Title
Metric Unfolding Revisited: Straight Answers to Basic Questions

Authors
Nakanishi, Masao
Cooper, Lee G.

Publication Date
2003

Metric Unfolding Revisited: Straight Answers to Basic Questions

Masao Nakanishi

Kwansei Gakuin University

and

Lee G. Cooper[1]

Anderson School at UCLA

May 2003

[1] Corresponding author: Lee Cooper, Anderson School at UCLA, 110 Westwood Plaza, Suite B518, Los Angeles, CA 90095-1481. Mobile: 310.339.8036, Work: 310.825.4488, Fax: 310.206.7422, E-Mail: lee.cooper@anderson.ucla.edu.

Metric Unfolding Revisited: Straight Answers to Basic Questions

Abstract

Marketing researchers commonly interpret joint-space solutions as if the distances between the points from different sets are meaningful. This is our practice despite appropriate warnings from the authors of joint-space methods that the origin (or metric) of the row objects is not the same as the origin (or metric) of the column objects – making inter-set distances meaningless. We develop a method of metric unfolding where, given only the inter-set judgments, we still retrieve a joint space in which inter-set distances are meaningful. We illustrate this method using: a) a classic car-preference data typically analyzed with MDPref, b) an example involving children's wear in which splitting the stimuli into two groups and collecting inter-set similarities substantially reduces the data collection burden, while providing a readily interpretable perceptual map, c) individual level inter-set judgments of soft drinks to obtain individual level perceptual maps, d) adjective-association data for athletic shoes to produce a joint space for brand image, and e) asymmetric switching data from the Japanese beer market to reflect clout and vulnerability. The ability to properly employ inter-set distances as simple distances greatly facilitates interpretation of these joint-space solutions.

Keywords: Multidimensional Scaling & Classification, Market Structure, Measurement.

INTRODUCTION

We believe that marketing practitioners and academics have lost interest in multidimensional scaling (MDS) in general, and joint-space solutions in particular, in recent years because neither marketing academics nor psychometricians have developed straightforward answers for the most compelling questions. Users look at a perceptual map and interpret the distance between objects as just that – a *distance*. Yet, with most current methods we cannot interpret the distance between an ideal point for an individual and a brand as a simple distance. The metric among brands it typically a distance, but the metric between brands and people is a function of squared distance or a function of angles between a personal preference vector and the brands. When we try to relate the words people use to describe brands to the positions of those brands in a perceptual space, the simple meaning of distance is lost. The origin for the relations between brands is not the same as the origin for the relations between adjectives, leaving the distance between brands and adjectives as undefined. Using MDS to represent brand switching patterns over purchase occasions has not been possible without linguistic gymnastics that leave even sophisticated investigators scratching their heads. The distance of a brand at time one from an average time-one profile compared to the distance of a brand at time two from the average time-two profile, is too convoluted to follow. Thus we take the shortcut, and inappropriately interpret the apparent distance as if it were a real distance.

These basic marketing-research questions inherently involve two sets of objects: people and brands, brands and adjectives describing those brands, or brands at time one versus brands at time two. For instance, when we collect consumers' judgments about their preferences for brands or how they describe brands, these are in reality *inter-set* judgments, since they relate one set of objects (adjectives) with another set of objects (brands). As we describe below, simple distance solutions for these basic inter-set data have never been developed. We want a common space for the objects in both sets, and a simple distance to properly represent what we see in that common spatial map. This work develops that common map, and illustrates its utility representing marketing data from both historic examples and new research.

Background

Most multidimensional scaling (MDS) methods are conceptually based on some measures of distance. Observed similarity (or proximity) measures between objects are assumed to be monotonically

decreasing functions of inter-object distances in some hypothetical space, and attempts are made to recover their coordinates. In those cases in which distances among a set of objects may be assumed to be Euclidean, the object coordinates can be easily recovered by forming the squared-distance matrix, double-centering to produce a matrix of "scalar products," then applying the Eckart-Young (1936) singular-value decomposition to it to find the coordinates of the objects in a Euclidean space (c.f., Torgerson 1958). We note that the major problem of MDS of course lies in the fact that it is difficult to define a clear-cut and accurate measure of distance for most MDS problems. But, since various models have been already developed to deal with problems of this type, we will not pursue them further here.

This paper addresses itself to the multidimensional scaling of partial (or "rectangular") distance matrices, rows and columns of which correspond to different sets of objects: what Carroll and Arabie (1980) classify as two-mode, two-way data. Suppose that Euclidean distances from a set of m objects (row points, ideal points, origins, etc.) to another set of n objects (column points, brands, destinations, etc.) are known. Is it possible to recover the coordinates of entire m+n objects from this partial distance matrix? This is the question that lies at the core of the so-called "unfolding" models of Coombs (1950, 1964). The literature, instead of directly answering it, mostly discusses the conditions (or "constraints") under which estimated object coordinates approximate the observed similarity data. For example, after discussing the constraints for symmetrical square matrices, Shepard (1972, p. 26) summarizes the constraints for rectangular (distance) matrices as follows.

> "Since so much data are missing from the implied 'complete' matrix, constraints are more difficult to formulate. In particular, the conditions having to do with the diagonal (a') and symmetry (b') are no longer available. However, it is still possible to formulate some weaker condition relating to the analog of the triangle inequality (c')".

(In the above paragraph, an implied "complete" matrix means a super-matrix formed by inter-object distances of all m+n objects.) Other writers take a similar stance (cf. Lingoes 1972). We suspect that this tendency to circumvent the question directly stems from a tacit belief that it is unanswerable in a general way.

Schönemann (1970) mounted the first systematic attempt at a metric unfolding solution. Coombs' (1950, 1964) original method was nonmetric (ordinal) unfolding in both the input dominance data and the output scaling space (i.e., the output solution simply ordered the objects and people on each dimension of the final space). While Schönemann laid out the basic algebra (equivalent to our

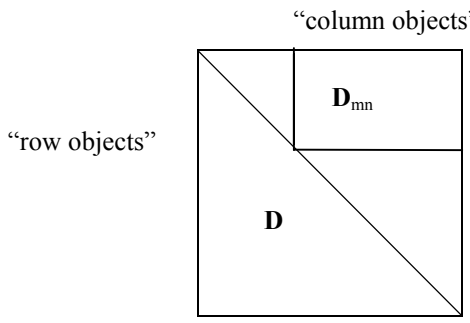
Proposition 1 and 3 below), his efforts fell short on two aspects. First, he only hints at the existence proof that we provide in Proposition 2. Our Proposition 2 provides the key linkages to a robust calibration procedure. And second, his calibration procedure for real data is cumbersome, at best, and is partially wrong, as we describe later.

To his credit, Schönemann approached the inter-set distance problem as a *distance* problem. Most attempts have used surrogates for distance, such as squared distance (DeSarbo and Rao, 1984), or other monotonic transformations of distance (cf. Carroll, 1980). The advantage of a *distance* solution to the inter-set distance problem is that the final joint space has an intuitive interpretation: the not only do the relative distances within each set make sense (object-to-object distances and person-to-person distances), but also the inter-set, person-to-object, distances make the same sense. All entities are related to a common origin with a common metric. The lack of a common origin and metric has been one of the basic obstacles to interpretation in correspondence analysis (Hoffman and Franke, 1986; Hoffman and de Leeuw, 1992). In correspondence analysis the origin for the spatial configuration of row objects is the average row profile across the columns. The origin for the columns is the average column profile across the rows. So while a general understanding is obtained for the relations of row objects to each other and column objects to each other, the inter-set comparison is meaningless.

In the following, we will show that all m+n object coordinates (and consequently all inter-object distances) are recoverable from a partial distance matrix under a condition on the dimensionality (or rank) of the implicit "complete" distance matrix. Even if this condition is not met, we will still be able to use the same theory to obtain the approximate coordinates for all objects.

To be more precise, let **D** be the "complete" distance matrix for all m+n points. **D** is (m+n) by (m+n). Suppose that only those elements of **D** in the first m rows and the last n columns are known. Let this partial distance matrix be called $\mathbf{D}_{mn}$. (See Figure 1.)

Figure 1 **D** and $\mathbf{D}_{mn}$



The main question in this paper is restated as follows: Is it possible to recover the coordinates of all m+n objects from $\mathbf{D}_{mn}$? How ever unlikely it may seem, we show that, by applying the singular value decomposition (SVD) to the "partial" scalar products of $\mathbf{D}_{mn}$ and linearly transforming the results, we can recover the coordinates of both row and column objects, provided that the dimensionality of the complete distance matrix is less than m or n.

In the next section we develop a series of propositions that are useful in proving the existence of linear transformations from the SVD results to object coordinates. We will then discuss some computational problems that are encountered in the recovery of object coordinates. Numerical and practical examples will be presented to show the applicability of the proposed approach.

THEORETICAL CONSIDERATIONS

Recovery of Coordinates from a Full Distance Matrix

We first demonstrate that singular-value decomposition of the scalar products of **D** yields the coordinates of all m+n objects. Form a squared-distance matrix, $\mathbf{D}_2$, by squaring each element of **D**. If we let the true coordinates of objects be **X***, then $\mathbf{D}_2$ may be written in a matrix notation as

$$\boldsymbol{D_2 = diag(X^*X^{*\prime})\, J + J\, diag(X^*X^{*\prime}) - 2X^*X^{*\prime}},$$

where **J** is an (m+n)×(m+n) matrix of all 1s and diag(**X*X***`) denotes the diagonal matrix that is formed by taking the diagonal elements of **X\*X\***`. Considering an origin at the centroid of all m+n point, the scalar product matrix, **S**, of **D** is obtained by double-centering $\mathbf{D}_2$, as follows:

$$\boldsymbol{S = -(I - J/(m+n))\, D_2\, (I - J/(m+n))/2 = (I - J/(m+n))\, X^*X^{*\prime}\, (I - J/(m+n))}, \qquad (1)$$

where **I** is an (m+n)×(m+n) identity matrix.[2]

Equation (1) suggests that (**I** – **J**/(m+n)) **X*** may be recovered by Eigen-decomposing **S** and letting

$$\boldsymbol{X = WE^{1/2}} \qquad (2)$$

where: $\boldsymbol{W}$ = Eigen vectors of **S**

$\boldsymbol{E}$ = the diagonal matrix of positive Eigen-values of **S**.

Obviously **S** = **XX**`. We shall call **X** the "Eigen" solution for the true coordinates. If h is the number of positive Eigen-values of **S**, then **X** is an (m+n)×h matrix. **X** is not necessarily identical to **X*,** nor

---

[2] This result was derived by Tucker, Green, and Abelson, and reported by Torgerson (1958), p. 258.

unique in the sense that they are determined up to an arbitrary orthogonal rotation. But, unless **X*** is known a priori, we are obliged to treat **X** as the "true" coordinates in the following sections**.**

Relationship between Full and Partial Scalar Products Matrices

Let $\mathbf{X}_1$ be the first m rows and $\mathbf{X}_2$ be the last n rows of **X**, respectively. **S** may be partitioned as

$$\boldsymbol{S} = \begin{matrix} \boldsymbol{X_1X_1}` & \boldsymbol{X_1X_2}` \\ \boldsymbol{X_2X_1}` & \boldsymbol{X_2X_2}` \end{matrix} .$$

The partial squared-distance matrix, $\mathbf{D}_{2mn}$, which is formed by squaring each element of $\mathbf{D}_{mn}$, may be re-written in a matrix notation as follows.

$$\boldsymbol{D_{2mn}} = diag(\boldsymbol{X_1X_1}`)\, \boldsymbol{J_{mn}} + \boldsymbol{J_{mn}}\, diag(\boldsymbol{X_2X_2}`) - 2\boldsymbol{X_1X_2}`, \qquad (3)$$

where $\mathbf{J}_{mn}$ is an (m×n) matrix of 1s. We define the partial scalar product matrix, S_{mn} , as follows.

$$\boldsymbol{S_{mn}} = -(\boldsymbol{I_m} - \boldsymbol{J_m}/m)\boldsymbol{D_{2mn}}(\boldsymbol{I_n} - \boldsymbol{J_n}/n)/2 \qquad (4)$$

where: $\boldsymbol{D_{2mn}}$ = an m×n matrix formed by squaring each element of $\mathbf{D}_{mn}$

$\boldsymbol{I_m}$, $\boldsymbol{I_n}$ = an m×m and n×n identity matrix, respectively,

$\boldsymbol{J_m}$, $\boldsymbol{J_n}$ = an m×m and n×n matrix of 1s', respectively.

Our next task is to show the relationship between **S** and $\mathbf{S}_{mn}$.

Proposition 1

$$\boldsymbol{S_{mn}} = (\boldsymbol{I_m} - \boldsymbol{J_m}/m)\, \boldsymbol{X_1X_2}`(\boldsymbol{I_n} - \boldsymbol{J_n}/n) .$$

Proof of Proposition 1: From equations (3) and (4), we have

$$\begin{aligned} \boldsymbol{S_{mn}} &= -(\boldsymbol{I_m} - \boldsymbol{J_m}/m)\boldsymbol{D_{2mn}}(\boldsymbol{I_n} - \boldsymbol{J_n}/n)/2 \\ &= -(\boldsymbol{I_m} - \boldsymbol{J_m}/m)[diag(\boldsymbol{X_1X_1}`)\, \boldsymbol{J_{mn}} + \boldsymbol{J_{mn}}\, diag(\boldsymbol{X_2X_2}`) - 2\boldsymbol{X_1X_2}`](\boldsymbol{I_n} - \boldsymbol{J_n}/n)/2 \\ &= (\boldsymbol{I_m} - \boldsymbol{J_m}/m)\, \boldsymbol{X_1X_2}`\, (\boldsymbol{I_n} - \boldsymbol{J_n}/n) . \end{aligned}$$

The last equation holds because

$$\boldsymbol{J_{mn}}\, (\boldsymbol{I_n} - \boldsymbol{J_n}/n) = (\boldsymbol{I_m} - \boldsymbol{J_m}/m)\, \boldsymbol{J_{mn}} = \boldsymbol{0}$$ (= a matrix of all 0s).

