MATHÉMATIQUES ET SCIENCES HUMAINES

W. J. HEISER
J. DE LEEUW

Multidimensional mapping of preference data

Mathématiques et sciences humaines, tome 73 (1981), p. 39-96
<http://www.numdam.org/item?id=MSH_1981__73__39_0>

Math. Sci. hum. (*19*^{e} année, n°73, *1981*, p.39-96)

MULTIDIMENSIONAL MAPPING OF PREFERENCE DATA

W.J. HEISER*
J. DE LEEUW*

I. INTRODUCTION

The analysis of preferential choice data has attracted the attention of methodologists in the social sciences for a long time. The *classical approach*, starting off from the work of Fechner on experimental esthetics (Fechner, 1871), and formulated as a theory of choice by Thurstone in his famous Law of Comparative Judgment (Thurstone, 1927, 1959), involves the assumption of an unidimensional utility continuum and normal distributions of utilities. Pairwise choice frequencies are then accounted for in terms of properties of the utility distributions: means, standard deviations and correlations.

A rigorous statistical treatment of Thurstonean scaling is given by Bock and Jones (1968). Further interesting developments in the theory of individual choice behaviour were made by Luce (1959, 1977), Tversky (1972) and Fishburn (1974); for an analysis of social choice behaviour and many related topics, see Arrow (1951) and Davis, De Groot and Hinich (1972). An authorative review of the method of Paired Comparisons is David (1963) and, more recently, Bradley (1976); an up-to-date bibliography on this method is given by Davidson and Farquhar (1976). These 'statistical' approaches will not interest us here, however. Instead, we will focus upon 'data-analytic' approaches that have been advocated in recent years. They are multidimensional in nature and emphasize the graphical display of data.

The first approach we will discuss is very much in line with Thurstonean theory. In fact, it starts from the general Law of Comparative Judgment

*Department of Datatheory, University of Leiden, Holland

but avoids the customary restrictive case V assumptions. It then aims at a *multidimensional analysis of the comparatal dispersions*. The second approach is made up of what we call *decomposition techniques*. Here we assume transitivity for each subject and a 'latent' cognitive or evaluative structure, common to all subjects. The individual utilities are then decomposed into the common structure and a set of points or vectors which represent the subjects. Finally, a third class of techniques will be discussed which tries to map the individual utilities into a known common structure straight away. We will call these *projection techniques* (Carroll (1972) uses the terms *internal* and *external* analysis for decomposition and projection respectively, but these terms do not clarify much the completely different nature of the techniques involved).

II. SOME TERMINOLOGY AND NOTATION

The several kinds of data that will be considered in the next sections can all be thought of to be derived from (or actually computed from) the central three-way datablock in figure 1.

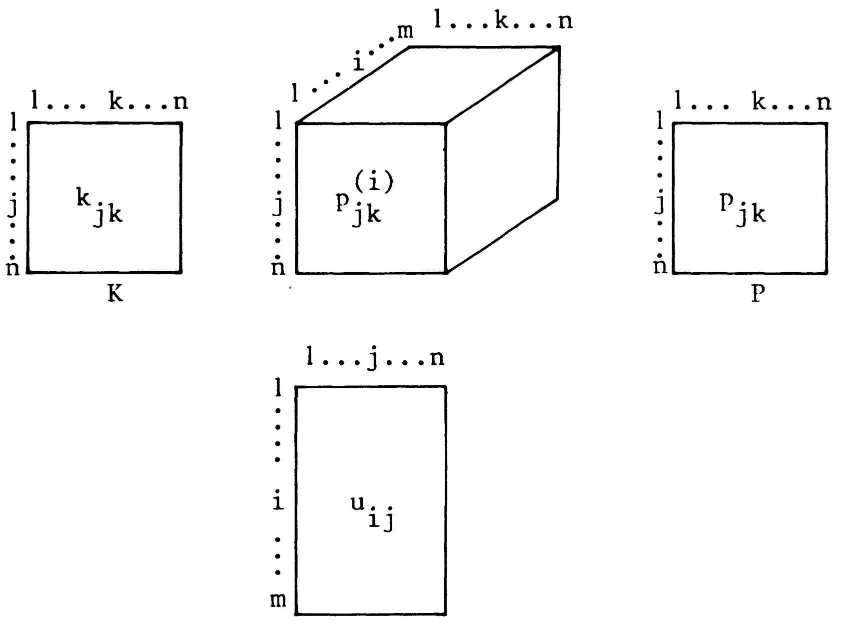

Figure 1. Central three-way datablock with three derived matrices

For convenience, we will interpret the general element $p_{jk}^{(i)}$ to be the *preference strength* of subject i with which he prefers object j to object k.

The collection of objects may be anything: odours, crimes, concepts, political parties, commodity bundles, persons etc. Moreover, we use the word preference in its broadest sense: any judgment or behaviour which indicates that, according to subject i, object j is nicer than, wilder than, heavier than, or more valuable than object k will be suited for our analyses. And of course, when we say that the 'third way' pertains to subjects, we do not want to imply that this couldn't be replications, occasions, conditions, groups or any other datasource.

So, we consider a set of n objects, a set of m subjects and a measure of preference strength $p_{jk}^{(i)}$. In many applications, the experimental set-up calls for a *preferential choice*, so that the individual $p_{jk}^{(i)}$ are dichotomous variables, simply indicating whether or not subject i prefers object j to object k. Sometimes however, we want to incorporate indifference judgments (*trichotomous* case) or quantitative measures of preference strength (*graded pair comparisons*). The marginal table P is defined as follows:

$$p_{jk} = \frac{1}{m} \sum_{i=1}^{m} p_{jk}^{(i)} \qquad (1)$$

So P is simply the mean preference strength and is the usual input to a Thurstonean or Bradley-Terry-Luce analysis. No attempt will be made to describe these procedures here in detail, since they are well documented and summarized elsewhere (Thurstone (1927), Mosteller (1951), Luce (1959), Bradley and Terry (1952), David (1963), Torgerson (1958), Bock and Jones (1968)). For a treatment of the trichotomous case, see Glenn and David (1960) and Greenberg (1965); for an analysis-of-variance approach to graded pair comparisons, see Scheffé (1952) or Bechtel (1976).

We shall discuss two types of generalized Thurstonean analysis in section III. There we also need the matrix K, which contains the so-called *comparatal dispersions*. The remainder of the paper will be devoted to techniques to analyse the matrix U, defined as follows:

$$u_{ij} = \frac{1}{n} \sum_{k=1}^{n} (p_{jk}^{(i)} - p_{kj}^{(i)}) \qquad (2)$$

Thus u_{ij} indicates to what extent subject i prefers j over the other objects (centered at zero). The table U will be referred to as the matrix of *utilities* and the u_{ij} as the *individual utilities* (these are sometimes called preference ratings or individual (affective) values). The fact that we reduce the $p_{jk}^{(i)}$ to u_{ij} does not imply that we are unwilling to accept intransitive choices; we just don't "model" them. If intransitivity is around, it will introduce ties or a decrease of variance in the rows of U. Of course, the individual utilities may be collected by any ordering or rating scale method right from the start. Transitivity is assured then and we might reconstruct the other matrices by the rule

$$p_{jk}^{(i)} = F(u_{ij} - u_{ik}) \;, \tag{3}$$

where F is a suitably chosen monotone increasing function.

III. MULTIDIMENSIONAL ANALYSIS OF COMPARATAL DISPERSIONS

Our concern with the multidimensional analysis of a pair comparison matrix is motivated by two lines of thought: in the first place, there is the theoretical argument that choice models should incorporate parameters which account for 'ambiguity' or 'similarity' of the objects, apart from their 'popularity'. This goes back to Thurstone's earlier work (for a specific discussion, see Thurstone (1945)) and was taken up by Sjöberg and his collaborators. The second, more data-analytical argument is that people are used to analyse correlation- and dissimilarity matrices with a lot more parameters than they tend to do with pair comparison matrices, without very compelling reasons to do so. This last type of motivation, including an alternative derivation of one of the models to be discussed here, may be found in Carroll (1980). But first, we focus on Thurstone and Sjöberg.

III.1. Why comparatal dispersions?

According to the Thurstonean model for pairwise choices, the set of objects corresponds with a *multivariate discriminal process* $\{U_1, U_2, \ldots , U_n\}$, with parameters μ_j (j = 1, ... , n) and σ_{jk} (j,k = 1, ... , n). Thus it is assumed that any two choice objects, j and k, give rise to partly overlapping normal utility distributions (see figure 2).

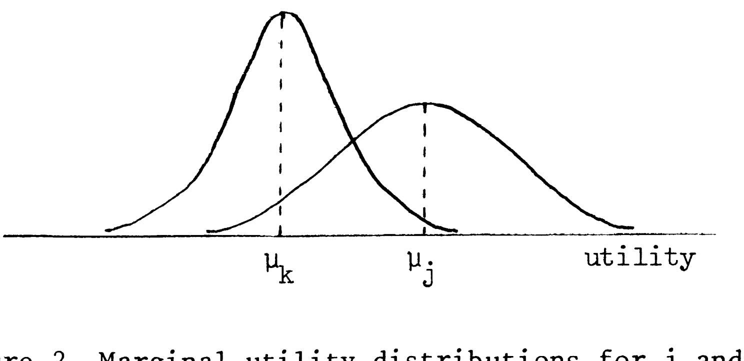

Figure 2. Marginal utility distributions for j and k.

Assume that, in a pairwise choice, the object with the larger utility is always preferred; more precisely, if j and k are compared, the subject samples from the process and prefers j to k if $U_j > U_k$. By standard statistical results, the difference process $\{U_j - U_k\}$ will be normally distributed with mean $\mu_j - \mu_k$ and variance

$$\kappa^2_{jk} = \sigma^2_j + \sigma^2_k - 2\sigma_{jk} \tag{4}$$

Following Gulliksen (1958), we call the κ_{jk} *comparatal dispersions* (the σ_j are called *discriminal dispersions*). Let the probability that the utility of j is larger than the utility of k be denoted by π_{jk}, then

$$\pi_{jk} = \Phi(\frac{\mu_j - \mu_k}{\kappa_{jk}}) , \tag{5}$$

where Φ is the univariate standard normal distribution function. Furthermore if p_{jk} estimates π_{jk} and $z_{jk} = \Phi^{-1}(p_{jk})$ is the corresponding unit normal deviate, we get

$$z_{jk} = \frac{\mu_j - \mu_k}{\kappa_{jk}} . \tag{6}$$

This is *Thurstone's Law of Comparative Judgment* (Thurstone (1927)). A basic difficulty in the model is that there are too many unknowns. We may attempt to resolve this difficulty in at least two different ways:

a. by imposing *restrictions on the parameters*, such as that all comparatal dispersions are equal (case V), or that all covariances are equal and the variances are almost equal (case IV).
b. by deriving *more equations* accounting for the same experimental data (using tetrachoric correlations between pairs) or slightly different equations accounting for slightly different experimental data (category judgments of size of difference).

The first approach is by far the most popular. In its usual form, however, it has two major drawbacks. As Mosteller (1951) and Torgerson (1958) have pointed out, statistical tests for the goodness-of-fit of these highly restricted models are insensitive to violations of the assumption of equal comparatal dispersions. Unequal dispersions may or may not cause high chi-square values. So we have nothing to evaluate the seriousness of faulty assumptions. In the second place, we might have theoretical and practical reasons to be interested in the comparatal dispersions themselves. This position has recently been advocated strongly in the work of Sjöberg (1975).

Remember that the probability of choosing one object over the other is a function of both the mean difference in utility and the standard deviation of the differences. Now, consider the distributions of utility differences $\{U_j - U_k\}$ in figure 3. In 3a, the mean utility difference $\mu_j - \mu_k$ is posi-

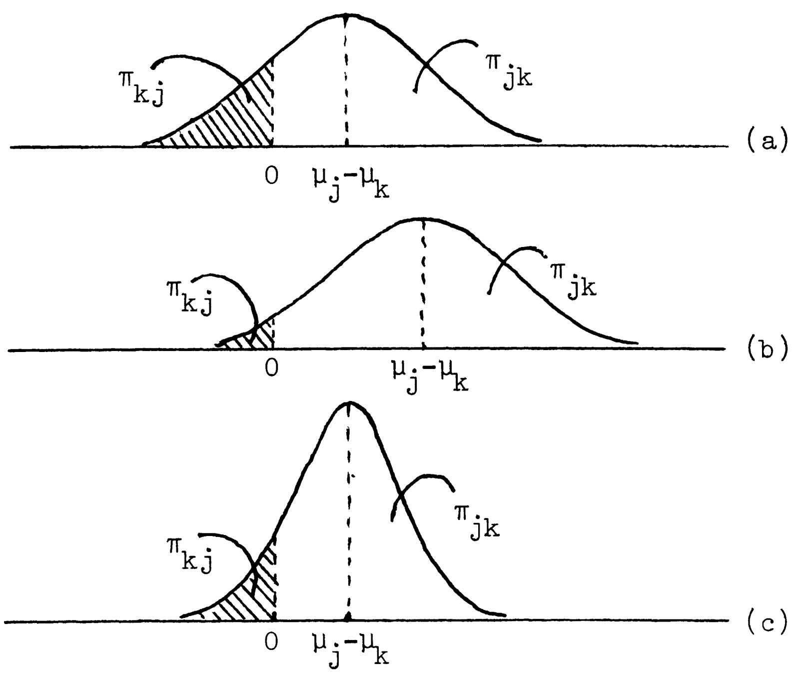

Figure 3. Three distributions of utility differences.

tive and the proportion of minority votes π_{kj} corresponds to the shaded area, the proportion of majority votes to the unshaded part. There are *two quite different mechanisms* whereby the relation between minority and majority votes may be changed. The first is illustrated in figure 3b. By a change in the mean utility difference the proportion of minority votes has decreased. Here the model simply says that the more popular an object gets, the more votes it will obtain. This is so close to common intuition that we can not expect to learn many qualitative new things from a case V analysis alone (nor could we from the various alternatives that have been proposed, which assume different distribution functions but stick to unidimensionality and, by the way, arrive at virtually indistinguishable estimates of the utilities, cf. Mosteller (1958) and Noether (1960)).

Next consider figure 3c. Again the proportion of minority votes has decreased, but for a completely different reason. The mean utility difference has remained the same, whereas the variance of the distribution has diminished. If we want to understand this effect, one possibility is to assume

$$\kappa^2_{jk} = \sigma^2_j + \sigma^2_k , \tag{7}$$

i.e., the usual case III assumption of zero correlations between discriminal processes. A possible interpretation of σ^2_j and σ^2_k is in terms of ambiguity and the model now says that the object in the majority will gain votes from a decreased ambiguity of one or both objects, whereas the object in the minority will lose by it. Certainly, this is only one possibility; we will treat others later. It is clear that we get a richer theory of preference behaviour if we do not restrict the more interesting parameters in the model so heavily.

The second approach is not using restrictions but deriving more equations. This is exemplified by the work of Sjöberg. At first, he suggested that the tetrachoric correlations between pairs contain information about the correlations between utility distributions (Sjöberg (1962)). Although it was found that they do give some useful information, Sjöberg (1967) remarks that his methods are cumbersome to use even with a moderate number of objects. He therefore switched over to an analysis of graded pair comparisons, which require the subject to give a response richer in information, and proposed a method which utilizes this increased information to obtain estimates of the comparatal dispersions up to a constant. In the next subsection we will review Sjöberg's method and some of his empirical findings. After that, we will discuss a new method which uses the restriction approach again (but with a more general class of restrictions).

III.2. Comparatal dispersions and similarity.

The procedure proposed by Sjöberg (1967) calls for preference ratings on all possible pairs of objects. It is not assumed that $p^{(i)}_{jk}$ and $p^{(i)}_{kj}$ add up to some constant for all j and k, as is usually done; this would ruin the possibility to estimate the dispersions. The experimental set-up typically runs as follows. Subjects are instructed to consider for each pair first which object they prefer. Then they are asked to check to what extent they like the chosen object better. A possibility of checking a category of equal preference is mostly provided. So, if seven categories of size of difference are used, the subject is asked to mark one of the figures in the string

j | 7 - 6 - 5 - 4 - 3 - 2 - 1 - 0 - 1 - 2 - 3 - 4 - 5 - 6 - 7 | k

If the raw data are collected in the matrix $\{p_{jk}^{(i)}\}$, then we may derive in this case several matrices $\{_{\ell}p_{jk}\}$ which indicate the proportion of times that j is preferred to k at least the amount t_{ℓ} (ℓ = 1, ... , r). The number of *'threshold' parameters* t_{ℓ} may be chosen in accordance with the presumed judgment accuracy of the subjects (in the example above, r = 7). To get a

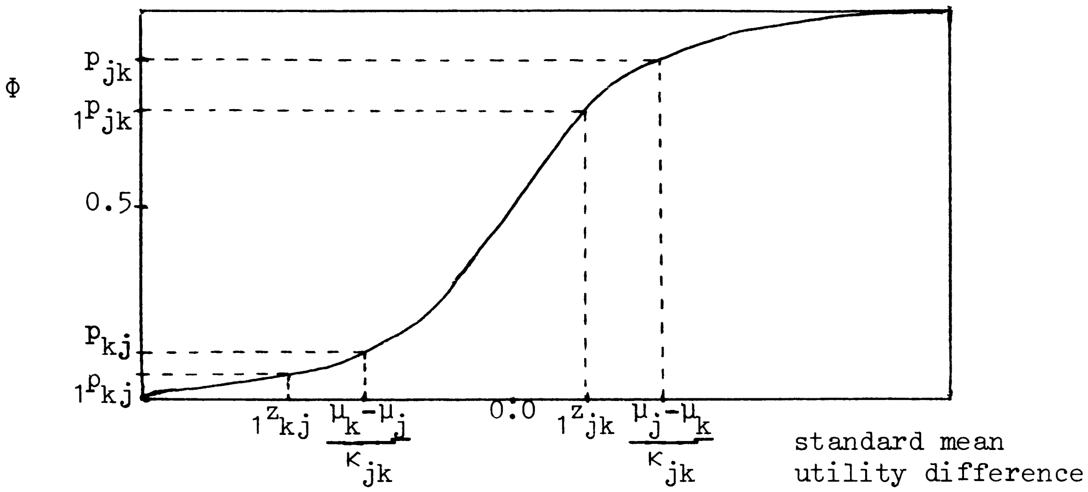

Figure 4. The effect of a threshold parameter.

rough idea about the way the method works, we will consider the case r = 1 (this is analogous to trichotomous pair comparison data, it allows for the judgments j > k, k > j and indifference).

