject profiles are connected by age-ordered line segments, separately for males (solid line) and females (dotted line).

TABLE 11
Results for the Shoplifting Data (Israels, 1987)

Constraint on Row	External Analysis	Internal Analysis: Dimension 1	Dimension 2	Dimension 1 + 2
(1.) Unconstrained (Simple CA)	.603	.350	.119	.470
	(100.0%)	(58.1%)	(19.8%)	(77.9%)
(2.) Age & Gender Additivity (CCA)	.546	.349	.115	.465
Percent out of 1	(90.6%)	(57.9%)	(19.1%)	(77.1%)
Percent out of 2	(100.0%)	(63.9%)	(21.1%)	(85.0%)
(3.) Residuals from Age & Gender (CALC)	.057	.038	.010	.048
Percent out of 1	(9.5%)	(6.3%)	(1.7%)	(7.0%)
Percent out of 3	(100.0%)	(66.7%)	(17.5%)	(84.6%)

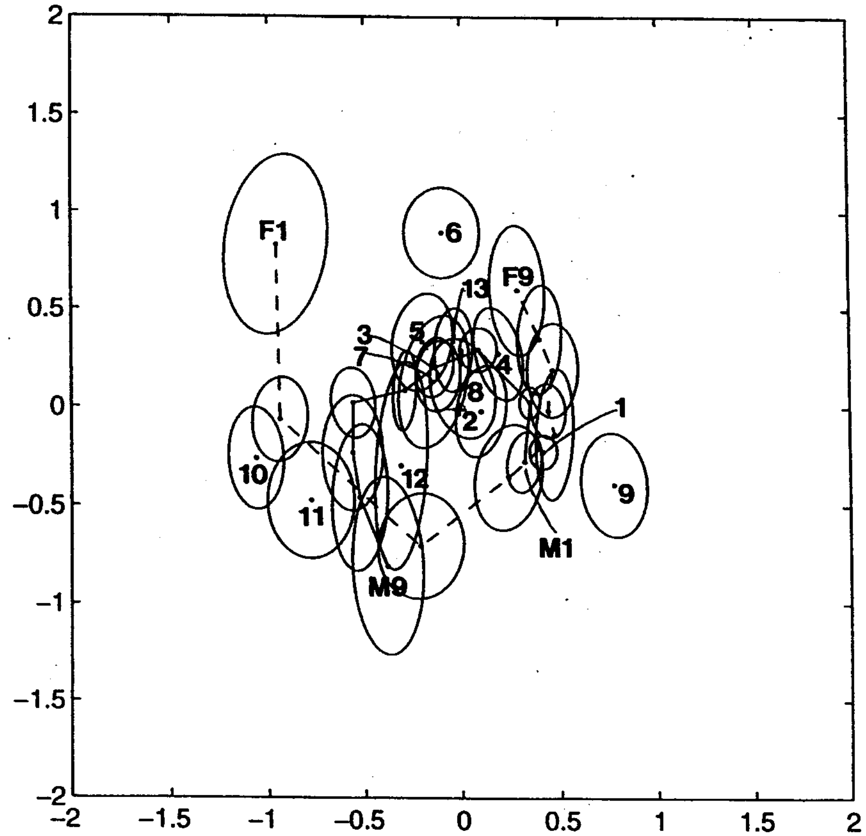

FIGURE 9. *Two-dimensional configuration of the row and column weights obtained from the unconstrained analysis of Israels' shoplifting data. Ellipses surrounding the points represent 99% bootstrap confidence regions.*

Figures 9 and 10 can be interpreted in a rather straightforward manner. In both, the dimensions represented are primarily the main effects of *A* (the horizontal direction) and *S* (the vertical direction). In Figure 11, which is the analysis of residuals, two dimensions of the *AS* interaction are represented. Dimension 1 (the horizontal direction) represents the interaction between *S* and the linear trend over *A*. Young females and adult males are at the left of the configuration, and young males and adult females are at the right of the configuration. Clothing (1) and toys (9) are stolen more often by the group located at the right, whereas jewelry (10), perfume (11), and hobbies and tools are stolen more by the group at the left. Dimension 2 (the vertical direction), on the other hand, represents the interaction between *S* and the quadratic trend over *A*. Combining the two dimensions, clothing (1) is stolen