The most important implication of Proposition 1 is that the "true" (or Eigen) coordinates $\mathbf{X}_1$ and $\mathbf{X}_2$ *cannot* be recovered by simply applying the SVD to $\mathbf{S}_{mn}$. Because $\mathbf{X}_1$ and $\mathbf{X}_2$ are separately "centered" (that is, the column means are subtracted from each value), the locations of $\mathbf{X}_1$ and $\mathbf{X}_2$ relative

to the true origin are unknown. This leads to the interpretability problem for the joint space in correspondence analysis.

Existence of Linear Transformations

Next we proceed to the proof of the proposition that $\mathbf{X}_1$ and $\mathbf{X}_2$ may be recovered by linear-transforming the SVD results of $\mathbf{S}_{mn}$. We write the SVD results as

$$\boldsymbol{S}_{mn} = \boldsymbol{ULV}`$$

where **U**, **V** = the left and right orthogonal vectors, respectively

L = the diagonal matrix of singular values.

We will call the following coordinates, which are computed from the SVD results, the "SVD" solution.

$$\boldsymbol{Y}_1 = \boldsymbol{UL}^{1/2}$$

$$\boldsymbol{Y}_2 = \boldsymbol{VL}^{1/2} \qquad (5)$$

Since only positive singular values in **L** are used in (5), $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are an m×k and an n×k matrix, respectively, where k is the number of positive elements in **L**. Clearly, $\mathbf{S}_{mn} = \mathbf{Y}_1\mathbf{Y}_2`$. Unfortunately, $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are not good substitutes for $\mathbf{X}_1$ and $\mathbf{X}_2$ for the present purposes, because in general they do not reproduce $\mathbf{D}_{mn}$ correctly.

In the following analysis we use the fact that $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are linear functions of $\mathbf{X}_1$ and $\mathbf{X}_2$ when k = h. If this were true, then the reverse must be also true. From Proposition 1, we have

$$\boldsymbol{S}_{mn} = \boldsymbol{(I_m - J_m/m)\, X_1 X_2`(I_n - J_n/n)}\ ,$$

and hence

$$\boldsymbol{S_{mn}S_{mn}`} = \boldsymbol{(I_m - J_m/m)\, X_1\, [X_2`(I_n - J_n/n)(I_n - J_n/n)\, X_2]X_1`(I_m - J_m/m)}\ .$$

Since the matrix within [ ] in the above equation is an h×h symmetric matrix, it can be factored as

$$\boldsymbol{X_2`(I_n - J_n/n)(I_n - J_n/n)\, X_2 = F_1F_1`},$$

where $\mathbf{F}_1$ is an h×h matrix. That

$$\boldsymbol{S_{mn}S_{mn}`} = \boldsymbol{(I_m - J_m/m)\, X_1\, [F_1F_1`]X_1`\, (I_m - J_m/m) = UL^2U`},$$

suggests that we may write

$$\boldsymbol{Y}_1 = \boldsymbol{UL}^{1/2} = \boldsymbol{(I_m - J_m/m)\, X_1F_1\, L^{-1/2}}\ . \qquad (6)$$

It is clear that $\mathbf{Y}_1$ is a linear function of $\mathbf{X}_1$.

Similarly, if we replace the h×h symmetric matrix $[\mathbf{X}_1`\ (\mathbf{I}_m - \mathbf{J}_n/m)\ (\mathbf{I}_n - \mathbf{J}_n/m)\ \mathbf{X}_1]$ by $\mathbf{F}_2\mathbf{F}_2`$, where $\mathbf{F}_2$ is an h×h matrix, we have

$$\boldsymbol{S}_{mn}\boldsymbol{`S}_{mn} = (\boldsymbol{I}_n - \boldsymbol{J}_n/n)\,\boldsymbol{X}_2\,[\boldsymbol{F}_2\boldsymbol{F}_2\boldsymbol{`}]\,\boldsymbol{X}_2\boldsymbol{`}\,(\boldsymbol{I}_n - \boldsymbol{J}_n/n) = \boldsymbol{V}\boldsymbol{L}^2\boldsymbol{V}\boldsymbol{`}.$$

One may write $\mathbf{Y}_2$ as a linear function in $\mathbf{X}_2$ as follows.

$$\boldsymbol{Y}_2 = \boldsymbol{V}\boldsymbol{L}^{1/2} = (\boldsymbol{I}_n - \boldsymbol{J}_n/n)\,\boldsymbol{X}_2\boldsymbol{F}_2\,\boldsymbol{L}^{-1/2} \qquad (7)$$

The following proposition is the consequence of the above derivations.

Proposition 2

There exist linear transformations from the SVD solutions, $\mathbf{Y}_1$ and $\mathbf{Y}_2$, to the "true" (or Eigen) coordinates, $\mathbf{X}_1$ and $\mathbf{X}_2$, provided that the number of columns in **Y** and **X** are equal (i.e., k = h).

Proof of Proposition 2: Rewrite equations (6) and (7) as,

$$\boldsymbol{Y}_1 = (\boldsymbol{X}_1 - \boldsymbol{J}_m\,\boldsymbol{X}_1/m)(\boldsymbol{F}_1\boldsymbol{L}^{-1/2})$$

$$\boldsymbol{Y}_2 = (\boldsymbol{X}_2 - \boldsymbol{J}_n\,\boldsymbol{X}_2/n)(\boldsymbol{F}_2\boldsymbol{L}^{-1/2})\ .$$

Taking inverse transformations from **Y**s to **X**s, we have the desired results.

$$\boldsymbol{X}_1 = \boldsymbol{Y}_1\,\boldsymbol{A}_1 + \boldsymbol{B}_1$$

$$\boldsymbol{X}_2 = \boldsymbol{Y}_2\,\boldsymbol{A}_2 + \boldsymbol{B}_2$$

where: $\boldsymbol{A}_1 = (\boldsymbol{F}_1\boldsymbol{L}^{-1/2})^{-1}$

$$\boldsymbol{A}_2 = (\boldsymbol{F}_2\boldsymbol{L}^{-1/2})^{-1}$$

$$\boldsymbol{B}_1 = \boldsymbol{J}_m\,\boldsymbol{X}_1/m$$

$$\boldsymbol{B}_2 = \boldsymbol{J}_n\,\boldsymbol{X}_2/n\ .$$

The proviso is necessary because the number of columns of **Y**, k, may be less than the number of columns of **X**, h, in many applications. When k < h, it is obvious that $\mathbf{X}_1$ and $\mathbf{X}_2$ cannot be exactly reproduced from $\mathbf{Y}_1$ and $\mathbf{Y}_2$. Those cases for which k < h (to be called "non-exact" cases) will be dealt with in a later section.

Proposition 2 gives us an important clue as to the recovery of the "true" (or Eigen) coordinates, **X**. Since we now know that $\mathbf{X}_1$ and $\mathbf{X}_2$ are linear functions in $\mathbf{Y}_1$ and $\mathbf{Y}_2$, we first obtain the SVD solution from $\mathbf{S}_{mn}$. Then we calibrate the linear parameters ($\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{B}_1$, and $\mathbf{B}_2$) in such a way that the partial distance matrix, $\mathbf{D}_{mn}$, is correctly reproduced. But, before we move to the calibration procedure in the next section, we will prove another proposition that is useful in reducing the number of parameters to be calibrated.

Constraints on Parameters

At present the total number of parameters in the linear transformations $\mathbf{A}_1$(k×k), $\mathbf{A}_2$(k×k), $\mathbf{B}_1$(m×k) and $\mathbf{B}_2$(n×k) is (2h+m+n)×k, where k is the number of positive singular values in **L**. (We are still assuming that k = h.) Even for a small k (3, say), the m×n elements of $\mathbf{D}_{mn}$ may be too small to calibrate all the parameters. Fortunately, there are two important constraints on $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{B}_1$, and $\mathbf{B}_2$ that work to reduce the number of parameters.

From Proposition 1, we have

$$(\boldsymbol{I}_m - \boldsymbol{J}_m/m)\, \boldsymbol{X}_1\boldsymbol{X}_2`(\boldsymbol{I}_n - \boldsymbol{J}_n/n) = \boldsymbol{S}_{mn} = \boldsymbol{Y}_1\, \boldsymbol{Y}_2` .$$

Hence the parameters are constrained in such a way that

$$(\boldsymbol{I}_m - \boldsymbol{J}_m/m)\, (\boldsymbol{Y}_1\boldsymbol{A}_1 + \boldsymbol{B}_1)\, (\boldsymbol{Y}_2\boldsymbol{A}_2 + \boldsymbol{B}_2)`\, (\boldsymbol{I}_n - \boldsymbol{J}_n/n) \;\equiv\; \boldsymbol{Y}_1\, \boldsymbol{Y}_2` . \qquad (8)$$

But

$$(\boldsymbol{I}_m - \boldsymbol{J}_m/m)\, \boldsymbol{Y}_1\boldsymbol{A}_1 = \boldsymbol{Y}_1\boldsymbol{A}_1$$

$$(\boldsymbol{I}_n - \boldsymbol{J}_n/n)\, \boldsymbol{Y}_2\boldsymbol{A}_2 = \boldsymbol{Y}_2\boldsymbol{A}_2$$

$$(\boldsymbol{I}_m - \boldsymbol{J}_m/m)\, \boldsymbol{B}_1 = (\boldsymbol{I}_n - \boldsymbol{J}_n/n)\, \boldsymbol{B}_2 = \boldsymbol{0}$$ (= a matrix of all 0s).

These relationships hold because column means of $\mathbf{Y}_1$ are 0 and all rows of $\mathbf{B}_1$ are equal. The same is true for $\mathbf{Y}_2$ and $\mathbf{B}_2$. After the simplification of the left-hand side, equation (8) becomes

$$\boldsymbol{Y}_1\boldsymbol{A}_1\boldsymbol{A}_2`\boldsymbol{Y}_2` \;\equiv\; \boldsymbol{Y}_1\, \boldsymbol{Y}_2` .$$

In order for the above constraint to hold, the following relationship must exist between $\mathbf{A}_1$ and $\mathbf{A}_2$.

$$\boldsymbol{A}_2 = (\boldsymbol{A}_1^{-1})`$$

This shows that we need to calibrate only one k×k matrix instead of two.

Also, there is a constraint placed on $\mathbf{B}_1$ and $\mathbf{B}_2$. Since the overall column sums of **X** are 0 (because the sums of Eigen-vectors are 0 in Equation (2)), it is necessary that the column sums of $\mathbf{X}_1$ and $\mathbf{X}_2$ must sum to 0, that is,

$$\boldsymbol{1}_m\, \boldsymbol{X}_1 + \boldsymbol{1}_n\, \boldsymbol{X}_2 = \boldsymbol{0}$$ (= a vector of all 0s)

where $\mathbf{1}_m$ and $\mathbf{1}_n$ are an (1×m) and an (1×n) matrix (row vector) of 1s, respectively. From the definition of the **B**s in Proposition 2,

$$\boldsymbol{1}_m\, \boldsymbol{B}_1 + \boldsymbol{1}_n\, \boldsymbol{B}_2 = \boldsymbol{1}_1\, \boldsymbol{J}_m\, \boldsymbol{X}_1/m + \boldsymbol{1}_n\, \boldsymbol{J}_m\, \boldsymbol{X}_2/n = \boldsymbol{1}_m\, \boldsymbol{X}_1 + \boldsymbol{1}_n\, \boldsymbol{X}_2 = \boldsymbol{0} \qquad (9)$$

because $\mathbf{1}_m\, \mathbf{J}_m /m = \mathbf{1}_m$ and $\mathbf{1}_n\, \mathbf{J}_n /n = \mathbf{1}_n$.

We summarize the above results in the following proposition.

Proposition 3

The parameters of linear transformations defined in Proposition 2 are constrained in the following manner.

(a) $\boldsymbol{A}_2 = (\boldsymbol{A}_1^{-1})`$

(b) $\boldsymbol{1}_m \boldsymbol{B}_1 = -\boldsymbol{1}_n \boldsymbol{B}_2$

Proposition 3 reduces the number of parameters to be calibrated greatly. The constraint on $\mathbf{B}_1$ and $\mathbf{B}_2$ suggests that we should let $\mathbf{b} = \mathbf{1}_{1m}\mathbf{X}_1$ and construct $\mathbf{B}_1$ such that it's every row is equal to $\mathbf{b}/m$. Similarly, let every row of $\mathbf{B}_2$ be $-\mathbf{b}/n$. In this manner constraint (9) is satisfied and the number of parameters in $\mathbf{B}_1$ and $\mathbf{B}_2$ to be calibrated is reduced from (m+n)×k to k. $\mathbf{A}_1$ contains k×k elements and $\mathbf{B}$ contains k elements, and hence the total number of parameters to be calibrated is now only (k+1)×k. In the next section we will discuss the calibration procedure using the m×n elements in $\mathbf{D}_{mn}$.

CALIBRATION PROCEDURE

Formulation of Problem: Exact Solution Case

We first consider the cases for which k = h, that is, the number of positive Eigen-values in $\mathbf{E}$ is equal to the number of positive singular values in $\mathbf{L}$. We will call those cases "exact," because Proposition 2 tells us that in those cases there exist exact linear transformations from the SVD solution to the Eigen solution. It remains for us to find the right linear transformations that reproduce $\mathbf{D}_{mn}$ correctly.

The calibration procedure for the exact solution case may be formally stated as follows.

