PSYCHOMETRIKA—VOL. 58, NO. 1, 7–35
MARCH 1993

POINTS OF VIEW ANALYSIS REVISITED: FITTING MULTIDIMENSIONAL STRUCTURES TO OPTIMAL DISTANCE COMPONENTS WITH CLUSTER RESTRICTIONS ON THE VARIABLES

JACQUELINE J. MEULMAN AND PETER VERBOON

DEPARTMENT OF DATA THEORY
UNIVERSITY OF LEIDEN

Points of view analysis (PVA), proposed by Tucker and Messick in 1963, was one of the first methods to deal explicitly with individual differences in multidimensional scaling, but at some point was apparently superceded by the weighted Euclidean model, well-known as the Carroll and Chang INDSCAL model. This paper argues that the idea behind points of view analysis deserves new attention, especially as a technique to analyze group differences. A procedure is proposed that can be viewed as a streamlined, integrated version of the Tucker and Messick Process, which consisted of a number of separate steps. At the same time, our procedure can be regarded as a particularly constrained weighted Euclidean model. While fitting the model, two types of nonlinear data transformations are feasible, either for given dissimilarities, or for variables from which the dissimilarities are derived. Various applications are discussed, where the two types of transformation can be mixed in the same analysis; a quadratic assignment framework is used to evaluate the results.

Key words: points of view, individual differences, group differences, nonlinear multivariate analysis, nonmetric multidimensional scaling, distance components, composite dissimilarities, variable clustering.

1. Introduction

Points of view analysis (PVA) as proposed by Tucker and Messick (1963) was one of the first methods that sought a compromise between two approaches to the multidimensional scaling (MDS) analysis of dissimilarity data for objects or stimuli obtained from different individuals (sources). One approach was based on group averages, where groups were chosen a priori; the other was an analysis on an individual level. The objective of a points of view analysis is to obtain different multidimensional spaces for different groups of individuals, each having a particular viewpoint about the object interrelationships; the groups have to be empirically derived. Before summarizing the Tucker and Messick approach, and the issues that were raised with respect to the procedure, some preliminary notation must be given.

Dissimilarity data are available for N objects, obtained from $M > 1$ different sources or individuals, $m = 1, \ldots, M$. The dissimilarity data may be represented in two different forms; the first is a symmetric matrix $\mathbf{\Delta}_m = \{\delta_{ijm}\}$, of order $N \times N$, containing the dissimilarities δ_{ijm} between pairs of objects $\{i, j\}$ according to source m. We assume $\mathbf{\Delta}_m$ is normalized so that the sum of squares of the off-diagonal elements is $2N$. The alternative representation is a vector $\boldsymbol{\delta}_m$, of order l, where $l = N(N - 1)/2$,

The research of the first author was supported by the Royal Netherlands Academy of Arts and Sciences (KNAW); the research of the second author by the Netherlands Organization for Scientific Research (NWO Grant 560-267-029). An earlier version of this paper was presented at the European Meeting of the Psychometric Society, Leuven, 1989. We wish to thank Willem J. Heiser for his stimulating comments to earlier versions of this paper, and we are grateful to the Editor and anonymous referees for their helpful suggestions.

Requests for reprints should be sent to Jacqueline Meulman, Department of Data Theory, Faculty of Social Sciences, University of Leiden, PO Box 9555, 2300 RB Leiden, THE NETHERLANDS.

0033-3123/93/0300-91338$00.75/0

containing the dissimilarities δ_{ijm} for each object pair for which $i < j$; here, the sum of squares of the elements is set equal to N.

The index set $\{1, \ldots, M\}$ is assumed to be partitioned into subsets J_s, $s = 1, \ldots, r$, giving the sources that form the s-th group; M_s indicates the number of indices in J_s. An important feature of the procedure is that the partitioning of the sources into groups is not known a priori. For each of the r groups a representation space $\mathbf{X}_s$ is sought; the dimensionality of $\mathbf{X}_s$ is assumed to be p_s, and the rows of $\mathbf{X}_s$ give the coordinates for the N objects in the representation space for group s.

A squared distance between a pair of objects $\{i, j\}$ in $\mathbf{X}_s$ is defined by

$$\mathrm{d}_{ij}^2(\mathbf{X}_s) = (\mathbf{e}_i - \mathbf{e}_j)'\mathbf{X}_s\mathbf{X}_s'(\mathbf{e}_i - \mathbf{e}_j),$$

where $\mathbf{e}_i$ is the i-th column of the $N \times N$ identity matrix $\mathbf{I}$. Applying the squared distance function, $\mathrm{D}^2(\cdot)$, to map coordinates $\mathbf{X}_s$, into squared distances gives the matrix formulation:

$$\mathrm{D}^2(\mathbf{X}_s) = \boldsymbol{\alpha}_s\mathbf{1}' + \mathbf{1}\boldsymbol{\alpha}_s' - 2\mathbf{X}_s\mathbf{X}_s',$$

with $\mathbf{1}$ an N-vector of all 1's and $\boldsymbol{\alpha}_s$ an N-vector containing the diagonal elements of $\mathbf{X}_s\mathbf{X}_s'$.

Given these preliminaries, the following steps can be distinguished in the original Tucker and Messick (1963) procedure. The dissimilarities are represented in the vectors $\boldsymbol{\delta}_m$, and are regarded as variables that can be subjected to a principal component analysis (PCA) to give principal component scores in r dimensions, $1 \leq r \leq M$, where r denotes the assumed number of different points of view. Each principal components gives a weighted average (linear combination) of the original M dissimilarity variables. Since the principal axes orientation may not be the most appropriate one for displaying different viewpoints, a rotation to simple structure is sought. This implies that only the s-th group of variables obtains high loadings on the s-th component, while all other groups get low loadings. Transforming the weights obtained in the PCA accordingly, gives a rotated component, denoted as

$$\boldsymbol{\theta}_s = M^{-1} \sum_{m=1}^{M} \mathrm{a}_{ms}\boldsymbol{\delta}_m,$$

$s = 1, \ldots, r$, where a_{ms} is an element of the transformed weight matrix $\mathbf{A} = \{\mathrm{a}_{ms}\}$ that represents simple structure. Next, the weights in $\mathbf{A}$ are used to obtain

$$\boldsymbol{\Theta}_s = M^{-1} \sum_{m=1}^{M} \mathrm{a}_{ms}\boldsymbol{\Delta}_m,$$

$s = 1, \ldots, r$, where each $\boldsymbol{\Theta}_s$ is a differently weighted average (linear combination) of the matrices $\boldsymbol{\Delta}_m$. The final step proposed in Tucker and Messick (1963) consisted of r separate multidimensional scaling analyses fitting Euclidean distances $\mathrm{D}(\mathbf{X}_s)$ to each of the $\boldsymbol{\Theta}_s$.

Ross (1966) criticized the method; among other things, he focuses on the possibility that a point of view might not be a linear combination of judgements of subjects. Cliff (1968), who reviews the method favorably, argues that this criticism can be refuted by realizing that Ross misinterprets the intention of Tucker and Messick with respect to what a point of view really is: It is not a way of looking at the objects, but it is a structure for the objects obtained by an MDS. Another objection by Ross concerns the fact that an arbitrary linear combination might give negative weights for some dissimilarity sources, and could result in a dissimilarity matrix for which no Euclidean solution exists. In the procedure that will be described below, it is guaranteed that the weights are always positive; moreover, they turn out to be Tucker's (1951) congruence coefficients between each separate dissimilarity source and the distances fitted in the points of view.

Carroll and Chang (1970) proposed the INDSCAL model as an alternative to points of view analysis; the model is also discussed in Horan (1969) and Bloxom (1978). This so-called weighted Euclidean model (we will consider this term and INDSCAL model as interchangeable) does not fit weights to the dissimilarity sources, but fits weights for the dimensions of an unknown space **X** common to all sources. Carroll and Chang state that PVA is little more powerful than doing separate scalings, and question the fact that no explicit assumptions are made about the possible communality of the multidimensional structures. Since its introduction, the weighted Euclidean model has become the dominant model to analyze individual differences. Yet, it is the purpose of the present paper to show there is still room for the PVA model, explicitly when one is interested in finding subsets of individuals (clusters of sources) that have the same frame of reference. Thus, PVA is truly different from doing separate scalings because it is not known a priori which sources belong to the same points of view; therefore, we will perform a clustering task that assigns the sources to different points of view.

In the weighted Euclidean model, individual differences are defined on dimensions of the common space; points of view are defined on the distances, and the spaces can, but need not, be interpreted in terms of dimensions. It is also possible to look at clusters, structures, or more and other directions than the principal dimensions. In the weighted Euclidean model, the dimensions in **X** are not constrained to be uncorrelated; in a points of view analysis, the distances in the several $\mathbf{X}_s$ are not restricted to be uncorrelated. Whether multiple points of view are really different or actually very similar can be investigated by using quadratic assignment procedures, as discussed, for example, in Hubert (1987).

2. A Comprehensive Objective Function for Points of View Analysis

In the previous section it was shown that points of view analysis deals with three different tasks: the first is to find principal components and weights applying the PCA model to given dissimilarity variables; the second is to find an optimal rotation to simple structure, and the final task is to find optimal representation spaces for the objects on the basis of the rotated components (the composite dissimilarities). In this section a least squares loss function will be introduced that integrates these different optimization tasks. The loss function is defined on the distances in the representation spaces, and thus fits into the STRESS framework, for which Kruskal (1964a, 1964b) and Guttman (1968) laid the foundation.

If $\|\cdot\|^2$ denotes a least squares discrepancy measure such that

$$\|a_{ms}\mathbf{\Delta}_m^* - D(\mathbf{X}_s)\|^2 = \text{tr}\,(a_{ms}\mathbf{\Delta}_m^* - D(\mathbf{X}_s))'(a_{ms}\mathbf{\Delta}_m^* - D(\mathbf{X}_s)),$$

the PVA loss function can be written as

$$\text{STRESS}(\mathbf{A};\mathbf{\Delta}_1^*,\ldots,\mathbf{\Delta}_M^*;\mathbf{X}_1,\ldots,\mathbf{X}_r) = M^{-1}\sum_{s=1}^{r}\sum_{m\in J_s}\|a_{ms}\mathbf{\Delta}_m^* - D(\mathbf{X}_s)\|^2, \quad (1)$$

which is a function of three sets of parameters. For the moment, consider $\Delta_1^*, \ldots, \Delta_M^*$ as given, so without loss of generality $\Delta_m^* = \Delta_m$; then the loss function has to be minimized over the weight matrix $\mathbf{A}$ and the points of view $\mathbf{X}_1, \ldots, \mathbf{X}_r$. Since a perfect, but trivial, solution is easily obtained by setting $\mathbf{A} = \mathbf{0}$ and $\mathbf{X}_s = \mathbf{0}$, it is required, without loss of generality, that $\mathrm{tr}(\mathbf{X}_s'\mathbf{X}_s) = 1$. The loss function must also be minimized over $\mathbf{A} = \{a_{ms}\} \in \Omega$, where Ω is the set of all *restricted* weight matrices that give some form of simple structure; explicitly, it is required that a source Δ_m^* contributes to a single point of view, so only one element of the m-th row of $\mathbf{A}$ is not equal to zero.

The objective function in (1) will be minimized by constructing a convergent algorithm using various components from the majorization approach to multidimensional scaling (de Leeuw, 1988; de Leeuw & Heiser, 1980; Meulman, 1986). In the following, the components of the algorithm will be discussed.

Fitting the Multidimensional Structures

The overall minimization problem in (1) can be partitioned into several parts. First of all, for fixed $\mathbf{A}$ and $\Delta_1^*, \ldots, \Delta_M^*$, the problem of finding $\mathbf{X}_1, \ldots, \mathbf{X}_r$ consists of r separate MDS tasks. For each group of sources one must minimize

$$\mathrm{STRESS}(\mathbf{X}_s) = \underline{M^{-1}} \sum_{m \in \mathrm{J}_s} \|a_{ms}\Delta_m^* - \mathrm{D}(\mathbf{X}_s)\|^2, \tag{2}$$

which is a function of $\mathbf{X}_s$ only. The objective function (2) can be simplified in the following way. Define the composite dissimilarity matrix for the sources that constitute the s-th group as

$$\Theta_s = M_s^{-1} \sum_{m \in \mathrm{J}_s} a_{ms}\Delta_m^*;$$

then (2) can be written as

$$\mathrm{STRESS}(\mathbf{X}_s) = M^{-1}\left[\sum_{m \in \mathrm{J}_s} \|a_{ms}\Delta_m^* - \Theta_s\|^2 + M_s\|\Theta_s - \mathrm{D}(\mathbf{X}_s)\|^2\right], \tag{3}$$

where M_s indicates the number of sources assigned to group s. The first term on the right-hand-side of (3) gives stress due to heterogeneity of sources within group s; the second term gives the group stress, with respect to the optimally aggregated dissimilarity matrix Θ_s.