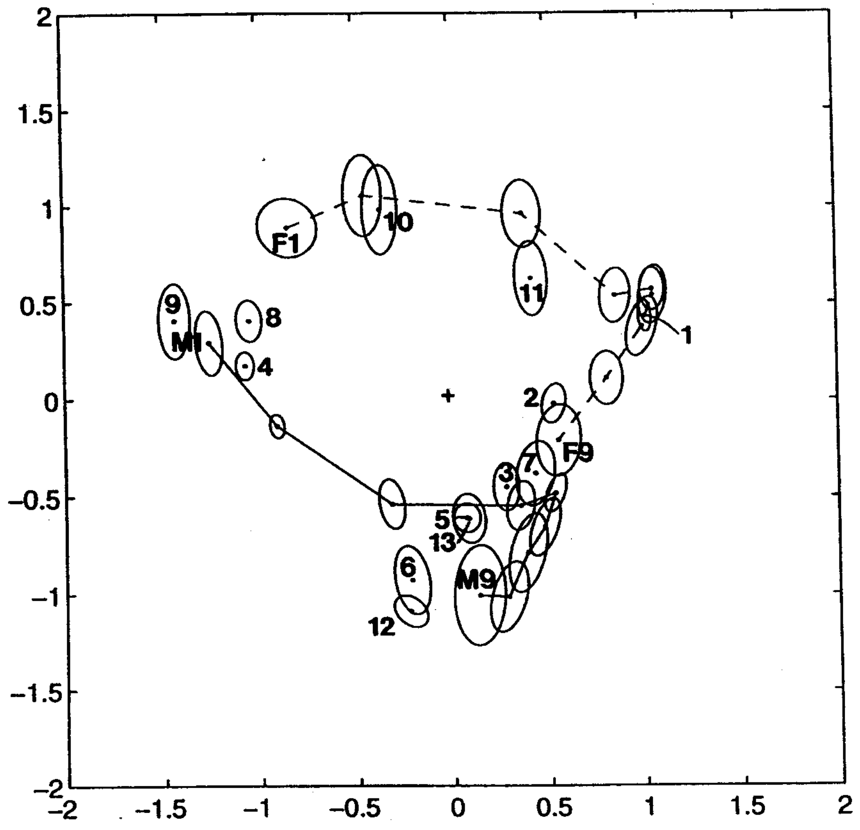

FIGURE 10. *Two-dimensional configuration of the row and column weights obtained under* age *and* sex *additivity constraints. Ellipses surrounding the points represent 99% bootstrap confidence regions.*

more often by adult females than by adult males, toys (9) by young males than by young females, jewelry (10) and perfume (11) by young females than by young males, and hobbies and tools (12) by adult males than by adult females. Note, however, that although the configuration in Figure 11 clearly reflects interaction effects, the two dimensions together account for only a relatively minor proportion (8.0%) of the total sum of squares.

Ellipses surrounding the points in Figures 9 through 11 represent 99% bootstrap confidence regions. These confidence regions were obtained by first generating 100 bootstrap samples, analyzing each of them by CPCA, and estimating variance-covariance matrices among the estimated coordinates (see Weinberg, Carroll, & Cohen, 1984). The confidence regions are tightest under the additivity constraint, implying that the additivity constraint (the age and sex main effects) is empirically