Minimization Problem

$$\text{Minimize } Q = \Sigma_i \Sigma_j (d_{ij} - \hat{d}_{ij})^2 \text{ with respect to } \boldsymbol{A} \text{ and } \boldsymbol{b} \qquad (10)$$

where: d_{ij} = the (i, j) th element of $\mathbf{D}_{mn}$

$\hat{d}_{ij}$ = square root of the (i, j) th element of $\mathbf{D}_{2mn}$

$$\boldsymbol{D}_{2mn} = diag(\hat{\boldsymbol{X}}_1\hat{\boldsymbol{X}}_1`)\,\boldsymbol{J}_{mn} + \boldsymbol{J}_{mn}\,diag\,(\hat{\boldsymbol{X}}_2\hat{\boldsymbol{X}}_2`) - 2\hat{\boldsymbol{X}}_1\hat{\boldsymbol{X}}_2`$$

$$\hat{\boldsymbol{X}}_1 = \boldsymbol{Y}_1\boldsymbol{A} + \boldsymbol{J}_{mk}\,diag(\boldsymbol{b}/m)$$

$$\hat{\boldsymbol{X}}_2 = \boldsymbol{Y}_2\,(\boldsymbol{A}^{-1})` - \boldsymbol{J}_{nk}\,diag(\boldsymbol{b}/n)$$

and J_{mk} and J_{nk} are an m×k and an n×k matrix of all 1s', respectively.

It may be possible to derive the first- and second-order derivatives of the objective function (10) with respect to **A** and **b**, and analytically solve for their optimal values. However, since Proposition 2 guarantees the existence of **A** and **b**, it is much simpler to use a numerical analysis algorithm (such as Excel Solver) and directly minimize Q to find the optimal values.

A numerical example is given to illustrate the proposed approach. Table 1 gives the 2-dimensional coordinates of seven objects and suppose that they represent the Eigen solution, **X,** of equation (2).

Table 1 True Coordinates(**X**)

	Dimensions	
Objects	1	2
1	-1.414	0.000
2	0.000	-1.414
3	1.414	0.000
4	-1.414	-1.414
5	-1.414	1.414
6	0.707	2.121
7	2.121	-0.707

Table 2 Partial Squared-Distances ($\mathbf{D}_{mn}$)

Objects	4	5	6	7
1	2	2	9	13
2	2	10	13	5
3	10	10	5	1

Table 3 Partial Scalar Products ($\mathbf{S}_{mn}$)

Objects	4	5	6	7
1	1.167	2.5	-0.167	-3.5
2	1.667	-1.0	-1.667	1.0
3	-2.833	-1.5	1.833	2.5

Table 2 gives the partial squared-distance matrix, $\mathbf{D}_{2mn}$ (rather than $\mathbf{D}_{mn}$), between the first 3 objects and the last 4 objects. We are to recover the coordinates in Table 1 from the knowledge of $\mathbf{D}_{2mn}$.

First we compute the partial scalar products matrix, $\mathbf{S}_{mn}$ (shown in Table 3), and derive the SVD solution from it (Table 4). Next we calibrate **A** and **b** by solving the minimization problem stated above. Table 6 gives the calibrated values of **A** and **b**. The recovered coordinates (Table 5) are indeed identical to the Eigen solution, except that the direction of both axes is reversed in the former. In practical applications this is not a problem because the Eigen solution itself is rotationally indeterminate. The important point is that we recovered a set of coordinates, which reproduce $\mathbf{D}_{mn}$ perfectly. We will call the resultant coordinates of **X** the "Linear" solution for obvious reasons.

Table 4 SVD Solution (**Y**)

	Dimensions	
Objects	1	2
1	1.732	-0.749
2	0.000	1.498
3	-1.732	-0.749
4	1.155	1.113
5	1.155	-0.668
6	-0.577	-1.113

Table 5 Linear Solution ($\hat{\mathbf{X}}$)

	Dimensions	
Objects	1	2
1	1.414	0.000
2	0.000	1.414
3	-1.414	0.000
4	1.414	1.414
5	1.414	-1.414
6	-0.707	-2.121

7	-1.732	0.668	7	-2.121	0.707

Table 6 Calibrated Parameters

	Dimensions			Dimensions	
A	1	2	B	1	2
	0.81650	0		0	1.41421
	0	0.62933			

Some Issues in Calibration

In obtaining the parameters in Table 6, we encountered a number of minor problems that could not be ignored entirely.

1. Recovery of Partial Scalar Products: Depending on the method used to obtain the SVD solution, $\mathbf{Y}_1$ and $\mathbf{Y}_2$ may not reproduce the partial scalar product matrix, $\mathbf{S}_{mn}$, correctly, and it may become necessary to change the signs of some columns of either $\mathbf{Y}_1$ or $\mathbf{Y}_2$. It is advisable to check the recovery of partial scalar products before other steps in the calibration procedure are taken.
2. Initial Values for Iteration: Almost all numerical optimization routines require initial values for parameters to be calibrated. It was found that the proposed minimization problem was rather sensitive to the choice of initial values, so much so that initial values determined the calibration outcomes. An easy choice for the initial values for the example problem would be an identity matrix for **A** and a zero vector for **b**, but these initial values did not lead to the global optimum (namely, 0) for the objective function. After an exhaustive grid search for the right combination of initial values, we discovered that they were limited to only a very small region around the optimum values of **A** and **b** in Table 6.

   A useful strategy in searching the right combination of initial values is as follows.

   (1) Fit **b** first, using an identity matrix for **A**.

   (2) Let the diagonal values of **A** be diag(**A**). Try various initial values for diag(**A**), holding the values of **b** obtained at the first step. The search range between 0 and 2 for the elements of diag(**A**) seems adequate in most cases.

   This strategy seems to work because the discrepancies between **X** and **Y** are due in large part to separate centering operations (i.e., subtraction of the respective means) applied to $\mathbf{X}_1$ and $\mathbf{X}_2$ to obtain $\mathbf{Y}_1$ and $\mathbf{Y}_2$.
3. Orthogonality of Recovered Coordinates: The dimensions of the recovered coordinates (i.e., the

Linear solution) must be orthogonal because those of the Eigen solution are orthogonal. Hence it is necessary to find an orthogonal rotation, **T**, such that $\mathbf{T}'(\hat{\mathbf{X}}'\hat{\mathbf{X}})\mathbf{T}$ is a diagonal matrix, where $\hat{\mathbf{X}}$ represents the coordinates computed from calibrated values of **A** and **b**. This is easily accomplished by Eigen-decomposing $\hat{\mathbf{X}}'\hat{\mathbf{X}}$ and taking the Eigen vectors as **T**. Post-multiplication of $\hat{\mathbf{X}}$ by **T** results in the Eigen solution. Note that post-multiplying $\hat{\mathbf{X}}$ by **T** does not affect the reproduction of $\mathbf{D}_{mn}$.

Non-exact Solution Case

It is unreasonable to think that the assumption that k = h always holds. In many applications the number of positive Eigen-values, h, (that is, the number of columns for the Eigen solution) may be greater than min(m, n). Furthermore, h is not usually known beforehand in practical applications. In those cases for which k is suspected to be less than h (i.e., in "non-exact" cases), we cannot expect that **X** will be exactly recovered from **Y**. However, we may still apply the calibration procedure developed for exact cases because the Linear solution, $\hat{\mathbf{X}}$**,** may be considered as a least-squares estimate of **X** in non-exact cases. In the following we discuss some issues that are associated with non-exact cases.

1. Since the Linear solution, $\hat{\mathbf{X}}$**,** represents at best an approximate solution when k < h, after the Linear solution is computed, adjustments will have to be made to the resultant $\hat{\mathbf{X}}$ in order to increase their fit to $\mathbf{D}_{mn}$. We propose that a useful procedure in this situation is to fit the coordinates of row objects and those of column objects to $\mathbf{D}_{mn}$ alternately. The objective function (Equation (10)) is the same as the minimization problem (i.e., Equation 10) but the variables to be optimized in this case are object coordinates. Q in (10) is minimized with respect to $\mathbf{X}_1$ first, holding the values of $\mathbf{X}_2$ constant. Then the optimal values of $\mathbf{X}_2$ are computed, holding the current values of $\mathbf{X}_1$ constant, and so on, until there is no more improvement in Q. Because of the structure of the objective function, this process amounts to the optimization of the coordinates of each object in turn, holding the coordinates of all other objects constant. The minimization task is not a great one and can be performed easily with a numerical optimization algorithm such as Excel Solver. Our experience suggests that this alternate minimization procedure converges to a solution within a finite number of cycles. We will call the resultant solution the alternating least-squares or "ALS" solution (cf. de Leeuw, Young, and Takane, 1976).
2. As was already pointed out, it is impossible to reproduce $\mathbf{D}_{mn}$ perfectly from the recovered

coordinates when k < h. This means that the optimal value of the objective function (10) in the minimization problem is no longer 0, and therefore searching for the right combination of initial values becomes all the more difficult, especially when min(m, n) is small. But, if min(m,n) is fairly large (over 5, say), the choice of initial values is not a serious problem, and it is advisable to use the Linear solution as the initial values for the alternate-least squares iterations.

3. It is recommended to rotate the resultant coordinates again to make them orthogonal to each other. By the adjustments made through the alternating least-squares procedure stated above, the coordinates tend to deviate from the orthogonality requirement. Also, we may want to rotate the coordinates later to increase interpretability, in which case it is better to begin with a set of coordinates that are known to be orthogonal.

PRACTICAL EXAMPLES

It would not be far fetched to state that the researchers in the field of asymmetrical MDS so far had little interest in metric unfolding, because their attention had been directed mostly toward non-metric methods. But there are many instances in business and marketing where, with only a simple transformation of original data, metric inter-set distances may be defined and measured. In the following we will present four examples to which metric unfolding is applied with a minimum number of additional assumptions. Metric unfolding solutions of these examples add much to the interpretation of original data, due mostly to the ability of metric unfolding to give joint-space graphic presentations of both row- and column-objects.

Car Preference Example

In a classic example of the internal analysis of preference (SAS Ver. 8 Sample Data Libraries), 25 judges rated 17 makes (or models) of passenger cars on a 0 to 9 scale in accordance with their preference to each make (0: no preference, 9: strong preference). The original data is listed in Table 7. The original preference scores were converted to distances using the following simple transformation.

Distance = 9 – Preference Score + 0.5

0.5 in the above transformation was added to avoid a distance of 0 between objects. The calibration procedure for non-exact cases was applied to the resulting distance measures. In order to indicate the degree of fit, we computed the Pearson correlation coefficients, r, and squared values of r, R^2, between

observed and calibrated distances from successive solutions. They are given in Table 8.

Table 8 Correlations between Observed and Computed Inter-set Distances

Solutions	r	R^2
SVD	0.691	0.478
Linear	0.863	0.745
ALS	0.899	0.791

The final (e.g., ALS) solution is shown in Table 9. Only three dimensions are shown here, though the final solution was 17-dimensional. The three dimensions cumulatively explain 86.8% of variations in the final solution. A three-dimensional graph of the coordinates is shown in Figure 1(a) and 1(b).

[Figure 1(a) and 1(b) about here.]

Since this car-preference example was created originally to be analyzed by MDPref (Chang and Carroll, 1969), a three-dimensional MDPref solution was obtained and shown in Table 10 for comparison purposes. Since distance-based metric unfolding solutions and vector-based MDPref solutions were not directly comparable, an attempt was made to reconstruct the preference scores from the coordinates in Table 9 and 10 and Pearson correlation coefficients were computed between the original preference scores and reconstructed scores.

Table 11 Correlation Coefficients between Original and Reconstructed Preference Scores

Methods	r	R^2
MDPref	0.788	0.621
Metric Unfolding	0.845	0.714

Metric Unfolding gives a slightly better fit to the original scores, but this is a heuristic comparison, and we basically assert that both methods are approximately equal in reproducing the original scores. Two-dimensional graphs for the metric unfolding and MDPref solutions are given in Figure 2 and 3, respectively.

[Figure 2 and 3 about here.]

Though the basic assumptions of those two models are greatly different, interpretations of those two graphs as to brand preferences are approximately the same: each judge's order of preferences can be approximately read from those two graphs.

More basic, however, than merely the increase in statistical fit, we believe that the ideal-point

model is a better representation of the psychological-judgment process. The MDPref model is almost a statistical convenience. Think of a two by two matrix, where the columns are Internal Analysis of Preferences (IAP) and External Analysis of Preferences (EAP), and the rows are Ideal-Point Model (IPM) and Vector Model (VM) below.

Table 12 Classification of Preference Analysis

	IAP	EAP
IPM	Metric Unfolding	LEAP-IM or PREFMAP-IP
VM	MDPref	LEAP-VM or PREFMAP-VM

Metric Unfolding is the concern of our current work. LEAP is Logistic External Analysis of Preferences (Cooper and Nakanishi, 1983). Three of these four cells are filled with models of human judgments that are distance based. Only one (MDPref) is scalar-products based. In the EAP Models, R^2 will necessarily increase from VM to IPM and a reduced model test will tell if the extra parameters of the IPM are statistically worthwhile. We can make this call because the human-judgment model in both cases is distance based. In judging between Metric Unfolding and MDPref, we have a distance-based model and a scalar-product-based model. These certainly are not nested. Simple R^2 comparisons only tell us these are in the same general range. So we have a more defensible model of human judgments that provides a more natural and interpretable visual map of preference relations without sacrifice in terms of fit to the original data.

Cognitive Mapping of Brands

Cognitive mapping of brands based on brand similarity measures had been in vogue in the 1960's and 70's, but fallen into disuse in recent years, perhaps because ranking of the similarity of m×(m-1) brand pairs (where m is the number of brands) imposed a heavy mental burden on the respondent when m was even moderately large. In this example, we will show that an interpretable result may be obtained by applying metric unfolding to a similarity data set collected by a method less taxing to the respondent.

A manufacturer of children's wears (brand X) surveyed 170 of its most valuable customers (who purchased this company's products in excess of $1600 per year) and asked how similar brand X was to six competing brands. Note that the respondents only answered six 5-point interval scales (0: extremely similar; 4: not similar at all) for this questionnaire.

It is theoretically possible to apply metric unfolding to the 170 by 6 matrix of similarity ratings, but too many dots for respondents would blot out six brand positions and make 3D maps difficult to read. In our example, we first clustered the respondents with the SAS Proc Cluster procedure (Ward method), before metric unfolding was applied to the data. Four clusters (X1~X4) were identified, and the average similarity scores for each cluster were computed (shown in the following table).

