A MODEL FOR THE ANALYSIS OF ASYMMETRIC DATA IN MARKETING RESEARCH*

RICHARD A. HARSHMAN, † PAUL E. GREEN, ‡ YORAM WIND‡ AND MARGARET E. LUNDY†

Over the last decade, numerous methods for the multidimensional scaling (MDS) of perceptions and preferences have been applied by researchers in marketing. However, one notable gap in MDS methodology has been the lack of suitable models for analyzing inherently asymmetric data relationships. Recently, Harshman (1978, 1982a) has proposed a new family of models—called DEDICOM (DEcomposition into DIrectional COMponents)—for analyzing data matrices that are intrinsically asymmetric.

In this article, the single-domain DEDICOM model is described and applied to two illustrative cases in marketing research. The examples demonstrate that DEDICOM solutions will sometimes make more substantive sense and provide significantly better fits to asymmetric data than solutions obtained by factor analysis or MDS. DEDICOM also provides a novel type of information—a description of asymmetric relations among dimensions or clusters. Such information will often have useful marketing implications.

(Multidimensional Scaling; Factor Analyses)

Introduction

Since the publication of Roger Shepard's pioneering articles on multidimensional scaling (Shepard, 1962), MDS methodology has received wide attention and application by researchers in the behavioral and administrative sciences. One of the most active applied fields has been marketing research; since the late 1960s, MDS has been used in hundreds of marketing studies in industry and government. For the interested reader, several general state-of-the-art reviews of MDS (Shepard, 1974; Carroll, 1976; Carroll and Arabie, 1980) have appeared. Also, at least one review directed specifically to marketing research (Green, 1975) has been published.

*Received January 1981. This paper has been with the authors for 3 revisions.

†University of Western Ontario, London, Ontario N6A 5B9, Canada.

‡The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104.

MARKETING SCIENCE
Vol. 1, No. 2, Spring 1982
Printed in U.S.A.

0732-2399/82/0102/0205$01.25

In MDS the prototypical data matrix that underlies the analysis of perceptual judgments is square and symmetric, representing a set of empirically obtained similarities or other kinds of proximities between pairs of objects (e.g., brands, sales representatives, etc.). The typical representation of such data is spatial, where pairs of objects perceived to be highly similar plot close to each other in some multidimensional space of relatively low dimensionality (e.g., two or three dimensions).

However, there is a broad class of data in which the data entries of an $n \times n$ matrix of relationships among pairs of n objects are *not* symmetric; that is, where the cell entry $\delta_{ij} \neq \delta_{ji}$ for $i, j = 1, 2, \ldots, n$; $i \neq j$. Examples include:

- The estimated probability of a consumer switching to brand j, given that brand i was bought on the last purchase occasion.
- The frequency with which brand j is incorrectly perceived to be brand i, in a tachistoscopic test of new package designs.
- Trade flows between countries and input/output distribution flows between industries.
- The subjective likelihood that subjects believe that a person has trait j, given that the person is described as having trait i.
- One's preference for brand j, given that one has already obtained his/her first choice, brand i.

Other examples could be mentioned as well. Suffice it to say that the collection of asymmetric data is a common phenomenon in marketing research and in the behavioral sciences generally.

Until quite recently, asymmetric data were usually treated as something of a nuisance. Such data were either to be avoided at the outset, if possible, or otherwise "symmetrized" in some way. Various procedures were used (including additive or multiplicative adjustments of rows and columns) to bring about the desired matrix symmetry if the asymmetries were assumed to be due to various kinds of response biases (e.g., Levin and Brown, 1979). In cases where the asymmetries were believed to reflect only noise, the easiest procedure was simply to average the counterpart off-diagonal proximity entries, δ_{ij} and δ_{ji}.

In still other cases—assuming that the asymmetric character of the data was to be preserved—rows and columns were treated as separate points, and an unfolding-type of MDS analysis was carried out, leading to a spatial configuration of 2n points (Gower, 1977; Carroll and Arabie, 1980). This required that separate row and column identities be distinguished and interpretation of the configuration often tended to be difficult.

Over the last few years or so, researchers in psychometrics have developed new approaches to the analysis of intrinsically asymmetric data (Chino, 1978; Constantine and Gower, 1978; Gower, 1977; Tobler, 1979; Young, 1974; for a more detailed review see Harshman, 1982a). One family of these models, developed by the first author (Harshman, 1978; 1982a), is DEDICOM (DEcomposition into DIrectional COMponents). Unlike previous attempts to

adapt distance-like methods of MDS to asymmetric relationships, DEDICOM takes a nonspatial approach at the outset (although subsequent spatial representation is possible).