Because the loss due to heterogeneity is a constant term with respect to $\mathbf{X}_s$, (3) is minimized by minimizing $\|\Theta_s - \mathrm{D}(\mathbf{X}_s)\|^2$ over $\mathbf{X}_s$. The latter can be done by using the majorization algorithm for MDS in its simplest form (e.g., see de Leeuw, 1988). For each representation space $\mathbf{X}_s$ compute, from a starting point $\mathbf{X}_s^o$, the Guttman transform $\tilde{\mathbf{X}}_s$:

$$\tilde{\mathbf{X}}_s = N^{-1}\mathrm{B}(\mathbf{X}_s^o)\mathbf{X}_s^o. \tag{4}$$

The elements of the $N \times N$ matrix $\mathrm{B}(\mathbf{X}_s^o)$ can be defined in terms of the elements of two auxiliary matrices: the $N \times N$ matrix, $\mathrm{B}^o(\mathbf{X}_s^o)$, whose elements are

$$b_{ij}^o(\mathbf{X}_s^o) = \frac{M_s^{-1} \sum_{m \in \mathrm{J}_s} a_{ms}\delta^*_{ijm}}{d_{ij}(\mathbf{X}_s)}, \quad \text{if } i \neq j \text{ and } d_{ij}(\mathbf{X}_s) \neq 0,$$

$$b_{ij}^o(\mathbf{X}_s^o) = 0, \quad \text{if } d_{ij}(\mathbf{X}_s^o) = 0,$$

and the $N \times N$ diagonal matrix $B^*(\mathbf{X}_s^o)$, with diagonal elements

$$b^*_{ii}(\mathbf{X}_s^o) = \mathbf{1}'B^o(\mathbf{X}_s^o)\mathbf{e}_i.$$

Then,

$$B(\mathbf{X}_s^o) = B^*(\mathbf{X}_s^o) - B^o(\mathbf{X}_s^o).$$

The theory of the majorization algorithm for MDS guarantees that

$$\|\mathbf{\Theta}_s - D(\bar{\mathbf{X}}_s)\|^2 \leq \|\mathbf{\Theta}_s - D(\mathbf{X}_s^o)\|^2.$$

Thus, by repeatedly computing the Guttman transform, a series of convergent configurations is obtained. When $\text{STRESS}(\mathbf{X}_s^o) - \text{STRESS}(\bar{\mathbf{X}}_s) \leq \varepsilon$, with ε some preset small value, the (possibly local) minimum of (3) will be achieved. In the actual minimization of the overall loss function (1), there is no need to converge to the minimum of (3) in each step; as long as the strict inequality $\text{STRESS}(\bar{\mathbf{X}}_s) < \text{STRESS}(\mathbf{X}_s^o)$ holds, a single Guttman transform $\bar{\mathbf{X}}_s$ suffices to decrease the loss in (1). When the overall $\text{STRESS}(\mathbf{A}; \mathbf{\Delta}_1^*, \ldots, \mathbf{\Delta}_M^*; \mathbf{X}_1, \ldots, \mathbf{X}_r)$ has attained its minimum with respect to ε, it must also be true that $\text{STRESS}(\mathbf{X}_s^o) - \text{STRESS}(\bar{\mathbf{X}}_s) \leq \varepsilon$.

Assigning the Sources to Different Points of View

The second step in the algorithmic scheme should minimize (1) over $\mathbf{A} \in \Omega$, for fixed $\mathbf{\Delta}_1^*, \ldots, \mathbf{\Delta}_M^*$ and $\mathbf{X}_1, \ldots, \mathbf{X}_r$. In this paper, we explicitly require that the index sets J_s are mutually exclusive, so that each dissimilarity source contributes to only one point of view, but other approaches are also possible. To solve for $\mathbf{A}$, first construct $\bar{\mathbf{A}}$, minimizing

$$\text{STRESS}(\mathbf{A}) = M^{-1} \sum_{s=1}^{r} \sum_{m=1}^{M} \|a_{ms}\mathbf{\Delta}_m^* - D(\mathbf{X}_s)\|^2,$$

over $\mathbf{A}$ unrestricted. By setting the partial derivatives with respect to a_{ms} equal to zero, one obtains

$$\bar{a}_{ms} = (2N)^{-1} \text{ tr } (\mathbf{\Delta}_m^* D(\mathbf{X}_s)). \tag{5}$$

(The dissimilarity matrices were assumed to be normalized so that the sum of squares of the elements is $2N$.) The estimates in (5) are positive by definition, and because $\text{tr }(\mathbf{X}_s'\mathbf{X}_s) = 1$, the sum of squares of the elements in $D(\mathbf{X}_s)$ is equal to $2N$, so that (5) gives Tucker's (1951) congruence coefficient between the individual dissimilarities and the distances in the s-th point of view. The congruence coefficient is defined on normed vectors of raw scores, it is an association coefficient for ratio scales, and reflects the degree to which two variables are identical up to a positive multiplicative transformation. Therefore, it may also be called the coefficient of proportionality (Zegers & ten Berge, 1985). When dissimilarities are represented in l-vectors $\mathbf{\delta}_m^*$ and distances in l-vectors $\mathbf{d}(\mathbf{X}_s) = \{d_{ij}(\mathbf{X}_s) \text{ for } i < j\}$ that contain the lower diagonal elements of the matrix $D(\mathbf{X}_s)$ in some predetermined order, then $\mathbf{\delta}_m^{*\prime} \mathbf{\delta}_m^* = \mathbf{d}(\mathbf{X}_s)'\mathbf{d}(\mathbf{X}_s) = N$. Now, the congruence coefficient between source m and point of view s is given by $N^{-1}\mathbf{\delta}_m^{*\prime} \mathbf{d}(\mathbf{X}_s)$.

When each source is assigned to only one group, the index set $\{1, \ldots, M\}$ must be partitioned into *nonoverlapping* subsets; in that case the constraints on the weight matrix $\mathbf{A}$ can be written in the form $\mathbf{A} = \mathbf{WG}$, where $\mathbf{W}$ is a diagonal matrix, of order $M \times M$, containing a single weight w_{mm} for each source $\mathbf{\Delta}_m^*$ on its diagonal, and $\mathbf{G}$ is a binary and orthogonal matrix, of order $M \times r$, that assigns each source $\boldsymbol{\Delta}_m^*$ to one of the r groups. So

$$\|\tilde{\mathbf{A}} - \mathbf{WG}\|^2,$$

is minimized over $\mathbf{W}$ and $\mathbf{G}$. This function, which finds nonoverlapping clusters of dissimilarity sources $\boldsymbol{\Delta}_m^*$, can be fitted row after row. The diagonal elements of $\mathbf{W}$ are found as $w_{mm} = \max(\tilde{a}_{m1}, \ldots, \tilde{a}_{mr})$; in the (unlikely) case that some values in the m-th row of $\tilde{\mathbf{A}}$ are exactly equal, we would need an "untie" procedure. Next, $\mathbf{G}$ is obtained by setting $g_{ms} = 1$ if $\tilde{a}_{ms} = w_{mm}$, and 0 otherwise, and the restricted weight matrix is set to $\mathbf{A} = \mathbf{WG}$. Because the weights $\tilde{\mathbf{A}}$ are a function of dissimilarities and distances (in (5)), the result of the allocation of sources to groups is invariant over rotation of the axes in the point of view spaces.

3. Nonlinear Generalizations

In the previous section it was described how the general loss function (1) is minimized over $\mathbf{X}_1, \ldots, \mathbf{X}_r$ and $\mathbf{A}$; in this section nonlinear generalizations (i.e., finding the optimal $\boldsymbol{\Delta}_m^*$) will be discussed. There are two different approaches. First, the relation with Gifi's (1990) approach to nonlinear principal components analysis will be considered; it will be shown that this approach, when applied to distance variables, reduces to nonmetric scaling. Next, a second form of transformation will be proposed, originating from the distance approach to nonlinear multivariate analysis (Meulman, 1986, 1992). Finally, the two possibilities will be combined.

In Gifi (1990) a system of nonlinear MVA techniques is developed that has the notion of homogeneity as starting point. In a principal components analysis, the $N \times M$ data matrix $\mathbf{Z}$ is analyzed, whose columns are defined by N-vectors $\mathbf{z}_m$, $m = 1, \ldots, M$, that contain observations on the variables assumed to have means of zero and sums of squares of one. The measurements on the objects for the M variables define the rows in $\mathbf{Z}$. In the Gifi system, a nonlinear PCA in r dimensions can be written in the form of the loss function:

$$\text{STRIFE}(\mathbf{q}_1, \ldots, \mathbf{q}_M; \mathbf{x}_1, \ldots, \mathbf{x}_r; \mathbf{A}) = M^{-1} \sum_{s=1}^{r} \sum_{m=1}^{M} \|a_{ms}\mathbf{q}_m - \mathbf{x}_s\|^2, \tag{6}$$

that has to be minimized over $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$, constrained so that $\mathbf{X}'\mathbf{X} = \mathbf{I}$, over $\mathbf{A} = \{a_{ms}\}$, and over $\mathbf{q}_1, \ldots, \mathbf{q}_M$, satisfying $\mathbf{q}_m'\mathbf{q}_m = 1$ and $\mathbf{q}_m \in \Gamma_m$, where Γ_m indicates the set of admissible transformations of the given variable $\mathbf{z}_m$.

The class of transformations may be defined differently for each variable $\mathbf{z}_m$, and includes nominal transformations (that preserve equal values in $\mathbf{z}_m$ by forcing ties in $\mathbf{q}_m$), monotonic transformations (that maintain the order of the elements of $\mathbf{z}_m$ in $\mathbf{q}_m$), and linear transformations (which implies setting $\mathbf{q}_m = \mathbf{z}_m$, since it was required that $\mathbf{q}_m'\mathbf{q}_m = 1$). The use of the loss function in (6) suggests that each weighted transformed variable $a_{ms}\mathbf{q}_m$ should resemble the unknown $\mathbf{x}_s$ as closely as possible, where $\mathbf{x}_s$ turns out to be the normalized s-th principal component (Gifi, 1990, chap. 3).

Transformation of the variables in PCA can be applied straightforwardly to points of view analysis, when the latter is regarded as a components analysis of dissimilarity variables. When $\boldsymbol{\delta}_m^*$ denotes the optimal transformation of a given dissimilarity variable $\boldsymbol{\delta}_m$, then (6) could be viewed as a nonlinear variety of the original PVA procedure (Verboon & van der Kloot, 1989), replacing $\mathbf{q}_m$ by $\boldsymbol{\delta}_m^*$ and $\mathbf{x}_s$ by $\boldsymbol{\theta}_s$. The weight matrix $\mathbf{A}$ will in general not give a simple structure, but as is remarked in Gifi (1990, chap. 4

& 10), it is possible to generalize (6) to restricted $\{a_{ms}\}$, requiring, for instance, some a_{ms} to be zero.

When the variables are dissimilarities, the components are optimal with respect to replacing the M variables $\boldsymbol{\delta}_m^*$ by a fewer number of (latent) dissimilarity variables. The particular linear combination, however, will in general not be optimal for the final step in a PVA (i.e., obtaining a low-dimensional space in which the distances between the objects resemble the composite dissimilarities in $\mathbf{x}_s$ as closely as possible). Therefore, we propose to find optimal distance components, minimizing

$$\text{STRESS}(\boldsymbol{\delta}_1^*,\ldots,\boldsymbol{\delta}_M^*;\mathbf{X}_1,\ldots,\mathbf{X}_r;\mathbf{A}) = M^{-1}\sum_{s=1}^{r}\sum_{m\in J_s}\|a_{ms}\boldsymbol{\delta}_m^* - \mathbf{d}(\mathbf{X}_s)\|^2, \qquad (7)$$

satisfying $\boldsymbol{\delta}_m^{*\prime}\boldsymbol{\delta}_m^* = N$ and $\boldsymbol{\delta}_m^* \in \Lambda_m$. Here, Λ_m denotes the set of admissible transformations of $\boldsymbol{\delta}_m$; when the variables are dissimilarities, Λ_m is typically chosen as a set of monotonic transformations. The vector $\mathbf{d}(\mathbf{X}_s) = \{d_{ij}(\mathbf{X}_s) \text{ for } i < j\}$ contains the lower diagonal elements of the matrix $D(\mathbf{X}_s)$ in some predetermined order, so (7) is a special case of the general loss function (1).