Consider figure 4. Here we have the unit normal cumulative distribution function; on the y-axis we have the usual proportions p_{jk} and p_{kj}, which are symmetric around 0.5, and on the x-axis we have their corresponding unit normal deviates $z_{jk} = (\mu_j - \mu_k)/\kappa_{jk}$ and $z_{kj} = (\mu_k - \mu_j)/\kappa_{jk}$, which are symmetric around 0.0. The key assumption is, that the effect of the threshold will be to decrease all utility differences $U_j - U_k$ by an amount t (this implies that utility differences have to be bigger to produce the same proportion of preference votes). This decrease doesn't affect the variance of the utility differences, but it does affect their mean. For the new normal deviates we get

$$_tz_{jk} = \frac{\mu_j - \mu_k - t}{\kappa_{jk}} , \tag{8a}$$

$$_tz_{kj} = \frac{\mu_k - \mu_j - t}{\kappa_{jk}} \tag{8b}$$

Subtracting (8b) from (8a) gives us:

$$t^{z}_{jk} - t^{z}_{kj} = \frac{2(\mu_j - \mu_k)}{\kappa_{jk}} \quad , \tag{9}$$

adding (8a) and (8b) gives

$$t^{z}_{jk} + t^{z}_{kj} = \frac{-2t}{\kappa_{jk}} \tag{10}$$

For identification purposes, we may set t = 1 and estimate the comparatal dispersions by

$$\mathrm{est}(\kappa_{jk}) = \frac{-\ 2}{t^{z}_{jk} + t^{z}_{kj}} \quad , \tag{11}$$

and the mean utility differences by

$$\mathrm{est}(\mu_j - \mu_k) = \frac{t^{z}_{kj} - t^{z}_{jk}}{t^{z}_{jk} + t^{z}_{kj}} \tag{12}$$

Once the differences are known, it is a routine matter to find the mean utilities themselves. In the general case, we add r parameters and find that the number of equations has been multiplied by 2r. This gives us a strongly overdetermined system, which is solvable by standard methods.

In line with the view that the estimation of comparatal dispersions wouldn't be of much theoretical interest if we couldn't connect them with other characteristics of the choice objects, Sjöberg and his collaborators (Sjöberg (1975a,b), Sjöberg and Capozza (1975), Franzén, Nordmark and Sjöberg (1972)) sought empirical evidence for the conjecture that *correlation between utility distributions correspond to rated subjective similarity*. This notion is motivated by the general argument that two objects which are considered to be very similar by many people are often found to be correlated in many attributes, so their utility distributions should be correlated too. Similarity judgments in some form are taken as a basic approach to finding a *'cognitive map'*, which in turn is supposed to influence the preferential choice process.

In the studies cited above, the estimated comparatal dispersions are taken to be "inversely related to the correlations". So instead of the case III assumption of constant covariances, as in (7), it seems that constant variances are assumed. I.e., if, according to (4),

$$\kappa^2_{jk} = \sigma^2_j + \sigma^2_k - 2\sigma_j\sigma_k\rho_{jk} \quad , \tag{13}$$

where ρ_{jk} is the correlation between the utility distributions of j and k, then some assumption like

$$\sigma_j^2 = \sigma_k^2 = \sigma^2 \tag{14}$$

is necessary to arrive at a (linear) inverse relation

$$\kappa_{jk}^2 = c(1 - \rho_{jk}) , \tag{15}$$

for some constant c. However, the way in which Sjöberg e.a. try to verify the conjecture suggests an alternative reparametrization of the κ_{jk}'s. For, they perform a multidimensional scaling analysis both on the estimated κ_{jk}'s and on the rated similarities, say s_{jk}. This means that they try to represent the choice objects as points in p-space, in such a way that small interpoint distances correspond to small comparatal dispersions (cq. large similarities), and larger distances correspond to larger dispersions (cq. smaller similarities). If Y is the p-dimensional cognitive map derived from the similarities data and X the p-dimensional representation of the choice objects derived from the comparatal dispersions, then the conjecture may be stated as "X = Y".

The alternative reparametrization thus would be, to write the dispersions as a function of X. This is possible, because the variance-covariance matrix $\{\sigma_{jk}\}$ may be decomposed into the form

$$\sigma_{jk} = \sum_{a=1}^{n} x_{ja}x_{ka} , \tag{16}$$

as can be done with any positive semi-definite matrix, and therefore

$$\kappa_{jk}^2 = \sum_{a=1}^{n} x_{ja}^2 + \sum_{a=1}^{n} x_{ka}^2 - 2\sum_{a=1}^{n} x_{ja}x_{ka} = \sum_{a=1}^{n} (x_{ja} - x_{ka})^2 = d_{jk}^2(X) \tag{17}$$

So the comparatal dispersions may be interpreted as distances between points in n-space. This new system is still unrestrictive, but as usual we throw away n - p dimensions. This means that that we replace the $\frac{1}{2}n(n-1)$ parameters κ_{jk}^2 by the np parameters x_{ja}. Thus the comparatal dispersions are accounted for by a p-dimensional representation X, which should resemble the p-dimensional cognitive map Y obtained from other data.

As an illustration, we take a study of political preference in Italy by Sjöberg and Capozza (1975). The choice objects were the seven political parties listed in table 1.

1:	PCI	(communist party)
2:	PSI	(socialist party)
3:	PSDI	(social democratic party)
4:	PRI	(republican party)
5:	DC	(christian democratic party)
6:	PLI	(liberal party)
7:	MSI	(national right wing)

Table 1. Choice objects from Sjöberg and Capozza (1975).

The subjects were 195 students of the university of Padua, Italy. The relevant experimental tasks were similarity rating on a seven-category rating scale and preference rating on a fifteen-category rating scale, both for all pairs of parties. The estimated comparatal dispersions are given in table 2.

	PCI	PSI	PSDI	PRI	DC	PLI	MSI
PCI	-	1.78	2.03	2.04	2.64	2.37	2.05
PSI	1.78	-	1.50	1.45	1.90	1.97	1.91
PSDI	2.03	1.50	-	1.10	1.21	1.36	1.47
PRI	2.04	1.45	1.10	-	1.16	1.05	1.12
DC	2.64	1.90	1.21	1.16	-	1.07	1.22
PLI	2.37	1.97	1.36	1.05	1.07	-	0.79
MSI	2.05	1.91	1.47	1.12	1.22	0.79	-

Table 2. Standard deviations of utility differences, total group (Sjöberg and Capozza, 1975).

The derived two-dimensional structure is given in figure 5a (the multidimensional scaling program TORSCA was used), and the two-dimensional cognitive map derived from the mean similarity judgments in figure 5b. Clearly, the

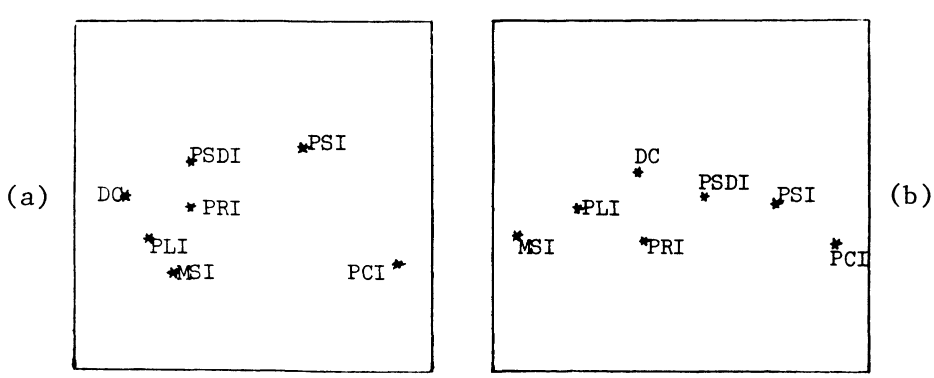

Figure 5. Two derived structures from Italian data (expl. see text).

the two structures are globally the same, as conjectured, although figure 5a appears to be a bit more 'bended' version of the usual political left-right dimension emerging in figure 5b.

III.3. A new formulation: restricted multidimensional scaling.

We now return to the restriction approach for the general Thurstonean model, using (17) or a similar assumption. We suppose the z_{jk} are given numbers satisfying $z_{jk} = - z_{kj}$, and we want to fit the model (cf. (6)):

$$z_{jk} = \frac{\mu_j - \mu_k}{d_{jk}(X)} , \tag{18}$$

where the d_{jk} are euclidean distances defined on the rows of X. Such a parametrization implies that the restrictions should be imposed on the x_{ja}, instead of directly on the σ_{jk}. This gives us the advantage that we obtain a much broader class of models compared with the classical 'cases'. In the first place, if X is any n × p matrix, we have a case very similar to the one in the last section and a model which is essentially equivalent to Carroll's (1980) 'wandering vector' model. Furthermore, if X is n × n and diagonal, we get

$$d^2_{jk}(X) = x^2_{jj} + x^2_{kk} , \tag{19}$$

corresponding to the usual case III assumption. And if we restrict X to be of the form

$$X = \begin{array}{c} \begin{matrix} 1 \ldots p & p+1 & \ldots & n+p \end{matrix} \\ \left[\begin{array}{c|ccccc} & x_1 & & & & \\ & & x_2 & & & \\ & & & \ddots & & \\ X_c & & & & x_j & \\ & & & & & \ddots \\ & & & & & x_n \end{array} \right] \end{array} \tag{20}$$

we obtain a model in which the comparatal dispersions are associated with both 'common' (X_c) and 'unique' (x_j) dimensions, i.e. something comparable with the factor-analytic model. With X of the form (21) we obtain a simplex model, different from the one proposed by Bloxom (1972), who defines an order relation on the utilities. In our case, if X has pattern (21) then the matrix of squared distances exhibits the pattern (22), for n = 4; i.e., if we move successively further from the main diagonal, the comparatal dispersions increase. Note that this may be a reasonable alternative for the

$$X = \begin{array}{c} \begin{array}{ccccccccc} 1 & & \cdot & \cdot & \cdot & \cdot & & n-1 & \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ x_1 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ x_1 & x_2 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ x_1 & x_2 & x_3 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & x_4 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & x_{j-1} & & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_1 & x_2 & x_3 & \cdot & \cdot & \cdot & \cdot & \cdot & x_{n-1} \end{bmatrix} \end{array} \tag{21}$$

$$\begin{bmatrix} 0 & x_1^2 & x_1^2 + x_2^2 & x_1^2 + x_2^2 + x_3^2 \\ x_1^2 & 0 & x_2^2 & x_2^2 + x_3^2 \\ x_1^2 + x_2^2 & x_2^2 & 0 & x_3^2 \\ x_1^2 + x_2^2 + x_3^2 & x_2^2 + x_3^2 & x_3^2 & 0 \end{bmatrix} \tag{22}$$

case III or the case VI assumption (cf. Bock and Jones (1968), p.65)

Many more special structures may be imposed on X, giving just as many 'cases' for Thurstone's Law. To fit this family of cases, we will first symmetrize (18) by taking absolute values; we define

$$\lambda_{jk}(\mu) = |\mu_j - \mu_k| \ , \tag{23}$$

i.e., $\lambda_{jk}(\mu)$ is the distance between the mean values on the utility continuum. Furthermore, we define a second transformation of the data:

$$v_{jk} = |z_{jk}| = |\Phi^{-1}(p_{jk})| \ . \tag{24}$$

This gives us the transformed model of preference strength

$$v_{jk} = \frac{\lambda_{jk}(\mu)}{d_{jk}(X)} \tag{25}$$

which still incorporates the two basic mechanisms mentioned in section III.1 Note that the choice objects are associated with two sets of parameters: μ and X. Increase in distance on the utility continuum (involving μ) heightens the preference strength, whereas increase of distance on the cognitive map (involving X) lowers it. A loose way of stating the relationship between these mechanisms is: *incomparables tend to be confused, even though their utilities may differ a lot, and things that are alike tend to be contrasted if one has to choose between them.*

The direction of preference strength only involves μ (we have to be a bit careful here, because the 'direction' of the one-dimensional continuum implied by (23) is not determined; mostly, however, a quick look at the end-points will suffice to identify the 'good' and the 'bad' side). For estimation purposes, we employ the least squares loss function

$$L(X,\mu) = \sum_{j=1}^{n} \sum_{k=1}^{n} (v_{jk}d_{jk}(X) - \lambda_{jk}(\mu))^2 \tag{26a}$$

which may also be written as

$$L(X,\mu) = \sum_{j=1}^{n} \sum_{k=1}^{n} v_{jk}^2(d_{jk}(X) - \frac{\lambda_{jk}(\mu)}{v_{jk}})^2 \tag{26b}$$

This is a function of two sets of parameters and we can use the *Alternating Least Squares (ALS) principle* to minimize it. The ALS principle is a general rule to tackle least squares problems. It says that we first have to find a partition of the total set of parameters into 'nice' subsets, such that the minimization of the loss function over each subset alone, with the remaining parameters regarded as fixed, is relatively simple. Then we may cycle through a series of simple least squares subproblems and repeat that process until convergence. For a general discussion of ALS in a somewhat different context, see de Leeuw, Young and Takane (1976).

In our case, the ALS principle suggests that we must alternate the minimization of two subproblems: minimization of $L(X,\mu)$ over μ for fixed X and minimization of $L(X,\mu)$ over X for fixed μ. For convenience, we suppress reference to the set of fixed parameters and state our two subproblems as:

$$\min_{\mu} \sum_{j=1}^{n} \sum_{k=1}^{n} (v_{jk}d_{jk} - \lambda_{jk}(\mu))^2 \tag{26c}$$

and

$$\min_{X} \sum_{j=1}^{n} \sum_{k=1}^{n} v_{jk}^2(d_{jk}(X) - \frac{\lambda_{jk}}{v_{jk}})^2 \tag{26d}$$

The first subproblem is the unweighted, metric, one-dimensional case of a Multidimensional Scaling problem; i.e., we want to find μ such that the (one-dimensional) distances $\lambda_{jk}(\mu)$ are as much as possible equal to the quantities $v_{jk}d_{jk}$. The second subproblem is a weighted, metric, *restricted* MDS problem; i.e., we want to find X such that it satisfies conditions like (19), (20) or (21) and at the same time the distances $d_{jk}(X)$ should be as much as possible equal to the quantities λ_{jk}/v_{jk}, where the deviations from perfect match are weighted by v_{jk}^2. Both subproblems can be conveniently solved by exploiting the general multidimensional scaling approach of de Leeuw and Heiser (1980).

To illustrate some of this, we use another set of data collected by Sjöberg (1967). It concerns the nine choice objects listed in table 3. These were judged by 106 psychology students on the attribute 'immorality' (graded pair comparisons on a 20-point scale). To remove the grading and the indifference judgments, we used (9) and got the proportions listed in table 4.

1. A drunken driver hit a person and left him in the road.
2. Foster-parents mistreated the four-year-old boy 'to teach him a lesson'.
3. A motorist refused to take a victim of a traffic accident to the hospital in his car.
4. A swindler sold the same house to eight persons.
5. A teenager smashed up a 'borrowed' car.
6. He made his living on moon-shining.
7. A congressman kept a watch he has found.
8. A farmer shot a deer out of season.
9. An elderly person stopped in a 'no-parking' zone to put a letter in a mail box.

Table 3. Nine choice objects from Sjöberg (1967).

	1	2	3	4	5	6	7	8	9
1	.500	.417	.488	.785	.972	.832	.976	.957	.984
2	.583	.500	.611	.879	.981	.936	.991	.989	.999
3	.512	.389	.500	.739	.913	.849	.960	.957	.989
4	.215	.121	.261	.500	.935	.779	.976	.949	.976
5	.028	.019	.087	.065	.500	.319	.719	.644	.928
6	.168	.064	.151	.221	.681	.500	.877	.811	.963
7	.024	.009	.040	.024	.281	.123	.500	.464	.849
8	.043	.011	.043	.051	.356	.189	.536	.500	.851
9	.016	.001	.011	.024	.072	.037	.151	.149	.500

Table 4. Proportion of times j was judged being more immoral than k (reconstructed).

Two analyses were done with the ad-hoc APL program PAIRS. One utilized assumption (18) with X two-dimensional ('case I'), the other (19) ('case III'). We have listed the obtained mean utility values in table 5, together with the values reported by Sjöberg and the ordinary case V values based on

	1	2	3	4	5	6	7	8	9
Sjöberg	0.927	1.000	0.902	0.700	0.373	0.516	0.220	0.270	0.000
Case V	0.866	1.000	0.840	0.722	0.356	0.533	0.214	0.264	0.000
Case III	0.910	1.000	0.897	0.678	0.416	0.547	0.303	0.326	0.000
Case I	0.913	1.000	0.904	0.439	0.163	0.321	0.093	0.124	0.000
σ_j^2	0.296	0.248	0.317	0.147	0.120	0.179	0.154	0.182	0.279

Table 5. Estimated mean utility values and dispersions.

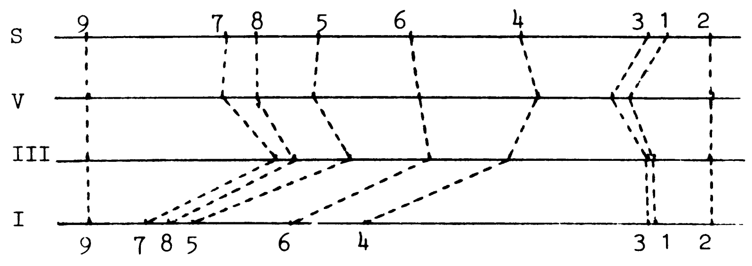

Figure 6. Utility scales from table 5.

table 4. For ease of comparison, all scales are linearly transformed such that the most immoral action gets a value of 1 and the least immoral one a value of 0. An alternative representation of these results is given in figure 6. The discriminal dispersions which we get from the analysis under the case III assumption are also listed in table 5, and the two-dimensional configuration X which best reproduces the comparatal dispersions is displayed in figure 7.

A global interpretation of these results is that all utility scales show the same order of actions, the discriminal dispersions seem to increase with extremeness of utility in both directions, and that the cognitive map contrasts *physical harm* with *material damage* on the one hand, and *reckless* actions with *intentional* actions on the other. There are some interesting details too. If we compare the case III utility values with those of case V, it seems as if the extremes have been pushed away. This is compensated for by higher values of the discriminal dispersions for these actions (which appear in the denominator of (25)). Maybe this gives us a somewhat nicer interpretation of the scale: actions 2, 1 and 3 are really bad, action 9 is no offence at all, but to which extent this is true is controversial among the subjects. A similar reasoning applies to the case I values, where we have, say, the unforgivable things against accepted offences, and correspondingly increased distances (between 2, 1, 3 and the others) on the cognitive map.