The purpose of this paper is to describe the simplest member of the DEDICOM family, the single domain, two-way DEDICOM model for square data sets, and discuss its application to two illustrative problems in marketing research. (For more extensive technical discussion of the entire family of DEDICOM models, see Harshman, 1978; 1982a; 1982b.) We first discuss the formal characteristics of the single domain model in the context of an applied problem in the analysis of free associations data. Following this, the model is applied to a set of brand switching data obtained in an actual marketing research study. The paper concludes with a discussion of some future methodological developments, and other kinds of applications where the model may prove useful. A technical appendix describes ways in which various versions of the basic model are solved.

The Single Domain DEDICOM Model

To motivate discussion, consider the small 8×8 matrix of free associations shown in Table 1. These data are drawn from a larger table, originally published in the *Journal of Marketing Research* (Green, Wind, and Jain, 1973). The table entries represent the frequencies (across people) with which each column phrase was evoked in the minds of respondents when the interviewer called out (in randomized order) each of the eight row phrases. All data are based on the responses of 84 female users of hair shampoo, who were 19 to 30 years of age. Multiple responses for each respondent were recorded.

The first thing to be noticed about Table 1 is that the entries are not symmetric; for example, "Body" evokes "Fullness" 44 times, while "Fullness" evokes "Body" only 22 times. It turns out that these asymmetries are more systematic than would be expected on the basis of chance. With DEDICOM we will be able to fit both a symmetric (factor analytic or MDS model) and a more general asymmetric model to the same data, and compare the results. A

TABLE 1

Word Association Frequencies

Stimulus Phrase	Evoked Phrase: Body	Fullness	Holds Set	Bouncy	Not Limp	Manageable	Zesty	Natural
Body	—	44	5	23	1	19	1	3
Fullness	22	—	5	3	1	9	1	2
Holds Set	17	21	—	5	0	17	0	5
Bouncy	15	12	3	—	1	5	0	14
Not Limp	28	27	4	18	—	4	1	7
Manageable	17	13	11	2	0	—	0	3
Zesty	7	9	2	22	0	4	—	13
Natural	4	9	1	2	0	7	1	—
Total	110	135	31	75	3	65	4	47

statistical test (to be described below) indicates that the asymmetric model provides a significantly ($\alpha = .05$) bigger improvement in fit than would be expected if the directions of the asymmetries were randomly determined.

In order to keep our example simple, let us assume that there are "really" only two different aspects of shampoo benefits which influence associations among the eight phrases. Each of the eight evoking and evoked phrases can then be described in terms of two numbers, which show how strongly the phrase is related to each underlying aspect. For example, (to anticipate the results actually obtained with these data) the two aspects might be expressed as THICKNESS and VIGOR. Phrases such as "Body" and "Fullness" would probably show high association with THICKNESS while "Not Limp" and "Zesty" might show high association with VIGOR. DEDICOM teases out such underlying aspects in a manner similar to factor analysis or MDS. But DEDICOM carries the analysis further. The two aspects, THICKNESS and VIGOR, are themselves assumed to show associative relationships, and these "latent" associative relationships may be asymmetric, just as the observed or "surface" associations are asymmetric.

DEDICOM provides us with a table describing the (usually asymmetric) associative relationships among these underlying aspects (dimensions or types), and shows us how to build up the original (observed) associations from linear combinations of the latent associations among the aspects. Unlike MDS, where observed non-spatial relationships are modeled in terms of underlying spatial relationships, DEDICOM assumes that both the observed and the latent relationships are of *the same kind.* The usefulness of the solution comes from the fact that the latent relationships are much simpler, and reveal the patterns present in the surface relationships.

While the aspects of phrases uncovered by the analysis can be treated as dimensions, and plotted so as to express relationships among phrases geometrically, such a plot reveals only part of the story. It shows the similarity of patterns for different stimuli, but not the important asymmetries. To understand how asymmetries are represented, we consider the algebra of the model.

The Algebra of the Single Domain DEDICOM *Model*

Let $\mathbf{X}$ denote the original $n \times n$ matrix of asymmetric relationships. A general entry of $\mathbf{X}$, denoted as x_{ij}, represents the directed relationship of object i to object j. The single domain DEDICOM model can be written as:

$$\mathbf{X} = \mathbf{ARA}' + \mathbf{E}. \tag{1}$$

In this DEDICOM model, $\mathbf{A}$ denotes an $n \times q$ (vertical) matrix of weights of the n observed objects on a relatively few ($q < n$) aspects or types of objects. $\mathbf{R}$, of order $q \times q$, is usually an *asymmetric* matrix, giving the directional relationships among the basic types. The matrix $\mathbf{A}'$ is the $q \times n$ transpose of $\mathbf{A}$, while E is simply a matrix of error terms.