Including general monotonic transformations of the dissimilarities brings us to the domain of nonmetric multidimensional scaling, originating from Shepard (1962a; 1962b) and Kruskal (1964a; 1964b). In fact, the KYST program (Kruskal, Young, & Seery, 1973) could be used to find a single point of view, because it fits a single $\mathbf{X}$ to several optimally transformed dissimilarity matrices, where the latter contribute differentially to the aggregated dissimilarity matrix. (In KYST, a_{ms} is absorbed in $\boldsymbol{\delta}_m^*$, i.e., the latter is not normalized.) So (7) can also be viewed as a generalization of KYST, where multiple configurations (points of view) are found by applying cluster restrictions to $\mathbf{A}$.

The transformed dissimilarities are called pseudo-distances, and are usually restricted to be monotonic with the given vector $\boldsymbol{\delta}_m$. When $\bar{\boldsymbol{\delta}}_m$ denotes the unrestricted estimate, obtained by setting partial derivatives with respect to $\boldsymbol{\delta}_m^*$ in (7) equal to zero,

$$\bar{\boldsymbol{\delta}}_m = \frac{\mathbf{d}(\mathbf{X}_s)}{a_{ms}} \text{ if } m \in J_s,$$

it is easy to show that we minimize (7) by minimizing

$$\|\bar{\boldsymbol{\delta}}_m - \boldsymbol{\delta}_m^*\|^2,$$

over $\boldsymbol{\delta}_m^* \in \Lambda_m$, satisfying $\boldsymbol{\delta}_m^{*\prime}\boldsymbol{\delta}_m^* = N$, where Λ_m denotes the set of admissible monotonic transformations of $\boldsymbol{\delta}_m$. Λ_m can be chosen as the set of general monotonic transformations as in Kruskal (1964a, 1964b), but another possibility is to choose Λ_m as the set of monotonic spline transformations of a particular degree, with a prechosen number of knots, as in Ramsay (1982a, 1982b).

The second nonlinear generalization is of a different nature. Considering the loss function (6) in the analysis of the data matrix $\mathbf{Z}$, instead of approximating $a_{ms}\mathbf{q}_m$ directly, one can approximate $D(a_{ms}\mathbf{q}_m)$, the set of distances generated by a weighted variable. This approach is consistent with Meulman (1992), and applied to PVA it implies the minimization of

$$\text{STRESS}(\mathbf{q}_1,\ldots,\mathbf{q}_M;\mathbf{X}_1,\ldots,\mathbf{X}_r;\mathbf{A}) = M^{-1}\sum_{s=1}^{r}\sum_{m\in J_s}\|D(a_{ms}\mathbf{q}_m) - D(\mathbf{X}_s)\|^2. \qquad (8)$$

Due to the homogeneity of the Euclidean distance function, $D(a_{ms}\mathbf{q}_m) = a_{ms}D(\mathbf{q}_m)$, and by setting $\boldsymbol{\Delta}^*_m = D(\mathbf{q}_m)$, (8) turns into another special case of (1).

The optimal transformations of the variables are obtained by the following procedure, derived from the majorization algorithm for MDS with restrictions on the configuration (de Leeuw & Heiser, 1980). In the minimization of (8), the representation spaces $\mathbf{X}_s$ generate target values $d_{ij}(\mathbf{X}_s)$ that have to be approximated, and $\mathbf{q}_m$ is considered a restricted one-dimensional configuration. When $\mathbf{q}^\circ_m$ denotes a starting point that satisfies the constraints, the unrestricted estimate $\bar{\mathbf{q}}_m$ is obtained by computing the so-called reversed Guttman transform (Meulman, 1986) defined by analogy with (4) as

$$\bar{\mathbf{q}}_m = N^{-1}B(\mathbf{q}^\circ_m)\mathbf{q}^\circ_m,$$

where

$$B(\mathbf{q}^\circ_m) = B^*(\mathbf{q}^\circ_m) - B^\circ(\mathbf{q}^\circ_m).$$

Here, the elements of $B^\circ(\mathbf{q}^\circ_m)$ are given by

$$b^\circ_{ij}(\mathbf{q}^\circ_m) = \frac{d_{ij}(\mathbf{X}_s)}{d_{ij}(\mathbf{q}^\circ_m)}, \text{ if } m \in J_s,\ i \neq j \text{ and } d_{ij}(\mathbf{q}^\circ_m) \neq 0;$$

$$b^\circ_{ij}(\mathbf{q}^\circ_m) = 0, \text{ if } d_{ij}(\mathbf{q}^\circ_m) = 0,$$

and the diagonal elements of $B^*(\mathbf{q}^\circ_m)$ as

$$b^*_{ii}(\mathbf{q}^\circ_m) = \mathbf{1}'B^\circ(\mathbf{q}^\circ_m)\mathbf{e}_i.$$

The basic theory of the majorization algorithm implies that

$$\|a_{ms}D(\bar{\mathbf{q}}_m) - D(\mathbf{X}_s)\|^2 \leq \|a_{ms}D(\mathbf{q}^\circ_m) - D(\mathbf{X}_s)\|^2.$$

Using the results from de Leeuw and Heiser (1980), it can be shown that also

$$\|a_{ms}D(\hat{\mathbf{q}}_m) - D(\mathbf{X}_s)\|^2 \leq \|a_{ms}D(\mathbf{q}^\circ_m) - D(\mathbf{X}_s)\|^2,$$

when

$$\hat{\mathbf{q}}_m = \text{argmin } \|\bar{\mathbf{q}}_m - \mathbf{q}_m\|^2,$$

where the minimization is over $\mathbf{q}_m \in \Gamma_m$, satisfying $\mathbf{q}'_m\mathbf{q}_m = 1$, with Γ_m the set of all admissible transformations of a given variable $\mathbf{z}_m$. As in (6) the transformations may be nominal (preserving ties), monotonic (preserving order), or linear (setting $\mathbf{q}_m = \mathbf{z}_m$). The normalization $\mathbf{q}'_m\mathbf{q}_m = 1$ is equivalent to $\boldsymbol{\delta}^{*\prime}_m\boldsymbol{\delta}^*_m = N$ and $\sum_i\sum_j \delta^{*2}_{ijm} = 2N$.

In this application of PVA, the sources $\boldsymbol{\Delta}^*_m$ are generated by the variables, the columns of the transformed multivariate data matrix $\mathbf{Q} = \{\mathbf{q}_1, \ldots, \mathbf{q}_M\}$, and different subsets of variables are assumed to generate different viewpoints about the interrelationships between the objects, the rows of $\mathbf{Q}$. The squared weights in $\mathbf{A}$ could be viewed as a replacement for the squared component loadings in an ordinary principal components analysis, and their means as equivalent to the eigenvalues. In fact, the STRESS in PVA equals $1 - \sum_m\sum_s a^2_{ms}$.

Combining (7), where dissimilarities are transformed, and (8), where variables are transformed, in the form of the general objective function (1), creates the possibility of analyzing mixtures of data that consist of dissimilarities and multivariate data for the same set of N objects. If a multivariate data matrix $\mathbf{Z}$ is available at the outset, one still has the choice of considering each $\mathbf{\Delta}_m^*$ either as a monotonic transformation of $\mathrm{D}(\mathbf{z}_m)$, or as $\mathrm{D}(\mathbf{q}_m)$, with either nominal, ordinal, or numerical transformations for $\mathbf{q}_m$.

4. Points of View Analysis Regarded as a Constrained Weighted Euclidean Model

The model discussed above has a very particular relation with the weighted Euclidean model. Due to the normalization $\sum_i \sum_j \delta_{ijm}^{*2} = \sum_i \sum_j \mathrm{d}_{ij}^2(\mathbf{X}_s) = 2N$, and the homogeneity of the distance function, STRESS can be written as

$$\mathrm{STRESS} = M^{-1} \sum_{s=1}^{r} \sum_{m \in \mathrm{J}_s} \|\mathrm{a}_{ms}\mathbf{\Delta}_m^* - \mathrm{D}(\mathbf{X}_s)\|^2$$

$$= M^{-1} \sum_{s=1}^{r} \sum_{m \in \mathrm{J}_s} \|\mathbf{\Delta}_m^* - \mathrm{D}(\mathrm{a}_{ms}\mathbf{X}_s)\|^2.$$

Furthermore, if one defines a $p_s \times p_s$ diagonal matrix $\mathbf{C}_{m(s)}$ with all diagonal elements equal to a_{ms}, then

$$M^{-1} \sum_{s=1}^{r} \sum_{m \in \mathrm{J}_s} \|\mathbf{\Delta}_m^* - \mathrm{D}(\mathrm{a}_{ms}\mathbf{X}_s)\|^2 = M^{-1} \sum_{s=1}^{r} \sum_{m \in \mathrm{J}_s} \|\mathbf{\Delta}_m^* - \mathrm{D}(\mathbf{X}_s\mathbf{C}_{m(s)})\|^2.$$

Now, for instance, consider the situation with two points of view, so $s = 1, 2$. Define the $N \times p$ matrix $\mathbf{X}$, where $p = (p_1 + p_2)$, with the dimensions from $\mathbf{X}_1$ and $\mathbf{X}_2$. Also, define $\mathbf{C}_m$ as a $p \times p$ diagonal matrix, with the first p_1 diagonal elements equal to a_{m1} and the next p_2 diagonal elements equal to a_{m2}. Then,

$$M^{-1} \sum_{s=1}^{r} \sum_{m \in \mathrm{J}_s} \|\mathbf{\Delta}_m^* - \mathrm{D}(\mathbf{X}_s\mathbf{C}_{m(s)})\|^2 = M^{-1} \sum_{m=1}^{M} \|\mathbf{\Delta}_m^* - \mathrm{D}(\mathbf{X}\mathbf{C}_m)\|^2,$$

where the term on the right-hand-side gives the weighted Euclidean model in the STRESS framework. By the definition of $\mathbf{A}$ in our PVA, either a_{m1} or $\mathrm{a}_{m2} = 0$, so the matrices $\mathbf{C}_m$ have a very particular structure. The diagonal elements of each $\mathbf{C}_m$ can be collected in the rows of an $M \times p$ matrix; for a hypothetical analysis with two points of view, each with two dimensions, the structure of the corresponding INDSCAL weights are given in Table 1.

So, individual weights are different within a group for which a separate point of view is fitted, but the dimension weights are equal (we assumed the configurations $\mathbf{X}_s$ to have a certain shape; if instead the dimensions would be given equal sums of squares, the corresponding weights would be proportional). A quite different model for groups using the weighted Euclidean framework has recently been proposed by Winsberg and De Soete (in press). In their so-called latent class model, the weights are equal within groups and different across dimensions. For a comparison with our restricted model, the corresponding INDSCAL weight structure is given at the right-hand-side of Table 1. Further research is needed to investigate whether it would be worthwhile to combine features from the two approaches.

5. Points of View Analysis in Action

To discuss the properties of PVA as presented in this paper in more detail, data will be analyzed that were obtained from a questionnaire study among the members of the

TABLE 1
Constrained INDSCAL Weight Patterns for Not-normalized Dimensions in X According to Two Different Group Models (Groups are indicated by G_1, G_2)

	Points of view approach					Latent class approach			
	Dimension					Dimension			
	1	2	3	4		1	2	3	4
	a_{11}	a_{11}	0	0		a_{11}	a_{12}	a_{13}	a_{14}
	a_{21}	a_{21}	0	0		a_{11}	a_{12}	a_{13}	a_{14}
		⋮					⋮		
G_1	a_{m1}	a_{m1}	0	0	G_1	a_{11}	a_{12}	a_{13}	a_{14}
		⋮					⋮		
	$a_{M_1 1}$	$a_{M_1 1}$	0	0		a_{11}	a_{12}	a_{13}	a_{14}
	0	0	$a_{(M_1+1)2}$	$a_{(M_1+1)2}$		a_{21}	a_{22}	a_{23}	a_{24}
		⋮					⋮		
G_2	0	0	a_{m2}	a_{m2}	G_2	a_{21}	a_{22}	a_{23}	a_{24}
		⋮					⋮		
	0	0	$a_{M_2 2}$	$a_{M_2 2}$		a_{21}	a_{22}	a_{23}	a_{24}

Second Chamber of the Dutch Parliament in 1979–1980 (the data were kindly made available by the Department of Political Science of the University of Leiden). In this study, 139 of the 150 members of the parliament (MP's) participated; they belong to 11 different political parties, and a short description is given in Table 2. From the extensive questionnaire, several variables were chosen for different applications of PVA, the data always pertaining to the same set of 139 MP's. The parties in Table 2 are ordered by using averages within parties derived from a variable that gives the position that the MP's assigned themselves to on a political left-right scale, with a range from 1 (extremely left) to 9 (extremely right).