Table 13 Average Similarities by Cluster

	Brands						Cluster
Cluster	A	B	C	D	E	F	Size
X0	1.982	1.207	2.729	2.345	1.170	2.098	66
X1	1.636	1.087	2.500	2.261	1.043	1.421	25
X2	2.000	1.238	2.650	2.350	1.150	1.667	23
X3	1.737	1.278	2.600	2.250	1.316	1.941	21
X4	1.964	1.414	2.517	2.286	1.143	1.842	35

Cluster X0 is a group of respondents who were excluded from the SAS procedure because of some missing data. (The SAS Cluster procedure excludes missing data list-wise.) The size of each cluster is listed in the last column of Table 13.

We may treat this table as a partial distance matrix and apply metric unfolding to it. The next table shows the final (linear-fitted and rotated) solution for this data. This four-dimensional solution happened to be an exact one; that is, a linear transformation of the SVD solution reproduced the original distance matrix (Table 13) perfectly. The final (ALS) solution is given in Table 14.

Table 14 ALS Solution

	Coordinates			
Objects	Dim_1	Dim_2	Dim_3	Dim_4
X0	-1.008	-0.430	-0.400	-0.132
X1	-0.744	-0.030	0.452	-0.230
X2	-0.833	0.823	0.013	0.099
X3	-0.658	-1.023	0.113	-0.165
X4	-0.795	0.649	-0.404	-0.288
A	0.863	-0.192	0.210	-0.132
B	-0.297	-0.131	0.061	0.675
C	1.651	0.094	-0.232	-0.410
D	1.268	-0.112	-0.406	0.337
E	-0.026	0.030	-0.147	0.228
F	0.579	0.323	0.738	0.019

[Figure 4 and 5 about here.]

Figures 4 and 5 give 2-dimensional and 3-dimensional maps of the final solution. The size of a red circle in those maps represents the relative size of the corresponding cluster. All cluster centroids (X0 ~

X4) are gathered in the left side of the maps, indicating that the best customers of brand X share a fairly common view of competitive brands. From these maps we can see that all best customers perceive brand B and E to be similar to brand X, but brand C and D, which happened to be the two largest competitors, are perceived to be dissimilar. Perhaps this is a competitively defensible position for brand X, since its best customers are not likely to be easily lured away to two largest competitors. Marketing strategies of Brand X will only have to be focused on two less powerful competitors.

There are several advantages of the above approach to cognitive mapping of brand similarity data. Firstly, the amount of respondent efforts is much less than the usual procedure of ranking or rating m×(m-1) pairs of objects in the order of similarity. Secondly, the resulting map becomes more legible by clustering respondents and by showing the relative sizes of clusters in the map. And finally, the clusters may be further analyzed for the possibility of segmentation by brand similarity perception. On the other hand, this approach is limited to "our brand against other brands" comparisons. It will be better to have some variations in brand similarity perceptions to give a solution (and the resulting map) greater interpretability.

Individual Level Similarity Judgments

In this section we show that metric unfolding may be applied to similarity judgment data at the individual level. Two subjects were asked to supply similarity judgments between pairs of beverages in both ranking and rating formats. Six kinds of beverages (Ginger Ale, Apple Juice, Grapefruit Juice, Coca Cola, Orange Juice, and Grape Juice) were divided into two groups, and nine pairs of objects were formed by selecting one kind from each group. The respondents were asked to rank them (to be called the "ordinal" data) and also to rate them on a 5-point interval scale (the "interval" data). The following tables give the responses to those two types of questions.

Table 15 Similarity Judgments from Two Subjects

Subject 1: Ordinal Data	Coca Cola	Orange Juice	Grape Juice	Subject 2: Ordinal Data	Coca Cola	Orange Juice	Grape Juice
Ginger Ale	1	6	7	Ginger Ale	2	6	5
Apple Juice	8	3	4	Apple Juice	9	3	4
Grapefruit Juice	9	2	5	Grapefruit Juice	8	1	7
Subject 1: Interval Data	Coca Cola	Orange Juice	Grape Juice	Subject 2: Interval Data	Coca Cola	Orange Juice	Grape Juice
Ginger Ale	1	5	5	Ginger Ale	1	5	4
Apple Juice	5	2	4	Apple Juice	3	2	3
Grapefruit Juice	5	1	4	Grapefruit Juice	4	1	2

The Pearson correlation coefficients between ordinal and interval data are 0.918 for subject 1 and 0.622 for subject 2. The final solutions for all four data sets are shown in the following figures. All solutions were two-dimensional. (The interval data for subject 2 yielded an exact solution.)

[Figure 6 and 7 about here.]

All the maps appear to have a degree of face validity since likely objects (Orange Juice and Grapefruit Juice [both are citrus] and Coca Cola and Ginger Ale [both are carbonated]) are grouped together in them. Though the two maps for subject 1 are in fair agreement, the maps for subject 2 are different (especially in the location of Apple Juice is very different). Those examples suggest a simple rule of thumb: if the correlation between ordinal and interval data for a subject were high, the final solutions would be also in good agreement.

The reliability of ordinal and interval scales probably depends on the task conditions for similarity judgment. Rating with an interval scale may be a much easier task for the subjects when the number of object pairs is large, except that the subjects may give less serious thoughts to rating (in which each pair of objects is rated independently from other pairs) than to straight ranking, which requires comparisons among pairs. We should expect a greater degree of internal consistency in distances computed from ranks, but internal consistency does not guarantee that a psychologically valid configuration of objects will be recovered. Thus we are faced with a rather difficult choice between reliability (ranking) and easy to use (rating). There seems no uniform answer to this choice problem.

Brand Association Example

In this section we describe an application in branding. A group of forty college students were given a list of 10 sneaker brand names and asked to write down as many adjectives as came to their minds in association with each brand name. Table 16 gives the frequencies of adjectives mentioned by the respondents. Of 78 adjectives mentioned, only top 20 adjectives are listed here.

Several factors affect the frequencies of adjectives associated with each brand name. Given the assumption that the human memory structure is associative, the strength of association between concepts in long-term memory may be modeled as the "distance" in a hypothetical space; that is, the greater the strength of association between two concepts, the shorter the distance between them. If this assumption is adopted as a working hypothesis, then, by estimating the distances from the frequencies, we may be able to find positions (coordinates) of both brands and adjectives in the common "semantic" space.

But, because of the data collection procedure used here, adjective frequencies are influenced by another factor that may be called the "Popularity" effect. Suppose that 30 respondents are aware of brand A and 10 respondents are aware of brand B. Suppose also that 15 of the former group and 10 of the latter group chose adjective "Economical". Do we conclude that this adjective is closer to brand A than to brand B in the semantic space? But, considering the fact that all of respondents who are aware of brand B and only 50% of those who are aware of brand A mentioned this adjective, we are led to conclude that brand B is closer to this adjective is than brand A. In order to estimate the distances between brands and adjectives, we will have to take account of such "popularity" effects from the frequencies. The number of respondents who are aware of each brand is given in Table 16. In this example the awareness level is very high for all brands, but the popularity effect may be a critical factor when the awareness levels are greatly different among brands.

To incorporate these two factors we propose the following multiplicative model for adjective frequencies.

$$E(n_{ij}) = n_i \, c \, d_{ij}^{-\lambda} \qquad (11)$$

where: $E(n_{ij})$ = expected number of the j-th adjective for brand i

n_i = number of respondents who are aware of brand i

d_{ij} = distance between brand i and adjective j

c, λ = parameters to be specified.

If we let $p_{ij} = n_{ij} / n_i$ and assume that $\lambda = 2$ for simplicity, we obtain estimated values of d_{ij}^2 as $1/p_{ij}$. (Since c is only a scaling constant, we may assume c=1 without the loss of generality.)

Now we turn to the task of finding the coordinates of both brands and adjectives in a common space. First, the squared-distance matrix was computed from Table 16 by using equation 11 and then the SVD solution was obtained. The number of positive singular values, k, was 9. The Linear and ALS solutions were successively computed from the SVD solution. The fit of reconstructed distances to the "observed" distances (calculated from the frequencies by model (11)) improved significantly as we moved toward the ALS solution. The r and R^2 values are given in Table 17. The ALS solution is given in Tables 18(a) and (b). Judging from the value of R^2, the ALS solution is able to account for about 94.1% of the total variance in observed distances.

Table 17 Correlations between Observed and Computed Inter-set Distances

Solutions	r	R^2

SVD	0.834	0.696
Linear	0.936	0.877
ALS	0.970	0.941

Figures 8 and 9 give the 2-dimensional and 3-dimensional representations of brands and adjectives in the "semantic" space.

[Figure 8 and 9 about here.]

Though it would be over reaching to derive definite conclusions from this one example, it appears that the center of the semantic space is occupied by those adjectives that are characteristic of the product category (sneakers in this case) and surrounded by brands. If an attribute is prototypical enough of the whole category to be shared by all brands in it, we would expect that attribute to be close to equidistant from all brands. The centroid of the brands is our closest approximation (in a least-squares sense) to such a point without increasing the dimensionality of the configuration. Adjectives that are not characteristic of all brands in the product category comprise the outer-most layer. When those adjectives are extreme enough they can help name the dimensions involved. In Figure 8 we see a masculine-feminine dimension characterizing the vertical alignment and a economical-stylish/fashionable dimension as characterizing the horizontal alignment. This three-layered structure seems to suggest that a brand's unique character is determined, not by the adjectives in the center, but by the adjectives in the outer fringes of the semantic space. This conjecture may be highly relevant to brand identity research, but we will refrain from making further comments because of the small and biased nature of the sample. This example illustrates the advantage of metric unfolding methods over correspondence analysis. With only a simple transformation of the original data, we obtain a joint-space representation of both row- and column-objects (brands and adjectives in this example) for which there is no interpretive problem as to inter-set comparisons.

Asymmetric Competitive Maps

Colombo and Morrison (1989) describe a simple procedure for estimating the proportion of potential switchers by comparing the *conquesting* ratios in the *southwest* boxes of the switching matrices, and noting the ultimate parameters, π_i , must sum to 1.0. Table 19 gives a brand-switching matrix for a group of Japanese beer brands. Different sizes for a brand are separately listed in this table. From this table one may compute the row-conditional, transition-probability matrix that shows the likelihood that a person purchasing the beer in the row on the last occasion will purchase the beer in the column on the

next. The diagonal entry in the transition probability matrix shows the repeat-purchase probabilities composed of both true loyals and potential switchers. Except for the diagonal entry, the entries in a column reflect the *draw* or *conquesting* potential of the column brand. The column average excluding the diagonal entry summarizes this draw or conquesting potential. Renormalizing the draw or conquesting probabilities provides our estimates of parameter π_i . The procedure for estimating π_i amounts to forming the row-conditional, transition-probability matrix, deleting the diagonals, finding the column averages as the raw draw of each brand, and renormalizing so that the sum is 1.0.

With these estimates of π_i we can estimate the proportion of true loyals, α_i , from Colombo Morrison equation (1):

$$p_{ii} = \alpha_i + (1 - \alpha_i)\,\pi_i$$

We set up a squared-error criterion and used Excel Solver to accomplish the estimation. With these estimates we can translate the number of repeat buyers in the diagonals of the original transition matrix into the number of true loyals and the number of potential switchers who simply didn't switch this time around. The *vulnerability* of each brand would be reflected in the sum of the entries in a row minus the number of true loyals for the row brand. The *clout* of each brand would be reflected in the sum of the actual switchers into the column brand from other brands, plus the true loyals to that brand.

The distance matrix is based on the original transition matrix, *T*, with the *true loyals* removed from the appropriate diagonal. Let *L* be a diagonal matrix with the number of true loyals in each appropriate diagonal. Let *V* be a diagonal matrix with the *vulnerability* measure on each diagonal. And let *C* be a diagonal matrix with the *clout* measure on each diagonal.

$$T - L = V \bullet [\, f(D) \,] \bullet C\,.$$

Or

$$f(D) = V^{-1} \bullet (T - L) \bullet C^{-1}.$$

In our example we used a simple function for D, that is, $f(d_{ij})=d_{ij}^2$ and applied metric unfolding to the resultant D matrix. Figure 10 summarizes the result of analysis.

[Figure 10 about here.]

Those three dimensions cumulatively accounts for approximately 75% of total variance in the D matrix. A red circle represents the vulnerability and a blue circle represents the clout of the corresponding brand/size combination. The relative size of those circles indicates the strength of either clout or vulnerability. The closer a brand's clout position is to another brands vulnerability position, the

more competitive pressure it exerts. What a brand manager would want to see in his or her brand is a large, centrally located bubble reflecting the brand's clout, and a tiny bubble reflecting that brand's vulnerability. This would reflect a brand exerting broad competitive pressure, while being insulated from the counterinfluence.

First, we see the competition is more within size groups than between size groups. Note that the smallest size, 135ml is a little over one-third of a standard 12oz. glass. These small serving sizes have no clout over the larger (350ml or 500ml) size offering, but are also not vulnerable to the pressures from those much larger sizes. Within the 135ml size offerings we see the distance between Asahi Super Dry (S135C) and Kirin Draft (D135V) is much smaller than the distance between Kirin Draft (D135C) and Asahi Super Dry (S135V). This reflects Asahi Super Dry (S135C) exerting more competitive pressure on Kirin Draft (D135V) than Kirin Draft (D135C) can return. Second, within the larger sizes the dominant competitive pressures come from Kirin Ichibanshibori (I350C) and Asahi Super Dry (S350C). Kirin Ichibanshibori gets more support in exerting competitive pressure from its large size (I500C) than Asahi Super Dry, since we see that 500ml Asahi Super Dry is remote (both S500C and S500V are at the far left of the map). Sapporo's Hokkaido brand has a small, relatively loyal following, but exerts little competitive pressure on others (H350C). It is aligned to be competitively vulnerable to Kirin Ichibanshibori (I350C), but its high loyalty (reflected in the small size of the vulnerability bubble of H350V) indicates not much will be lost. Sapporo's Hokkaido needs to work to maintain that loyalty, since without it the brand is positioned to be very vulnerable to Kirin Ichibanshibori. We also see that the Kirin Lager offerings (L500, L350, and L250) tend to be on the right side of this map, while the Asahi Super Dry offerings (S500, S250, and S350) tend to be on the left. These brands compete very little with each other.