We might compare this model to familiar factor analytic and MDS models. The **A** matrix is analogous to the factor loading matrix in factor analysis, and to the stimulus projection matrix in multidimensional scaling. It relates the observed entities (e.g., the phrases) to the latent entities (e.g., the aspects of phrases). The DEDICOM **R** matrix might be considered analogous to Φ, the matrix of correlations among factors in oblique factor analysis, or to the cosines of angles among dimensions in multidimensional scaling with oblique perceptual axes. The central matrix in all these cases describes the interrelations among the latent entities. DEDICOM generalizes factor analysis and MDS by allowing these latent relationships to be asymmetric. In fact, DEDICOM analysis reduces to factor analysis when the input data are symmetric (covariance or correlation-like) and the option to ignore the diagonal is used. (It reduces to principal components analysis if such data are input and the option to ignore the diagonal is not used.) Similarly, DEDICOM reduces to metric multidimentional scaling if the data are symmetric and are originally like Euclidean distances, but are preprocessed to become scalar product-like (see Harshman, 1982a). Conceptually, however, DEDICOM might be considered to diverge substantially from factor analysis and MDS, since it does not describe the observed nonspatial relations in terms of a spatial analog but, instead, describes the observed nonspatial relations in terms of latent nonspatial relations of the same kind as the surface relations.

The interesting feature of the single domain DEDICOM model (which is the only version that we shall employ in this article) is that while **R** is allowed to become asymmetric, the left and right hand **A** matrices are still required to be identical. This provides a description of the data in terms of asymmetric relations among a *single* set of things, rather than envisioning a different set of aspects for the eight words in their "evoking" role than they have in their "evoked" role. This second possible approach, where the left **A** may differ from the right **A**, is called the dual domain DEDICOM model, and is discussed in Harshman (1982a).[1]

Like factor analysis and MDS, DEDICOM requires that the user refer to some external criterion to determine the desired dimensionality (the number of columns of **A** and rows and columns of **R**). As in MDS, we examine fit-vs.-dimensionality curves to look for an "elbow," after which fit values only increase gradually. We also consider the interpretability of the successive solutions. (Significance tests for the fit increases due to each added dimension are also being explored.)

Our first exampe (Table 1), involves a very small data set, containing only 56 distinct values. (The diagonals were undefined, and will be ignored in the analysis by estimating their value from the model in an iterative procedure, identical to that sometimes used for communality estimation in factor analysis.) Given the small number of data values, we are reluctant to extract more than two dimensions. The fit-vs.-dimensionality curve for these data is given

[1] The dual domain model is also described briefly in the technical Appendix of this article.

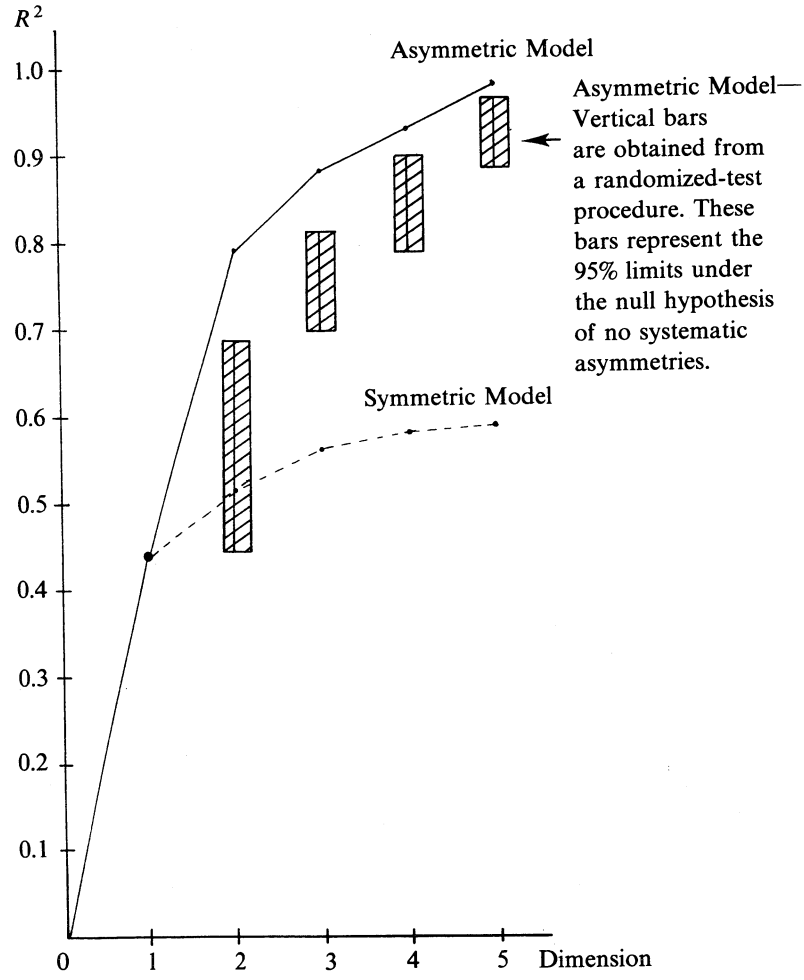

FIGURE 1. Fit Versus Dimensionality—Word Association Frequencies

by the solid line in Figure 1. (The dashed line represents fit values obtained with a symmetric model, and will be discussed later.) The elbow at two dimensions suggests that there are only two major dimensions, although there may be a suggestion of a third dimension. To keep this first example simple, we decide to focus on the two-dimensional solution.