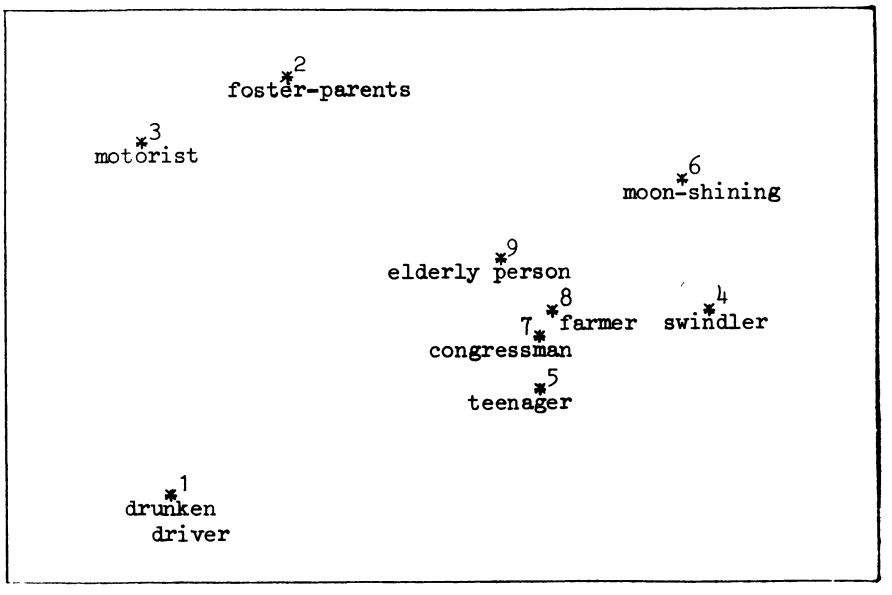

Figure 7. Cognitive map obtained from the immorality judgments in table 5 ($L(X,\mu) = 0.0511$).

A more subtle interpretation arises here if we consider two pairs which are about equally distant on the utility scale. Compare for example the pairs 1,2 (*drunken driver* vs *foster-parents*) and 4,6 (*swindler* vs *moon-shining*). Action 2 is a bit worse than 1, as is 4 compared with 6; but the proportion of times that *swindler* has been judged worse than *moon-shining* is much greater (.779) than the proportion for *foster-parents* and *drunken driver* (.583), due to the fact that *swindler* and *moon-shining* are much more com-

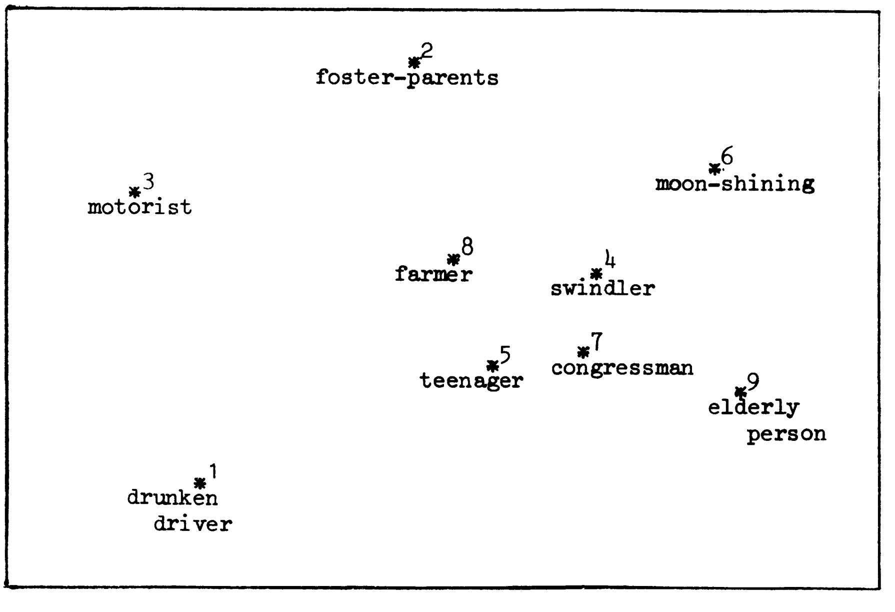

Figure 8. Cognitive map obtained from Sjöberg's estimates of the dispersions (stress = .17244).

parable on the cognitive map. Similarly, *motorist* and *drunken driver* are about as much worse than *teenager*, but *drunken driver* is judged more unanimously worse (p_{15} = .972 versus p_{35} = .913), because for both *drunken driver* and *teenager* about the same recklessness is involved. These conclusions do not change if we look at fitted proportions instead of observed ones.

Finally, we want to compare figure 7 with figure 8, which shows another cognitive map, derived from the comparatal dispersions as estimated by Sjöberg (for this purpose we used SMACOF-1, a metric multidimensional scaling program described in de Leeuw and Heiser (1977). The reckless versus intentional contrast seems to be the same, but this time we do not have physical harm contra material damage, but something like *physical harm - material damage - no damage*. Note that the extreme position of *elderly person* corresponds with an increased distance between 9 and the others on the Sjöberg-scale of figure 6, compared with the case I scale.

IV. DECOMPOSITION TECHNIQUES

It is not always plausible to assume that the individual subjects in a preference study essentially all sample from the same underlying process; to put the matter more strongly, sometimes we are convinced that individual choices are not alike because individual utilities are not alike in some fundamental sense.

Imagine a group of friends who have decided to go to the wintersports together. On their first preparatory meeting, they settle upon the characteristics of the ideal skiing resort: it should be high, but not too high; there should be at least 70 kilometres of skiing tracks; the place should be cosy, not too crowded, cheap, sunny and there should be other skiing possibilities in the immediate neighbourhood. They also want to stay in a comfortable chalet, close to the centre of the village, not too expensive, etc. Where to go? One of them then asks several travel agencies for information and comes out with eight possibilities, none of which is completely satisfactory, of course. To make up their mind, they all compare all resorts in pairs and perform a Thurstonean analysis.

This is a perfectly sensible thing to do. The objects here are selected and seen as imperfect approximations to one ideal. Consequently, the subjects are supposed to utilize the same *appropriateness-for-the-wintersports* continuum, on which each resort gets a scale value indicating its distance from the ideal. In fact, their task is to estimate and combine all kinds of subtle differences ; the use of a probabilistic choice model reflects the expectation, that these subtle differences are estimated differently by different subjects.

A completely different situation arises if we consider the preference behaviour of, say, all customers of one particular travel agency on one particular day, who ask for information regarding wintersports. Suppose that the travel agency gives them all the same traveller's guide, which comprises information about eight selected wintersport resorts. Moreover, suppose that we ask the customers to read and think a while and after that to give us all their pairwise preferences. Certainly, the present subjects are a much less homogeneous group and the present objects will show a much less restricted range of characteristics compared with those in the first situation. Some people prefer sophisticated places to simple ones, others don't; some want to make fast descents, others primarily want to make tours; some like 'curling' and do not intend to ski at all, others like skiing in virgin snow and do not intend to stay in the village at all; for some, the more disco's the better, for others the other way round, etc. All these different requirements will result in different preferences. How can we decribe these individual differences?

IV.1. The concept of a multidimensional joint utility space.

We might conveniently imagine that each object can be represented by an appropriately selected point in a space of one, two, three or more dimensions. We don't know yet, how many dimensions this space should have and where the points representing objects will be located; we want the data to give us a clue to that. To capture the individual differences in the model, we use the notion of an *isochrest* (this is in analogy with 'isobar' or 'isotherm'; Carroll (1972) uses the word 'isopreference contour', an somewhat unfortunate name as preference is usually understood in terms of pairs of points). An isochrest is a curve which connects points of equal utility.

Consider the psychological map of eight resorts presented in figure 9. We will not pay attention to the way in which this particular location of points was chosen (it certainly does not correspond to a geographical map). An imaginary subject told us, that among these eight places his first choice would be: Selva or Obergurgl; his second: Saas Fee, Cervinia or les deux Alpes; his last choice would be: Gerlos, Kitzbühel or Chamonix (for ease of presentation, we gratefully acknowledge the presence of ties). The isochrests labelled 1, 2 and 3 represent these choices. Of course, for another subject we would have to draw other curves.

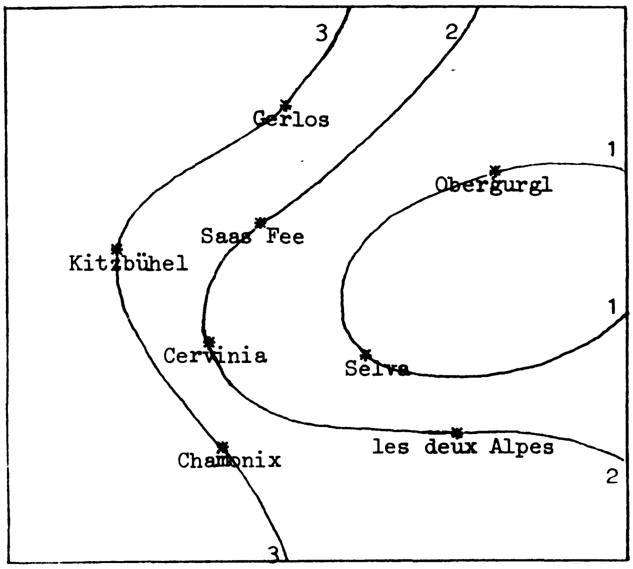

Figure 9. Psychological map of winter-sport resorts, with isochrests

In general, we could take any arbitrary location of points and represent any series of utility values by drawing a set of isochrests. This would portray the data, but in a disorderly and trivial way. So we want to tighten up the requirements, such that they impose restrictions on the data. This can be done in several ways, usually called 'models'. All of them involve the idea that the isochrests should be a family of regular curves, defined on one unique configuration of points:

a. *the vector model:* each subject is represented by a vector and his isochrests are parallel lines (planes, hyperplanes) perpendicular to his vector, in the order (and spacing) of his utilities.
b. *the unfolding model:* each subject is represented by a point and his isochrests are concentric circles (spheres, hyperspheres) around this point, in the order (and spacing) of his utilities.
c. *the weighted unfolding model:* each subject is represented by a point and his isochrests are concentric ellipses (ellipsoids) around this point, in the order (and spacing) of his utilities.
d. *the compensatory distance model:* each subject is represented by a point and his isochrests are parallel lines (planes, hyperplanes) perpendicular to the line connecting this point with the origin of the space. This time the utilities are reflected by the distances between the subject point and the parallel lines.

We will call the joint space of object points $\{y_{ja}\}$ and subject points (or vectors) $\{x_{ia}\}$ *(multidimensional) joint utility space* (Coombs (1964), who

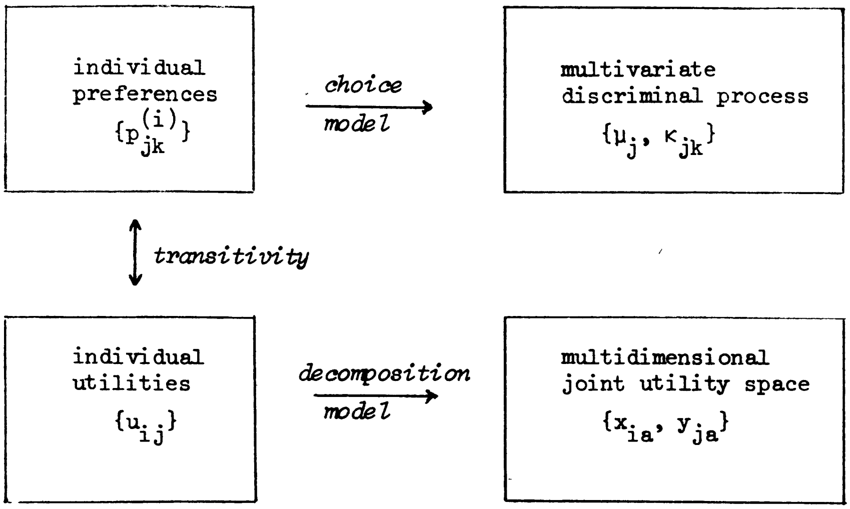

Figure 10. The mapping of preference and utility.

developed the concept into its present form, just calls it 'joint space' or 'joint scale'; in his pioneering 1946 paper, Guttman talks about 'quantifying comparisons', confining himself to the vector model). For the first two models it is possible to devise a decomposition technique, which reconstructs from a given table of utilities a multidimensional joint utility space such that the requirements of the model are as much as possible fulfilled. We will discuss this in more detail in sections IV.2 through IV.5. We will describe the models as alternative ways to map the data and will not address the intricate question whether or not one model is more 'true' than the other. Also, we will refrain from technicalities. The weighted unfolding model is discussed in Carroll (1972) and the compensatory distance model in Coombs (1964) and Roskam (1968). For both, however, no reliable decomposition techniques are available and we omit any further discussion.

Our development is summarized in figure 10. As Bechtel (1976) has pointed out, the representation of subjects and objects in multidimensional joint utility space is in the testtheoretic tradition of joint or dual parametrization, which emphasizes rather than obliterates information about individual and intergroup differences. The decomposition models as we treat them here do not contain probabilistic notions, they are in a sense just 'the deterministic bridge' between individual utilities and joint utility space. If we were willing to accept distributional assumptions, we could try to connect joint utility space directly with the individual preferences (cf. Zinnes and Griggs (1974)). Another interesting approach we will not discuss is to partition the subjects into *classes*, with homogeneity of utility within classes (summarized in one or more 'local' preference orders) and heterogeneity between classes (cf. Hayashi (1964), Lemaire (1977)).

IV.2. The vector model.

In the vector model, the subjects are represented by vectors, each of which supports a set of parallel isochrests. This is illustrated for one subject in figure 11. From this figure we may infer that this particular subject (in

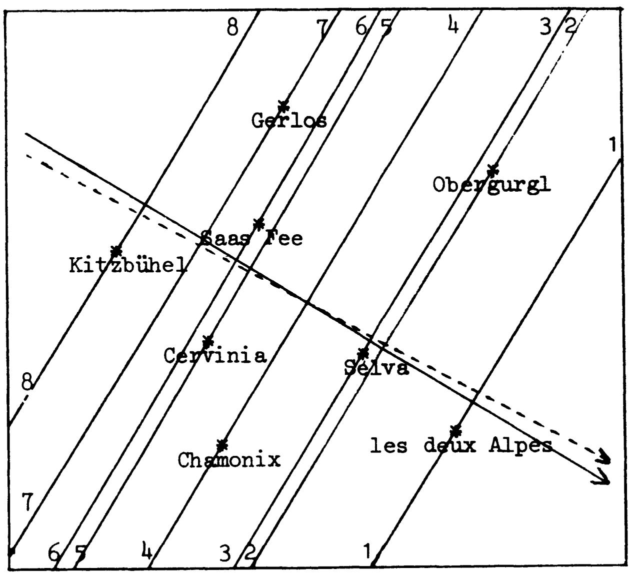

Figure 11. Wintersport map with parallel isochrests

the same map as before) orders the wintersport resorts as: les deux Alpes, Obergurgl, Selva, Chamonix, Cervinia, Saas Fee, Gerlos and, finally, Kitzbühel. The subject vector not only implies a particular order among the object points, but also a specific spacing between each of them, which corresponds with the distance between the isochrests along the subject vector. If we move the vector in figure 11 a bit, we get the same order but a different spacing (e.g., with respect to the dotted vector, Selva and Obergurgl are more separated, whereas Cervinia and Saas Fee nearly coincide). In fact, if we go on moving around the vector (keeping the map fixed), we will encounter 56 different orders, but an infinite number of differently spaced orders. More generally, the number of different orders that can be 'explained' by the vector model (with any non-degenerate configuration of points) is finite, depends on the dimensionality of the space and the number of points we want to accomodate in it, and is very small compared with the total number of different orders that may be formed (Bennett (1956), Coombs(1964)).

Let's now look at the structure of the model more closely. Mathematically, the decomposition may be expressed as

$$u_{ij} = \sum_{a=1}^{p} x_{ia}y_{ja} \tag{27}$$

where u_{ij} denotes the utility of subject i for object j, the $\{x_{i1},\ldots, x_{ip}\}$ are the coordinate values of the vector representing subject i and the $\{y_{j1},\ldots, y_{jp}\}$ the coordinate values of the point representing object j. To simplify the discussion, we will confine ourselves now to two dimensions and consider one subject only, with utilities $\{u_1,\ldots, u_j,\ldots, u_n\}$. This simplifies (27), and we get the system

$$\begin{aligned} u_1 &= x_1y_{11} + x_2y_{12} \\ &\vdots \\ u_j &= x_1y_{j1} + x_2y_{j2} \\ &\vdots \\ u_n &= x_1y_{n1} + x_2y_{n2} \end{aligned} \tag{28}$$

where $\{x_1,x_2\}$ is the subject vector. To see how the isochrests come in, it is convenient to transform (28) into

$$\begin{aligned} y_{12} &= u_1' - v\,y_{11} \\ &\vdots \\ y_{j2} &= u_j' - v\,y_{j1} \\ &\vdots \\ y_{n2} &= u_n' - v\,y_{n1} \end{aligned} \tag{29}$$

which is a series of n parallel straight lines through the points $\{y_{j1},y_{j2}\}$ with slope $v = x_1/x_2$ and shift $u_j' = u_j/x_2$. Clearly, if for two objects we have $u_j = u_k$, then $u_j' = u_k'$ and they must be situated on the same straight line.

In a similar way, another subject is associated with another set of parallel straight lines or, equivalently, with another direction in utility space. If we regard the coordinate axes as fixed 'psychological dimensions', we may say that each subject *weights* these dimensions differently to arrive at his utilities. In this reasoning, all subjects 'use' all dimensions of joint utility space, but in a different way (or to a different degree). This implies the idea of *compensation*: two objects may be far apart, but if this happens to be in (a) direction(s) perpendicular to the vector of subject i, that particular subject still gives them equal utility (compare in figure 11 Selva with Obergurgl, which have nearly equal utility, with Selva and les deux Alpes which are closer together but more saliently differentiated). On the other hand, suppose two objects get the same utility ($u_j = u_k$), then

$$y_{j2} + v\, y_{j1} = y_{k2} + v\, y_{k1} \tag{30}$$

using (29), and this implies

$$\frac{y_{j2} - y_{k2}}{y_{k1} - y_{j1}} = v \tag{31}$$

In words, (31) says that a dominance of j over k on the second dimension is compensated for by a dominance of k over j on the first one.