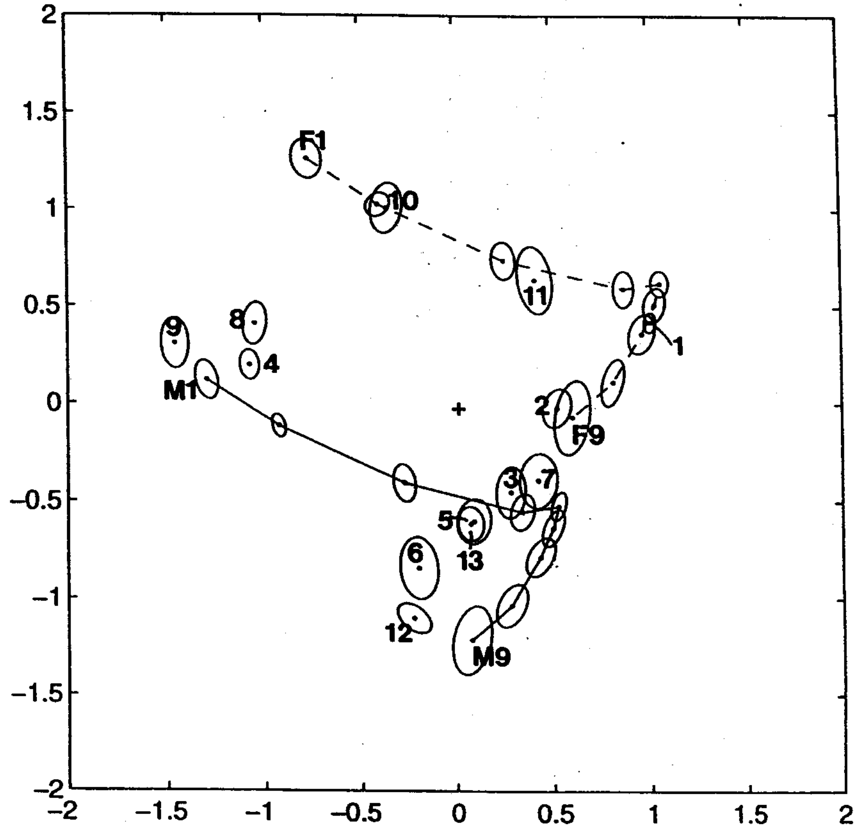

FIGURE 11. *Two-dimensional configuration of the row and column weights obtained with the* age *and* sex *additivity constraints eliminated. Ellipses surrounding the points represent 99% bootstrap confidence regions along with 99% bootstrap confidence regions.*

validated. The points are close together with overlapping confidence regions in the residual analysis, suggesting an unreliable result. This directly reflects the small proportion of the sum of squares in the residual.

Conclusion

This article illustrates various uses and extensions of CPCA through examples. The method is very general, encompassing a wide range of existing techniques, and following the extensions discussed in this article and in Takane and Hunter (2001) promises to expand into a variety of new principal component based applications. For example, beyond the analysis of contingency tables, the use of metric matrices in CPCA allows for the possibility of scale invariant component analysis, of incorporating measurement error into component analysis, and of improving P-technique factor analysis by using serial correlations from single-subject multivariate time series data as a metric. Similarly, the ability to incorporate higher order structures makes it possible to use CPCA for multilevel analysis (Takane & Hunter, 2002), for evaluating factor structure invariance in between-group designs, and for path analysis with formative variables (i.e., components rather than common factors).

Indeed, as shown in Takane and Hunter (2001), CPCA is comparable in its generality to structural equation modeling with structured means. The major difference is that CPCA models data directly, whereas structural equation modeling models means and covariances derived from the data. One consequence is that CPCA and structural equation modeling handle measurement error differently (e.g., CPCA would incorporate reliability of measurement via metric matrices). Another consequence is that computationally, common factor-based structural equation modeling is iterative, whereas CPCA involves only projection and singular value decomposition (or generalized singular value decomposition; see Takane & Hunter, 2001), both of which can be carried out noniteratively. Moreover, unlike in latent variable structural equation modeling, there is no factorial indeterminancy problem in CPCA. Component scores arising out of CPCA are uniquely determined. Structural equation modeling is, however, more flexible in that it can accommodate common (reflective) factors as well as principal components (formative factors).

CPCA is primarily descriptive. However, statistical inference (assessment of reliability, hypothesis testing, etc.) is possible, as was demonstrated in the analysis of Israels' (1987) data. Variance-covariance estimates among estimated parameters may be obtained by the bootstrap method or by analytic means (e.g., the delta method; Rao, 1973). This information can then be used to draw confidence regions that indicate the stability of estimated parameters.