Possible Transformations

Before we can determine specific numerical values for the **A** and **R** matrices in Equation (1), there are several indeterminacies of scale and "rotation" which need to be resolved. Like factor analysis and MDS, DEDICOM dimensions can be linearly transformed into alternative dimensions with no loss of fit to the data. In general, if **T** is any nonsingular q by q transformation matrix, we can define an alternative **A** matrix, $\overset{*}{\mathbf{A}}$, by letting $\overset{*}{\mathbf{A}} = \mathbf{AT}$ and de-

fining the associated **R** matrix $\overset{*}{\mathbf{R}}$ as $\mathbf{T}^{-1}\mathbf{R}\mathbf{T}'^{-1}$. Which **T** we select can depend on which perspective we wish to take toward the interpretation of **A** and **R**. Although in general we might like to seek some kind of approximate simple structure for the columns of **A**, this can be done in different ways, corresponding roughly to different orthogonal or oblique rotation criteria in factor analysis.

In this article, we will adopt as standard a procedure which applies VARIMAX to the columns of **A** *after* they have been scaled to have equal sums of squares; this keeps the columns of **A** mutually orthogonal. (This columnwise orthogonality is not to be confused with orthogonality of the dimensions in the space spanned by **A**; "rotating" **A** so that the columns remain orthogonal produces solutions which are usually oblique in the factor-analytic sense, i.e., the matrix of factor intercorrelations—or its analog, **R**—has fairly large off-diagonal elements.)

But even after deciding on the conventions used to determine the orientation of the axes, a further decision needs to be made concerning the scale of the **A** and **R** matrices. For example, when the data to be predicted contain large values, as here, either **A** or **R** (or both) must contain large entries to predict these observed values. One can standardize **R** to some small scale (e.g., unit sums of squares) and let **A** contain large entries, or standardize **A** (e.g., to have columns whose entries sum to 1.0) and let the entries in **R** be large to predict the observed values. In either case, we will have a solution which in some respect looks different from what we are used to seeing in the factor analysis of correlations, or the (analog) multi-dimensional scaling of scalar products. (A similar problem arises when one does factor analysis of covariances, unless the covariances are standardized, i.e., turned into correlations.)

In DEDICOM, several different options are available for scaling **A** and **R** in various ways. When **R** is standardized (and there are several ways to do this), we gain the benefit of being able to compare across the columns of **A** to see the relative sizes of the contributions that different dimensions make to a given stimulus. When the columns of **A** are standardized, we lose this simple comparability across columns of **A**, but we gain a benefit which, for our own current purpose, is even greater. The **R** matrix can then be interpreted as expressing relationships among the dimensions in the same units as the original data. That is, the **R** matrix can be interpreted as a matrix of the same kind as the original data matrix **X**, but describing the relations among the latent aspects of the phrases, rather than the phrases themselves. (For a complete discussion of scale and rotation in DEDICOM solutions, see Harshman, 1982b.)

In the analyses presented in this article, we have adopted the convention that the columns of **A** are standardized so that they sum to 1.0. This is feasible because we are dealing with data which contain all positive values, and our columns of **A** will contain mostly positive entries. With other data, we might adopt the convention that the sum of *squares* for each column of **A** sums to 1.0, which is obviously feasible regardless of the sign pattern of the data or of the entries in **A**.

When we adopt the convention that the column sums of first powers of entries equal 1.0, we will obtain entries in **A** that are smaller than those usually found in factor analysis and MDS. We should remember that it is the *relative* size of an entry, compared to other entries in the same column, which we must evaluate in deciding which loadings are "large." If we choose to think of each dimension as an aspect, or perhaps a "fuzzy cluster" of the observed phrases, then the loading which occurs in a given column represents the proportion of the total influence or effects of the aspect or cluster which is attributable to that particular phrase. In cluster terms, it is the proportion of the total cluster that is "occupied by" the phrase in question. When we have several different phrases participating in a given cluster, naturally no one phrase will usually occupy more than a small proportion of the cluster. Instead of looking for loadings greater than 0.3, as is traditional in factor analysis, we might look for numbers greater than $(1.5/N)$ or $(2/N)$ where N is the number of rows in **A**. Such entries would be more than 1.5 or 2 times the mean loading for that column.