Another interpretation of the decomposition could be that each subject selects just one direction in joint utility space and disregards all p - 1 other ones. According to this point of view, we need not to commit ourselves to an interpretation in terms of projections on some set of (arbitrary) coordinate axes, but may look at 'the picture as a whole' and use notions like *contiguity* vs *separation (clustering)* and *circular ordering* (cf. Lingoes and Borg (1977)) as well. Consider for example another possible map for the ski-

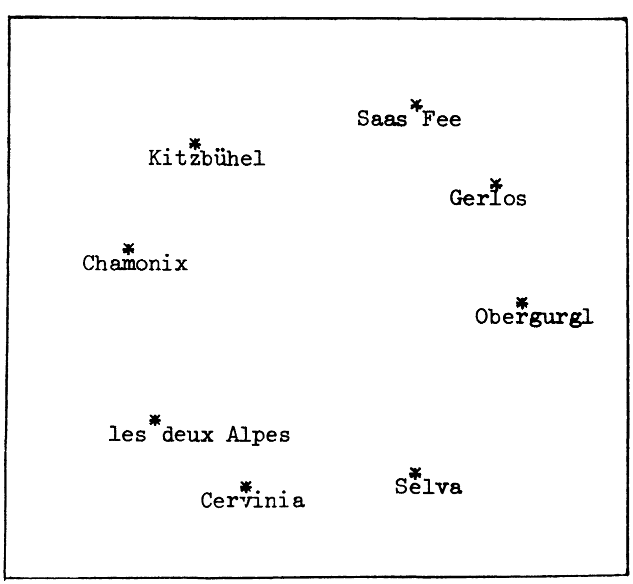

Figure 12. Alternative wintersport map showing a circular ordering (circumplex).

ing resorts in figure 12. This structure could very well arise in practice. After all, Chamonix in many respects resembles Kitzbühel but it is more stylish and you can ski there in summertime; les deux Alpes is less expensive than Chamonix and more attractive for the 'young' jet-set; Cervinia does not have that many 'après-ski' possibilities and is less attractive for beginning skiers than les deux Alpes; in Selva you don't have summerski possibilities as in Cervinia, but better touring possibilities; Obergurgl is attractive for people of all ages, but there are less 'cross-country' possibilities than in Selva; Gerlos is more attractive for beginners than Obergurgl, but less sporting; in Saas Fee there are less touring possibilities compared with Gerlos, but more skilifts; and again Kitzbühel is bigger and has more 'cross-country' than Saas Fee, but is also more expensive. Of course, these qualifications should not be taken too seriously, they only try to illustrate the idea of a circular order (circumplex structure): neighbouring points share many aspects and differ in a few; if two points are far away along the circle, they share very few aspects and differ in a lot. In a case like this, it is not so natural, or even very difficult, to pick out two orthogonal psychological dimensions for interpretation, whereas the *ordering without beginning or end* may be perfectly convincing on its own. The vector model here says, that each subject may have an *ideal combination of correlated aspects* somewhere upon the circle, and that his utility decreases evenly in both directions.

We now turn to the matter of estimation. We want to be brief about it. Many techniques have been proposed, primarily differing in generality and elegance of presentation (see Guttman (1946), Slater (1960), Tucker (1960), Carroll and Chang (1964), Hayashi (1964), Benzécri (1967), de Leeuw (1968), Bechtel (1969), Carroll (1972), de Leeuw (1973) and Nishisato (1978)). For our purposes, it suffices to say that they all seem to amount to minimizing the loss function

$$L(X,Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} (u_{ij} - \sum_{a=1}^{p} x_{ia}y_{ja})^2 \qquad (32)$$

under certain normalization requirements. Thus, for given $\{u_{ij}\}$ we want to find both $\{x_{ia}\}$ and $\{y_{ja}\}$ such that L(X,Y) is as small as possible. In its simplest form, the problem can be solved by routine methods (a truncated singular value decomposition); the solution for X and Y will be unique up to a joint rotation, which mostly will not bother us. A more general approach to handle the problem, where each row of U may be transformed with any monotone transformation to maximize the fit, can be found in van Rijckevorsel and de Leeuw (1979). Unfortunately, there are cases where these generalized algorithms converge to undesirable solutions (unique axes for certain individuals), so they are really not yet suited for standard usage.

IV.3. Application of the vector model.

A number of succesful applications of the simple vector decomposition have been published. We name a few scattered examples. In the area of marketing research, Green and Rao (1972); in political science, Daalder and Van de Geer (1977) or de Leeuw (1973); in experimental psychology, Mc Dermott (1969); in clinical psychology, Slater (1965); in esthetics, Fênelon (1971); in the area of population studies, Delbeke (1968); in sociology, Seligson (1978). We will analyse a fresh example here, adopted from Dijkstra (1978), and illustrate some aspects of data manipulation which are not really a part of the method but nevertheless important from a data-analytical point of view.

The data concern the motivation to work in an academic setting; each of 47 subjects from the Department of Philosophy and Social Science of the Technical University Eindhoven indicated their preference order among ten aspects of job satisfaction (these are summarized in table 6). We analysed

1:	PART	(participation)
2:	SECU	(security prospects)
3:	IMPO	(important work)
4:	PROV	(provocative work)
5:	RESP	(heavy responsability)
6:	HEAD	(right department head)
7:	PLEA	(pleasant work setting)
8:	MAKE	(possibility to make one's way)
9:	SALA	(good salary)
10:	WELF	(right welfare facilities)

Table 6. Ten aspects of job satisfaction.

these utilities with the popular program MDPREF (Carroll and Chang (1968)). Using standard options, we get the two-dimensional solution of figure 13. The subject vectors extent over a range of about 215 degrees, due to the fact that WELF and PLEA are generally being judged low and never chosen first. Still, there is considerable interindividual variation, but we have to be careful: the vectors just indicate *directions* in utility space (MDPREF standardizes all vectors to have equal length) and don't tell us whether or not a particular subject fits in well or badly. Therefore, we adjusted the

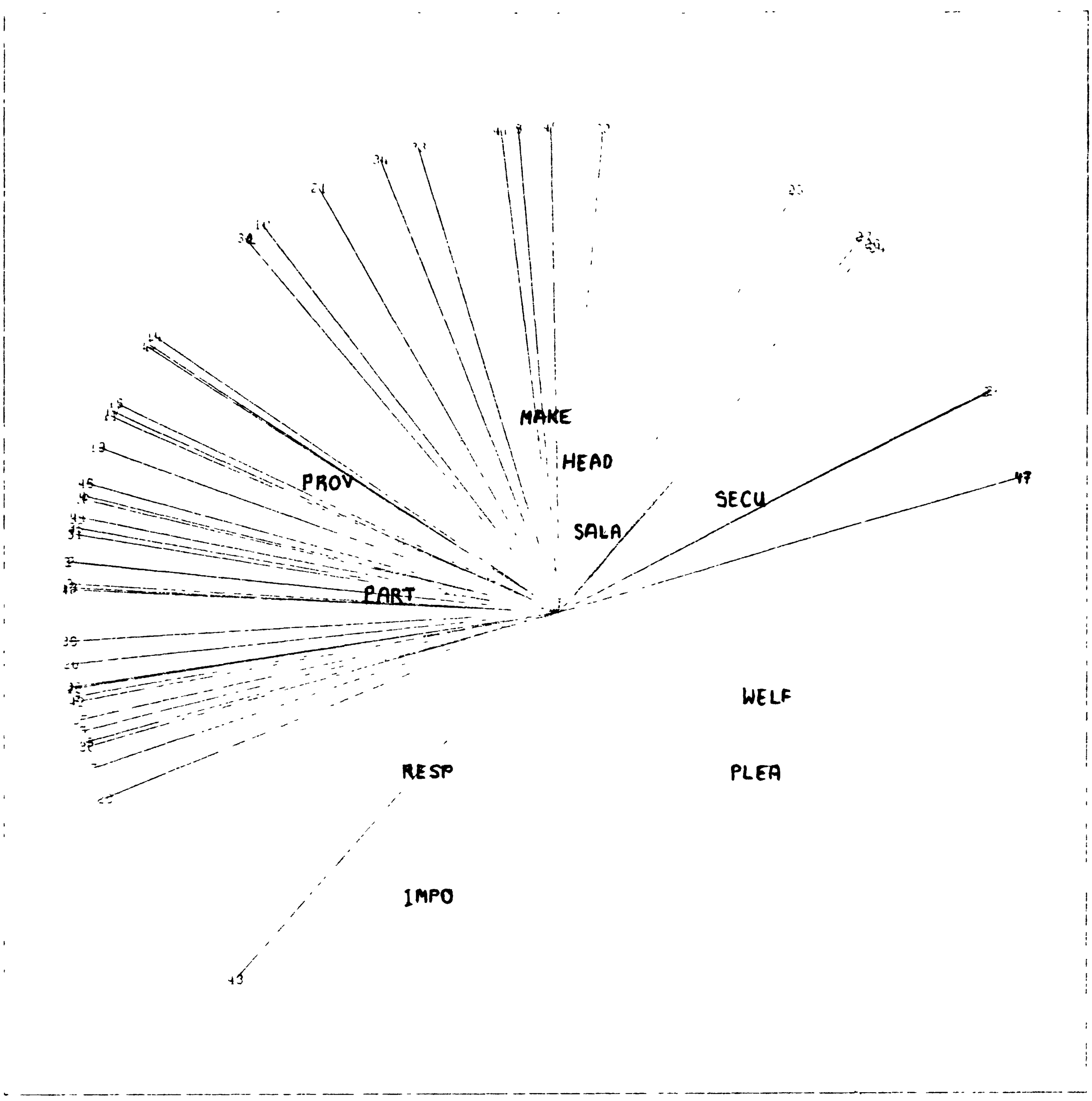

Figure 13. Two-dimensional MDPREF solution for job satisfaction data ('explained' variance 58%).

the length of the vectors in such a way that they reflect the goodness-of-fit of each individual with respect to the two-dimensional solution (their 'communality'). This is plotted in figure 14.

Because the program standardizes the configuration of object points such that their projections on each axis have sum of squares one and are mutually orthogonal, we have the important property that solutions in different dimensionalities are *nested*: the first p axes of the (p+q)-dimensional solution are identical to the p-dimensional solution. We may interpret the adjusted subject vectors as projections out of a high-dimensional space into p-space such that their *average* length is as large as possible. This implies that we should evaluate 'short subjects' differently from 'long subjects'.

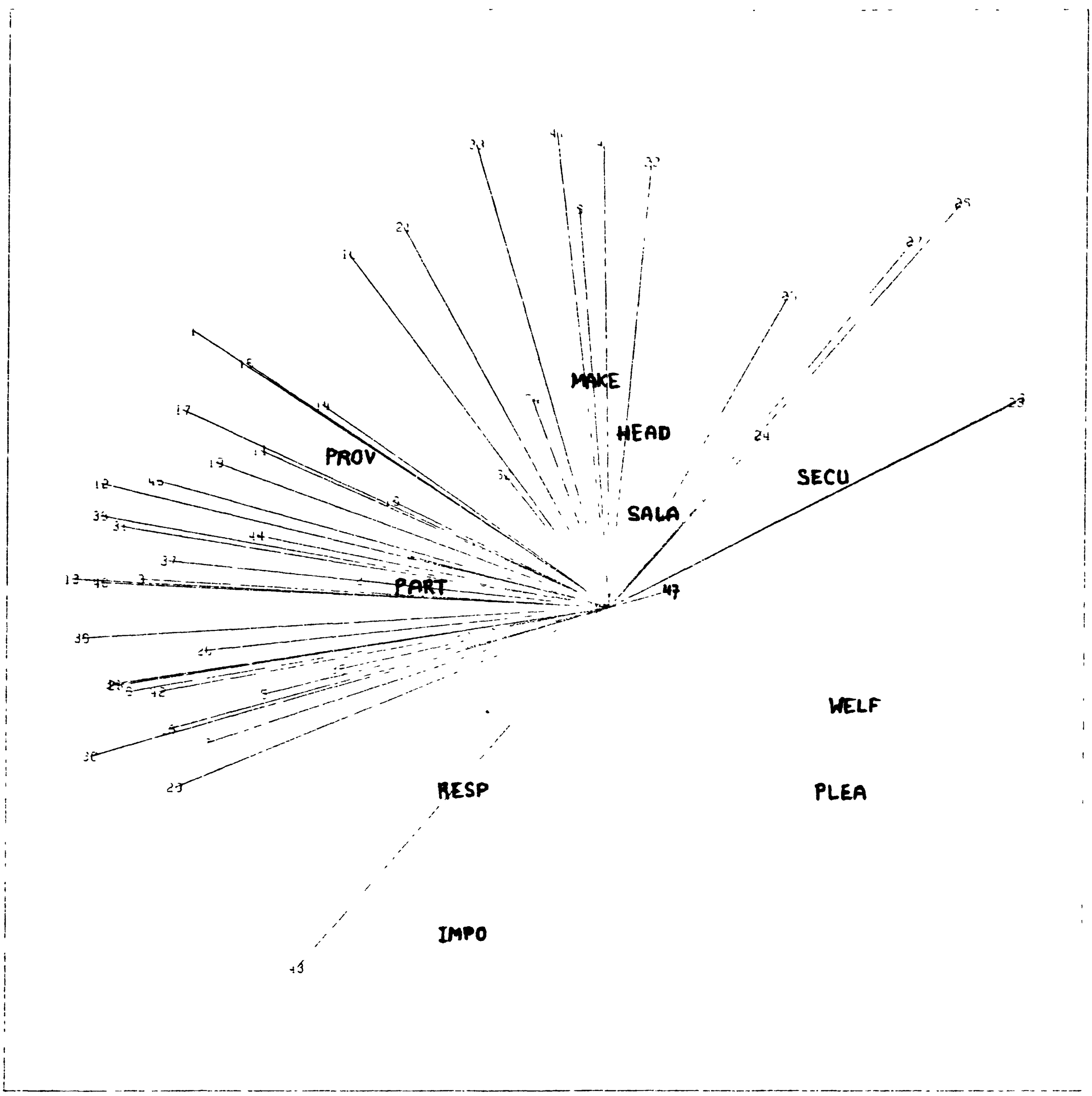

Figure 14. Two-dimensional MDPREF solution for job satisfaction data with adjusted vector length.

Thus we see that subject 47 does have quite a different point of view, but most of all he is pointing in still another direction out of this two-space (his first choices are SALA, WELF and SECU, but his last three are HEAD, MAKE and PLEA; fitted isochrests would dramatically reveal here the mismatch with the data). For similar reasons, subjects 24, 34, 32, 16, 4, 2, 5, and 35 apparently also don't fit in well.

Dijkstra (1978) suggests that PART, IMPO, PROV, RESP and MAKE are *intrinsic*, whereas SECU, HEAD, PLEA, SALA and WELF are *extrinsic* motivations. This dichotomy comes out nicely along the horizontal axis. The vertical axis could be something like *oriented to the future* (MAKE, PROV, HEAD), versus *oriented to the present* (IMPO, RESP, PLEA). This is not very satisfacfactory and we may look at something else.

Note that the configuration of object points vaguely exhibits a *horseshoe* form: we could look at it as a curved dimension, on which the points are ordered as PLEA, WELF, SECU, SALA, HEAD, MAKE, PROV, PART, RESP, IMPO. Also note that the subject vectors predominantly point into the direction (north-)west. It turns out that the mean utility values of the objects (computed here as mean rank numbers, and very closely related to Thurstone case V values) along the horseshoe are: 7.9, 7.7, 6.5, 5.0, 5.3, 4.8, 3.3, 3.9, 5.0, 5.5. Thus, starting with PLEA (7.9) and going counter-clockwise, the mean values first decrease down to the most popular PROV (3.3) and then rise again. We may argue that this direction of mean utility (a direction in space approximately going from PROV to PLEA) certainly represents something important (common opinion, norm, academic hypocrisy), but also obscures the typical nature of the individual differences.

We can remove the effect of mean utility by taking deviations from column means. Analysis of these deviation scores (again with adjusted subject length) gives us the result in figure 15. The intrinsic-extrinsic dichotomy is still there, but there are changes on the vertical axis (IMPO and RESP are more differentiated, as well as PROV and PART; WELF and SALA are closer together, as are PLEA and HEAD). The pleasing thing about the distribution of subject vectors is, that they now cover the whole range of directions. We have marked four regions in figure 15 which divide the total group into four typical subgroups:

The modest (I): SECU, WELF and SALA are evaluated relatively high in this group, IMPO and PROV relatively low. These people are just making a living and some of them probably have settled down in this university for the rest of their lives.

Opportunists (II): here MAKE is relatively high and RESP is relatively low. This group is more ambitious than group I, but they want to keep away from duties.

The hopeful (III): typically, PROV and IMPO are important and 'material things' are not. They are eager to make their own way in science.

Managers (IV): here RESP, PART (and IMPO) are dominant, whereas SECU and MAKE are not. These probably are the people in high positions (or a certain class of paid students?).

It is possible to accomodate most subjects in one of these four groups. Some of the subjects which don't fit in very well actually conform to the general norm (cf. subjects 1, 10, 11, 21 in figure 14); others are in fact differen-

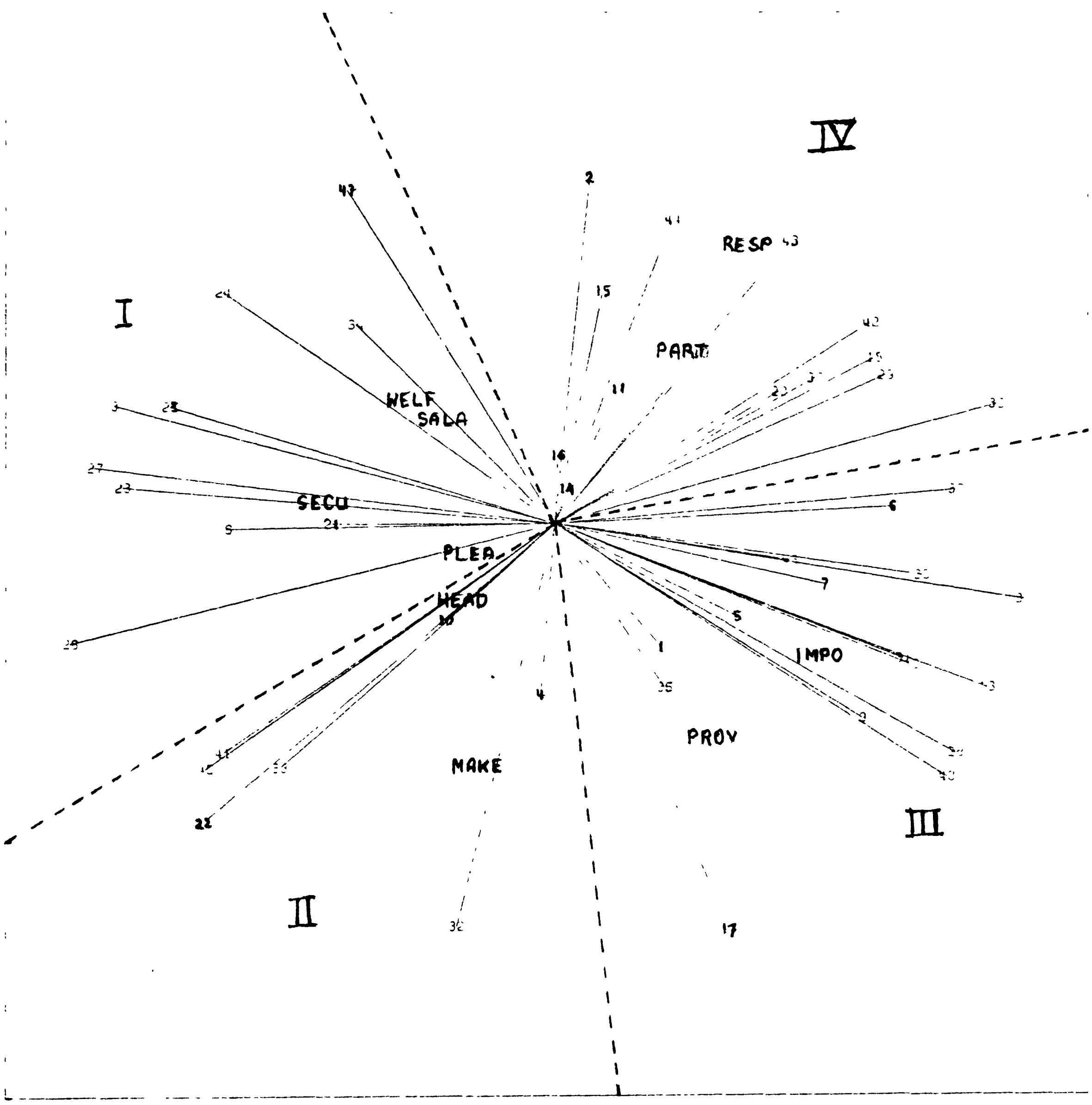

Figure 15. Job satisfaction data in deviation from column means (MDPREF, 'explained' variance 51%).

tiated in the third dimension of the deviation scores solution (this direction contrasts IMPO, SALA, MAKE with PROV, HEAD, PART and the two little subgroups are subjects 4 and 5, the *autonomous careerists*, versus 14 and 16, the *dedicated scholars*).