CONCLUSIONS

We believe that researchers and managers are going to interpret perceptual maps in terms of distances whether or not the underlying analysis allows such interpretation. Squared distances are not distances. Scalar-product distances are not distances. Non-comparable inter-set distances are just that – *non-comparable*. The joint-space maps that reflect the basic and compelling problems in marketing applications of perceptual mapping need to have the simple distance interpretation users of these techniques expect, or misinterpretation will result. We could try to change the behavior of all the users

who need these maps, or we could develop a method that allows for the straightforward interpretation of these maps. We chose the later approach.

By revisiting the basic problem of metric unfolding we have derived a method that provides a simple distance interpretation of joint spaces, even when only the inter-set distance judgment exist. We have developed widely available estimation procedures using Excel Solver, and applied this method to classic examples of internal analysis of preference, cognitive mapping of brands, individual similarity judgments, brand association data, and brand switching data. In each of those examples, we made a simple – albeit naïve – assumption to derive distance measures from the original data. The simplicity of assumptions was intentional. We wanted to demonstrate that metric unfolding may be applied to those diverse data with a minimum number of additional assumptions, and yet yield rich and managerially meaningful interpretations. We hope this effort will rekindle interest in and proper use of multidimensional scaling for addressing marketing problems.

REFERENCES

Carroll, J. Douglas (1980), "Models and Methods for the Multidimensional Analysis of Preferential Choice (or other Dominance) Data." In *Proceedings of the Aachen Symposia on Decision Making and Multidimensional Scaling*, ed. E. D. Lantermann, and H. Feger, Berlin: Springer-Verlag.

Carroll, J. Douglas, and Phipps Arabie (1980), "Multidimensional Scaling," *Annual Review of Psychology***.** M. R. Rosenzweig and L. W. Porter, eds., vol. 31, 1980.

Coombs, Clyde H. (1950), "Psychological Scaling without a Unit of Measure," *Psychological Review*, 57, 148-158.

Coombs, Clyde H. (1964), *A Theory of Data*, New York: Wiley.

de Leeuw, Jan, Forrest W. Young, and Yoshio Takane (1976), "Additive Structure in Qualitative Data: An Alternating Least Squares Method with Optimal Scaling Features," Psychometrika, 41 (4), 471-504.

DeSarbo, Wayne S., and Vithala Rao (1984), "GENFOLD2: A Set of Models and Algorithms for the

GENeral unFOLDing Analysis of Preference/Dominance Data," *Journal of Classification*, 1 (Winter), 147-186.

Eckart, Carl, and Gale Young (1936), "The Approximation of One Matrix by Another of Lower Rank," *Psychometrika*, 1, 211-218.

Hoffman, Donna L., and George R. Franke (1986), "Correspondence Analysis: Graphical Representation of Categorical Data in Marketing Research," *Journal of Marketing Research*, XXIII (August), 213-217.

Hoffman, Donna L., and Jan de Leeuw (1992), "Interpreting Multiple Correspondence Analysis as a Multidimensional Scaling Method," *Marketing Letters*, 3, 3 (July), 259-272

Schönemann, Peter H. (1970), "On Metric Multidimensional Unfolding," *Psychometrika*, 35, 349-366.

Torgerson, Warren S. (1958), *Theory and Method of Scaling*, New York: Wiley.

Table 7 Preference Ratings for Automobiles 17 Makes - 25 Judges

NAME	j01	j02	j03	j04	j05	j06	j07	j08	j09	j10	j11	j12	j13	j14	j15	j16	j17	j18	j19	j20	j21	j22	j23	j24	j25
Eldorado	8	0	0	7	9	9	0	4	9	1	2	4	0	5	0	8	9	7	1	0	9	3	8	0	9
Chevette	0	0	5	1	2	0	0	4	2	3	4	5	1	0	4	3	0	0	3	5	1	5	6	9	8
Citation	4	0	5	3	3	0	5	8	1	4	1	6	1	6	4	3	5	4	4	7	4	7	7	9	5
Malibu	6	0	2	7	4	0	0	7	2	3	1	2	1	3	4	5	5	4	5	6	6	8	6	5	8
Fairmont	2	0	2	4	0	0	6	7	1	5	0	2	1	4	4	3	5	3	0	6	4	8	6	5	5
Mustang	5	0	0	7	1	9	7	7	0	5	0	2	1	1	0	1	8	5	0	6	5	7	5	5	5
Pinto	0	0	2	1	0	0	0	3	0	3	0	3	0	2	0	1	5	0	0	5	1	4	0	7	8
Accord	5	9	5	6	8	9	7	6	0	9	6	9	9	9	5	2	9	9	8	9	7	5	0	7	8
Civic	4	8	3	6	7	0	9	5	0	7	4	8	8	8	5	2	5	6	7	7	6	5	0	7	5
Continental	7	0	0	8	9	9	0	5	9	2	2	3	0	4	0	9	9	6	2	0	9	1	9	0	9
Gran Fury	7	0	0	6	0	0	0	4	3	4	1	0	1	1	0	7	3	3	3	4	5	8	7	0	8
Horizon	3	0	0	5	0	0	5	6	3	5	4	6	1	3	0	2	4	4	4	6	7	5	6	5	5
Volare	4	0	0	5	0	0	3	6	1	4	0	2	1	6	0	2	7	5	4	4	7	6	5	5	5
Firebird	0	1	0	7	8	9	5	6	1	3	2	0	1	2	0	6	9	5	8	2	6	5	9	0	7
Dasher	4	8	5	8	6	9	6	5	0	8	8	7	7	7	9	5	3	7	7	8	9	5	0	0	0
Rabbit	4	8	5	8	5	0	9	7	0	9	6	9	5	7	9	5	4	8	7	8	8	5	0	0	0
Saab DL	9	9	8	9	9	9	8	9	0	9	9	9	9	9	8	7	9	8	9	9	1	9	0	0	0

Models	Dim 1	Dim 2	Dim 3
Eldorado	3.083	2.919	-0.462
Chevette	1.447	-4.093	-3.056
Citation	1.013	-3.402	-0.183
Malibu	2.727	-1.595	-0.555
Fairmont	1.891	-3.355	1.708
Mustang	2.164	-0.764	3.028
Pinto	2.199	-4.967	-1.819
Accord	-1.864	0.336	0.231
Civic	-2.566	-0.881	0.968
Continental	3.149	2.905	-0.524
Gran Fury	4.005	-0.139	-0.448
Horizon	1.905	-2.198	1.293
Volare	2.564	-1.935	2.644
Firebird	2.289	2.179	0.979
Dasher	-2.932	1.512	0.010
Rabbit	-2.999	0.140	0.738
VolvoDL	-2.290	0.782	-0.073

Judges	Dim 1	Dim 2	Dim 3
Judge 1	1.071	2.147	-0.248
Judge 2	-5.076	1.105	0.118
Judge 3	-3.873	-1.931	-3.094
Judge 4	0.270	1.483	0.656
Judge 5	-0.489	3.141	-0.801
Judge 6	0.533	3.881	0.352
Judge 7	-2.875	-0.643	2.413
Judge 8	0.323	-0.849	0.828
Judge 9	5.117	3.518	-2.279
Judge 10	-2.321	-0.435	0.666
Judge 11	-3.696	1.405	-2.027
Judge 12	-2.817	-0.573	-0.486
Judge 13	-4.737	0.590	-0.029
Judge 14	-2.240	0.478	0.456
Judge 15	-4.331	-0.501	-0.871
Judge 16	1.963	2.736	-1.578
Judge 17	1.388	1.195	0.841
Judge 18	-0.874	1.829	0.948
Judge 19	-2.362	0.900	-0.289
Judge 20	-1.637	-1.696	0.256
Judge 21	1.222	1.559	0.748
Judge 22	0.743	-1.663	0.573
Judge 23	4.804	0.493	-0.012
Judge 24	0.716	-5.128	-0.623
Judge 25	3.395	-0.484	-0.996

Table 9 Metric Unfolding Solution

Figure 1 (a) Joint-Space Presentation of Metric Unfolding Solution

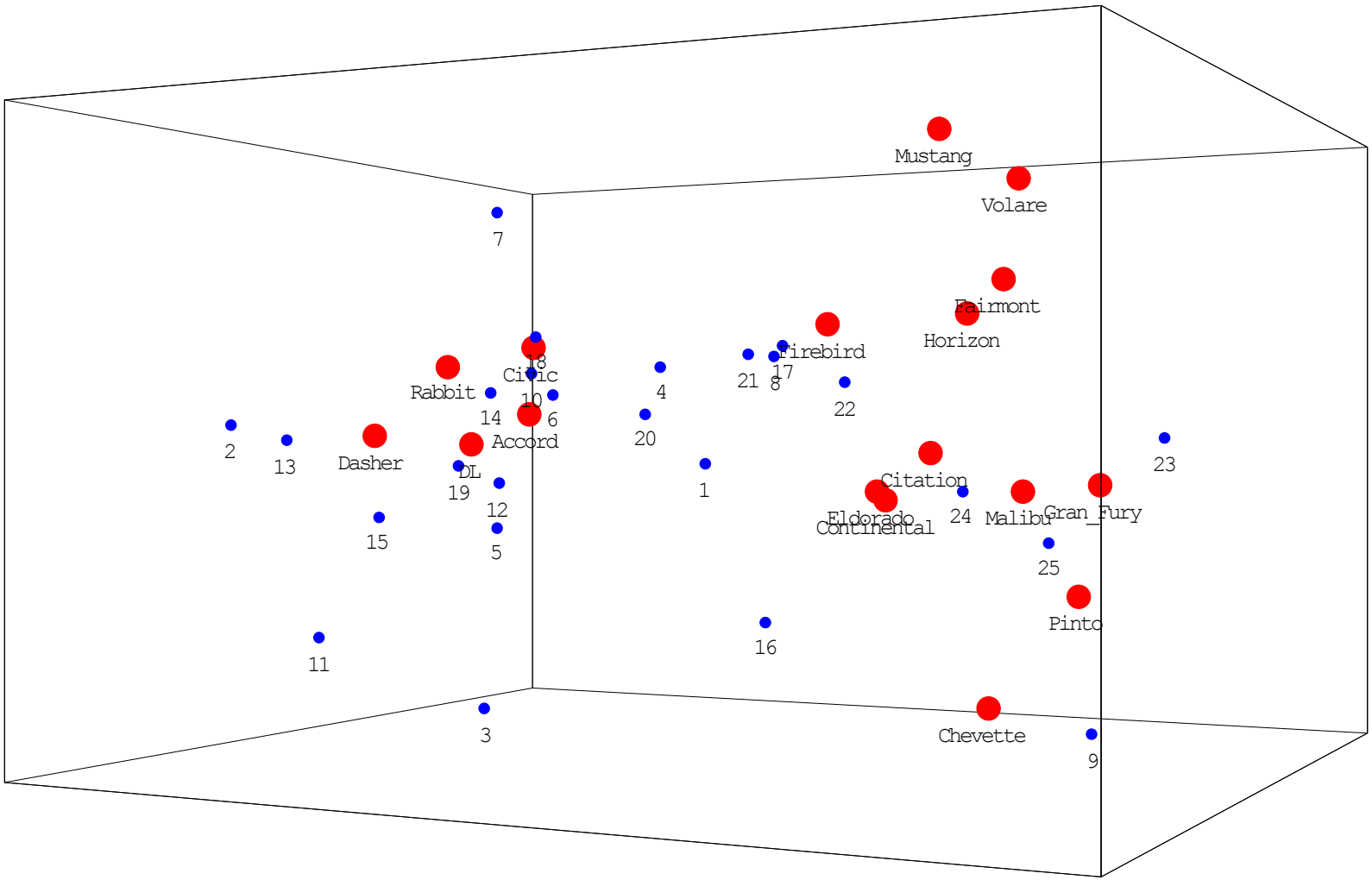
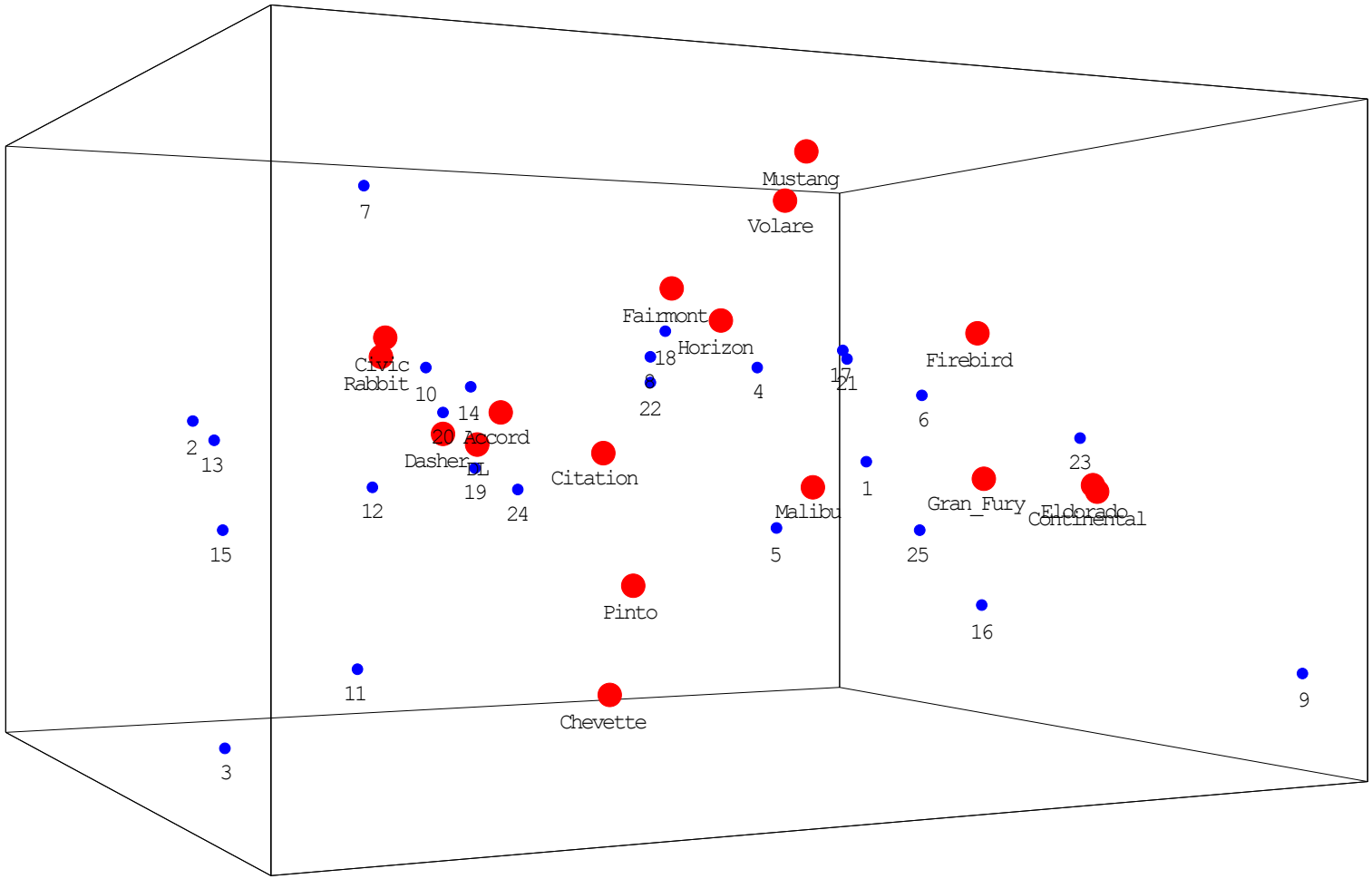