When the columns of **A** are standardized to have sums equal to 1.0, then the **R** matrix can be thought of as a compressed or miniature version of the original **X** matrix. The sum of all the elements in **R** is equal to the sum of all the elements in $\hat{\mathbf{X}}$ (the part of **X** fit by the model). This is easily proven, since

$$\sum_i \sum_j \hat{x}_{ij} = \sum_i \sum_j \left[\sum_s \sum_t a_{is} r_{st} a_{jt} \right] = \sum_s \sum_t r_{st} \sum_i a_{is} \sum_j a_{jt} \tag{2}$$

and since $\sum_i a_{is} = 1.0$ and $\sum_j a_{jt} = 1.0$ by our adopted convention, then:

$$\sum_i \sum_j \hat{x}_{ij} = \sum_s \sum_t r_{st}(1.0)(1.0) = \sum_s \sum_t r_{st} . \tag{3}$$

Furthermore, the sum of all the contributions to $\hat{\mathbf{X}}$ arising from the relationship of aspect or cluster s to aspect or cluster t can be written as:

$$\sum_i \sum_j (\hat{x}_{ij} \text{ due to } s \to t) = \sum_i \sum_j a_{is} r_{st} a_{jt} = r_{st} \sum_i a_{is} \sum_j a_{jt} = r_{st} . \tag{4}$$

Hence, each element of **R** gives the sum of the influence acting from cluster or aspect s to cluster or aspect t.

Analysis of the Word Associations Data

A two-dimensional DEDICOM analysis of the shampoo word association data is presented in Panel A of Table 2. The **A** matrix has been rotated by VARIMAX approximation to simple structure, keeping the columns of **A** orthonormal, and then the columns of **A** have been scaled to have sums of 1.0. The first column might be interpreted as representing THICKNESS of hair, since the phrases "Body" and "Fullness" have high loadings. The second

TABLE 2

A *and* **R** *Matrices from the Two-Dimensional Solutions for the Word Associations Data*

Panel A Asymmetric DEDICOM Model			Panel B Symmetric (Factor Analytic) Model		
	Dimension			Dimension	
	1	2		1	2
A MATRIX	(THICK-NESS)	(VIGOR)	**A** MATRIX	(THICK-NESS)	(BOUNCY)
1 Body	.299	.252	1 Body	.305	.081
2 Fullness	.355	−.158	2 Fullness	.286	−.071
3 Holds Set	.041	.213	3 Holds Set	.134	−.064
4 Bouncy	.172	.004	4 Bouncy	.002	.647
5 Not Limp	−.048	.420	5 Not Limp	.097	.096
6 Manageable	.150	.013	6 Manageable	.154	−.063
7 Zesty	−.043	.248	7 Zesty	−.002	.224
8 Natural	.074	.010	8 Natural	.022	.151
R MATRIX			**R** MATRIX		
1	248	44	1	372	68
2	216	24	2	68	77

column might be labeled VIGOR, since "Not Limp" has a high loading and "Zesty" and "Holds Set" also have sizable loadings. We should also note that "Body" has a high loading on this dimension, suggesting that "Body" has two different aspects of meaning. Weaker loadings can also be interpreted, although this should be done with caution since they represent less overlap of meaning. For example, the phrases "Bouncy" and "Manageable" seem primarily related to THICKNESS, but quite a bit of their meaning apparently resides outside these two dimensions (or they are generally less effective as stimuli) since they have smaller loadings overall. If we are to interpret the columns of **A** as "fuzzy clusters" then we must consider them "additive" clusters (in the sense defined by Carroll and Arabie, 1980). That is, a given object can participate in more than a single cluster. The phrase "Body" seems to be related to both THICKNESS and VIGOR. It contributes 29.9% of the effects attributable to THICKNESS and 25.5% of those attributable to VIGOR. The phrase "Fullness" also has loadings in both columns, but has a negative loading in the second column, indicating that it is associated with less VIGOR rather than more VIGOR. The actual size of these effects, however, depends on whether one considers the ability of the two aspects (or clusters) to *evoke* responses, vs. the tendency of the two aspects (or clusters) to *be evoked* as responses. The **R** matrix reveals that both clusters are almost equally potent as elicitors but THICKNESS occurs much more often than VIGOR as an evoked response.

The **R** matrix in Panel A of Table 2 can be interpreted as giving the frequency with which each of the aspects or clusters evokes each of the other aspects or clusters, and themselves. We can imagine a hypothetical experiment wherein each of the eight stimulus phrases was replaced by the stimuli THICKNESS or VIGOR (or some combination of the two), and the 84 users

of shampoo were instructed to respond with either THICKNESS OR VIGOR (or some combination of the two). The entries in the **R** matrix represent the frequencies with which each stimulus is predicted to evoke itself or the other stimulus. The sum of the entries in **R** is 532. This equals the sum of the predicted frequencies of association between the original phrases, as given by $\hat{\mathbf{X}}$. Thus, **R** breaks the total pattern of associations into four additive components, each arising from the relation of one cluster or aspect to the other cluster or aspect (or to itself). (In one sense our hypothetical experiment is a bit misleading; the diagonal cells of **R** are better interpreted as the strength of association among members of the same cluster, or the strength of association arising from a stimulus and response sharing the same aspect of meaning, rather than in terms of things eliciting themselves.)