This rather exhaustive interpretation of the obtained joint utility space would require validation through careful examination of background characteristics of the subjects. Also, reanalysis on subsets of the set of objects or on particular subgroups of subjects could prove to be useful, but this would lead us outside the scope of this paper.

IV.4. The unfolding model.

In the vector model, the family of isochrests was characterized by vectors pointing in different directions. This kind of representation has its roots in the long-winded Spearman/Thurstone factor-analytic tradition within psychology or, quite independently, in the French data-analytic tradition (Benzécri (1976)). It yields a very strong kind of model and there have been several attempts to generalize it. One of these was to specify the family of isochrests as a set of parallel *curves* (Carroll (1972)), but the properties of this *polynomial* model have never been worked out in detail.

A completely different proposal originated with Coombs (1952, 1964), although here too Thurstone's pioneering work on attitude scaling is relevant. The key notion is, that the dimensions of joint utility space should correspond to fundamental dilemma's. If we were to consider a collection of cars which differ on two attributes only, say price and safety, any 'rational' man would choose a car which is cheap and safe over an expensive and unsafe one. But the very thing which generates individual differences and which may be of practical or theoretical interest, is the *trade-off between opposing benefits* (Coombs and Avrunin (1977) discuss this in terms of so-called approach-approach, approach-avoidance and avoidance-avoidance conflicts).

The assumption that people do make different trade-off's (which in the example is reflected by the amount of money they are willing to pay for safety) leads to the concept of a point of maximum subjective utility or *ideal point*. An ideal point corresponds with an imaginary object which would be preferred to all other available ones. This subjective ideal need not to be ideal in an absolute sense, but it represents the best possible *compromise*. And joint utility space, as conceived here, will not reflect all attributes which characterize the objects at hand, but only the ones that are relevant in the sense of urging people to make different trade-off's. Furthermore, if a particular subject is confronted with two objects, he will prefer the one which is closest to his own ideal; put differently, joint utility space is constructed such that the individual utilities are reproduced by the distances between the object points and the ideal point.

In two dimensions, the model implies that the family of isochrests consists of sets of concentric circles. In figure 16, one subject is put in with an ideal point very close to Selva; the order of his utilities apparently is: Selva, les deux Alpes, Obergurgl, Saas Fee, Cervinia, Chamonix, Gerlos, Kitzbühel. If we move around the ideal point a little bit, the order

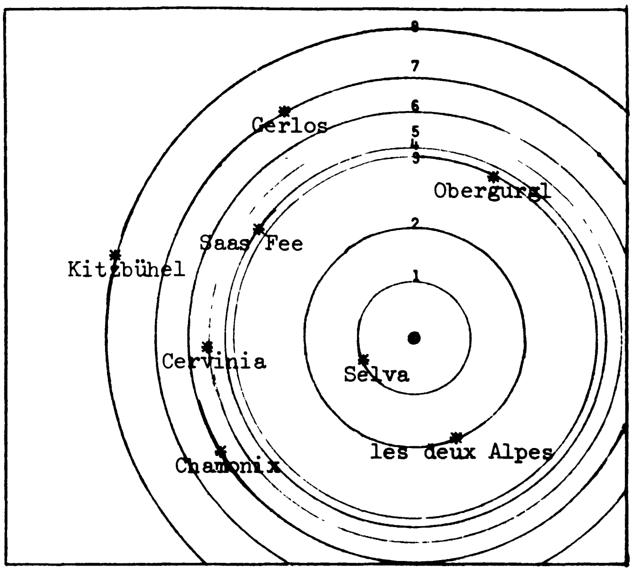

Figure 16. Wintersport map with circular isochrests.

doesn't change but the intervals between reproduced utilities do. A set of points which all generate the same order of utilities is called an *isotonic region*. With enough object points relative to the number of dimensions, isotonic regions tend to be compact and very small, at least in the interior of space. On the exterior, they are fan-shaped and extend to infinity. This gives us a geometrical hint that the unfolding model is a generalization of the vector model. For, if we imagine an ideal point moving outwards, then beyond a certain limit circular isochrests generate the same order of utili-

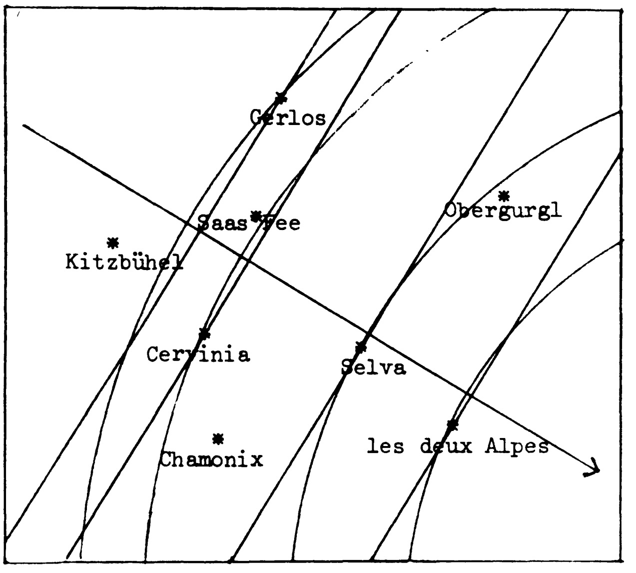

Figure 17. Equivalence of circular and straight isochrests when the subject point moves outwards.

ties as do parallel straight lines perpendicular to a vector pointing at the ideal point. This is illustrated in figure 17. For clarity, only four isochrests have been drawn. Note, however, that this notion of ideal points freely moving outwards is a bit overoptimistic: for in practice, we have to standardize the joint configuration of object- and subject points to some fixed value and as a consequence, the more subject points are moving outwards, the more the configuration of object points will relatively *shrink*. In the limit, all object points coincide, a situation which is intuitively undesirable but frequently encountered in practice, especially with badly fitting data.

Some results concerning the maximum number of preference orders generated by n objects in r dimensions can be found in Coombs (1964) or, more completely, in Good and Tideman (1977). In our example, with n=8 and r=2, this number turns out to be 351. Thus the unfolding model accomodates much more preference orders than the vector model does, but still a lot less than the number of possible orders (n!).

We have said that in the present model utilities are reproduced by distances. More specifically, the decomposition rule is mostly assumed to be *euclidean*:

$$\delta_{ij} = \{ \sum_{a=1}^{p} (x_{ia} - y_{ja})^2 \}^{\frac{1}{2}} \tag{33}$$

where δ_{ij} denotes the *disutility* of object j according to subject i, related to the original utility value by

$$\delta_{ij} = H(u_{ij}) \tag{34}$$

where H is a suitably chosen monotone decreasing function. In *metric* unfolding, we usually take

$$\delta_{ij} = \max_{i,j} (u_{ij}) - u_{ij} \tag{35}$$

In sympathy with the general strategy of Coombs (1964) to treat all data in the social sciences under the weakest possible assumptions, unfolding virtually has been equated with row-conditional non-metric unfolding: one wants to reproduce the rankorder of the utilities only, and moreover is not willing to assume intersubjective comparability of these ranknumbers. This leads to a modification of (34) into

$$\delta_{ij} = h_i(u_{ij}) \tag{36}$$

where h_i are optimally chosen monotone decreasing functions. Although there have been many attempts to find satisfactory algorithms for this relaxed version of the model, these do not seem to have been very succesful (cf. Kruskal and Carroll (1969), Heiser and de Leeuw (1978)). For the one-dimensional case, there are even weaker versions of unfolding which only try to find an 'ordered metric scale' for the objects (cf. Phillips (1971), Mc Clelland and Coombs (1975)). Here, we will consider the more tractable metric multidimensional case only.

The assumption of euclidean distance is vital to arrive at circular isochrests; i.e., if we look at all points for which (dis)utility is constant, (33) tells us that in two dimensions

$$c^2 = (x_{i1} - y_{j1})^2 + (x_{i2} - y_{j2})^2 \tag{37}$$

which is the general equation of a circle with centre $\{x_{i1}, x_{i2}\}$ and radius c. In case we had chosen a non-euclidean decomposition rule, such as the so-called *city-block* metric

$$\delta_{ij} = \sum_{a=1}^{p} |x_{ia} - y_{ja}| \ , \tag{38}$$

a set of square instead of circular isochrests would have come out. This kind of metric may have theoretical plausibility (cf. Coombs (1964), p. 206) but the algorithmic problem it poses is largely unsolved.

In the unfolding model, the idea of compensation no longer holds, since y_{j2} in (37) cannot be regarded as a single-valued function of y_{j1}, as could be done in (29). Furthermore, (im)popularity of objects is represented as *eccentricity*: popular objects will be close to the centroid of the ideal points, controversial ones nearby the edge. To see this, consider the mean squared disutility of object j:

$$p_j = \frac{1}{m} \sum_{i=1}^{m} \delta_{ij}^2 \ , \tag{39}$$

which we may take as a measure of impopularity. Applying (33), we get

$$p_j = \frac{1}{m} \sum_{i=1}^{m} \sum_{a=1}^{p} (x_{ia} - y_{ja})^2 \ . \tag{40}$$

We will now assume without losing generality that the configuration of subject points is centered, i.e. $\sum x_{ia} = 0$, and that its sum of squares $\sum\sum x_{ia}^2$ equals some unknown value $m.\xi$. We get

$$p_j = \frac{1}{m} \sum_{i=1}^{m} \sum_{a=1}^{p} x_{ia}^2 + \sum_{a=1}^{p} y_{ja}^2 - \frac{2}{m} \sum_{a=1}^{p} y_{ja} \sum_{i=1}^{m} x_{ia}$$
$$= \xi + \sum_{a=1}^{p} y_{ja}^2 \; . \qquad (41)$$

Thus if an object is very popular, p_j will be low and according to (41) the sum of squares of its coordinate values will be relatively small; if an object gets more controversial, its p_j value will be higher and its distance to the origin increases, etc. If there happens to be an object which is always dominated by almost all other objects, it usually 'needs a dimension on its own' (it might be better to discard it for further analysis).

Two kinds of decomposition techniques have been proposed to estimate the parameters of the metric unfolding model. One of these uses an algebraic analysis of the squared distances implied by (33). This approach goes back to a conjecture of Coombs and Kao (1960); other contributors are Ross and Cliff (1964), Schönemann (1970) and Gold (1973). The second kind of technique tries to minimize the least squares badness-of-fit function

$$L(X,Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} (\delta_{ij} - \{ \sum_{a=1}^{p} (x_{ia} - y_{ja})^2 \}^{\frac{1}{2}})^2 \qquad (42)$$

over jointly normalised X and Y, by means of a specialized multidimensional scaling algorithm (cf. Heiser and de Leeuw (1979), who also compare several different algebraic methods as to their suitability to provide a good initial configuration for the iterative minimization of (42)).

We conclude this section with the remark that the name 'unfolding' plastically describes the problem which the decomposition techniques have to solve: imagine joint utility space depicted on a handkerchief; pick it up in some ideal point i and pull it through a ring. The object points will come out in the order of the utilities of subject i; thus an individual preference order is produced by joint utility space, *folded* at point i. Obviously, the decomposition problem is to unfold all preference orders simultaneously.

IV.5. Applications of the unfolding model.

Unfortunately, not many applications have been reported in the literature.

There are some small pioneer studies such as in Coombs (1964), Roskam (1968) and Schönemann (1970), and some more substantial ones, such as Daalder and Rusk (1972), Green and Rao (1972), Levine (1972), Davison (1977) and Delbeke (1978). Some of the reported results exhibit suspiciously 'degenerate' clusterings of points, probably due to the fundamental weakness of the non-metric unfolding approach. We think this disappointing state of affairs can be remedied to some extent by adopting a metric approach or by imposing restrictions on the parameters of the model. We will discuss two analyses performed with the metric program SMACOF-3 (Heiser and de Leeuw (1979)), using data from Gold (1958) and Delbeke (1978).

The first example concerns the evaluation of power characteristics by eight different groups of middle-class american children. Among other things Gold's study yielded the datamatrix reproduced as table 7. The groups are labelled A-H; the details of data collection and group composition do not bother us here. The 17 row objects represent possible properties of children

	A	B	C	D	E	F	G	H
1. SMART: Smart at school	13.5	13	17	15	16	17	17	16
2. IDEAS: Good ideas how to have fun	1	17	13	6	6	10	9	4
3. MAKIN: Good at making things	13.5	6.5	12	15	13	12.5	15	14
4. GAMES: Good at games with running and throwing	16.5	3	14	17	17	11	13	13
5. FIGHT: Knows how to fight	12	4	11	15	14	15	16	12
6. STRON: Strong	9.5	13	15	13	12	16	14	15
7. FRIEN: Acts friendly	2	15.5	3	3	3	5	2	2
8. GOPER: A good person to do things with	9.5	1	4	11	9	6	6	9
9. ASKIN: Asks you to do things in a nice way	5.5	5	1	4	4	2	1	5
10. NOTEA: Doesn't start fights and doesn't tease	5.5	11	7.5	1	7	1	4	7
11. HOWTO: Knows how to act so people will like him	15	13	5	2	2	8	5	3
12. PLAYS: Plays with you a lot	3	8.5	9	10	8	9	11	6
13. LIKES: Likes to do the same things you like to do	5.5	6.5	10	5	1	7	8	1
14. NICEL: Nice looking	11	10	7.5	12	15	14	10	17
15. HAVIN: Has things you'd like to have	16.5	15.5	16	7	10	12.5	12	10
16. GIVIN: Gives you things	8	8.5	6	9	11	3	7	11
17. DOING: Does things for you	5.5	2	2	8	5	4	3	8

Table 7. Ranks of items by per cent of times they were rated 'very important' (low value: most important); from Gold (1958).

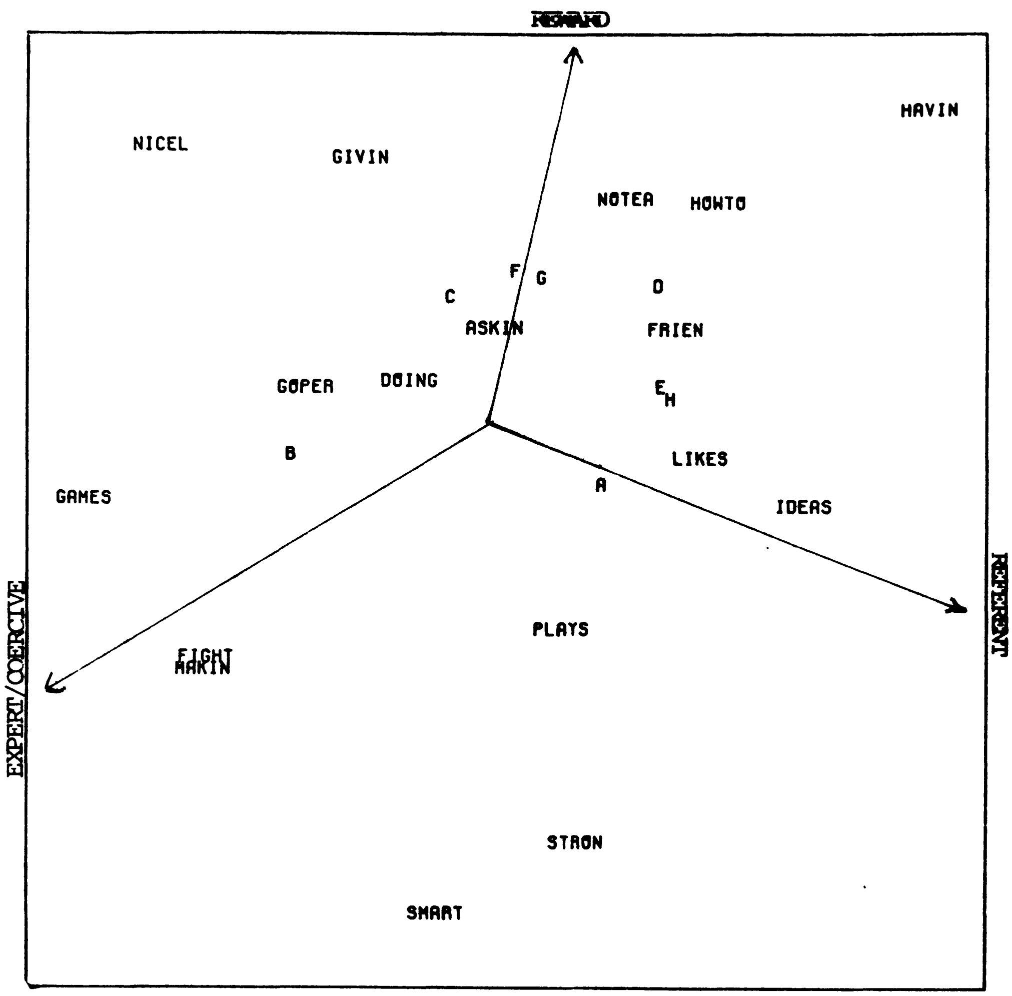

Figure 18. Map of power resources (Gold, 1958), obtained with SMACOF-3 (stress = 0.0299).

which, when valued highly in a group, supposedly contribute to the social power of children who possess them. Thus, IDEAS, FRIEN and PLAYS are very important to exercise power in group A, whereas GOPER, DOING and GAMES are required in group B, etc. Note that SMART and STRON are never appreciated very much and we expect that they will turn up at the edge of utility space.

The result of the SMACOF-3 analysis is presented in figure 18. As expected, SMART and STRON are far away from the centroid of the group points. Furthermore, the so-called *social-emotional* resources FRIEN, GOPER, ASKIN, NOTEA, HOWTO and DOING are all close to the centroid of the group points (with NOTEA, HOWTO and FRIEN far away from B and GOPER and HOWTO far away from A). A tentative interpretation of the configuration of object points might be based on the concepts of French (1956) and French and Raven (1959). *Reward power* is based on the ability of the actor to administer positive valences and to remove or decrease negative valences. Clearly, HAVIN, GIVIN, NOTEA and HOWTO exemplify this. *Referent power* is based on a liking or identification relationship; LIKES, PLAYS and IDEAS are typical (but SMART and STRON also). The third direction indicated in the figure concerns *expert-* and *coercive power*, based on the belief that someone has greater resources (knowledge or information) within a given area (SMART, MAKIN, GAMES) or can mediate punishments (FIGHT, STRON). The obtained joint utility space could be used to check whether children which are independently characterized as powerful within their group do exhibit group-typical power properties.