In the first application, data are analyzed that give the MP's position with respect to 8 political issues, measured by self-ratings on a 9-point scale. The lower and upper end of the scales for the political issues is given in Table 3, as well as the marginal frequencies of the categories. In the interpretation of the results, the skewness of some of the distributions of the MP's over the categories should be taken into account. In the second application, data are used that were expressed in values on so-called sympathy scales, ranging from 0 (extremely unappealing) to 100 (extremely appealing): each MP gave a score to each of the parties residing in Parliament in 1979 (described in Table 2).

Analysis of the Political Issues

For these data a points of view analysis is applied that accommodates the two different types of nonlinear transformations. Two natural questions arise when analyzing political issues: first, whether they can be captured adequately in a single point of

TABLE 2
Political Parties in the Dutch Parliament in 1979-1980,
Party Membership of the Respondents, and Mean Self Rating within Parties
on the Left (1)-Right (9) Scale in the Questionnaire

Label	Party	Description	Number of Respondents	Mean Left-Right
	CPN	Communists	0	
2	PSP	Pacifistic socialists	1	1.00
3	PPR	Radicals	3	1.33
4	PvdA	Social democrats	53	2.70
5	DS70	Social democrats (economically conservative)	1	3.00
6	D66	Liberals (economically progressive)	8	3.25
7	ARP	Protestants	12	4.00
8	KVP	Catholics	24	4.21
9	CHU	Protestants	9	5.00
10	VVD	Liberals (economically conservative)	25	5.04
11	GPV	Very conservative Calvinists	1	7.00
12	SGP	Very conservative Calvinists	2	8.50
	BP	Farmers Party	0	
	CDA	Merger of ARP, KVP and CHU		

view (dominated by the left-right dimension), or do they require more than one; secondly, whether parties are homogeneous with respect to the positions that their MP's have on the various issues. To include party membership in the analysis, an indicator matrix **B** was constructed, with $N = 139$ rows and 11 columns, indicating for each MP to which of the 11 parties (s)he belonged. From the indicator matrix **B**, dissimilarities between the MP's were derived; although any dissimilarity measure could have been considered, the chi-square distance was selected that is also used in homogeneity analysis (or multiple correspondence analysis). The particular use here is similar to Gifi's (1990) approach to principal components analysis, being a mixture of PCA as in (6) and homogeneity analysis. The squared χ^2-distance between two MP's i and j is defined by

$$\chi^2_{ij}(\mathbf{B}) = (\mathbf{e}_i - \mathbf{e}_j)'\mathbf{B}\mathbf{M}^{-1}\mathbf{B}'(\mathbf{e}_i - \mathbf{e}_j) = d^2_{ij}(\mathbf{B}\mathbf{M}^{-1/2}), \tag{9}$$

where the matrix $\mathbf{M}^{-1} = (\mathbf{B}'\mathbf{B})^{-1}$ is a diagonal matrix that has the inverse of the column marginals of **B** on its diagonal.

Because all dissimilarities between MP's that belong to the same party are zero, and all dissimilarities between the MP's of two different parties are equal, many ties exist in $D(\mathbf{B}\mathbf{M}^{-1/2}) = \{\chi^2_{ij}(\mathbf{B})\}^{1/2}$. Therefore, the dissimilarities were transformed monotonically with Kruskal's primary approach to ties that allows ties in the data to become untied in the transformation. Thus, within-party pseudo-distances remain smaller than between-parties pseudo-distances, and the latter are monotonic with the original chi-square distances.

By contrast, the 8 political issue variables were treated as follows. We chose to define $\mathbf{\Delta}^*_m$ as $D(\mathbf{q}_m)$ as in (8) to find a transformation of the variables; although other monotonic transformations of the data in $\mathbf{z}_m$ could have been considered, to obtain smooth transformations, Γ_m was defined as the set of monotonic spline transformations (as in the approach to PCA in Ramsay, 1989; Winsberg & Ramsay, 1983). Second-

TABLE 3
Political Issues in the Questionnaire

	Issue	Lower End (1).........................	(9) Upper End
1	Development aid	The government should spend much more money on development aid	The government should spend much less money on development aid
2	Abortion	The government should prohibit abortion under all circumstances	Every woman has the right to decide for herself on this matter
3	Law & Order	The government takes too rigorous measures against disturbances of the Queen's peace	The government should take even more rigorous measures
4	Income Differences	The differences in income should remain as they are	The differences in income should become much smaller
5	Employees' Participation	Only management should decide in matters that concern the company	Employees should have their say in decisions
6	Taxes	Taxes should be raised so that more money will become available for public provisions	Taxes should be lowered so that everybody can decide for him/herself
7	Armies	The government should insist on reducing armed forces, even if this would imply risk	The government should maintain strong armed forces.
8	Nuclear Energy	The number of nuclear power plants should be increased rapidly	Nuclear power plants should not be built at all

Marginal Frequencies of Categories

Issue	1	2	3	4	5	6	7	8	9
1	29	35	25	20	24	2	3	1	0
2	3	14	6	7	11	7	16	28	47
3	6	11	23	23	46	14	12	3	1
4	1	4	1	9	12	23	31	34	24
5	0	1	0	3	9	10	28	45	43
6	11	11	29	23	25	19	14	5	2
7	16	13	17	21	25	12	16	7	10
8	2	3	10	18	24	12	20	27	23

degree splines with two interior knots were used, which fixes the number of parameters fitted for each variable to four.

Figure 1 gives the transformation of the issue variables. It is important to scrutinize

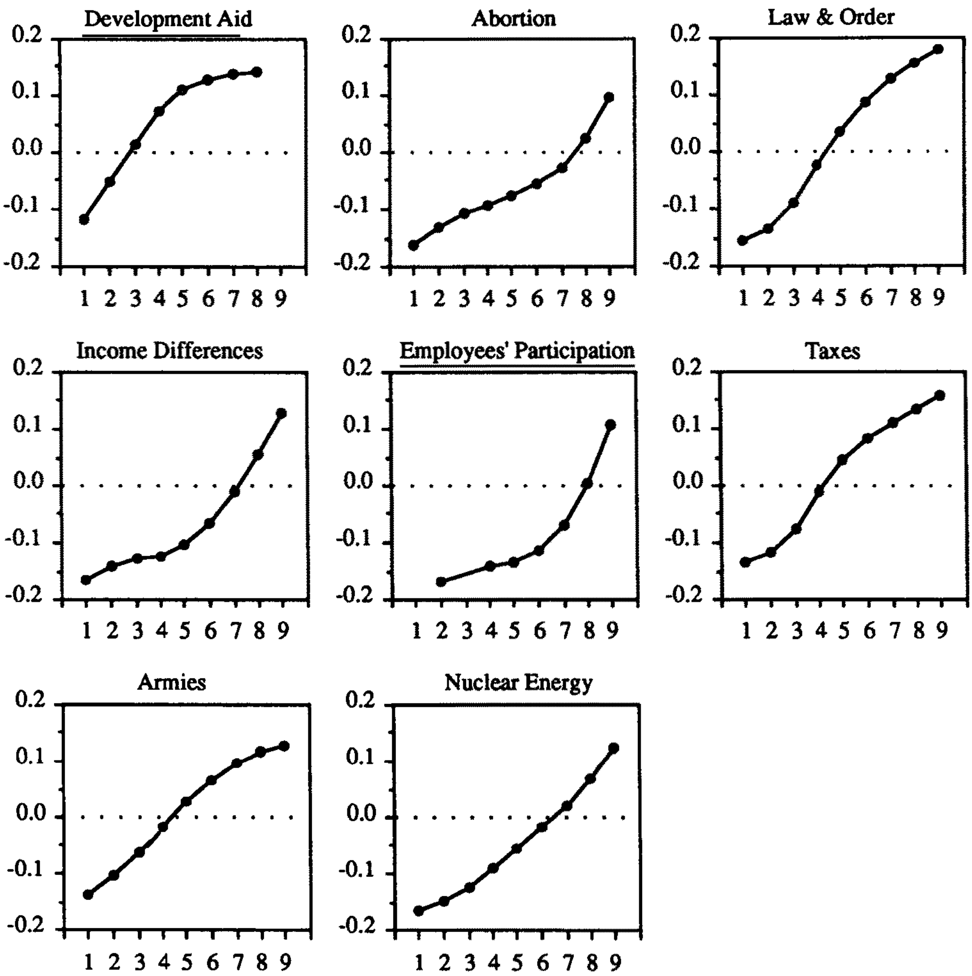

FIGURE 1.
Analysis of political issues and party membership: Optimal monotonic spline transformations for 8 political issues.

these, because the values on the original scales were equally-spaced, but generally this may not be true after the monotonic spline transformations. When the transformation is a concave function as for Development Aid, the lower end of the scale (much more money) is emphasized, while the upper end (less money) is de-emphasized. When the transformation is a convex function as for Employees' Participation, the lower end (only management should decide) is de-emphasized; the upper end (employees should have their say too) is emphasized, probably because this is not a very extreme statement. When the transformations are viewed together with the marginal frequencies of the categories in Table 3, the curves are steep when the associated marginal frequencies are large, while the curves are flat when the marginal frequencies are small (compare the concave function for Development Aid, with marginal frequencies 29, 35, 25 for the categories 1, 2, 3 and marginal frequencies 2, 3, 1 for the categories 6, 7, 8); in short, it turns out that the optimal spline transformations follow the cumulative frequency distributions very closely.

Two points of view were considered; the first displays the left-right dimension, but in the second, the positions of the MP's vis-a-vis political issues do not correlate well with the left-right continuum. Table 4 gives the weights and the badness-of-fit values for

TABLE 4
Weights and Loss for Each Source, and
Partitioned Loss: Total Stress = Heterogeneity + Group Stress

	First Point of View	Second Point of View	Loss in Applicable Point of View	
Source	Weight	Weight	Stress	Heterogeneity
Law & Order	0.900		0.190	0.144
Income Differences	0.893		0.203	0.149
Employees' Participation	0.829		0.313	0.202
Taxes	0.902		0.187	0.148
Armies	0.896		0.197	0.145
Nuclear Energy	0.893		0.202	0.147
Development aid		0.895	0.199	0.164
Abortion		0.901	0.189	0.158
Party Membership		0.933	0.129	0.068
	Total Stress	Hetero-geneity	Group Stress	Mean Weight
First point of view	0.144	0.104	0.040	0.886
Second point of view	0.057	0.043	0.014	0.907
Overall	0.201	0.147	0.054	0.894

each source. The overall stress value, described in (1), is the sum of the partitioned stress values given in (3). The mean square of the weights (replacing the sum of the eigenvalues in an ordinary PCA) is 0.799, which is 1—Overall Stress; the Group Stress in (3) is 0.040 for the first viewpoint and 0.014 for the second viewpoint, and the Stress due to Heterogeneity within groups is 0.104 and 0.043, respectively.

The two points of view found by the analysis are as follows (the weights are given in parenthesis): the first is formed by the variables Law & Order (0.90), Income Differences (0.89), Employees' Participation (0.83), Taxes (0.90), Armies (0.90) and Nuclear Energy (0.89), and the second by Development Aid (0.90) and Abortion (0.90); the monotonically transformed dissimilarities derived from party membership were fitted (weight: 0.93) with the second point of view.

The first point of view is represented in Figure 2; the 139 MP's are represented by points with the labels from Table 2 that indicate their party membership. Also, points are given for the parties: these are the centroids of individual MP's who belong to the same party. Finally, the political issues that constitute the first point of view are represented as vectors, whose coordinates are the correlations between the transformed variables and the two dimensions.

The space is in principal axes position; since the eigenvalues are 0.85 and 0.15, there is a very dominant first dimension. When the order of the parties along this dimension is compared with the ordering from left to right in Table 2, it might be considered closely related to the assumed left-right continuum (in contrast with the self rating, DS70 is positioned on the conservative side). MP's that are positioned left from the origin are more likely to feel that income differences should be smaller, and nuclear power plants should not be built when compared to MP's right from the origin; also, they feel stronger that armies should be reduced, and that the government takes too rigorous action against disturbances of the peace. To a smaller extent, MP's right from the origin felt that with respect to Employees' Participation, only Management should decide, while those to the left endorsed Taxes being raised for public provisions.