Because the entries in **R** represent summed numbers of associations among clusters or composites of stimuli, they are considerably larger than the individual entries in **X**. We might try to reduce **R** to more familiar size by noting that each of the clusters or aspects embraces roughly half the stimuli. So we might surmise that the frequencies in **R** represent composites of roughly four things, and hence should be divided by four to restore familiar-size row sums. When this is done, we obtain:

$$\mathbf{R} = \begin{array}{|c|c|}\hline 62 & 11 \\ \hline 54 & 6 \\ \hline \end{array}$$

This reduces our frequencies appropriately so that now we have roughly comparable row sums to the raw data. However, the columns have also been compressed from 8 to 2, so in order to get cell frequencies on the same scale as in the raw data; we should divide by 4 an additional time, giving 15, 3, 14, 2 as cell frequencies.

We note that there is a large asymmetry in **R**. Words from the THICKNESS cluster elicit words from the VIGOR cluster only 44 times, where VIGOR elicits THICKNESS 216 times. This striking asymmetry reflects observed asymmetries in the data. (Compare, for example, how often "Not Limp" elicits "Body" vs. "Body" eliciting "Not Limp.") In general, we see that words in the VIGOR cluster evoke both VIGOR and THICKNESS associations but do not get evoked (even by members of the same cluster). This could be useful information for someone writing advertising copy. It might, for example, suggest the importance of stating VIGOR-type benefits explicitly, while THICKNESS could often be simply implied.

Individual Data Points

Further insight into the model can be gained by considering how individual data points are represented. Equation (1) can be written in scalar form as:

$$x_{ij} = \sum_{i=1}^{q} \sum_{j=1}^{q} a_{is} r_{st} a_{jt} + e_{ij} . \tag{5}$$

We can either interpret this expression in cluster terms, or in aspect (dimensional) terms. In cluster terms, a_{is} is the proportion of cluster s identified with manifest phrase i, r_{st} is the total number of associations in the data arising from cluster s stimuli evoking cluster t responses, and a_{jt} is the proportion of cluster t identified with manifest phrase j. In aspect terms, a_{is} is the proportion of the total associations involving aspect s which can be attributed to manifest phrase i, r_{st} is the total number of associations in the data that can be attributed to the s-aspect of the stimulus evoking the t-aspect of the response, and a_{jt} is the proportion of the total number of associations involving aspect t which can be attributed to manifest phrase j.

In either case, our understanding of the predicted frequency for x_{ij} is basically the same. The degree to which phrase i evokes phrase j entails the sum of q^2 components (in this case four components), where each component is a triple product of the kind shown in Equation (5). Each triple product expresses a directional link between an aspect of i and an aspect of j. (In cluster terms, each product expresses a directional link arising from the participation of i in a certain cluster and the participation of j in the same or some other cluster.)

For example, the predicted extent to which "Body" evokes "Not Limp" is the sum of how much each of the two aspects in "Body" evokes each of the two aspects in "Not Limp": (0.299) (248) (−0.048) + (0.299) (44) (0.420) + (0.252) (216) (−0.048) + (0.252) (24) (0.420). This equals (−3.7) + (5.5) + (−2.7) + (2.5) or 1.3, in total. Despite round-off errors this comes close to our observed frequency of 1. Likewise, the predicted extent to which "Not Limp" evokes "Body" is the sum of how much each of the two aspects of "Not Limp" evokes each of the two aspects of "Body." These four links total 25.8, a value not far from the observed frequency of 28.

We see that each observed association has been decomposed into (four) directional components. These components are small in the case of the association from "Body" to "Not Limp," but there is one big one in the case of "Not Limp" to "Body." We can interpret the DEDICOM account of the asymmetry as follows: There is little association between these words arising from shared meaning or within-cluster bonds. This is because the two phrases do not overlap in the aspect of THICKNESS (since "Not Limp" has virtually none of this aspect). There is some overlap in terms of VIGOR, but this turns out to contribute little because the "self-associations" of the VIGOR cluster are weak (only 1/10 those of THICKNESS). Thus the associative relation between "Body" and "Not Limp" is determined primarily by the connections between *different* aspects. "Body" has primarily THICKNESS and "Not Limp" has almost exclusively VIGOR. As the **R** matrix indicates, the association from THICKNESS to VIGOR is weak, whereas the association from VIGOR to THICKNESS is strong. Thus the association from "Body" to "Not Limp" is weak, but from "Not Limp" to "Body" is strong.