The second example is a reanalysis of Delbeke's (1978) data concerning preferences for family composition. The objects here are all combinations of *number of sons* and *number of daughters*, ranging from 0 to 3. Thus (2,1) indicates two sons and one daughter, (0,3) no sons and three daughters, and so on. In the theory regarding family composition preferences (cf. Coombs, Mc Clelland and Coombs (1973), two new variables are defined in terms of the old ones, viz. *number of children* and *sex bias* (see table 8). The theory now says, that each subject employs two unimodal utility functions over the natural order of these characteristics and that his overall utility for family

		number of children						
		0	1	2	3	4	5	6
	3				(3,0)			
boy bias	2			(2,0)		(3,1)		
	1		(1,0)		(2,1)		(3,2)	
	0	(0,0)		(1,1)		(2,2)		(3,3)
	-1		(0,1)		(1,2)		(2,3)	
girl bias	-2			(0,2)		(1,3)		
	-3				(0,3)			

Table 8. Family composition: change of variables.

	0	1	2	3	4	5	6
3				13			
2			10		9		
1		15		11		4	
0	16		8		7		2
-1		14		3		1	
-2			6		5		
-3				12			

Table 9. Example of a perfect order for family types.

types may be obtained by simple summation (up to a monotonic transformation). So, if a subject has a sex bias -1 and number bias 5, he might order the

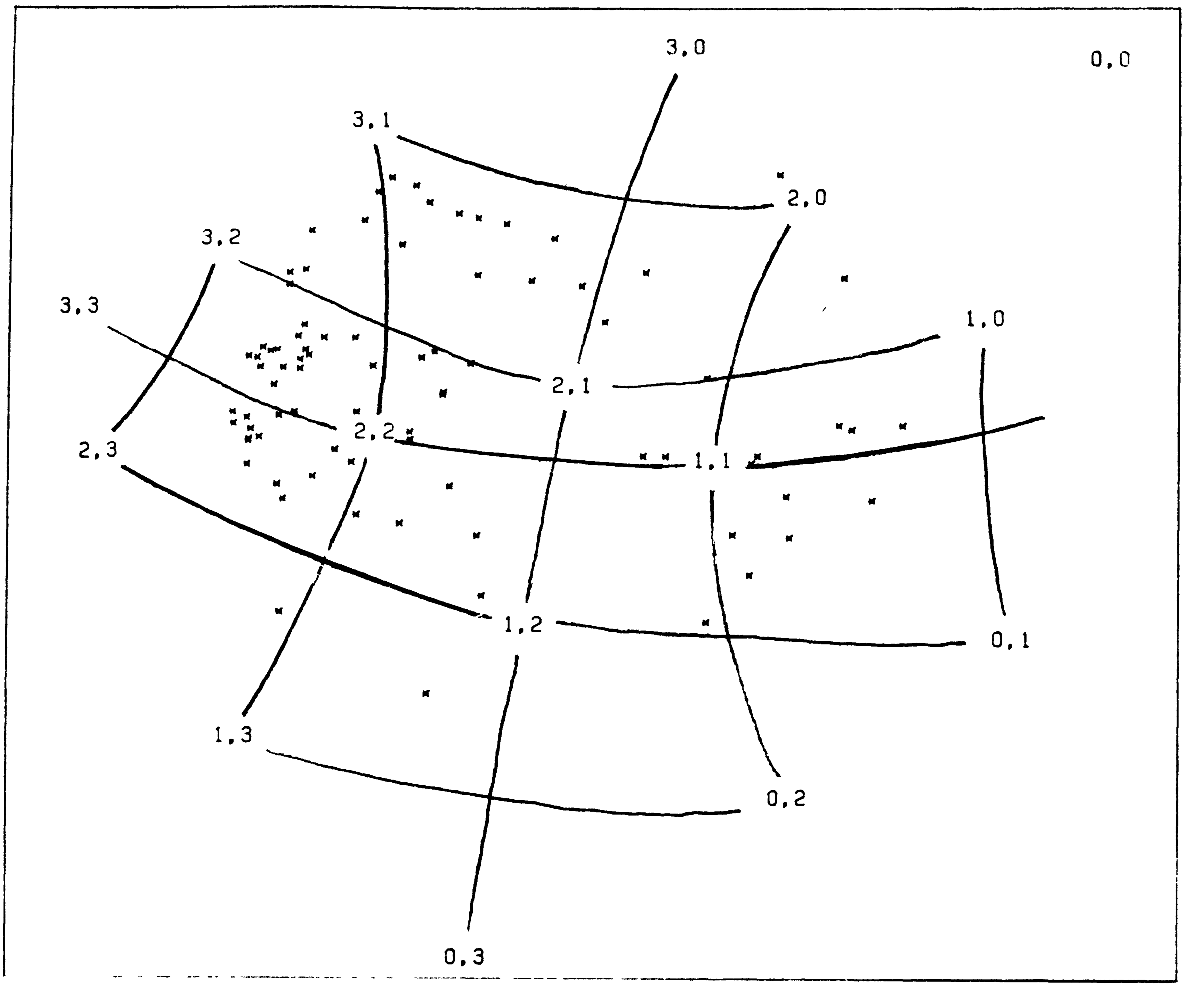

Figure 19. Map of family compositions, obtained with SMACOF-3.

family types as in table 9. Note that there are several 'perfect' orders possible, depending on the scale of the two utility functions. In each column of the table, disutility decreases (and eventually rises again, as in column 3); the same is true for each row.

For this kind of data, we expect unfolding to do well if we are willing to assume that differential weighting of dimensions is neglectable. Delbeke used 82 psychology students at Leuven University as subjects. The result of the SMACOF-3 analysis (stress = 0.1669) is plotted in figure 19. In this figure we have connected the object points with *isobias*- and *isosizecontours*. The expected grid comes out well, except for the point (0,0), which is very unpopular among these belgian students (only 3 first choices of male biased persons). Overall, there is a bias towards larger families and towards males.

We may compare this with the results obtained by Delbeke with the non-metric program MINIRSA (Roskam (1975)), given in figure 20. Here the grid doesn't come out at all; most subjects are clustered together inside

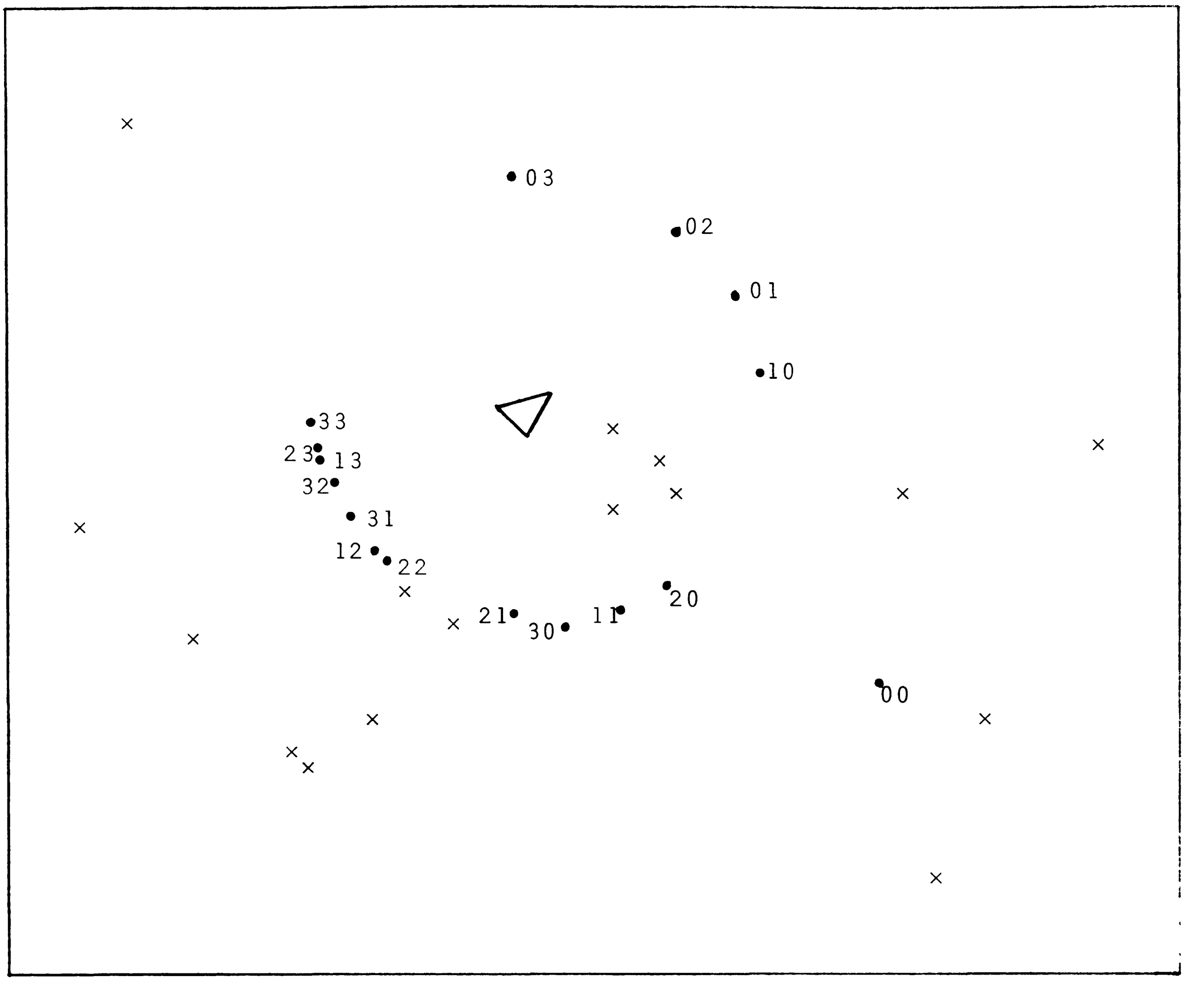

Figure 20. Map of family compositions, obtained with MINIRSA.

the triangle in the middle of the plot and their utilities are (monotonically!) transformed into constants for all objects except (0,0). Stress approaches zero in this case, but the solution is not very informative.

V. PROJECTION TECHNIQUES

The techniques from section IV all tried to map the complete table of utilities simultaneously into two sets of entities, according to some decomposition rule. Our task would be greatly simplified if one of both sets was known in advance. In Thurstonean attitude scaling, for instance, scale values for the object points are determined by a separate experimental procedure and the problem of finding attitude scores for the subjects, given their list of 'endorsements' or utilities, can be simply solved by a weighted least squares procedure. In general, the task that remains is to *project* a vector of utility values into some prescribed subspace. Apart from classical two-step attitude scaling, in which situations do we expect to be interested in connecting utilities with some known configuration?

V.1. Mapping the utilities into a known structure.

Although any respectable researcher 'should know something' about the objects under study, frequently this something is not enough to specify the exact position of the object points in p-dimensional space. That's why we have specific kinds of applications in mind, such as:

- *trade-off studies*: suppose we have a collection of objects known to differ on two negatively correlated desirable traits, e.g. a set of insurance policies different in prize and in cover. We now may want to characterize subjects in terms of *safety bias* on the basis of their utilities.
- *multidimensional psychophysics*: suppose we have a collection of objects chosen as to differ on two physical attributes, e.g. a set of taste mixtures, say alanine and glutamic acid combined in various concentrations, which are to be judged as to their *sweet-sourness*; or a set of odour mixtures, say jasmin and bergamot in various concentrations, to be judged on their *hedonic tone*.
- *impression formation studies*: here the objects are varied on psychological attributes; typically, one confronts the subject with hypothetical 'stimulus persons', differing on, say, intelligence and dominance and asks for a judgment of overall *likeableness*. A large amount of research has been dedicated to the discovery of the rule by which a subject combines different pieces of information into one final impression (see Rosenberg (1968), van der Kloot (1975)).

In all these applications we need not necessarily to assume monotonicity of utility with each of the independent (i.e., varied or selected by the experimenter) variables. Moreover, we will be primarily interested in questions like: "what is the psychological effect of simultaneous variation?" or "what kind of individual differences will turn up under simultaneous variation?". Applications of a slightly different type are:

- *discriminant and convergent validation studies*: suppose we have at our disposal a psychological or cognitive map of the objects (e.g., derived from a previous multidimensional scaling analysis of judged overall similarity); we now may ask ourselves how well the utilities can be connected with this particular configuration (cf. Abelson (1955), Jaspars e.a. (1972)). Another possibility is, that we have a previously derived joint utility space and we want to connect it with background variables of the subjects, or with actual characteristics of the idealized objects (e.g., in the Gold study (see figure 18), we may ask whether children which are independently chosen to be powerful within group B are indeed better at games with running and throwing, at making things, etc. and no good in knowing how to act so that people will like them).

- *cross-validation studies*: we may split up our original sample into two randomly chosen subsamples (or consider two independent samples right away). We then derive a joint utility space for the first (sub)sample and and regard the obtained configuration of object points as fixed for the second one (cf. Bechtel (1976), p. 74 - 77). This provides us with a check whether the obtained configuration does indeed accomodate all possible individual utilities.

In the next section we will consider some elementary techniques for displaying individual utilities in two-dimensional space. After that, we will discuss how to fit some specific models in possibly more dimensions by regression techniques.

V.2. Elementary techniques in two-dimensional space.

The most obvious way to display individual utilities in a known configuration of points is to label all points according to their corresponding utility value. A somewhat nicer representation is obtained if we plot isochrests. Note that, in contrast with the situation in section IV.1, this is no longer trivial as the configuration of points is no longer free to vary. Would the isochrests show up in a disorderly or criss-cross way, this would simply mean that our conjecture about the coherence between utilities and object map is falsified.

To illustrate this procedure, we take the data of one particular subject from a study by van Asten (1979) about the attitude towards differentiation and delegation of tasks in primary schools. The relevant tasks are summarized in table 10. There are two kinds of data: similarities between pairs of tasks and utility ratings for all tasks separately (also given in table 10, for one subject). First we computed an individual cognitive map with the program SMACOF-1, plotted in figure 22; we then drew isochrests, aided by computing for a lot of points, regularly spaced on a grid, the interpolated utility

$$u_k = \sum_{j=1}^{n} \frac{u_j}{a(b + d_{jk}^2)} \tag{43}$$

where a and b are suitably chosen constants. In words, (43) says that the utility of an arbitrary point k in the map may be obtained by a weighted average of the utilities of the fixed points, with the weights inversely re-

BB:	Blackboard writing	(3)
DA:	Domestic care	(5)
DB:	Stimulating desired behaviour	(2)
EA:	Looking after educational appliances	(5)
ET:	Showing expression techniques	(3)
GR:	Building up good relations	(1)
HD:	Practising, hearing and drilling	(6)
IF:	Informing pupils	(1)
IS:	Giving instructions	(2)
KB:	Doing administrative duties	(7)
KO:	Taking measures of order	(6)
LA:	Arranging learning activities	(1)
LI:	Doing library activities	(5)
LM:	Collecting learning material	(1)
LP:	Judging learning achievements	(4)
MA:	Treating post	(7)
SL:	Improving oral language	(2)
ST:	Setting tasks	(1)
SU:	Checking tasks	(4)
TE:	Making phone calls	(7)
TP:	Correcting test papers	(4)
WA:	Watching after the pupils	(6)

Table 10. Tasks used in van Asten (1979); utilities of one subject in parentheses.

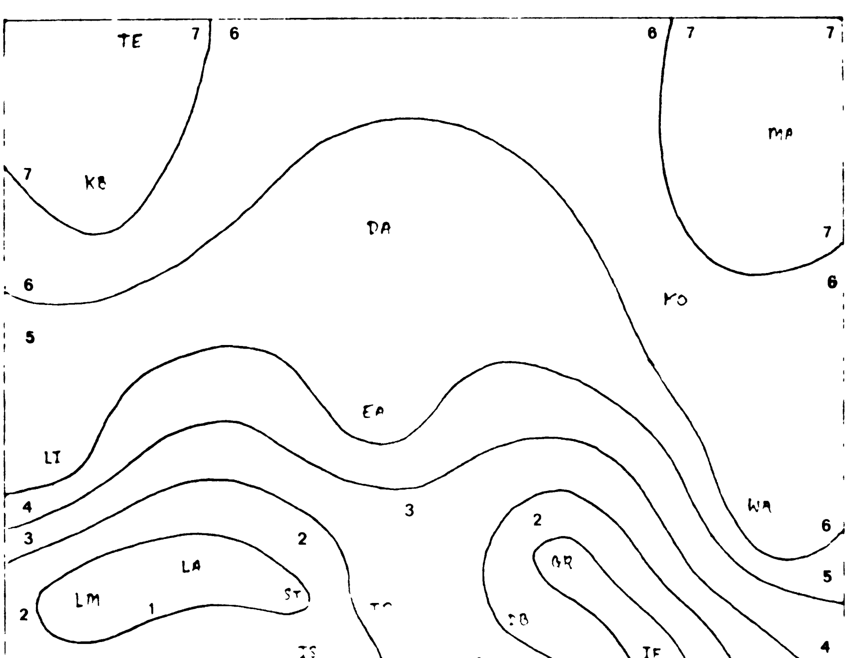

Figure 22. Individual cognitive map of educational tasks.

lated to the distances to these points.

There are several things to note about this rather exhaustive description of an individual case. In the first place, the north/south direction seems to contrast *non-professional* versus *professional* tasks; those on the right/below involve all kinds of *supervising* activities, those on the left/above all kinds of *preliminaries*. The tasks in the centre (LA, ST, IS, SL, TP, SU, DB, GR) apparently are seen as the core of the job. Furthermore, the isochrests indicate that the professional/non-professional distinction is primarily responsible for the differences in utility, but not monotonically (HD and LP may be typical, but not very pleasant). Note that, although the psychological distances between the pairs (LM,LI) and (LM,ST) or (IF,WA) and (IF,GR) are roughly the same, their utility differences are very different; this may be seen as a possible source of *stress* or *cognitive dissonance*. Finally, note that TP and SU are 'out of place'; they lie in an area of utility 3, whereas their actual value is 4, 'disharmony' again.

Whether or not this kind of representation, though plausible, has any practical or theoretical value is an open question. It certainly needs replication, to arrive at reliable maps and stable isochrests.

The second technique we want to demonstrate is to *delineate isotonic regions*, i.e. regions in the map which account for one particular rankorder of utilities. We will utilize a smaller set of dissimilarities and utilities borrowed from Jaspars e.a. (1972); for another secundary analysis of this material, see Bechtel (1976). The purpose of the Jaspars e.a. study was to clarify the development of national stereotypes and attitudes in children, with notions from Heider's theory of cognitive balance. We will use only part of their data here, in an attempt to represent it more thoroughly.