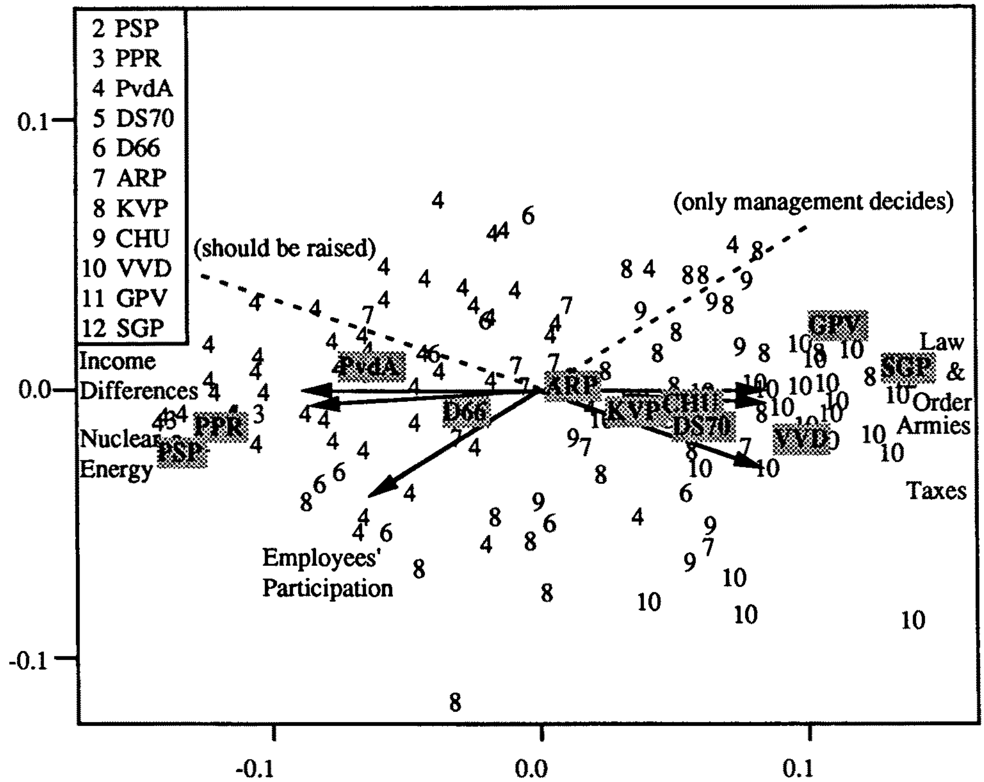

FIGURE 2.
Analysis of political issues and party membership: Multidimensional structure according to Employees' Participation and Taxes versus Income Differences, Nuclear Energy, Law & Order, and Armies.

The second point of view separates Abortion and Development Aid from the other political issues; Figure 3 gives the positions of the MP's. The configuration is certainly not one-dimensional (the eigenvalues are 0.52 and 0.48). Abortion separates the denominational parties GPV, SGP, CHU, KVP, and ARP (that feel that abortion should be prohibited) from the parties that feel that every woman has the right to decide for herself: these parties are the left-wing PSP, PPR, PvdA, and D66, but also the economically conservative VVD and DS70. The extreme positions towards Development Aid are taken by the PPR and PSP (much more money), and DS70, SGP, and GPV (less money); a small distinction is found between the ARP, CHU, KVP (more money), and PvdA and D66 (somewhat more). None of the two directions given by the issue variables display the left-right dimension. And the compromise, the first dimension, positions the (conservative) VVD to the left of ARP, CHU, and KVP, which is considerably different from its position on the left-right dimension in Figure 2.

As is clear from Figure 3, there is quite some heterogeneity within parties; this heterogeneity can be inspected more closely through the transformed dissimilarities according to party membership. Because ties were allowed to be untied, within-party pseudo-distances will differ from the original zero dissimilarities, and large discrepancies from zero indicate large heterogeneity. The within-party pseudo-distances were grouped into seven classes, and the frequencies for each class computed. The cumulative frequency distributions (expressed in proportions) are displayed in Figure 4 (the small parties were omitted from this graph). Compared to KVP and CHU, PvdA and VVD are more homogeneous and ARP is less so, while D66 displays a remarkable heterogeneity.

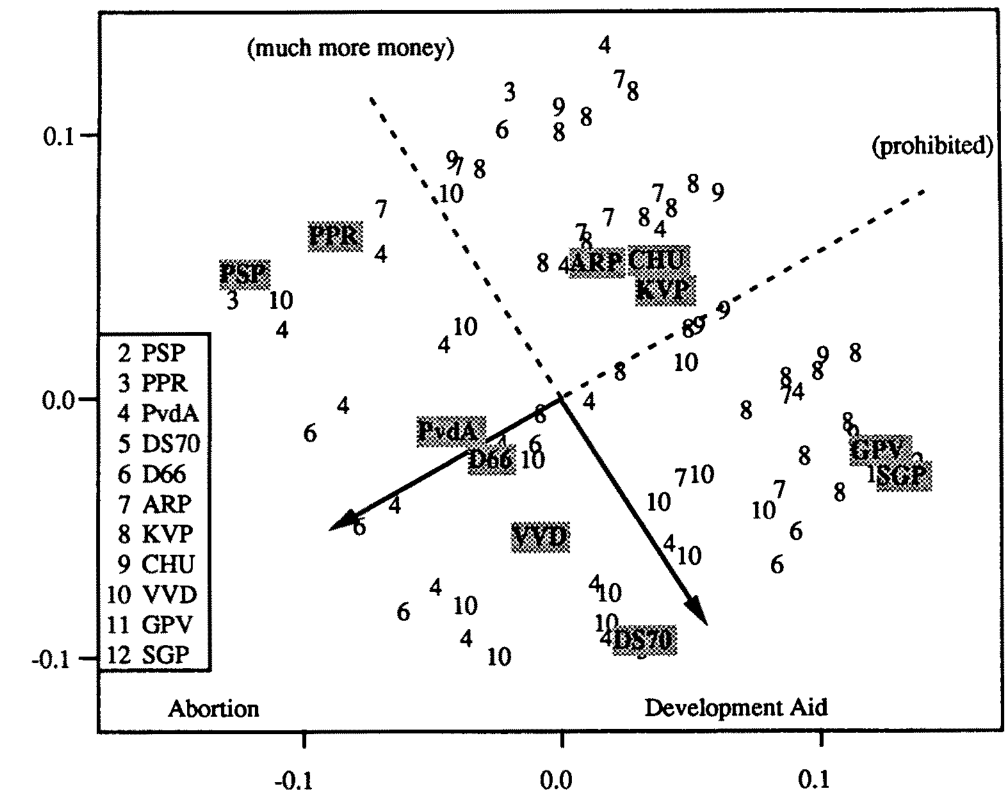

FIGURE 3.
Analysis of political issues and party membership: Multidimensional structure according to Abortion, Development Aid, and Party Membership.

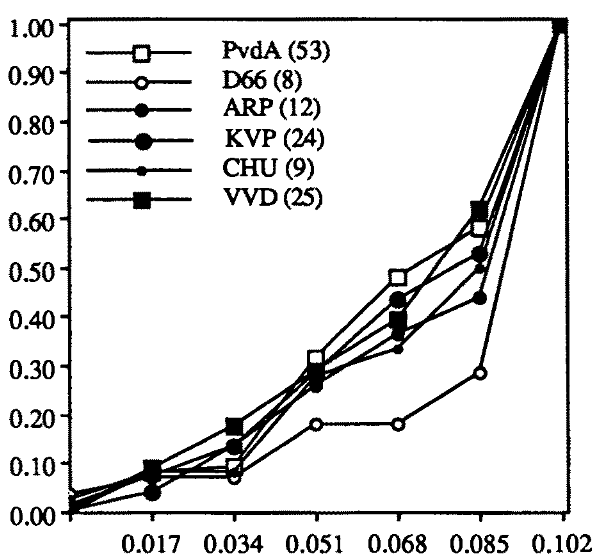

FIGURE 4.
Cumulative frequency distributions displaying heterogeneity within parties, derived from grouped within-party pseudo-distances. Proportions (vertical axis) versus pseudo-distances (horizontal axis).

TABLE 5
Distribution of MP's over Two Points of View, Mean Weights and Partitioned Loss: Total Stress = Heterogeneity + Group Stress

	First Point of View				Second Point of View			
Source	#	Mean Weight	Mean Stress	Hetero-geneity	#	Mean Weight	Mean Stress	Hetero-geneity
PSP	1	0.886	0.215	0.170	0			
PPR	3	0.930	0.135	0.112	0			
PvdA	52	0.930	0.133	0.113	1	0.948	0.101	0.106
DS70	0				1	0.883	0.221	0.157
D66	7	0.857	0.263	0.199	1	0.846	0.284	0.191
ARP	0				12	0.884	0.216	0.165
KVP	0				24	0.914	0.162	0.135
CHU	0				9	0.927	0.140	0.124
VVD	1	0.843	0.289	0.216	24	0.850	0.275	0.197
GPV	0				1	0.930	0.134	0.118
SGP	0				2	0.806	0.348	0.235

	Total Stress	Hetero-geneity	Group Stress	Mean Weight
First point of view	0.070	0.058	0.012	0.920
Second point of view	0.113	0.087	0.026	0.887
Overall	0.183	0.145	0.038	0.901

Analysis of the Sympathy Scales

The second application concerns the sympathy data. Here we wish to investigate whether members of parliament being in different positions in the political spectrum possibly have a different system of sympathies towards the other parties. To explore this question, the MP's are considered judges of the interrelationships between the political parties, so one has to consider the MP's as the columns of the data matrix or the variables, and the parties as the rows or the objects. This implies there are 139 dissimilarity matrices, one for each MP, of order 14×14, since there are 14 different parties judged (there are 3 more parties than the number of parties for which we have MP's, because the CPN and BP MP's did not participate in the questionnaire, and CDA only acts as a stimulus party, since it is the result of a merger between ARP, KVP, and CHU). In this application, the sympathy scales were treated in a similar way as the issue variables in the previous application: not the dissimilarity variables, but the given variables were optimally transformed to give distances $D(\mathbf{q}_m)$, using second-degree monotonic splines with one interior knot, which fixes the number of parameters fitted for each sympathy scale to three.

Two points of view were considered; a single point of view clearly did not fit the data, and two points of view were considered sufficient in terms of goodness-of-fit. The mean squared weight is 0.817 (1—Overall Stress), the Group Stress in (3) is 0.012 for the first viewpoint and 0.026 for the second viewpoint, and the Stress due to Heterogeneity within groups is 0.058 and 0.087, respectively. Table 5 gives the distribution of the MP's over the two points of view, and weights and badness-of-fit values, which have been averaged over sources that belong to the same political party. The overall stress value, described in (1), is the sum of the partitioned stress values given in (3), and is equal to the weighted sum over parties, divided by M, the total number of sources.

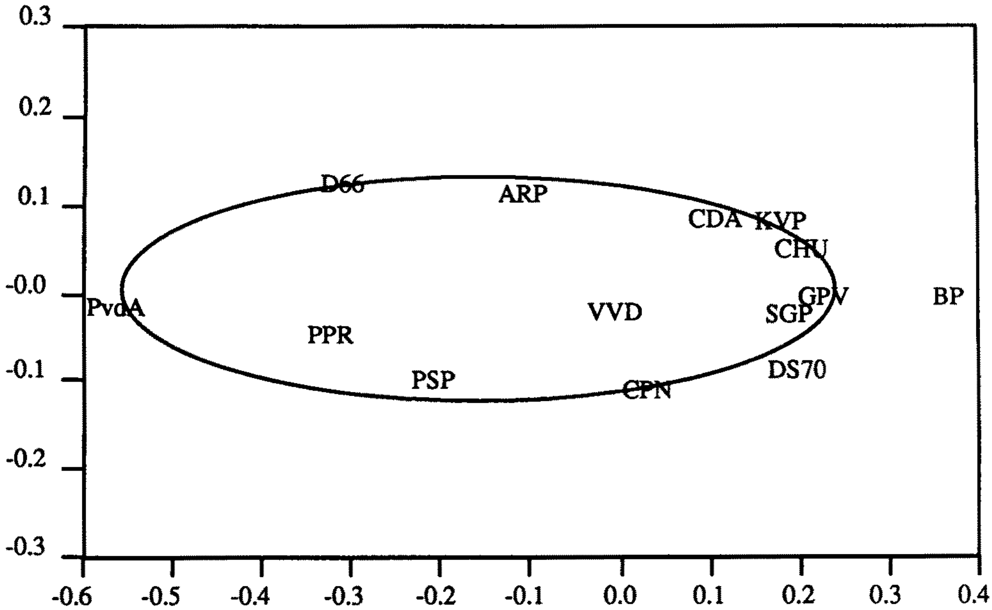

FIGURE 5.
First viewpoint in analysis of political sympathy scales: Left-wing point of view.