If our data set were larger than 8 by 8, we might try to break these two broad clusters up into finer components, by examining higher-dimensional solutions. As we shall see with a subsequent example, one can obtain a tree

structure similar to that provided by hierarchical clustering, which can shed light on the clusters which emerge at any given level by reference to their common "ancestors" at lower-dimensional levels, and their divergent "children" at higher-dimensional levels. We have looked at the higher-dimensional solutions for these data, and find that in three dimensions, "Bouncy" splits off to dominate a new dimension by itself with "Natural" and "Zesty" only weakly related. At four dimensions, a dimension of CONTROL emerges, which relates to "Holds Set" and "Manageable" and to a lesser degree, to "Natural." While this may seem like a reasonable dimension, there is some indication that we have extracted more than the data can support, particularly in the **R** matrix where a number of substantial negative entries emerge. All told, we conclude that the higher-dimensional solutions may be suggestive but are not adequately justified in the light of the small data set and the fit-vs.-dimensionality curve.

Comparison with a Symmetric Model

It is interesting to compare the DEDICOM analysis which incorporates asymmetry with one which does not, in order to assess the benefits (if any) of the more general asymmetric model. Certain earlier versions of the program included a capability for constraining the **R** matrix to be symmetric. This results in a model which only generates symmetric predictions (i.e., $\hat{\mathbf{X}}$ is necessarily symmetric). The fit of this model can then be compared to the fit of the asymmetric version.

It turns out, however, that fitting a symmetric DEDICOM to asymmetric data gives exactly the same **A**, **R**, and $\hat{\mathbf{X}}$ as fitting the asymmetric model to data which have been "symmetrized" by averaging the corresponding x_{ij} and x_{ji} entries. [This equivalence follows as a consequence of the fact that the symmetric and skew-symmetric components of the data are orthogonal to one another (see Harshman, 1982a).] Consequently, we can evaluate the advantages of including asymmetry in the model by simply comparing a DEDICOM analysis of the full asymmetric data matrix with a DEDICOM analysis of the symmetrized data. The fit-vs.-dimensionality curve for the "symmetric model" shown in Figure 1 was obtained by fitting the general DEDICOM model to symmetrized data, then computing the correlation between the $\hat{\mathbf{X}}$ (from that symmetric solution) and the original asymmetric data. This gives the same values that would be obtained by correlating the $\hat{\mathbf{X}}$ of a model with **R** constrained to be symmetric, and the full asymmetric data, when such a constrained model is fit to asymmetric data.

We see that there is no difference between the two models for the one-dimensional solution. This is because the one-dimensional solution is necessarily symmetric in either model. For two and higher-dimensionality solutions, however, there is a very large difference in fit values (e.g., for two dimensions, the symmetric model has an r^2 of 0.51, whereas the asymmetric model has an r^2 of 0.79).

In addition to comparing the fit values, it is useful to compare the solutions obtained by the asymmetric and symmetric models. Table 2 displays the

solutions obtained in the two-dimensional case. In the symmetric model's **A** matrix (Panel B), the first dimension resembles dimension one of the asymmetric analysis; it also seems to stress "Fullness" and "Body," and to a weaker extent "Manageable." The second dimension, however, is quite different from its asymmetric counterpart. The phrase "Bouncy" dominates this dimension (and no longer loads on dimension one). Dimension two also has a weak relationship to "Zesty" and perhaps "Natural." On the other hand, the asymmetric DEDICOM solution placed "Bouncy" in with the other indicators of THICKNESS, in part because the asymmetries displayed by "Bouncy" are similar to those displayed by other indicators of THICKNESS. Since this information is lost in the symmetric analysis, "Bouncy" appears to emerge as a dimension of its own. (In higher-dimensional solutions, both symmetric and asymmetric analyses show a dimension dominated by "Bouncy," but differ in other aspects; such details will not, however, be discussed here.) We conclude that the two-dimensional asymmetric solution has a more interesting and informative interpretation.

What is "Real" and What is Not?

Having obtained an interesting DEDICOM solution for the word associations data, we would like to assess the stability and generalizability of our results, and to determine which characteristics arise from systematic relationships in the data and which are simply reflections of chance fluctuations. The most straightforward approach to this type of evaluation is replication of the analysis on a new sample of data. A well-designed study would anticipate the need to test replicability, and gather enough data for two sizable split-half samples to be constructed. This allows us informally to assess the degree to which any conclusions we would like to draw will tend to generalize across other respondent samples. If generalizability to new stimuli is an issue, then splitting the stimulus set, or replication of the study with new stimuli, is called for. Since here we are analyzing data from other sources (e.g., published commercial studies), we are not free to collect our own replication samples; hence we will not demonstrate this important checking procedure here. However, such split-half checks are recommended as a powerful way of determining, for example, how many dimensions can be extracted before a solution becomes too unstable. Our lack of rigor here is mitigated by the fact that our focus of attention is primarily on the comparison of symmetric and asymmetric representations of the data, rather than the substantive conclusions arising from the examples.