CPCA is still under active development. Although the theoretical basis for many types of multivariate analysis appears in a companion paper (Takane & Hunter, 2001), putting this theory into practice and comparing the results with alternative methods awaits future work. For applications such as those presented here in which metric matrices are not required, a computer program (Hunter & Takane, 1998) is available from the first author at http://web.uvic.ca/psyc/software. Those examples requiring metric matrices were run using a newer program available by emailing the second author at: takane@takane2.psych.mcgill.ca.

Notes

[1]Figures 1a and 1d are complementary and should be interpreted together. Axes (bases) labels refer to the source of variation from which the components were obtained. Total 1 and Total 2 denote axes for the first two principal components obtained from variation in the total data. Between 1 and Between 2 denote axes for the first two principal components obtained from variation in the criterion data that is predictable from the between-groups constraints. Within 1 and Within 2 denote axes for the first two principal components of variation in the criterion data that is unrelated to group differences. Predictor 1 and Predictor 2 denote axes for the first two principal components of the predictor data that are paired with Between 1 and Between 2 (just as one obtains pairs of variates in canonical correlation analysis).

[2]These coefficients were derived using weights that would average and subtract contrasting items, then multiplied by a constant to eliminate fractions. The system used is exactly like the one used to compare groups in ANOVA designs. For example, given four groups of subjects, G1, G2, G3, and G4, one could compare the first group with the remaining groups by contrasting the first with the average of the remaining three. That is, one tests the hypothesis

$G1 = (G2 + G3 + G4)/3$, which can be rewritten as,
$G1 - (G2 + G3 + G4)/3 = 0$, and finally, eliminating fractions,
$3G1 - (G2 + G3 + G4) = 0$
$3G1 - G2 - G3 - G4 = 0$
$3G1 + (-1)G2 + (-1)G3 + (-1)G4 = 0$

Thus, the contrast coefficients would be 3, −1, −1, −1. Contrasts C1 through C5 were derived in precisely the same manner.

[3]In contrast, the canonical redundancies for the first two canonical variates were 18.5 and 21.2, respectively, and the canonical correlations were .92 and .68. Combined, these results indicated that higher canonical correlations do not necessarily correspond to higher redundancy (predictability).

[4]The canonical loadings for the criterion variables have a pattern more similar to the unconstrained loadings than do the CPCA criterion loadings (in reverse order). The canonical predictor loadings are similar for the abuse grouping factor, but indicate a contrast between dismissing and preoccupied attachment styles corresponding to the first canonical criterion variate, and between secure and fearful styles for the second canonical criterion variate.

[5]Interactions represented in Figure 7 are defined in exactly the same way as in loglinear models, although due to differences in estimation procedure (maximum likelihood versus least squares), their estimates are different. Let A, E, G, Q_1 and Q_2 represent the effects of Age, Education, Gender, Question 1 and Question 2, respectively. Then using standard notation from loglinear modeling, the solution obtained with the row additivity constraints but without the column additivity constraints represents $[AQ_1Q_2][EQ_1Q_2][GQ_1Q_2]$, whereas the solution obtained with the column additivity constraints represents $[AQ_1][AQ_2][EQ_1][EQ_2][GQ_1][GQ_2]$. In this notation expressions such as $[AQ_1Q_2]$ denote the three-way interaction among A, Q_1, and Q_2, and all lower order interactions involving Age, Question 1, and Question 2.

References

Bartholomew, K. (1990). Avoidance of intimacy: An attachment perspective. *Journal of Social and Personal Relationships, 7,* 147–178.

Bartholomew, K., & Horowitz, L. M. (1991). Attachment styles among young adults: A test of a four-category model. *Journal of Personality and Social Psychology, 61*, 226–244.

Bowlby, J. (1982). *Attachment and loss, vol. 1: Attachment (2nd Ed.)*. New York: Basic Books.

Böckenholt, U., & Böckenholt, I. (1990). Canonical analysis of contingency tables with linear constraints. *Psychometrika, 36*, 633–639.

Briere, J. (1996). Psychometric review of the Trauma Symptom Checklist-40. in B. H. Stamm (Ed.). *Measurement of stress, trauma, and adaptation.* Lutherville, MD: Sidran Press.