Figure 1 (b) Joint-Space Presentation of Metric Unfolding Solution



Models	Prin 1	Prin 2	Prin 3
Eldorado	-0.832	2.072	-0.738
Chevette	-0.804	-1.429	-1.568
Citation	-0.059	-0.872	0.634
Malibu	-0.410	-0.037	0.970
Fairmont	-0.430	-0.894	1.142
Mustang	-0.396	-0.194	1.613
Pinto	-1.068	-1.384	-1.328
Accord	1.365	0.106	-0.959
Civic	0.927	-0.453	-1.014
Continental	-0.853	2.176	-0.762
Gran Fury	-0.923	0.137	0.717
Horizon	-0.357	-0.415	-0.035
Volare	-0.502	-0.225	0.652
Firebird	-0.489	0.873	0.497
Dasher	1.354	0.230	-0.867
Rabbit	1.419	-0.070	-0.173
VolvoDL	2.058	0.379	1.219

Judges	Prin 1	Prin 2	Prin 3
Judge 1	0.288	0.622	0.299
Judge 2	0.955	0.046	-0.217
Judge 3	0.771	-0.318	-0.155
Judge 4	0.507	0.732	0.320
Judge 5	0.482	0.709	-0.258
Judge 6	0.272	0.695	0.000
Judge 7	0.792	-0.163	0.230
Judge 8	0.516	-0.046	0.718
Judge 9	-0.499	0.726	-0.212
Judge 10	0.932	-0.206	0.044
Judge 11	0.854	0.125	-0.280
Judge 12	0.833	-0.104	-0.365
Judge 13	0.942	-0.020	-0.151
Judge 14	0.832	0.181	-0.088
Judge 15	0.856	-0.189	-0.088
Judge 16	0.000	0.802	0.040
Judge 17	0.114	0.655	0.267
Judge 18	0.730	0.600	0.086
Judge 19	0.783	0.113	-0.043
Judge 20	0.764	-0.583	0.125
Judge 21	0.086	0.662	-0.123
Judge 22	0.237	-0.452	0.750
Judge 23	-0.752	0.359	0.278
Judge 24	-0.174	-0.749	-0.176
Judge 25	-0.763	0.168	-0.253