The first group is formed by 64 sources and seems to represent the MP's who, according to their self ratings, are left from the center in the political spectrum. There are 4 exceptions: one member of the PvdA and one member of D66 are in the second group and (as in the first example), so is the MP of DS70. But, one member of the VVD, whose other 24 MP's are in the second group, is in the first group. In the configurations for the point of views, the objects are represented by 14 political party points.

The first point of view is displayed in Figure 5; it is fitted to the group of sources that consists of the MP's of PvdA (52 out of the total 64), and the MP's of PSP, PPR and D66, and could be called "the left-wing point of view". The object point for the PvdA party is represented at the left-hand-side of the Figure: on the average, the largest sympathy in the first point of view is for the PvdA. Next, moving to the right, there are two parties that are separated from each other in the second dimension: the PPR which is more left-wing than PvdA, and D66 which is more to the center. More distant are the even more left-wing PSP, in the second dimension close to the PPR, and the more center ARP, in the second dimension closer to D66; next follow the CPN and the VVD. Distances become quite large between the PvdA on the one hand, and the denominational parties KVP, CHU, GPV, and SGP, on the other. The CDA is located in between the three participating parties (but closer to the KVP and CHU than to the ARP). There is also little sympathy for DS70, which is understandable since it is a conservative secession from the PvdA, and there is no sympathy at all for the very right-wing BP. The overall configuration can be captured in an elliptical structure (as drawn in Figure 5); starting at the point for the CPN, and moving clockwise along the ellipse in the direction of the PSP, the left-right order is recovered (see Table 2). The VVD does not fit on the ellipse; there is more sympathy for the VVD in the left-wing point of view than can be explained from the left-right scale. From a substantive point of view, the political structure can easily be understood by looking at the distances between the political parties, but the dimensions cannot be given a politically relevant interpretation.

Inspecting the data of the MP of the VVD who is in the first group, it turns out that this MP ranks the parties almost perfectly in reverse order compared to the average MP

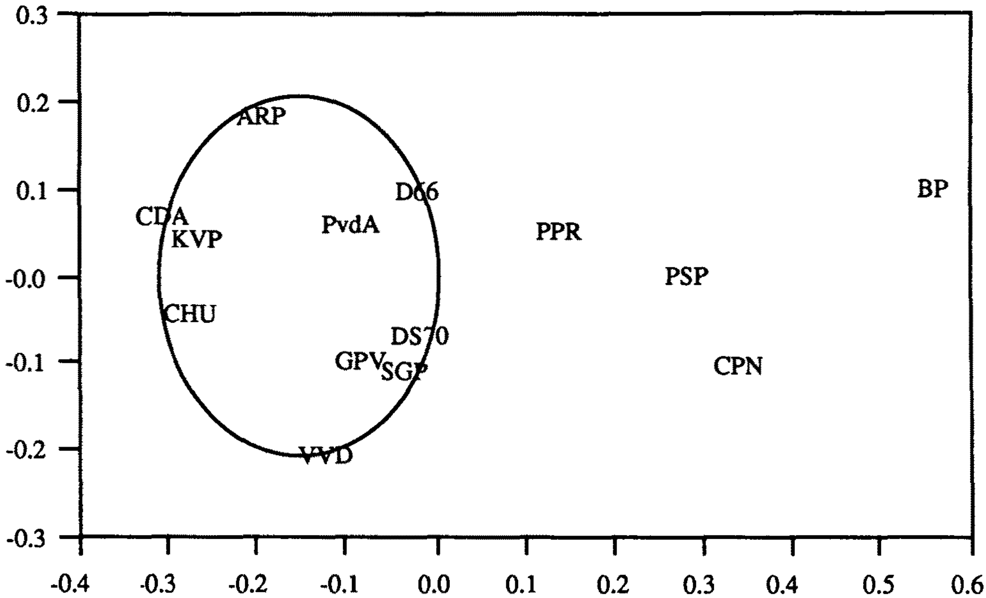

FIGURE 6.
Second viewpoint in analysis of political sympathy scales: Center-right-wing point of view.

of the PvdA. Compared to the average MP of the VVD, this MP has much less sympathy for PvdA, D66, and ARP, and much more for GPV and SGP. The MP of the PvdA who does not fit into the left-wing point of view, ranks the parties with decreasing sympathy as (CDA CHU KVP ARP) (GPV SGP) (VVD D66 PvdA) DS70 PPR (PSP CPN BP), where the parentheses indicate tied values. A first conclusion might be that the point of view analysis perfectly discovered a coding error in the party membership variable; however, this is unlikely considering other data available, so perhaps it should be concluded that this MP is changing his or her political affiliation.

Figure 6 shows the second point of view; it could be called the "center-right-wing point of view". The majority of the MP's that adhere to this viewpoint belong to the parties that merged into CDA (Christian Democrats); MP's of parties that are more conservative are also allocated to this viewpoint. One part of the Christian democrats has great sympathy for parties that are more left-wing (PvdA and D66), while the others have more sympathy to parties that are more to the right (VVD, GPV, SGP, and DS70). There is little sympathy for the small extreme left-wing parties PPR, PSP, and CPN, and the extreme right-wing BP. The latter political parties fall outside the ellipse that orders the parties from left to right when we start at the D66 point and move counterclockwise towards the ARP; in the center-right-wing point of view there is more sympathy for the PvdA than can be explained from the left-right scale.

The major agreement between the two viewpoints seems to be the antipathy towards the extremely right-wing BP. But, in the left-wing point of view, the distances between the BP and CHU, KVP, CDA, SGP, GPV, and DS70 are small, while in the center-right-wing point of view they are large. The same is true for the distance between CDA and CPN. It is exactly the other way around for the distances between PvdA and CHU, KVP, CDA, SGP, GPV, and DS70; in the left-wing point of view they are large, and in the center-right-wing point of view they are small.

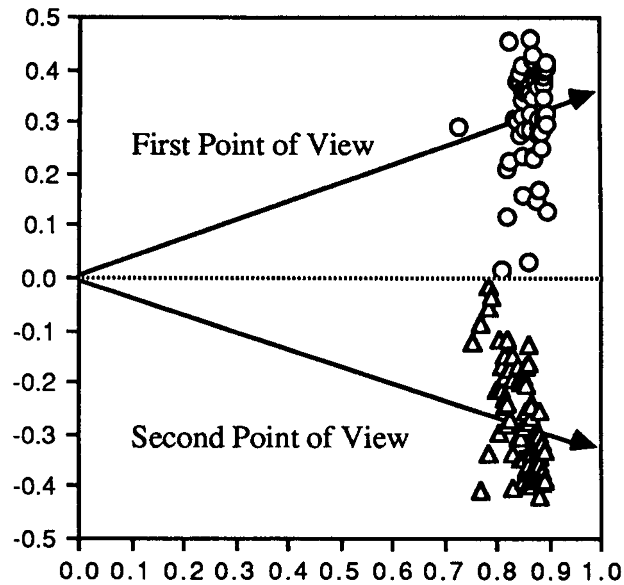

FIGURE 7.
Ordinary principal components analysis of dissimilarity variables: Component loadings for sources in two dimensions.

Application of Principal Components Analysis to Dissimilarity Variables

To see how the points of view analysis has behaved, the obtained dissimilarity variables δ^*_m were also analyzed in an ordinary principal components analysis in two dimensions. The component loadings are displayed in Figure 7, and the sources are labeled according to the group structure found in the PVA.

Basically, the same two clusters are obtained (the arrows that symbolize the two points of view are drawn through the average loadings for each group). Close to the horizontal axis, however, several sources are found that would have been difficult to allocate to one of the two groups on the basis of their loadings.

Application of the weighted Euclidean model to the Sympathy Scales

To study the properties of the PVA analysis compared to the original weighted Euclidean model, the sympathy scales were also analyzed by minimizing the loss function

$$\text{STRESS}(\mathbf{X}; \mathbf{C}_1, \ldots, \mathbf{C}_m) = M^{-1} \sum_{M=1}^{m} \|\mathbf{\Delta}^*_M - \text{D}(\mathbf{X}\mathbf{C}_M)\|^2,$$

over the common space $\mathbf{X}$ and the diagonal weight matrices $\mathbf{C}_1, \ldots, \mathbf{C}_M$ for given $\mathbf{\Delta}^*_m = \text{D}(\mathbf{q}_m)$, with the latter again obtained from the points of view analysis. A special purpose algorithm was developed, based on the majorization approach detailed in Heiser and Stoop (1986). There is some freedom to choose from different, but coherent, normalizations; because we wish the common space to have an explicit shape, the weights were normalized so that $M^{-1}\sum_m \mathbf{C}^2_m = \mathbf{I}$ (this is equivalent to normalized dimensions in $\mathbf{X}$, having sum of squares of one, and non-normalized weights).

Since there were 2×2 dimensions in the points of view analysis, the INDSCAL model was first fitted with 4 common dimensions; the stress was 0.145, so the 139 individual spaces (each using four parameters) fit the data only slightly better than the

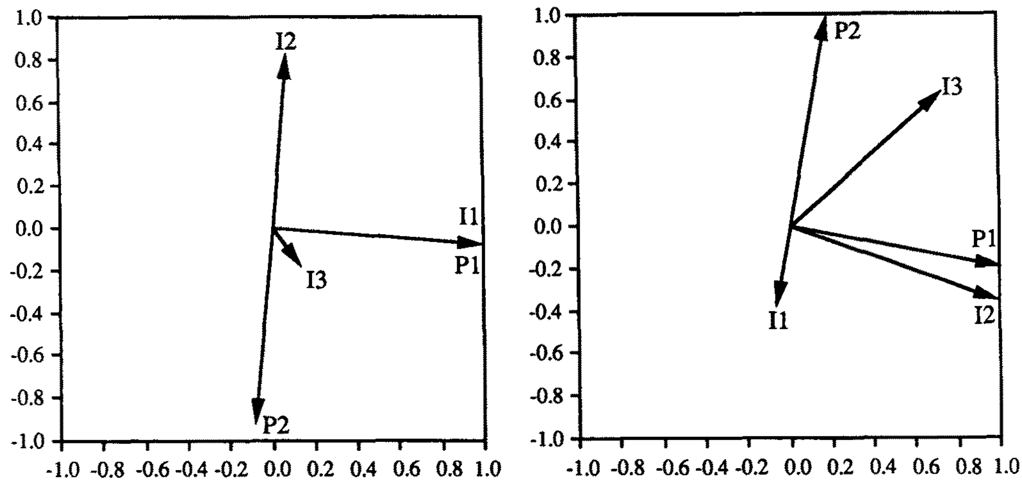

FIGURE 8.
Two canonical correlation analyses of the three INDSCAL dimensions with the two dimensions in the first (on the left) and second (on the right) point of view, respectively. Correlations of the original dimensions with the canonical variates (linear combinations of the INDSCAL dimensions).

2 point of view spaces (using 1 parameter per source, with the total stress 0.183). However, two of the dimensions of the INDSCAL solution turned out to be highly correlated. Therefore, a three-dimensional INDSCAL model was fitted; the stress of the solution was 0.158, which is acceptable compared to the four-dimensional solution since fewer parameters are fitted.

To see how the three-dimensional INDSCAL solution relates to each of the two-dimensional points of view, two canonical correlation analyses were carried out. The correlations of the dimensions with the canonical variates are depicted in Figure 8. At the left-hand-side, the relation with the first point of view (POV1) is given; the right-hand-side provides the relation with the second point of view (POV2). It is clear that the first dimension (P1) in POV1 is the first INDSCAL dimension (I1), and the second dimension (P2) of POV1 is the (reversed) second INDSCAL dimension (I2). With respect to the second point of view, it turns out that I2 is close to P1 in POV2, but the third INDSCAL dimension (I3) is in between P1 and P2. Because I2 and I3 are not highly correlated, the choice was made to depict the INDSCAL solution in two figures: Figure 9 gives I2 versus I1, and Figure 10 displays I3 versus I2.