The essential ingredients are again a SMACOF-1 scaling solution and a rankorder of utilities (see figure 23). The objects are: the Netherlands (N) England (E), the United States (A), France (F), the USSR (R) and Germany (G) and the subjects are second-grade Dutch children; their overall rankorder of utility is: N - E - A - F - R - G. According to Jaspars e.a., *nationalism* implies that one's own country is perceived as closely resembling the ideal country. If this is true, it follows that the more a country is perceived as different from one's own country, the less it is preferred over the other countries. This conjecture was checked by computing the correlation between the utility of the five 'other' countries (Thurstone case V values) and the distance from the Netherlands in the cognitive map. For this particular subgroup the correlation is close to zero, which need not surprise us in view

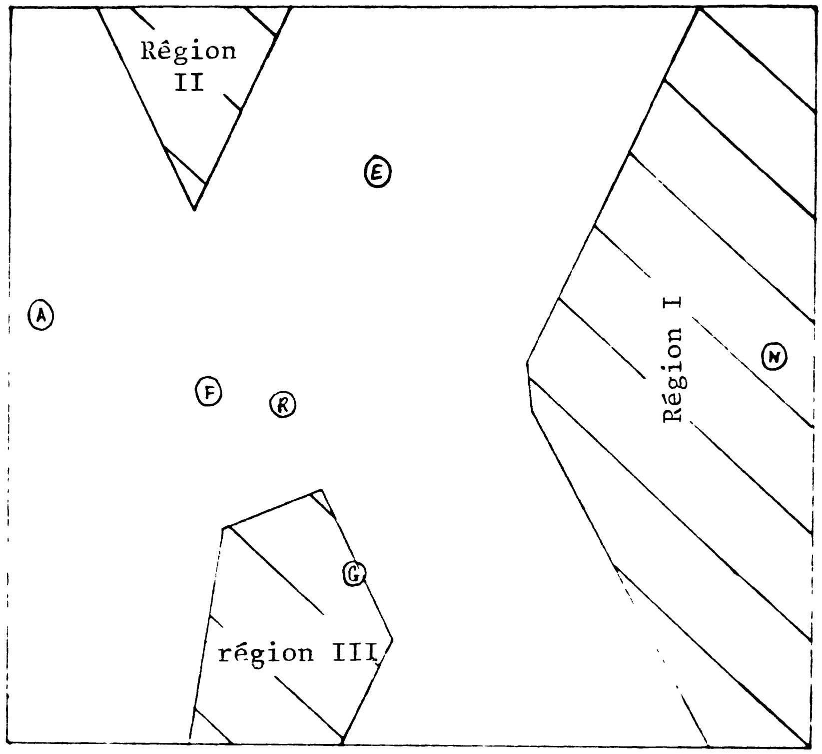

Figure 23. Second-grade children's cognitive map of nations with three isotonic regions

of figure 23. If we just want to describe the rankorder N - E - A - F - R - G in terms of any ideal point or an isotonic region in the cognitive map, we should look somewhere in region I, which is the set of points which are closer to the Netherlands than to any other country. But the remaining part of the utility order (E - A - F - R - G) can be represented perfectly by all points in region II, which is *disjunct* from region I; so we may not hope for a good representation of the complete rankorder (we could say: disregarding the Netherlands, which is chosen first anyway, an ideal point can be located anywhere in region II, reflecting a World War II direction). Alternatively, we could look for a 'perfect' *anti-idealpoint*; it turns out that every point in region III (including Germany) has a rankorder of distances G - R - F - A - E - N, precisely the reverse utility order. This implies that the utility order doesn't reflect nationalism, but *anti-Germanism*.

For the other subgroups in the Jaspars e.a. study, comparable conclusions can be reached with this kind of approach. For larger problems (more objects, more subjects, or both) or if our cognitive map has more dimensions the method gets bothersome and we need a mathematical formulation of the problem.

V.3. Fitting a family of isochrests by linear regression.

In this section we will first discuss the problem of finding one ideal point in a fixed p-dimensional configuration of object points; the procedure can be repeated for any number of ideal points. After that, we will briefly indicate the possibilities of fitting other families of isochrests and discuss some applications in section V.4. A complete account of the present topic can be found in Carroll (1972), who introduced it under the name *external analysis of preference data*, and Bechtel (1976); also, see Davison (1976a,b).

We start with an assumption like (36) in section IV.4, which says that the distance between ideal point i and object point j is a monotone decreasing function (specific for subject i) of the utility of object j. As we are dealing with just one subject here, we suppress reference to the subscript i and specify as our decreasing function:

$$\delta_j = \{ \frac{1}{\alpha} (\beta - u_j) \}^{\frac{1}{2}} \tag{44}$$

where α and β are arbitrary constants (provided that $\alpha > 0$ and $\beta \geq \max (u_j)$). The choice of this particular decreasing function is no coincidence; it allows us to write

$$u_j = \beta - \alpha \, \delta_j^2 \tag{45}$$

and to get rid of the square root in the euclidean decomposition rule (33), by which we get

$$\begin{aligned} u_j &= \beta - \alpha \sum_{a=1}^{p} (x_a - y_{ja})^2 \\ &= \beta - \alpha \{ \sum_{a=1}^{p} x_a^2 + \sum_{a=1}^{p} y_{ja}^2 - 2 \sum_{a=1}^{p} x_a y_{ja} \} \end{aligned} \tag{46}$$

Here the u_j and y_{ja} are known and the α, β and x_a are the unknown parameters. Now the second basic trick in this approach is to introduce the *change of variables:*

$$z_{j0} = 1 \tag{47a}$$

$$z_{ja} = -2 \, y_{ja} \, , \qquad a = 1, \ldots , p \tag{47b}$$

$$z_{j(p+1)} = \sum_{a=1}^{p} y_{ja}^2 \tag{47c}$$

for all j = 1, ... , n and the *reparametrization*:

$$\gamma_0 = \beta - \alpha \sum_{a=1}^{p} x_a^2 \qquad (48a)$$

$$\gamma_a = - \alpha x_a , \qquad a = 1, \ldots , p \qquad (48b)$$

$$\gamma_{p+1} = - \alpha \qquad (48c)$$

which makes it possible, using (46), to arrive at the transformed model

$$u_j = \gamma_0 + \sum_{a=1}^{p} \gamma_a z_{ja} + \gamma_{p+1} z_{j(p+1)} = \sum_{a=0}^{p+1} \gamma_a z_{ja} \qquad (49)$$

This is a set of n non-homogeneous linear equations in p+2 unknowns, which in general has no solution, but may be approximately solved (resorting to the least squares principle again) by standard multiple regression methods. Once estimates of the γ's have been found, (48) may be invoked to find estimates of the original parameters.

We can imagine the z_{ja} as n points in (p+2)-space, in which we want to find a direction (vector) γ such that the projections of the $\{z_{ja}\}$ onto γ are approximately equal to the $\{u_j\}$; i.e., we want to fit the 'vector model' to a fixed set of transformed coordinate values. This way of looking at the approach immediately suggests how we would fit in a vector for each subject instead of an ideal point: by not transforming coordinate values.

Now consider a family of isochrests consisting of concentric ellipses instead of circles: in one direction utility decreases faster than in the other one. In terms of a decomposition rule:

$$\delta_j^2 = \sum_{a=1}^{p} w_a (x_a - y_{ja})^2 \qquad (50)$$

Carroll calls this *the weighted unfolding model*: each subject may weight the axes differently. We can use a change of variables and the corresponding reparametrization again to transform the problem into the form (49); instead of two extra variables we will now get p+1 extra variables in the regression. An even more general model is obtained if we allow each subject to have his own orientation of ellipses: *the general unfolding model*. It is tempting to call this 'Carroll case I', as the number of parameters here easily outgrows the number of independent datavalues. Indeed, Carroll emphasizes the fact that the various models form a *hierarchy*, in which each simpler model is a special case of all the more general ones, obtained by imposing restrictions on their parameters.

This brings us to a final remark concerning the problem of connecting a set of utilities with a known configuration of object points. In general it is possible to combine a decomposition technique with a projection technique: our objective would be to decompose a table of utilities under specific restrictions upon the configuration of points. It seems to us that this approach is the most promising one, because that way we stay in one and the same (weighted least squares) framework, instead of using different ad-hoc methods for different variants of analysis. An example is given at the end of the next section.

V.4. Applications of projection techniques.

Carroll's hierarchy of models is implemented in the program PREFMAP (Carroll and Chang (1967)). Applications (sometimes using other programs) include Green and Rao (1972), Funk e.a. (1974), Bechtel (1976), Davison and Jones (1976), Nygren and Jones (1977), Bechtel (1976), Delbeke (1978) and van Asten (1979).

We will first discuss fitting in vectors, using data from Funk e.a. (1974), concerning stereotypes about ethnic groups in the U.S.A. For this purpose, we use a cognitive map obtained with SMACOF-1, which appears to be more informative than the one obtained by Funk e.a. (an essentially three-cluster structure). The map is presented in figure 24, together with seven directions, computed with PREFMAP, which represent the independently obtained rating scale data in table 11. Fourty-nine University of North Carolina

	activist	affluent	aggressive	emotional	industrious	intelligent	patriotic
AN - Anglo	2.4	3.2	3.2	2.5	3.2	2.3	2.6
BL - Black	3.0	1.4	3.1	2.6	2.1	1.4	1.9
CH - Chinese	1.1	1.9	1.7	1.4	2.8	1.5	2.3
GE - German	1.4	[illegible]	[illegible]	1.8	2.9	1.9	2.8
IN - Indian	1.9	0.7	1.9	1.6	1.9	1.5	1.9
IR - Irish	1.5	2.1	2.3	2.6	2.6	1.9	2.1
IT - Italian	1.6	1.9	2.3	2.9	2.4	1.9	2.0
JA - Japanese	1.3	[illegible]	1.6	1.4	3.2	1.6	2.4
JE - Jewish	1.6	3.2	[illegible]	2.5	3.1	1.7	2.8
ME - Mexican	[illegible]	0.8	2.1	2.4	1.7	1.3	1.6
NE - Negro	2.8	1.3	2.7	2.6	2.1	1.6	1.8
PO - Polish	1.3	1.6	1.6	1.8	2.3	1.7	1.8
PU - Puerto-Rican	1.5	0.9	2.3	2.3	1.8	1.4	1.6

Table 11. Mean ratings of ethnic groups on 7 attributes

students were subjects in this study. The seven attributes were selected so as to cover a wide range of personal impressions. In the figure, the length of the vectors is again proportional to the goodness-of-fit (indicated by multiple correlations here). The attributes seem to fall into two groups

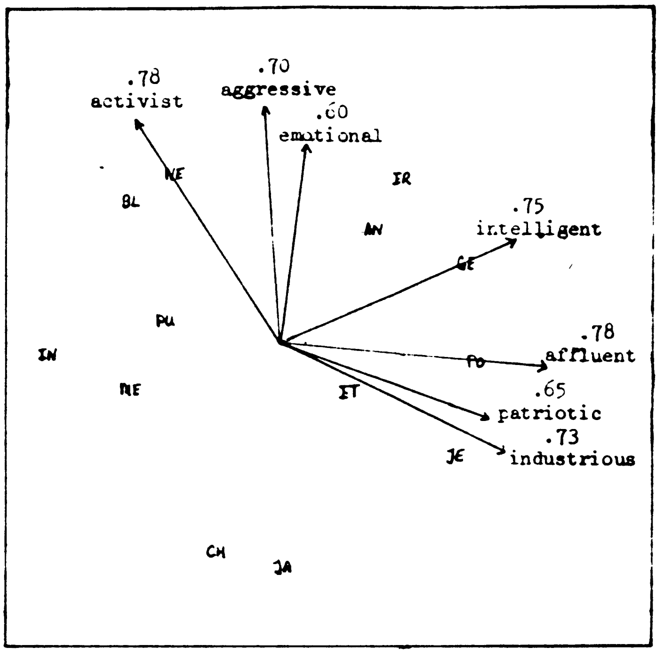

Figure 24. Cognitive map of ethnic groups with seven attribute vectors.

(*activist, aggressive, emotional*) and (*affluent, patriotic, industrious*) in opposite directions, with *intelligent* in between. Note, however, that some groups are high on all attributes (AN), others low everywhere (CH, IN). If we interpret this as an overall judgment affect, we may take deviations from the row means (after standardization of columns to make the scales comparable). The deviation scores are given in table 12. In an attempt to improve fit and interpretability, we use the three-dimensional SMACOF-1 solution (see figure 25). Note that AN for example is typified most strongly now by

	activist	affluent	aggressive	emotional	industrious	intelligent	patriotic
AN - Anglo	-.130	.074	.122	-.229	-.013	.255	-.079
BL - Black	.502	-.219	.401	.163	-.270	-.354	-.224
CH - Chinese	-.225	.166	-.177	-.303	.317	-.039	.260
GE - German	-.322	.165	-.141	-.323	.126	.136	.359
IN - Indian	.249	-.173	.015	-.110	-.088	.041	.066
IR - Irish	-.236	.028	-.047	.165	.002	.173	-.085
IT - Italian	-.182	-.038	-.043	.338	-.102	.177	-.149
JA - Japanese	-.198	.164	-.305	-.374	.461	-.005	.258
JE - Jewish	-.241	.258	-.144	-.044	.118	-.190	.243
MC - Mexican	.281	-.156	.110	.322	-.213	-.186	-.158
NE - Negro	.434	-.222	.207	.194	-.239	-.113	-.262
PO - Polish	-.092	.096	-.199	-.043	.082	.206	-.048
PU - Puerto-Rican	.160	-.142	.201	.244	-.181	-.103	-.180

Table 12. Deviation scores from table 11.

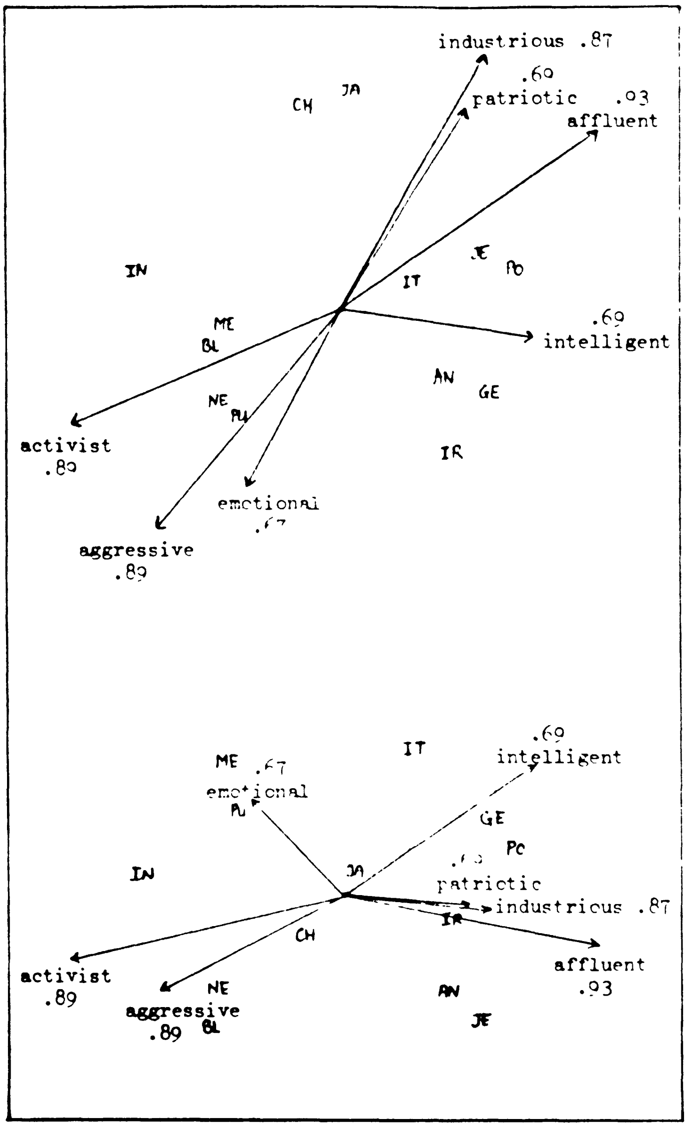

Figure 25. Three-dimensional solution for Funk e.a. data (first dimension horizontal, second vertical above, third vertical below). Vectors: table 12.

intelligent and not so much by *affluent* or *aggressive*, on which others have high scores too. The first two dimensions of the three-dimensional solution are roughly the same as the two-dimensional one, as is true for the attribute directions. The third dimension contrasts ME and PU with NE and BL on the 'coloured' side of space with a typical difference on the attribute *emotional*. It contrasts IT and GE (PO) with AN and JE on the 'white' side, with accompanying effects of *intelligent* and *affluent*.

Our last example concerns fitting in ideal points, following a restricted multidimensional scaling approach (a general treatment of this can be found in de Leeuw and Heiser (1980)). Thus combining a decomposition technique with a projection technique, this approach is in between an 'internal' and an 'external' analysis. We consider Delbeke's data regarding preference for family types again, and utilize the special factorial structure of the objects to reduce the number of free parameters. This is done by adopting the following simple-minded model:

$$F(u_{i(k,\ell)}) = (x_{is} - z_{ks})^2 + (x_{id} - z_{\ell d})^2 \tag{51}$$

with $F(u) = (16 - u)^2$. Here x_{is} is the ideal point of subject i on the 'son axis' and x_{id} his ideal point on the 'daughter axis'; z_{ks} is the common value of having k sons and $z_{\ell d}$ the common value of having ℓ daughters. F(u) is a fixed monotone decreasing transformation for $1 \leq u \leq 16$. Instead of the general theory of section IV.5 we have: "each subject employs two quadratic (single-dipped) disutility functions over some quantification of the variables 'number of sons' and 'number of daughters', and his overall utility for family types may be obtained by simple summation (up to the function F)".

Note that, although all monotonic parts of the general formulation have been replaced by simple specific functions, we have *not* constrained the order of z_{ks} and $z_{\ell d}$; i.e., we do not require

$$z_{0s} \leq z_{1s} \leq z_{2s} \leq z_{3s} \tag{52a}$$

$$z_{0d} \leq z_{1d} \leq z_{2d} \leq z_{3d} \tag{52b}$$

It will be clear that we can fit (51) by performing a metric two-dimensional unfolding analysis with *equality constraints* on the object coordinate values. More specifically, we must constrain Y to have elements

$$y_{(k,\ell)s} = z_{ks} \qquad \text{for all } \ell \tag{53a}$$

$$y_{(k,\ell)d} = z_{\ell d} \qquad \text{for all } k \tag{53b}$$

From the general SMACOF algorithm model it follows that we can achieve this by projecting each unrestricted update (from the basic SMACOF-3 iteration step) onto the space of permissible configurations defined by (53a,b); this space can be made smaller by requiring in addition (52a,b).