Table 10 MDPref Solution

Figure 2 Joint-Space Presentation of Metric Unfolding Solution – 2 Dimensions

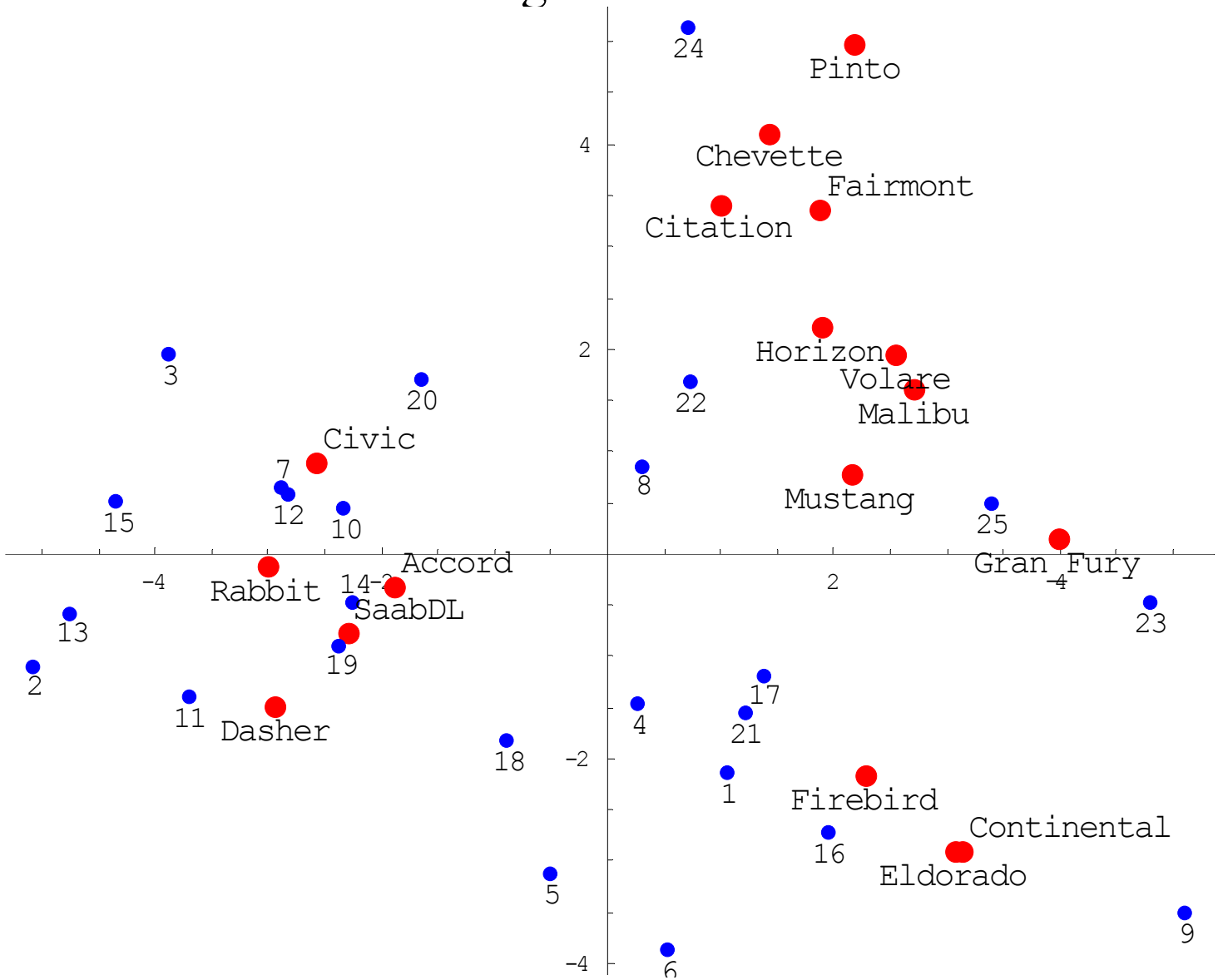


Figure 3 Joint-Space Presentation of MDPref Solution – 2 Dimensions

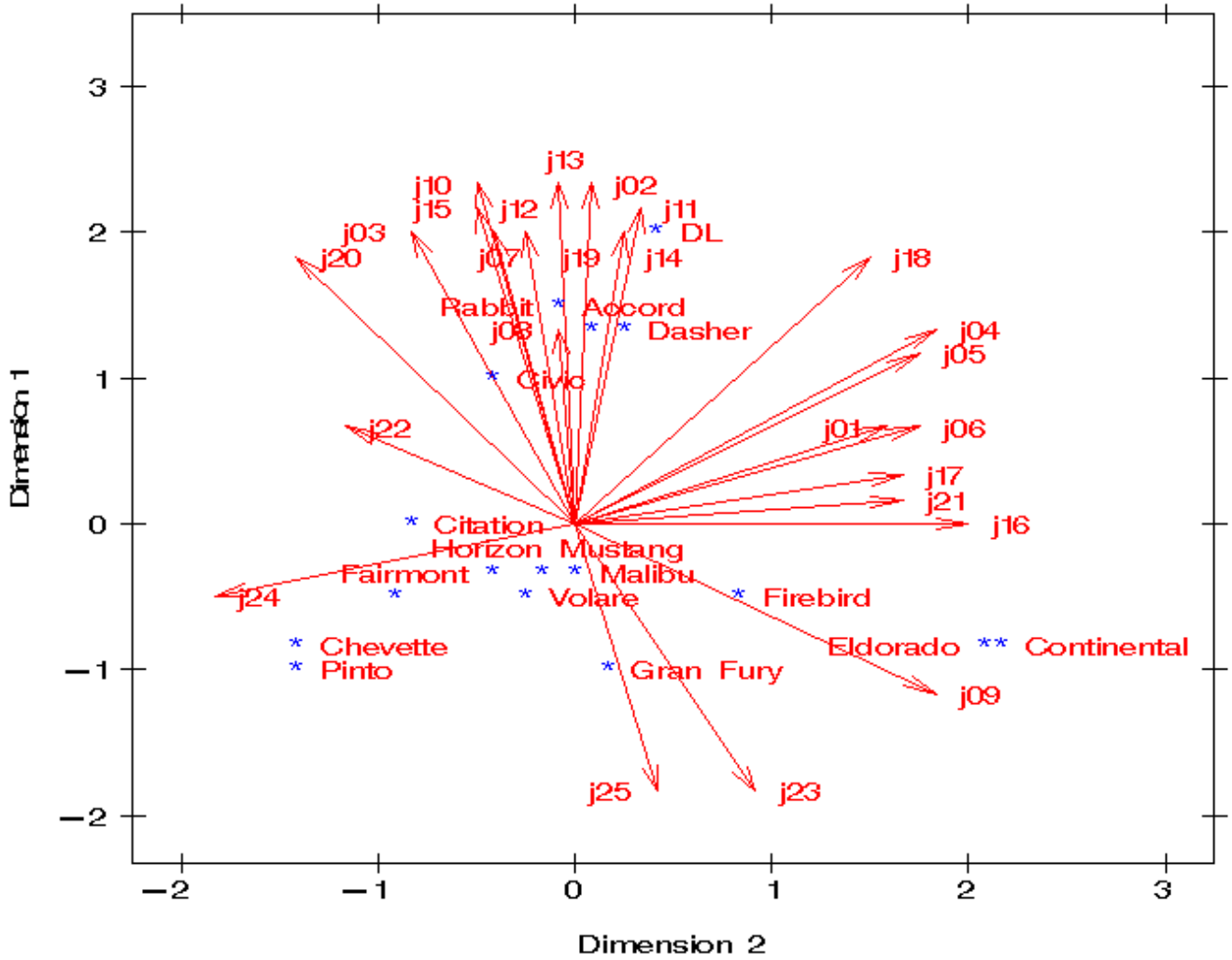


Figure 4 Brand Cognitive Map -2 Dimensions-

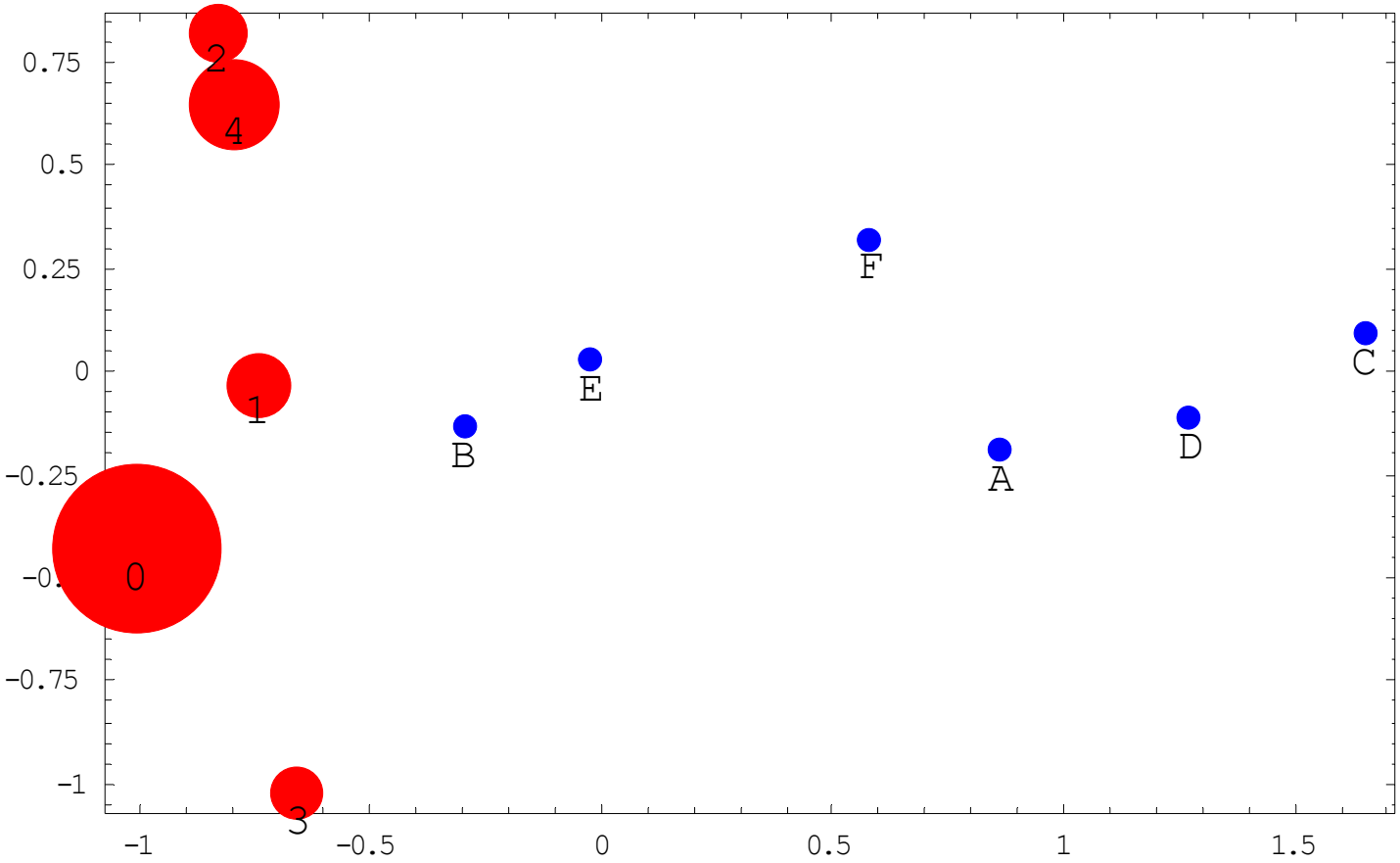


Figure 5 Brand Cognitive Map -3 Dimensions-

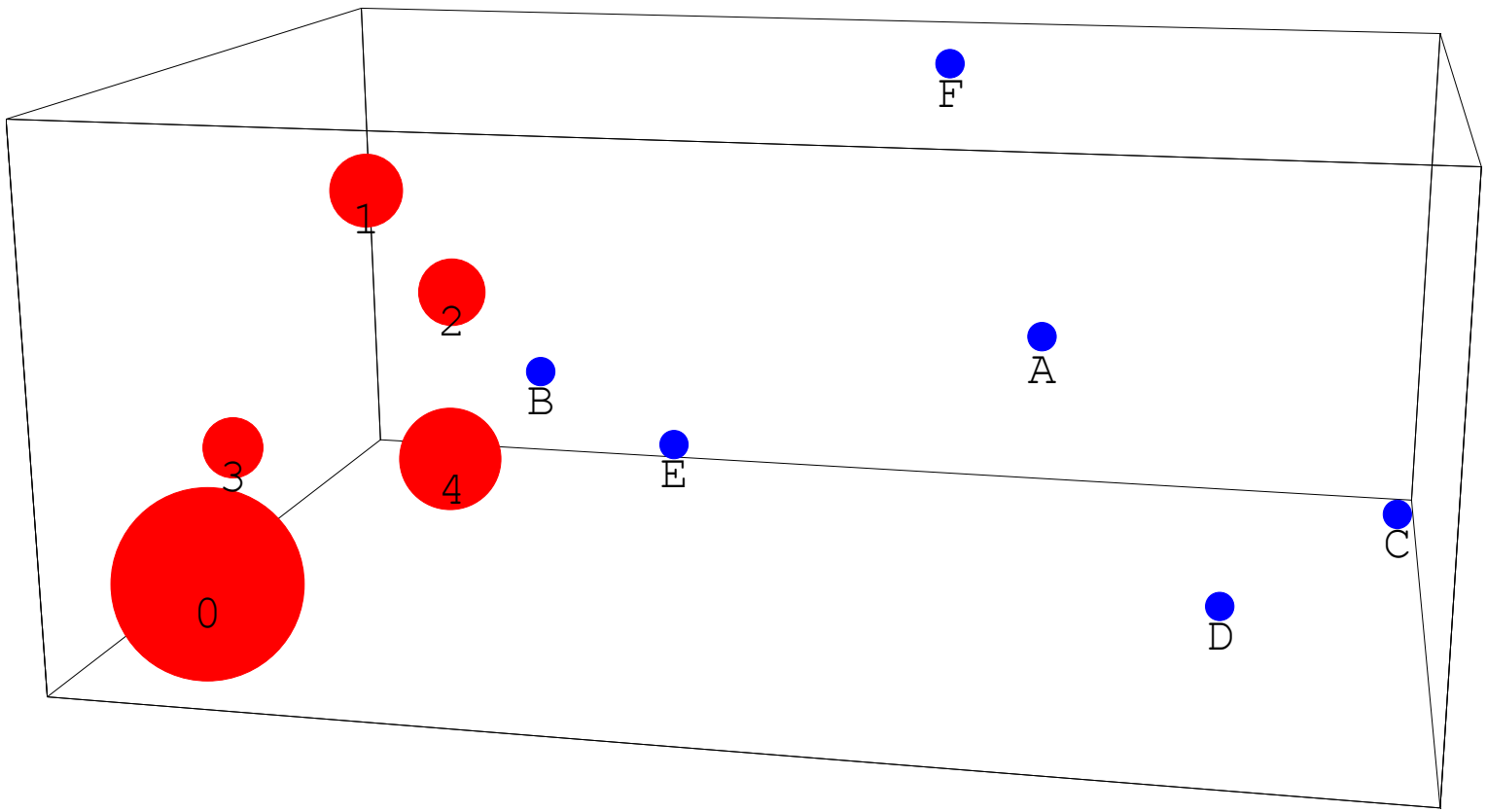


Figure 6 Individual Similarity Perception
-Subject 1-

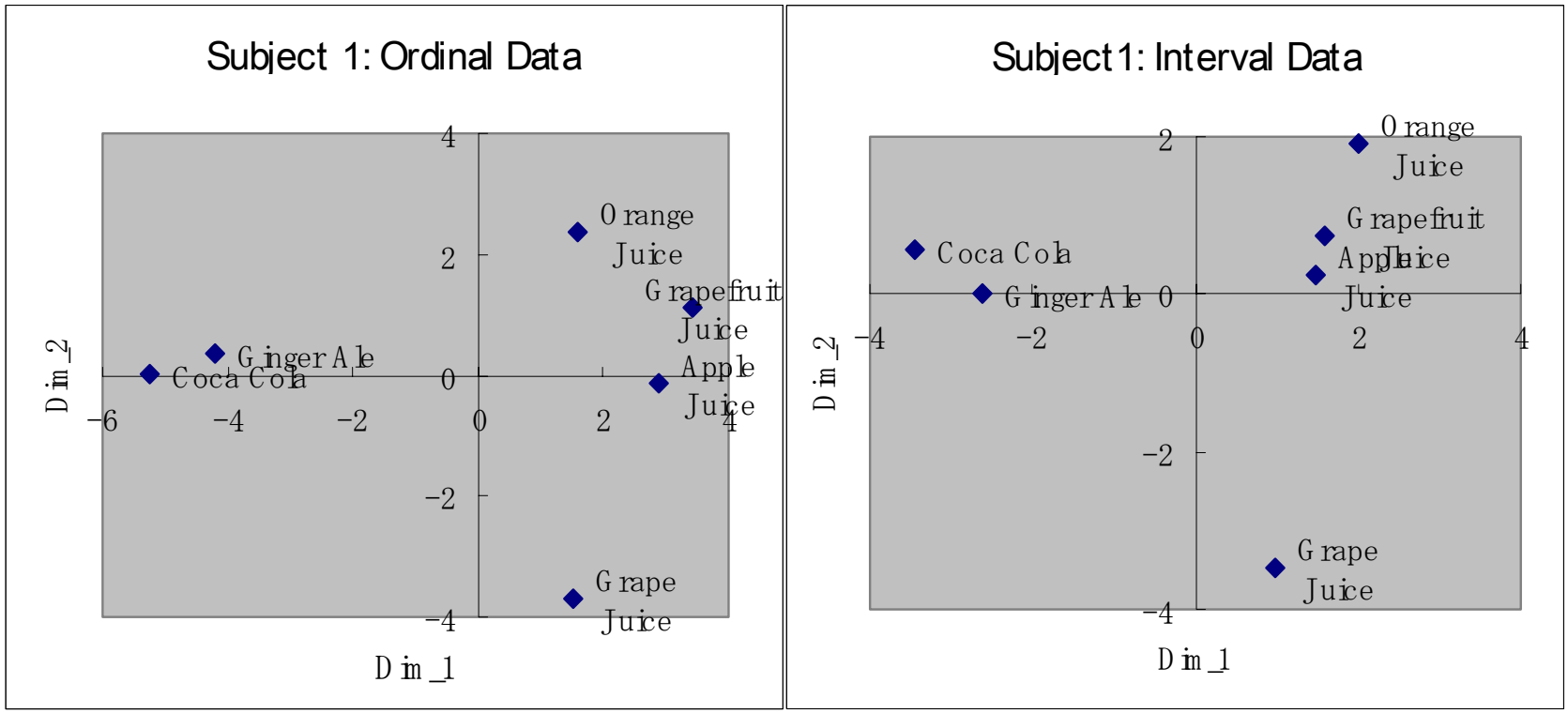


Figure 7 Individual Similarity Perception
-Subject 2-

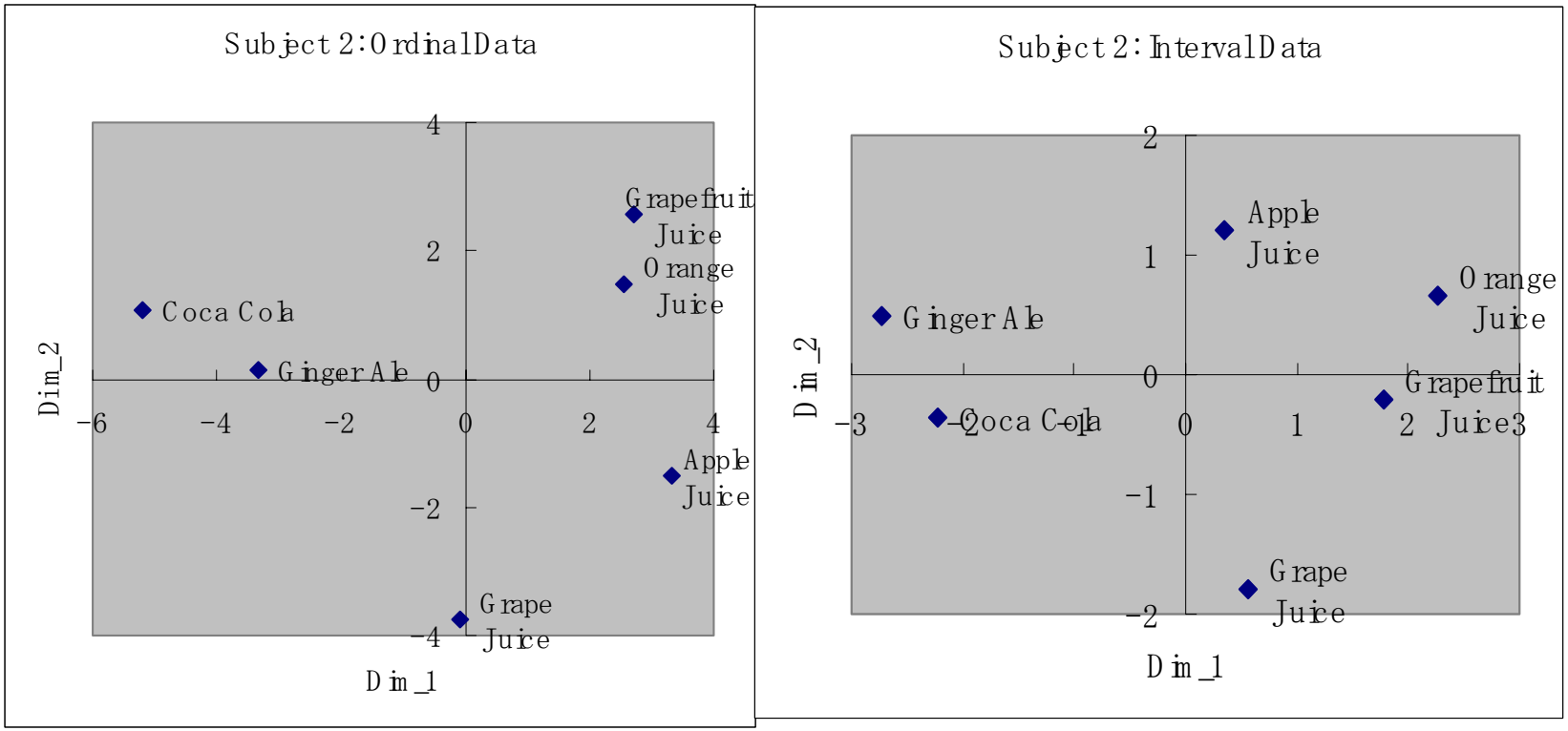


Table 17 Adjective Frequencies by Brands

	Adjectives																				No.
Brand Names	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Aware
NIKE	8	9	13	18	7	9	11	11	12	6	12	12	5	6	2	2	1	4		4	36
ADIDAS	11	7	7	8	7	5	7	6	6	5	5	3	5	3	2	4	5	2	1	3	37
NEW BALANCE	8	5	13	4	7	5	5	2	7	3	6	4	3	6	2	5	3	3	4	3	36
CONVERSE	15	5	8	6	7	6	4	4	1	5	3	2		7	5	10	2	4	4	1	35
PUMA	14	6	1	3	4	4	3	5	3	3	4	3	4	3	3	3	3	2		1	38
AIRWALK	2	2	3	7	2	2	5	4	5	3	3	6	4	1				1		1	28
ASICS	7	6	7			8	1	2		2		1	4				4	1		3	32
FILA	7	6	2	2	6	2	3	1	2	3	2	1		2	1		1	3	7	2	36
REEBOK	8	5			3	1	2	3	1	4		2	5	1	3		3	1	2	1	36
PRO KEDS	11	4			2	2			1	2			1	1	9		1	1	2		35