Because the sum of squares in the first and second dimension of the INDSCAL common space is 0.384 and 0.326, respectively, Figure 9 does not show the clearly dominant first dimension of POV1 in Figure 5 (the canonical analysis ignores the size of the dimensions). Apart from the size of the second dimension, a very obvious difference between the two Figures is the position of the extreme left-wing CPN with respect to the extreme right-wing BP. In Figure 9 they are very close; in Figure 5 they are separated by the calvinist SGP-GPV cluster. Comparing Figure 10 with Figure 6, we especially notice the different position of PvdA: in Figure 6 it is close to D66, and in Figure 10 it is close to SGP, DS70 and GPV; the latter is not easy to explain politically.

The weights from the INDSCAL analysis are depicted in Figure 11; the points are labeled with the group structure from PVA. In the graph at the left, the sources in the first group are separated from the other group in the first dimension. The sources that form the second group are scattered along the second dimension. The graph at the right shows the relation between dimension 3 versus 2. Sources in the second group that have small weights on the second dimension have large weights on the third dimension,

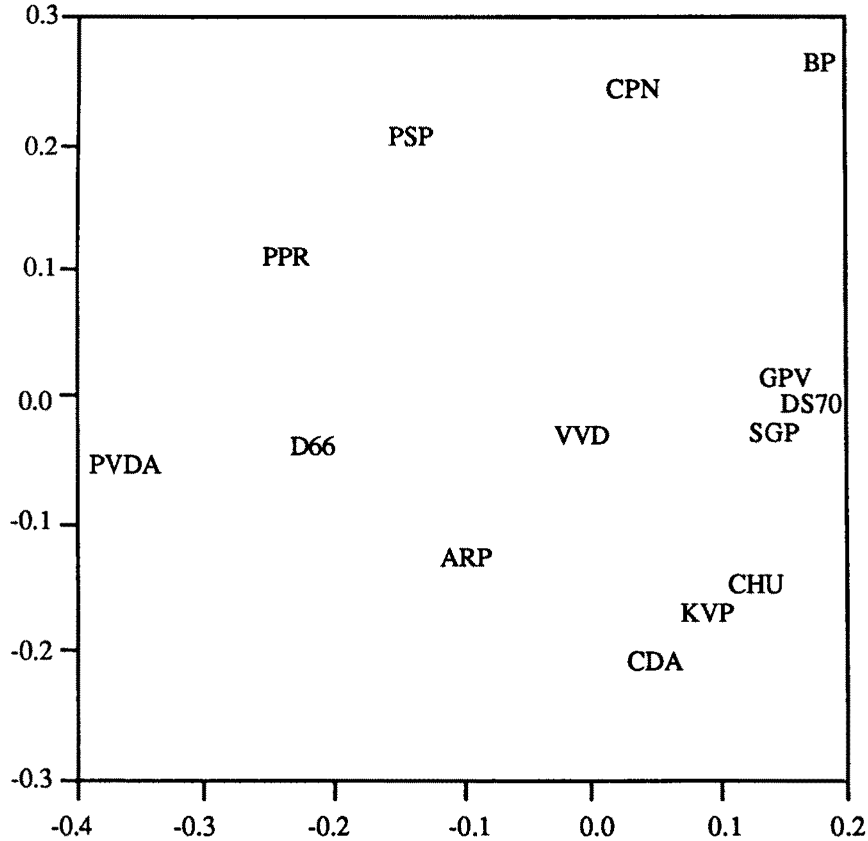

FIGURE 9.
INDSCAL analysis of political sympathy scales: Parties in dimension 2 (vertical axis) versus dimension 1 (horizontal axis).

and vice versa, while the first group has small weights for both the second and third dimension. There are a few weight patterns that are peculiar, given the PVA group structure. Going back to the data, it turned out that these sources did not fit very well in the solution. Nevertheless, inspecting the Pearson correlations for both the transformed sympathy scales and the derived dissimilarity variables, the average correlation (ignoring the sign) for these sources with the other sources in the group to which they were assigned was larger than the average correlation with the other group.

To see whether our analysis of PVA as a constrained weighted Euclidean model holds, diagnostics have been computed using the group structure from PVA. Table 6 partitions the total variance $\sum_m \|\mathbf{XC}_m\|^2$ into groups $\times$ dimensions; also, the average weights are given. The normalization of the weights is dependent on the normalization of the configuration **X**. The sum of squares of **X** is given in the third column; the first two columns of weights apply to unnormalized **X**; the last two columns apply to a normalized **X**. The latter columns clearly show that the average weights in the first two PVA dimensions (0.887 and 0.246) are close to the average INDSCAL weights (0.870 and 0.207) for the first group in Dimensions 1 and 2. PVA Dimensions 3 and 4 (average weights 0.820 and 0.338) resemble the average INDSCAL weights for the second group (0.713 0.396) in Dimensions 2 and 3. The weights in PVA that are zero by definition, are in INDSCAL 0.143 (for the first group on dimension 3) and 0.204 (for the second group on dimension 1). So, PVA viewed as a constrained weighted Euclidean model seems to

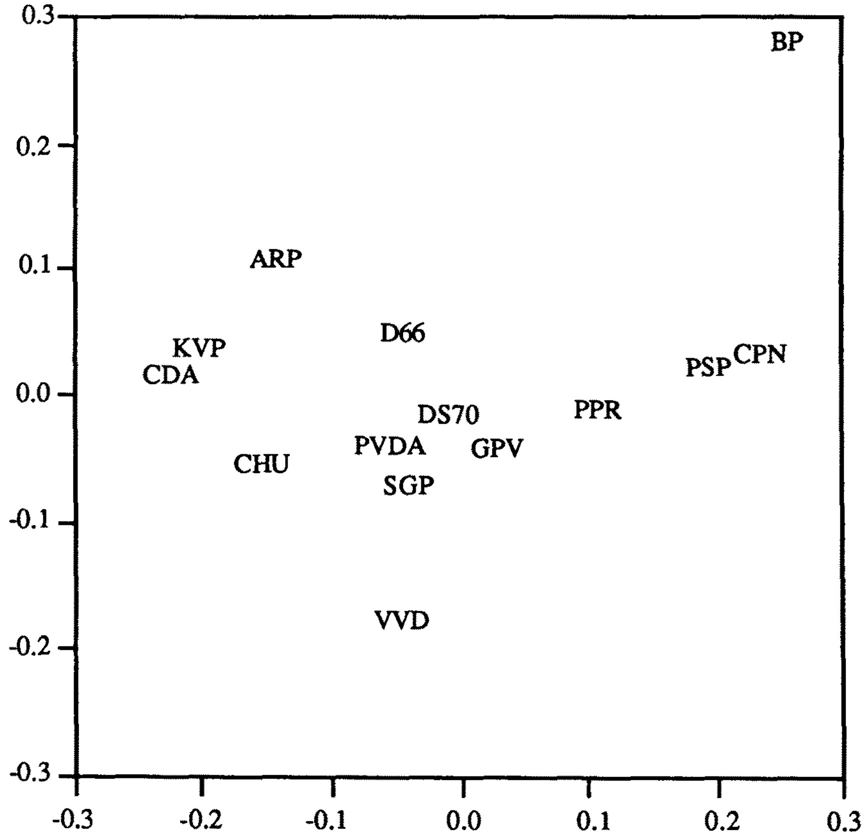

FIGURE 10.
INDSCAL analysis of political sympathy scales: Parties in dimension 3 (vertical axis) versus dimension 2 (horizontal axis).

do quite well considering the fact that fewer parameters are fitted ($139 + 14 \times 4$ instead of $139 \times 3 + 14 \times 3$).

Are Multiple Points of View Really Different?

In an ordinary INDSCAL analysis, the distinction between different groups can be investigated afterwards on the basis of the weights pattern (Jones, 1983; Shiffman, Reynolds, & Young, 1981). When instead two or more different points of view have been obtained, their possible similarity can be investigated by the use of quadratic assignment procedures (see Hubert, 1987, for an extensive review). In the analysis of the sympathy scales, the similarity between the two spaces is captured in the correlation between the two distance vectors $\mathbf{d}(\mathbf{X}_1)$ and $\mathbf{d}(\mathbf{X}_2)$, which turns out to have a very small value of 0.042. To estimate the probability of a correlation as large or larger than 0.042 occurring by a random displacement of the points, the coordinates in $\mathbf{X}_2$ were permuted, with the number of random permutations set to 5000. The Monte Carlo distribution is given in Table 7; with a p-value of 0.314, there is no evidence of communality between the two viewpoints.

For the PVA of the political issues, a choice was made to define the similarity between the two point of view spaces on the similarity between the party points. In the separate spaces, the party points are the centroids of the MP's who belong to the same party. Because the marginal frequencies (the weights assigned to the centroids) are very

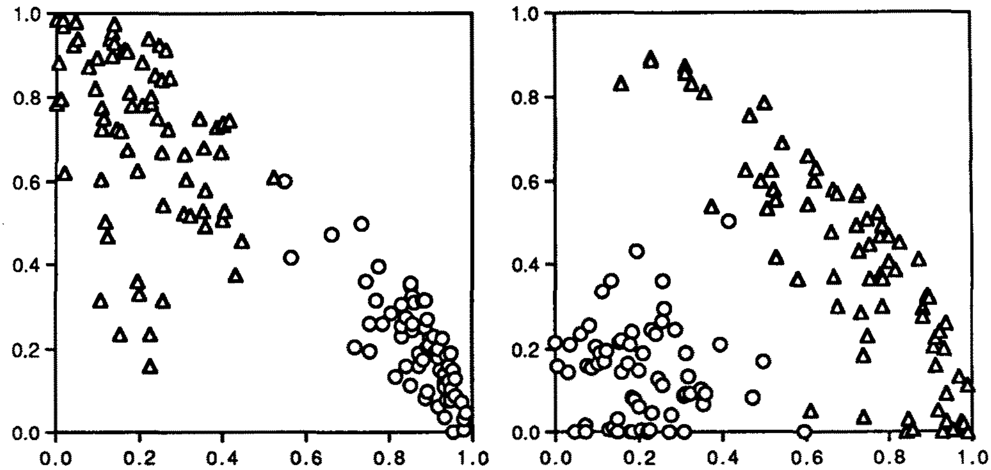

FIGURE 11.
INDSCAL analysis of political sympathy scales: On the left: weights in dimension 2 (vertical axis) versus dimension 1 (horizontal axis). On the right: weights in dimension 3 (vertical axis) versus dimension 2 (horizontal axis).

different, the correlation should take these different frequencies into account. When the centroids are given in $\mathbf{Y}_1 = \mathbf{M}^{-1}\mathbf{B}'\mathbf{X}_1$ with $\mathbf{Y}_2 = \mathbf{M}^{-1}\mathbf{B}'\mathbf{X}_2$, with $\mathbf{B}$ and $\mathbf{M}$ as defined in (9), the correlation is considered between $\mathbf{d}(\mathbf{BY}_1)$ and $\mathbf{d}(\mathbf{BY}_2)$. The observed correlation is 0.287; the Monte Carlo distribution of the grouped correlations is given in Table 8; the p-value is 0.109, so there is no strong evidence that the two sets of centroids are giving the same structure.

A Bootstrap Study of PVA and INDSCAL

To study the stability of both the PVA and the INDSCAL solutions, a bootstrap was performed, with 100 bootstrap samples for each technique, where the random sampling was done with respect to the $M = 139$ sources in the data. Table 9 summarizes the results for object points and dimensions. First, the sum of squared distances of each object point to the average of its corresponding 100 bootstrap points is given

TABLE 6
Diagnostics for PVA and INDSCAL by Partitioning into Groups × Dimensions Across Groups and Dimensions: Total Variance/M = 1 - STRESS

		Variance		SSQ Coordinates	Weights SSQ Unequal		Weights SSQ = 1.00	
Dimension		G1	G2	in X	G1	G2	G1	G2
	1	50.451	0.000	0.929	0.920	0.000	0.887	0.000
PVA	2	3.871	0.000	0.071	0.920	0.000	0.246	0.000
	3	0.000	50.638	0.855	0.000	0.855	0.000	0.820
	4	0.000	8.608	0.145	0.000	0.855	0.000	0.338
	1	48.949	4.344	0.383	1.405	0.330	0.870	0.204
INDSCAL	2	3.660	41.350	0.324	0.364	1.253	0.207	0.713
	3	2.137	16.627	0.135	0.389	1.078	0.143	0.396

TABLE 7
Monte Carlo Distribution of Grouped Correlations Between Distances in First Permutation Study
Sample size of 5000, Observed Correlation is 0.042; p-value = 0.314

Class			f	cf	f/N	cf/N
.641	-	.739	5	5	0.001	0.001
.491	-	.641	22	27	0.004	0.005
.342	-	.491	150	177	0.030	0.035
.192	-	.342	408	585	0.082	0.117
.042	-	.192	987	1572	0.197	0.314
−.108	-	.042	2194	3766	0.439	0.753
−.258	-	−.108	1230	4996	0.246	0.999
−.290	-	−.258	4	5000	0.001	1.000

(the variance of the two-dimensional coordinates for each object). The total of these within-bootstrap-group variances was also partitioned dimensionwise.