With an ad-hoc adapted version of SMACOF-3 we get the result of figure 26. Our restrictions have forced the family points to lie on a grid, but the length of the spikes was still free to vary. The stress (0.2026) is only slightly higher than in the unrestricted case (0.1669), although the number

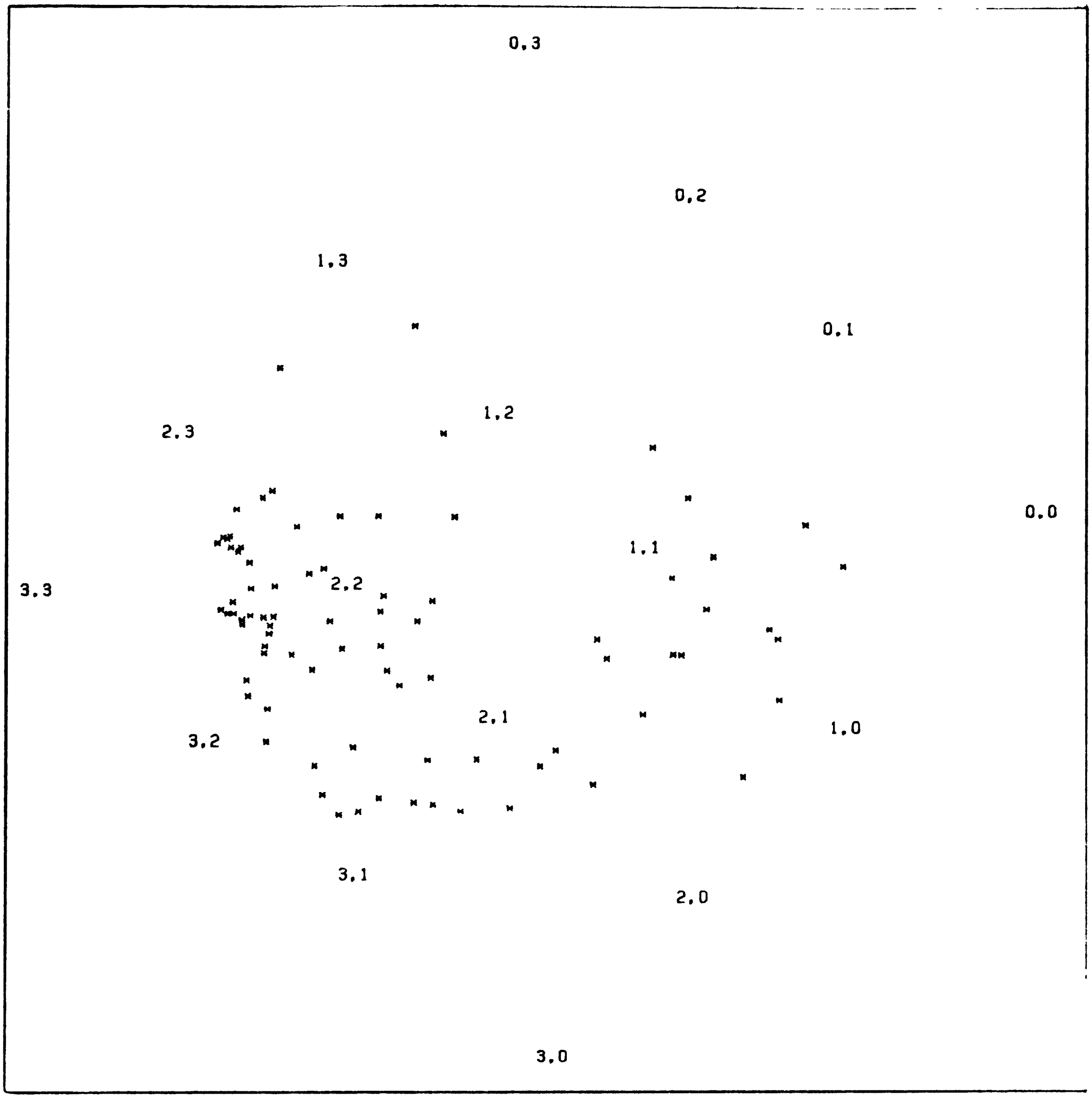

Figure 26. Restricted unfolding of family compositions.

of free parameters for the object points is 8 instead of 32. At the end of the iterative process, the program rotates the joint configuration to its principal axes, so that we may conclude that the subjects differ most on their preference for family size. Just like in figure 19, the son and daughter quantifications turn up in their natural order without being constrained that way. Of course, a similar analysis could be done with 'sex bias' and 'number bias' as basic variables.

We hope that this type of 'confirmatory' scaling analysis, which lacks the rigour of mathematical statistics but enjoys the flexibility which is needed so badly in data analysis, will help researchers in the social sciences to develop closer connections between 'theory' and 'method'.

BIBLIOGRAPHIE

(1) ABELSON R.P., "A technique and model for multidimensional attitude scaling", Public Opinion Quarterly, 19 (1955), 405-418.

(2) ARROW K.J., Social choice and individual values, New York, Wiley, 1970.

(3) ASTEN A. van, Delegeren of niet?, Leiden, Vakgroep Onderwijskunde, 1979.

(4) BECHTEL G.G., Individual differences in the linear multidimensional scaling of choice, mimeograph, Oregon Research Institute, 1969.

(5) BECHTEL G.G., Multidimensional Preference Scaling, The Hague, Mouton, 1976.

(6) BENZECRI J.P., Sur l'analyse des préférences, mimeograph, Institute de statistique de l'université de Paris, 1967.

(7) BENZECRI J.P. e.a., L'analyse des données, tome II. L'analyse des correspondances, Paris, Dunod, 1976.

(8) BLOXOM B., "The simplex in pair comparisons", Psychometrika, 37 (1972), 119-136.

(9) BOCK R.D. and JONES L.V., The measurement and prediction of judgment and choice, San Francisco, Holden-Day, 1968.

(10) BRADLEY R.A., "Science, Statistics and Paired Comparisons", Biometrics, 32 (1976), 213-232.

(11) BRADLEY R.A. and TERRY M.E., "The rank analysis of incomplete block designs, I. The method of paired comparisons", Biometrika, 39 (1952), 324-345.

(12) CARROLL J.D., "Individual differences in multidimensional scaling", In: Shepard R.N., Romney A.K. and Nerlove S.B. (eds.), Multidimensional Scaling: Theory and Applications in the behavioral sciences, vol. I, Theory, New York, Seminar Press, 1972.

(13) CARROLL J.D., "Models and methods for multidimensional analysis of preferential choice (or other dominance) data", In: Lantermann E.D. and Feger H. (eds.), Proceedings of Aachen symposia on decision making and multidimensional scaling, Berlin, Springer Verlag, 1980.

(14) CARROLL J.D. and CHANG J.J., "Nonparametric Multidimensional Analysis of paired comparisons data", Paper presented at the annual meeting of the Psychometric Society, Niagara Falls, 1964.

(15) CARROLL J.D. and CHANG J.J., "Relating preference data to multidimensional scaling solutions via a generalization of Coombs' unfolding model", Paper presented at the annual meeting of the Psychometric Society, Madison, Wisconsin, 1967.

(16) CARROLL J.D. and CHANG J.J., How to use MDPREF, a computer program for multidimensional analysis of preference data, Unpublished report, Bell Telephone Laboratories, 1968.

(17) COOMBS C.H., "A theory of Psychological Scaling", Engineering Research Institute Bulletin, no 23, Ann Arbor, University of Michigan Press, 1952.

(18) COOMBS C.H., A theory of data, New York, Wiley, 1964.

(19) COOMBS C.H. and AVRUNIN G.S., "Single-peaked functions and the theory of preferences", Psych. Rev., 84 (1977), 216-230.

(20) COOMBS C.H. and KAO R.C., "On a connection between factor analysis and multidimensional unfolding", Psychometrika, 25 (1960), 219-231.

(21) COOMBS C.H., MC CLELLAND G.H. and COOMBS L., "The measurement and analysis of family composition preferences", Michigan Mathematical Psychology Program, no 73-5, 1973.

(22) DAALDER H. and VAN DE GEER J.P., "Partijafstanden in de Tweede Kamer", Acta Politica, 3 (1977), 289-345.

(23) DAALDER H. and RUSK J.G., "Perceptions of party in the Dutch Parliament", In: Patterson S.C. and Wahlke J.C. (eds.), Comparative Legislative Behavior: frontiers of research, New York, Wiley, 1972.

(24) DAVID H.A., The method of paired comparisons, London, Griffin, 1963.

(25) DAVIDSON R.R. and FARQUHAR P.H., "A bibliography on the method of Paired Comparisons", Biometrics, 32 (1976), 241-252.

(26) DAVIS O.A., DEGROOT M.H. and HINICH M.J., "Social preference orderings and majority rule", Econometrica, 40 (1972), 147-157.

(27) DAVISON M.L., "Fitting and testing Carroll's weighted unfolding model for preferences", Psychometrika, 41 (1976), 233-247 (a).

(28) DAVISON M.L., "External analyses of preference models", Psychometrika, 41 (1976), 557-558 (b).

(29) DAVISON M.L., "On a metric, unidimensional unfolding model for attitudinal and developmental data", Psychometrika, 42 (1977), 523-548.

(30) DAVISON M.L. and JONES L.E., "A Similarity-Attraction Model for Predicting Sociometric Choice from Perceived Group Structure",

J.of Pers.and Soc.Psych., 33 (1976), 601-612.

(31) DELBEKE L., Construction of preference spaces, Leuven, Studia Psychologica, 1968.

(32) DELBEKE L., Enkele analyses op voorkeuroordelen voor gezinssamenstellingen, Leuven, Centrum voor Mathematische Psychologie en Psychologische Methodologie, 1978.

(33) DE LEEUW J., "Canonical analysis of relational data", Research Notes, no RN010-68, Leiden, Department of Datatheory, 1968.

(34) DE LEEUW J., Canonical analysis of categorical data, Leiden, Psychological Institute, 1973.

(35) DE LEEUW J. and HEISER W.J., "Multidimensional scaling with restrictions on the configuration", In: Krishnaiah P.R. (ed.), Multivariate Analysis V, New York, Academic Press, 1979.

(36) DE LEEUW J., YOUNG F.W. and TAKANE Y., "Additive structure in qualitative data: an alternating least squares method with optimal scaling features", Psychometrika, 41 (1976), 471-504.

(37) DIJKSTRA L., Ontvouwing. Over het afbeelden van rangordes van voorkeur in ruimtelijke modellen, Assen, van Gorcum, 1978.

(38) FECHNER G.T., "Zur experimentellen Aesthetik", Abhandlungen der Königliche Sächsische Gesellschaft der Wissenschaften, Mathematisch-physische Klasse, 9 (1871), 555-635.

(39) FENELON J.P., "Un cas d'analyse des préférences en esthétique", In: Ordres totaux finis, Paris, Gauthier-Villars et Mouton, 1971.

(40) FISHBURN P.C., "Lexicographic orders, utilities and decision rules: a survey", Management Science, 20 (1974), 1442-1471.

(41) FRANZEN O., NORDMARK J. and SJOBERG L., "A study of pitch", Göteborg Psychological Reports, no 12, 1972.

(42) FRENCH J.R.P., "A formal theory of social power", Psych.Rev., 63 (1956) 181-194.

(43) FRENCH J.R.P. and RAVEN B.H., "The basis of social power", In: Cartwright D. (ed.), Studies in social power, Ann Arbor, Michigan, University of Michigan Press, 1959.

(44) FUNK S.G., HOROWITZ A.D., LIPSHITZ R. and YOUNG F.W., "The perceived structure of american ethnic groups: the use of multidimensional scaling in stereotype research", Pers.and Soc.Psych. Bulletin, 1 (1974), 66-68.

(45) GLENN W.A. and DAVID H.A., "Ties in paired-comparison experiments using a modified Thurstone-Mosteller method", Biometrics, 16 (1960), 86-109.

(46) GOLD M., "Power in the classroom", Sociometry, 21 (1958), 50-60.

(47) GOLD E.M., "Metric unfolding: data requirements for unique solution and clarification of Schönemann's algorithm", Psychometrika, 38 (1973), 555-569.

(48) GOOD I.J. and TIDEMAN T.N., "Stirling numbers and a geometric structure from voting theory", J.of Comb.Theory (A), 23 (1977), 34-45.

(49) GREEN P.E. and RAO V.R., Applied Multidimensional Scaling, New York, Holt, Rinehart and Winston, 1972.

(50) GREENBERG M.G., "A modification of Thurstone's Law of comparative judgment to accomodate a judgment category 'equal' or 'no difference'", Psych.Bull., 64 (1965), 108-112.

(51) GULLIKSEN H., "Comparatal dispersion, a measure of accuracy of judgment", Psychometrika, 23 (1958), 137-150.

(52) GUTTMAN L., "An approach for quantifying paired comparisons and rank order", Ann.Math.Statist., 17 (1946), 144-163.

(53) HAYASHI C., Multidimensional quantification of the data obtained by the method of paired comparison", Ann.Inst.Statist.Math., 16 (1964), 231-245.

(54) HEISER W.J. and DE LEEUW J., How-to-use SMACOF-1, Leiden, Department of Datatheory, 1977.

(55) HEISER W.J. and DE LEEUW J., "Metric multidimensional unfolding", Paper presented at the annual meeting of the Psychometric Society, Uppsala, 1978.

(56) HEISER W.J. and DE LEEUW J., How-to-use SMACOF-3, Leiden, Department of Datatheory, 1979.

(57) JASPARS J.M.M., VAN DE GEER J.P., TAJFEL H. and JOHNSON N., "On the development of national attitudes in children", Eur.J.of Soc. Psych., 1972, 347-369.

(58) KRUSKAL J.B. and Carroll J.D., "Geometric models and badness-of-fit functions", In: Krishnaiah P.R. (ed.), Multivariate Analysis II, New York, Academic Press, 1969.

(59) KRUSKAL J.B. and SHEPARD R.N., "A non-metric variety of linear factor analysis", Psychometrika, 39 (1974), 123-157.

(60) LEMAIRE J., "Agrégation typologique de données de préférences", Math. Sci.Hum., 15 (1977), 31-50.

(61) LEVINE J.H., "The sphere of influence", Am.Soc.Rev., 37 (1972), 14-27.

(62) LINGOES J.C. and BORG I., "Indentifying spatial manifolds for intertation", In: Lingoes J.C. (ed.), Geometric representations of

relational data, Ann Arbor, Michigan, 1977.

(63) LUCE R.D., Individual Choice Behaviour, New York, Wiley, 1959.

(64) LUCE R.D., "The choice axiom after twenty years", J.of Math.Psych., 15 (1977), 215-233.

(65) MC CLELLAND G.H. and COOMBS C.H., "ORDMET: a general algorithm for constructing all numerical solutions to ordered metric structures", Psychometrika, 40 (1975), 269-290.

(66) MC DERMOTT B.J., "Multidimensional analyses of circuit quality judgments", J.of the Acoust.Soc.of Am., 45 (1969), 774-781.

(67) MOSTELLER F., "Remarks on the method of paired comparisons I, II and III", Psychometrika, 16 (1951), 3-11, 203-206, 207-218.

(68) MOSTELLER F., "The mystery of the missing corpus", Psychometrika, 23 (1958), 279-289.

(69) NISHISATO S., "Optimal scaling of paired comparison and rank order data: an alternative to Guttman's formulation", Psychometrika, 43 (1978), 263-271.

(70) NOETHER G.E., "Remarks about a paired comparison model", Psychometrika, 25 (1960), 357-367.

(71) NYGREN T.E. and JONES L.E., "Individual differences in perceptions and preferences for political candidates", J.of Exp.Soc.Psych., 13 (1977), 182-197.

(72) PHILLIPS J.P.N., "A note on representation of ordered metric scaling", Brit.J.Math.Stat.Psych., 24 (1971), 239-250.

(73) ROSENBERG S., "Mathematical models of social behaviour", In: Lindzey G. and Aronson E. (eds.), The Handbook of Social Psychology, vol I, Reading, Mass., Addison Wesley, 1968.

(74) ROSKAM E.E., Metric Analysis of ordinal data in Psychology, Voorschoten, VAM, 1968.

(75) ROSKAM E.E., "Non-metric data analysis", Report 75-MA-13, Nijmegen, Department of Psychology, 1975.

(76) ROSS J. and CLIFF N., "A generalization of the interpoint distance model", Psychometrika, 29 (1964), 167-176.

(77) SCHEFFE H., "An analysis of variance for paired comparisons", J.Am. Stat.Acc., 47 (1952), 381-400.

(78) SCHONEMANN P.H., "On metric multidimensional unfolding", Psychometrika, 35 (1970), 349-366.

(79) SELIGSON M.A., "Prestige among Peasants: A Multidimensional Analysis of Preference Data", Am.J.Soc., 83 (1978), 632-652.

(80) [illegible]ERG L., "The Law of Comparative Judgment: a case not assuming

equal variances and covariances", Scandinavian J. of Psych., 3 (1962), 219-225.

(81) SJOBERG L., "Succesive intervals scaling of paired comparisons", Psychometrika, 32 (1967), 297-308.

(82) SJOBERG L., "Choice frequency and similarity", Göteborg Psychological Reports, 23 (1975), University of Göteborg (a).

(83) SJOBERG L., "Uncertainty of comparative judgments and multidimensional structure", Multiv.Beh.Res., 10 (1975), 207-218 (b).

(84) SJOBERG L. and CAPOZZA D., "Preference and cognitive structure of italian political parties", Italian J. of Psych., 2 (1975), 391-402.

(85) SLATER P., "The analysis of personal preferences", Brit.J.Stat.Psych., 13 (1960), 119-135.

(86) SLATER P., "The use of the repertory grid technique in the individual case", Brit.J.Psychiatry, 111 (1965), 965-975.

(87) THURSTONE L.L., "A Law of Comparative Judgment", Psych.Rev., 34 (1927), 278-286.

(88) THURSTONE L.L., "The prediction of choice", Psychometrika, 10 (1945), 237-253.

(89) THURSTONE L.L., The measurement of values, Chicago, Ill., University of Chicago Press, 1959.

(90) TORGERSON W.S., Theory and Methods of Scaling, New York, Wiley, 1958.

(91) TUCKER L.R., "Intra-individual and inter-individual multidimensionality", In: Gulliksen H. and Messick S. (eds.), Psychological Scaling: Theory and Applications, New York, Wiley, 1960.

(92) TVERSKY A., "Elimination by aspects: a theory of choice", Psych.Rev., 79 (1972), 281-299.

(93) VAN DER KLOOT W.A., A cognitive structure approach to person perception, Meppel, Krips, 1975.

(94) VAN RIJCKEVORSEL J. and DE LEEUW J., An outline of PRINCALS, Leiden, Department of Datatheory, 1979.

(95) ZINNES J.L. and GRIGGS R.A., "Probabilistic multidimensional unfolding analysis", Psychometrika, 39 (1974), 327-350.