List of Adjectives:

1	Popular	6	Easy to Use	11	Cool	16	Easy to Fit
2	Active	7	Stylish	12	Innovative	17	Traditional
3	Comfortable	8	Tasteful	13	Masculine	18	Refreshing
4	Fashion-Conscious	9	Fashionable	14	Vivid	19	Feminine
5	Bright	10	Lively	15	Economical	20	Durable

Table 18(a) ALS (Rotated) Solution for Brands

Dimension	1	2	3	4	5	6	7	8	9
Brand Names:									
NIKE	2.7644	-0.4184	0.0682	0.5461	1.0620	1.7654	-1.0838	-1.1789	-0.7205
ADIDAS	1.7273	0.8508	-0.1413	-0.8700	0.2434	-0.8848	-0.1122	-0.7626	1.9988
NEWBALANCE	0.6497	-1.0363	-0.4741	-0.2280	-0.5884	-1.4055	1.9569	-1.2032	-0.9118
CONVERSE	-1.0309	-2.1658	1.5124	-2.7451	-1.6952	0.8009	-0.5808	0.2096	-0.0202
PUMA	0.5504	1.3283	2.6393	-0.6375	1.6856	-1.4937	-0.3080	1.0562	-0.6643
AIRWALK	2.2593	0.7287	0.3261	1.5351	-1.5012	1.0944	1.0715	1.4957	0.2380
ASICS	-0.5865	2.0330	-3.7833	-1.5828	0.1451	0.1046	-0.3503	0.6497	-0.4925
FILA	-0.6183	-3.3499	-1.3648	1.9399	0.1891	-1.2390	-1.0610	0.6376	0.1915
REEBOK	-2.6414	2.5667	0.9546	1.7168	-1.4832	-0.2357	-0.8218	-0.9024	-0.1400
PRO KEDS	-3.0741	-0.5371	0.2628	0.3254	1.9430	1.4933	1.2894	-0.0016	0.5210

Table 18(b) ALS (Rotated) Solution for Adjectives

Dimesion	1	2	3	4	5	6	7	8	9
Adjectives:									
Popular	-1.6241	0.2258	-0.2902	0.1242	0.9801	0.0644	-0.5069	0.3317	0.1577
Active	-1.2729	0.2840	-0.5707	0.3841	0.8634	0.1134	-0.7804	0.3748	0.0365
Comfortable	1.0407	-0.7985	-2.4574	-1.1275	-1.6015	1.3107	0.3415	-0.4422	0.4145
Fashion-Conscious	1.6580	-0.9326	0.7758	-0.1452	-0.3838	0.3485	0.1092	0.8993	0.7596
Bright	-0.5577	-0.8706	1.0662	1.0248	0.0619	-0.3694	-0.4327	-0.6320	0.8361
Easy to Use	0.2127	-0.2745	-1.0331	-1.4739	1.3723	0.7036	0.4562	1.4229	-0.4021
Stylish	0.8597	-0.4620	0.8359	1.0022	-1.0529	-0.3757	-0.6864	-0.0219	0.4682
Tasteful	0.5813	1.8540	0.7632	-0.5353	-0.8705	0.9618	-1.1365	0.3684	-0.2277
Fashionable	1.9973	-0.0171	-0.0893	1.7249	1.1657	-0.7742	1.2891	0.2613	0.3177
Lively	-0.7483	0.2439	0.0228	0.6820	-0.2473	0.5971	-0.7675	0.4291	0.5594
Cool	1.5152	-0.8060	0.7142	0.0235	0.2513	-0.3325	0.3038	0.5346	-0.0835
Innovative	1.0526	1.0734	1.2411	0.6544	-0.7784	0.8218	0.7139	-0.5256	-0.8443
Masculine	0.5749	3.0930	-0.6068	1.0667	0.0094	-0.2270	0.1018	-0.1217	-0.8456
Vivid	0.5402	-1.3701	0.6419	-0.9965	0.4765	-0.7423	-0.6337	-0.7995	-0.8766
Economical	-2.3074	0.5813	1.7323	-0.4310	1.0448	1.5311	0.8819	-0.8182	0.4256
Easy to Fit	0.2656	-0.3131	1.1304	-2.1544	-0.0782	-0.6959	0.7675	-0.6767	0.0921
Traditional	-1.3599	2.1684	-0.5187	-1.3993	-0.8578	-2.1849	0.2720	0.3037	0.5711
Refreshing	0.1628	-1.6613	0.1363	-0.2471	0.3369	-0.4659	-1.1264	-0.1015	-0.8191
Feminine	-3.0197	-1.9652	-0.5806	1.2648	-1.5417	-0.2386	1.0122	0.5111	-0.7979
Durable	0.4289	-0.0529	-2.9132	0.5585	0.8498	-0.0461	-0.1789	-1.2977	0.2584

Figure 8 2D Configuration of Adjectives and Brands

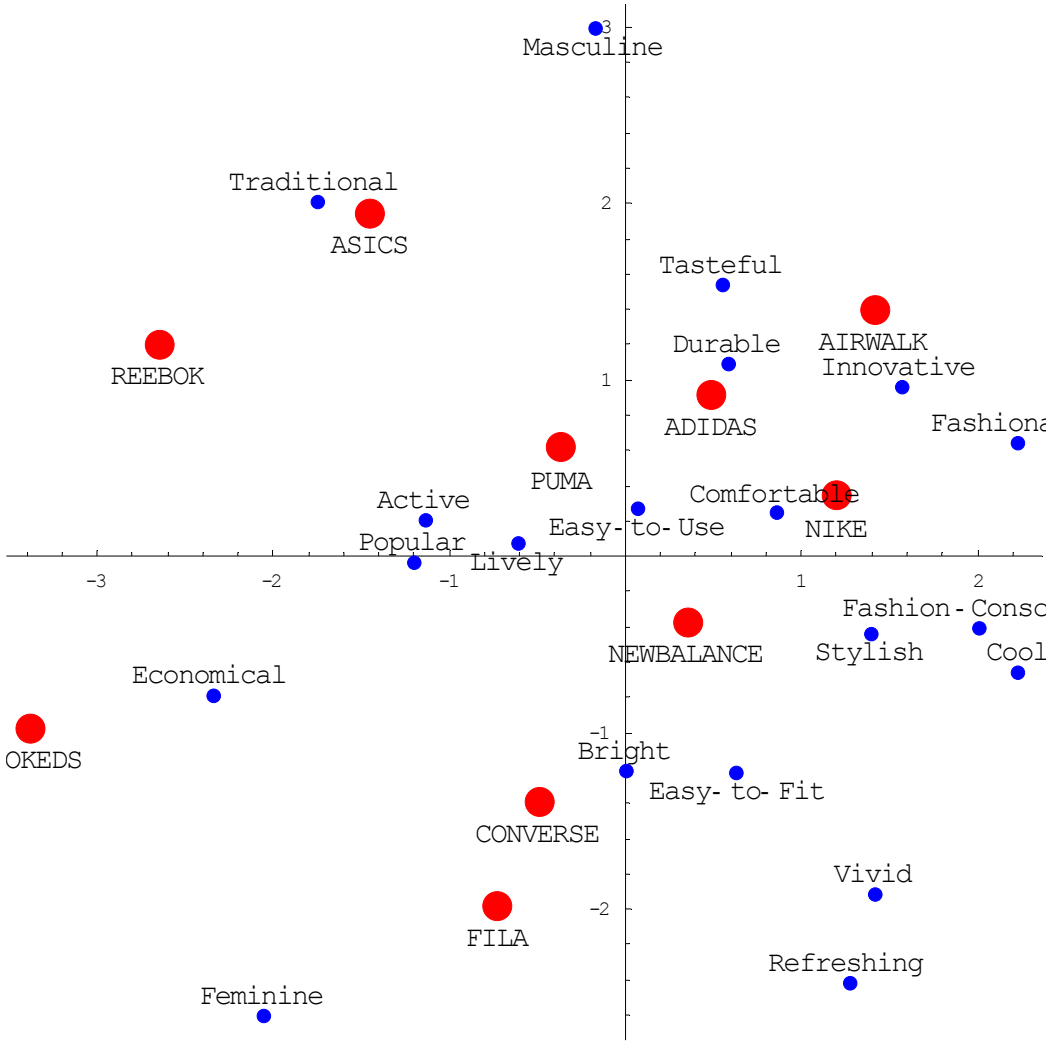


Figure 9 3D Configuration of Adjectives and Brands

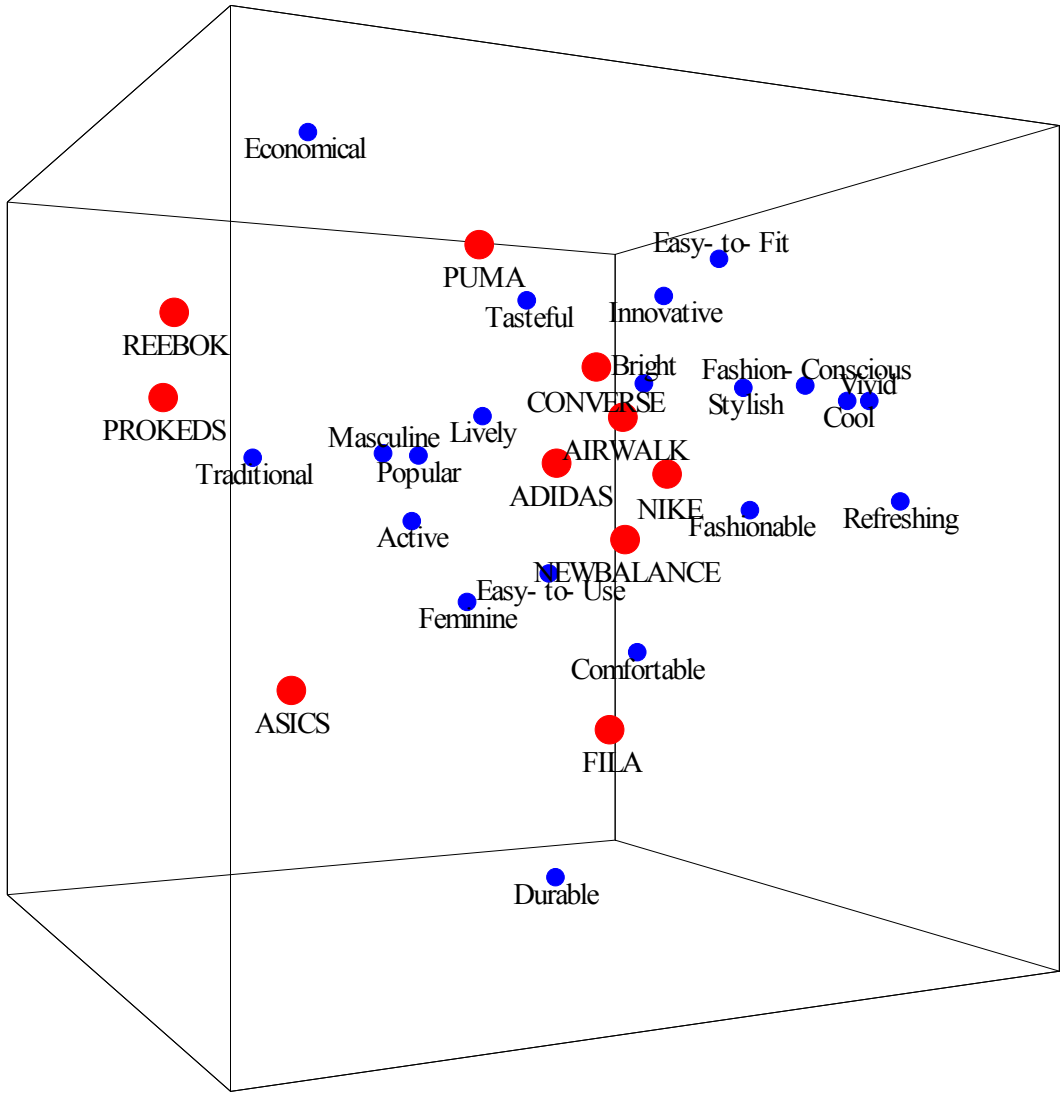


Table 19 Brand Switching Matrix for Japanese Beer Brands/Size Combinations (20 weeks from August 1990 to January 1991)

	I350	I500	S350	S500	S135	S250	L500	L350	L250	D135	H350
I350	440	189	29	16	13	14	19	52	21	9	19
I500	172	360	23	29	8	5	17	8	16	4	9
S350	33	22	430	134	7	47	22	27	11	4	8
S500	17	31	138	297	3	28	8	3	2	1	4
S135	16	11	8	5	141	22	3	5	3	7	4
S250	18	3	52	31	20	146	1	4	13	4	3
L500	17	22	28	8	3	3	336	93	26	3	2
L350	46	6	29	3	5	2	96	240	58	9	2
L250	18	11	8	2	6	25	34	51	245	17	5
D135	12	7	6	2	16	2	1	9	17	75	5
H350	13	5	9	8	3	1	4	7	6	4	105

Abbreviations: I=Ichibanshibori (Kirin), S=Super Dry (Asahi), L=Lager (Kirin), D=Draft (Kirin), H=Hokkaido (Sapporo)

From A. Inoue (1992), "A System for Identifying Competitive Groups with Population Heterogeneity," *JIMS Marketing Science*, Vol. 1, No. 1・2, 12-37.

Figure 10 Asymmetric Competitive Map for Japanese Beer Brands/Sizes

- Clout (Blue) and Vulnerability (Red) Positions -