It turns out that object points are more stable in the points of view dimensions than in the INDSCAL dimensions, with the third INDSCAL dimension being especially unstable. Perhaps it should be concluded that the INDSCAL analysis should be restricted to two dimensions only; reanalyzing the data gives two dimensions that correlate almost perfectly with the first two dimensions of the three-dimensional solution. The stress is 0.182, which is virtually equal to the stress in PVA, while INDSCAL involves more parameters ($139 \times 2 + 14 \times 2$) than PVA ($139 + 14 \times 4$). Because the shape of the two-dimensional INDSCAL common space is completely spherical, it gives us no information about the different viewpoints that different groups of MP's have on the structure of the political parties in parliament (of course, the individual spaces will show the variation between individual MP's). For displaying group differences, we would need additional information, as provided in the PVA model.

6. Discussion

The primary purpose of this paper has been to show that the concept of points of view analysis is worthwhile for the analysis of heterogeneous sources on a group level.

TABLE 8
Monte Carlo Distribution of Grouped Correlations Between Distances in Second Permutation Study
Sample size of 5000, Observed Correlation is 0.287; p-value = 0.109

Class			f	cf	f/N	cf/N
.496	-	.683	79	79	0.016	0.016
.287	-	.496	466	545	0.093	0.109
.078	-	.287	1154	1699	0.231	0.340
−.131	-	.078	1834	3533	0.367	0.707
−.339	-	−.131	1386	4919	0.277	0.984
−.473	-	−.339	81	5000	0.016	1.000

TABLE 9
Bootstrap Study of Point of View Analysis and Weighted Euclidean Model (INDSCAL) 100 Bootstrap Analyses For Each Model

Sum of squared distances to average bootstrap coordinates				
	Point of views		INDSCAL dimensions	
Source	1	2	1,2	2,3
CPN	0.020	0.019	0.043	0.126
PSP	0.021	0.017	0.034	0.158
PPR	0.013	0.043	0.030	0.204
PvdA	0.011	0.031	0.021	0.024
D66	0.013	0.027	0.020	0.052
ARP	0.013	0.030	0.020	0.082
KVP	0.008	0.016	0.017	0.047
CHU	0.010	0.009	0.010	0.040
CDA	0.010	0.014	0.008	0.062
VVD	0.022	0.024	0.014	0.033
DS70	0.075	0.047	0.099	0.061
GPV	0.017	0.028	0.032	0.029
SGP	0.020	0.019	0.029	0.033
BP	0.006	0.021	0.039	0.163
Total	0.259	0.341	0.451	1.114
Dimension 1	0.096	0.144	0.127	0.324
Dimension 2	0.163	0.197	0.324	0.790

Since sources may be homogeneous within groups and heterogeneous between groups, the strength of the points of view analysis concept is in its parsimonious display of objects in $r \geq 2$ points of view.

Tucker and Messick's original procedure was called an individual differences model; we agree with Carroll and Chang (1970) that if the objective of analysis is individual differences scaling, application of the weighted Euclidean or INDSCAL model would be more appropriate, since individual viewpoints are displayed in separate spaces and each individual source has the possibility to be distinct from all other sources. As an analysis for groups, the INDSCAL common space may not be the best way to display their differences; in fact, the common space does not need to fit any group or individual. More restricted INDSCAL models have been proposed, requiring, for example, that each individual source may use only $t < p$ dimensions, where p denotes the dimensionality of the common space. In this line of thought, the PVA procedure proposed in this paper is more restricted: sources that are found to be a homogeneous group must use exactly the same t dimensions (with dimension weights that are proportional). A mixture of the two models (the weighted Euclidean model within different points of view) remains a topic for further research.

Kiers (1989) discusses the relationship between various approaches to three-way scaling from a different perspective. He notes there is resemblance in lay-out between Tucker and Messick's original approach and the French method STATIS (based on Escoufier, 1973), but finds them clearly different in several respects. From our perspective, the similarity is more obvious. Both deal with (dis)similarity matrices, treated as variables, from which linear combinations are formed that are subsequently subjected to a secondary analysis. A basic problem in both the Tucker and Messick procedure and STATIS is how to avoid negative weights in a subsequent linear combination when the first composite matrix has been taken out (a problem that is not encountered in the procedure proposed in this paper). A related approach applied directly to given variables is in Escoufier (1988), where it is proposed to find non-overlapping subsets of variables to obtain different composites, which is a similar objective pursued by imposing the particular constraints on the weights in our procedure.

There is also a relationship with what is called "homogeneity analysis as a first step" in Gifi (1990, chap. 3). Here, categorical variables are quantified in r different ways. Next, these r sets of optimally scaled variables are used to obtain r principal components solutions in p_r dimensions. Related work on quantifying categorical variables in a three-way framework is Saporta (1975), and the extensive review in Kiers (1989).

The procedure described in this paper could be generalized to allow sources to be assigned to t points of view, where $1 \leq t < r$ (giving overlapping clusters). When variables are optimally transformed and assigned to more than one point of view, one has the choice between identical or possibly different optimal transformations with respect to the different points of view. This extension, however, needs further investigation with respect to its data analytical merits. Another possible extension would allow the given matrices $\mathbf{\Delta}_m$ to be asymmetric.

In the procedure described, the (dissimilarity) variables were assumed to constitute different groups for which points of view were fitted. Algebraically, the fitting procedure involved a differential aggregation over homogeneous sources that gave different composite matrices that were optimal with respect to the Euclidean distances to be fitted. It is important to realize that the basic idea of differential aggregation is very general, and can be applied to any other multivariate analysis technique, either in a two-way or three-way framework.

The present technique gives overall dissimilarity measures in $\mathbf{\Theta}_s$, the result of the optimal aggregation over a group of homogeneous sources. When the sources $\mathbf{\Delta}_m^*$ are homogeneous at the outset, we choose s equal to 1, and our procedure will be identical to an analysis with the multidimensional scaling program KYST. Other work in the area of finding a composite matrix is given in Escoufier (1980) and Gower (1971). A related procedure of differential weighting of variables to find an optimal representation is found in De Soete, DeSarbo, and Carroll (1985), who simultaneously estimate variable importance weights and the corresponding ultrametric tree. The procedure proposed in the present paper combines aspects of the work just referenced.

The individual dissimilarity sources in the analysis may be given directly or may be derived. The latter may be from numerical variables, ordinal variables (for which a monotonic transformation is fitted) or nominal variables, in which only unordered classes of observations are considered. Unordered classes can either be represented through new scale values, preserving the class structure, that give a single transformation of the categorical variables, or by a more-dimensional representation in the form of class-centroids of object points.

The PVA procedure finds clusters of sources, and weights the sources differentially. As a bonus, it finds composite dissimilarity matrices. Although a composite matrix $\mathbf{\Theta}_s$ is optimal for least squares Euclidean distance fitting, it could also be analyzed afterwards by other techniques, for example by a cluster analysis.

References

Bloxom, B. (1978). Constrained multidimensional scaling in N spaces. *Psychometrika, 43*, 397–408.

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

Cliff, N. (1968). The "idealized individual" interpretation of individual differences in multidimensional scaling. *Psychometrika, 33*, 225–232.

de Leeuw, J. (1988). Convergence of the majorization algorithm for multidimensional scaling. *Journal of Classification, 5*, 163–180.

de Leeuw, J., & Heiser, W. J. (1980). Multidimensional scaling with restrictions on the configuration. In P. R. Krisnaiah (Ed.), *Multivariate analysis, Vol. V* (pp. 501–522). Amsterdam: North-Holland.

De Soete, G., DeSarbo, W. S., & Carroll, J. D. (1985). Optimal variable weighting for hierarchical clustering: An alternative least-squares algorithm. *Journal of Classification, 2*, 173–192.

Escoufier, Y. (1973). Le traitement des variables vectorielles [The treatment of vector-valued variables]. *Biometrics, 29*, 751–760.

Escoufier, Y. (1980). Exploratory data analysis when data are matrices. In K. Matusita (Ed.), *Recent developments in statistical inference and data analysis* (pp. 45–53). Amsterdam: North-Holland.

Escoufier, Y. (1988). Beyond correspondence analysis. In H. H. Bock (Ed.), *Classification and related methods of data analysis* (pp. 505–514). Amsterdam: North-Holland.

Gifi, A. (1990). *Nonlinear multivariate analysis*. Chichester: Wiley. (First edition 1981, Leiden: The University of Leiden, Department of Data Theory)

Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics, 27*, 857–871.

Guttman, L. (1968). A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika, 33*, 469–506.

Heiser, W. J., & Stoop, I. (1986). *Explicit SMACOF algorithms for individual differences scaling* (Research Report RR-86-14). Leiden: Department of Data Theory.

Horan, C. B. (1969). Multidimensional scaling: Combining observations when individuals have different perceptual structures. *Psychometrika, 34*, 139–165.

Hubert, L. J. (1987). *Assignment methods in combinatorial data analysis*. New York: Marcel Dekker.

Jones, C. L. (1983). A note on the use of directional statistics in weighted Euclidean distances multidimensional scaling models. *Psychometrika, 48*, 473–476.

Kiers, H. A. L. (1989). *Three-way methods for the analysis of qualitative and quantitative two-way data.* Leiden: DSWO Press.

Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika, 29*, 1–28.

Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychometrika, 29*, 115–129.

Kruskal, J. B., Young, F. W., & Seery, J. B. (1973). *How to use KYST, a very flexible program to do multidimensional scaling and unfolding*. Murray Hill: Bell Laboratories.

Meulman, J. J. (1986). *A distance approach to nonlinear multivariate analysis*. Leiden: DSWO Press.

Meulman, J. J. (1992). The integration of multidimensional scaling and multivariate analysis with optimal transformations. *Psychometrika, 57*, 539–565.

Ramsay, J. O. (1982a). Some statistical approaches to multidimensional scaling data. *Journal of the Royal Statistical Society, Series A, 145*, 285–312.

Ramsay, J. O. (1982b). *MULTISCALE II manual*. Montréal: McGill University, Department of Psychology.

Ramsay, J. O. (1989). Monotone regression splines in action. *Statistical Science, 4*, 425–441.

Ross, J. (1966). A remark on Tucker and Messick's "Points of view" analysis. *Psychometrika, 31*, 27–32.

Saporta, G. (1975). Pondération optimale de variables qualitatives en analyse des données [Optimal weighting of qualitative variables in 'analyse des données']. *Statistique et Analyse des Données, 3*, 19–31.

Shepard, R. N. (1962a). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika, 27*, 125–140.

Shepard, R. N. (1962b). The analysis of proximities: Multidimensional scaling with an unknown distance function. II. *Psychometrika, 27*, 219–246.

Schiffman, S. S., Reynolds, M. L., & Young, F. W. (1981). *Introduction to multidimensional scaling*. New York: Academic Press.

Tucker, L. R. (1951). *A method of synthesis of factor analysis studies* (Personnel Research Section Report No. 984). Washington, D. C.: Department of the Army.

Tucker, L. R., & Messick, S. (1963). An individual differences model for multidimensional scaling. *Psychometrika, 28*, 333–367.

Verboon, P., & van der Kloot, W. A. (1989). *On the relation between INDSCAL and nonlinear points of view analysis* (Research Report PRM-89-02). Leiden: University of Leiden, Department of Psychometrics.

Winsberg, S., & De Soete, G. (in press). *A latent class approach to fitting the INDSCAL model. Psychometrika.*

Winsberg, S., & Ramsay, J. O. (1983). Monotone spline transformations for dimension reduction. *Psychometrika, 48*, 575–595.

Zegers, F. E., & ten Berge, J. M. F. (1985). A family of association coefficients for metric scales. *Psychometrika, 50*, 17–24.

File received 5/14/91
Final version received 4/6/92