Carroll, J. D., Green, P. E., & Shaffer, C. M. (1989). Reply to Greenacre's commentary on the Carroll-Green-Shaffer scaling of two-way correspondence analysis solutions. *Journal of Marketing Research, 26*, 366–368.

Cohen, J. (1982). Set Correlation as a general multivariate data-analytic method. *Multivariate Behavioral Research, 17*, 301–341.

de Leeuw, J., & Van der Heijden, P. (1991). Reduced rank models for contingency tables. *Biometrika, 78*, 229–232.

Greenacre, M. J. (1984). *Theory and applications of correspondence analysis.* London: Academic Press.

Hagenaars, J. A. (1990). *Categorical longitudinal data.* Newbury Park, CA: Sage Publications.

Hunter, M. & Takane, Y. (1998). CPCA: Principal component analysis with external information on subjects and variables. *Behavior Research Methods Instruments and Computers, 30*, 506–516.

Israels, A. (1987). *Eigenvalue techniques for qualitative data.* Leiden, The Netherlands: DSWO Press.

Joreskog, K. G. (1970). A general method for analysis of covariance structures. *Biometrica, 57*, 239–251.

McDonald, R. P. (1978). A simple comprehensive model for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology, 31*, 59–72.

Mezzich, J. E. (1978). Evaluating clustering methods for psychiatric diagnosis. *Biological Psychiatry, 13*, 265–281.

Nishisato, S. (1980). *Analysis of categorical data: Dual scaling and its applications.* Toronto: University of Toronto Press.

Nishisato, S. (1988). Assessing quality of joint graphical display in correspondence analysis and dual scaling. In E. Diday, et al. (Eds.), *Data analysis and informatics V* (pp. 409–416). Amsterdam: North-Holland.

Rao, C. R. (1964). The use and interpretation of principal component analysis in applied research. *Sankhya, 26*, 329–358.

Rao, C. R. (1973). *Linear statistical inference and its application.* New York: Wiley.

Reinsel, G. C., & Velu, R. P. (1998). *Multivariate reduced regression models.* New York: Springer.

Roche, D. N., Runtz, M. G., & Hunter, M. A. (1999). Adult attachment and child sexual abuse. *Journal of Interpersonal Violence, 14*, 184–207.

Rodgers, J. L. & Young, F. W. (1981). Successive unfolding of family preferences. *Applied Psychological Measurement, 5*, 51–62.

Takane, Y., & Hunter, M. A. (2002). Dimension reduction in hierarchical linear models. In S. Nishisato, Y. Baba, H. Bozdogan, & K. Kanefuji (Eds.), *Proceedings of ICMMA in Banff, Alberta, Canada.* Tokyo: Kluwer Publishing Co.

Takane, Y., & Hunter, M. A. (2001). Constrained principal component analysis: A comprehensive theory. *Applicable Algebra in Engineering, Communication and Computing, 12*, 391–419.

Takane, Y., & Shibayama, T. (1991). Principal component analysis with external information on both subjects and variables. *Psychometrika, 56*, 97–120.

Takane, Y., Yanai, H., & Mayekawa, S. (1991). Relationships among several methods of linearly constrained correspondence analysis. *Psychometrika, 56*, 667–684.

ter Braak, C. J. F. (1986). Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis. *Ecology, 67*, 1167–1179.

Van den Wollenberg, A. L. (1977). Redundancy analysis: An alternative for canonical correlation analysis. *Psychometrika, 42*, 207–219.

Van der Heijden, P. G. M., de Falguerolles, A., & de Leeuw, J. (1989). A combined approach to contingency table analysis using correspondence analysis and log linear analysis. *Applied Statistics, 38*, 249–292.

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

Authors

MICHAEL HUNTER is Associate Professor, Psychology Department, University of Victoria. Victoria, B.C., V8W 2Y2, Canada; mhunter@uvic.ca. His research specialties include: cognitive aging, statistics and methodology.

YOSHIO TAKANE is Adjunct Professor, Psychology Department, University of Victoria. Victoria, B.C., V8W 2Y2, Canada; takane@takane2.psych.mcgill.ca. His research specialties include psychometrics, statistics, and methodology.

Manuscript Received April 2000
Accepted December 2001