A question of particular importance is the "reality" of structure in the asymmetries of a given data set. We would like to know whether the attention which we have paid to the asymmetries is justified. Are these asymmetries systematic, or are they simply chance fluctuations from a symmetric structure? A method has been developed for testing the statistical significance of the asymmetric variance fit by the DEDICOM model (Harshman, 1982a). It is based on a comparison of the fit value obtained by the asymmetric DEDICOM model with the fit value obtained by the constrained symmetric

DEDICOM model. As we have noted earlier, the DEDICOM model with symmetric **R** is equivalent to the oblique factor analytic model; though we scale **A** and **R** differently, this does not affect the structure or fit of the obtained solution. Another way of stating our comparison is that we are concerned with the statistical significance of the improvement in fit obtained by going from a factor analysis of the symmetrized data to a DEDICOM analysis of the raw data.

To test the statistical signficance of the difference in fit between the two models, a randomization-test procedure was employed. The logic of this procedure is described more fully in Harshman (1982a) but we summarize it briefly here. We formulate a null hypothesis that the deviations of the data from symmetry are due to independent random perturbations of an underlying symmetric structure. The size of these perturbations may be related to the size of the symmetric observations which they perturb, but the *direction* of the perturbation is assumed to be random. According to this null hypothesis, there is no systematic structure to the direction of the asymmetries and hence the improved fit provided by the asymmetric DEDICOM model is simply capitalization on error. In other words, the fact that x_{ij} is larger (say) than x_{ji} is due to chance, and this relation has a 50 percent probability of being reversed in a new sample. To determine the fit values to be expected on the basis of an asymmetric DEDICOM capitalizing on error, 19 modified versions of the data set were constructed. In each modified version a randomly selected half of the data entries was transposed (x_{ij} became x_{ji} and vice versa, for that particular randomly selected value of i and j). This procedure randomly scrambled the asymmetries, while leaving the symmetric component of the data unaltered. (With this method of scrambling, any relationship between the size of the asymmetry and the size of the symmetric part—if there is such a relation—is preserved; it is only the *direction* of the asymmetry which is randomly scrambled.)

These 19 modified data sets were each analyzed in one to five dimensions by the asymmetric DEDICOM model, and the obtained fit values were compared to the 20th or unaltered version. If the unaltered version had a higher fit value than any of the 19 versions with randomly scrambled asymmetries, it was concluded that we had observed a relative improvement in fit due to the use of the asymmetric model which would occur by chance 5 percent of the time or less, and hence was "statistically significant" at the 0.05 level.

The range of the 19 randomized-asymmetry fit values obtained at each dimensionality is shown in Figure 1 by vertical bars. On the basis of these results it is concluded that, for the word association data, the observed improvement in fit provided by the asymmetric over the symmetric DEDICOM model is significant at the 0.05 level; the null hypothesis of no systematic relationships in the direction of the asymmetries can be rejected.

Not only does this test provide evidence for systematic variation in the asymmetries, but is also demonstrates that they are of a kind that can be fit (at least in part) by the DEDICOM model. Thus, this approach is preferable to

(or at least "stronger" than) simply showing that the asymmetries are non-randomly distributed. It demonstrates that a statistically significant amount of the asymmetric variance can be fit with a DEDICOM representation.

The Car Switching Data

Free associations data represent only one class of marketing research data suitable for the application of DEDICOM. An additional example should serve to show the flexiblity of the model and some of DEDICOM's preprocessing options. Each year Rogers National Research, a marketing consulting firm in Toledo, Ohio, collects car trade-in data for a large sample (tens of thousands) of U.S. buyers of new cars. The data are collected by mail questionnaires asking recent buyers of new cars to indicate both the newly purchased model and the old model (if any) disposed of at the time of purchase. The data analyzed here consisted of 1979-model new car purchases.

The auto industry typically classifies cars into 16 segments, with the specific (1979) models within each segment shown in Table 3. Based on the Rogers data, a 16×16 brand switching matrix was prepared showing the frequency with which any car owner in segment i switched to a new (1979 model) car in segment j ($i, j = 1, 2, \ldots, 16$). The resulting data matrix is shown in Table 4.

The car switching frequencies of Table 4 were subjected to a DEDICOM analysis. For comparison, a symmetric DEDICOM (factor analytic) analysis was also performed. The resulting fit values are shown in Figure 2. A fairly high fit value is obtained in one dimension, and since the one-dimensional solution is necessarily symmetric, the two models naturally have the same fit value at this dimensionality. A strong divergence occurs at two dimensions, however, and continues thereafter up to eight dimensions. Once again, the asymmetric model provides a considerably better fit (e.g., $r^2 = 0.92$ vs. 0.76 in the two-dimensional solution). The fit-vs.-dimensionality curve for the asymmetric model suggests that there are two major dimensions, but careful examination reveals that there are further increments for dimensions three and four, followed by gradual increases thereafter. This suggests that a maximum of four dimensions are present. For the symmetric model, on the other hand, the fit-vs.-dimensionality curve shows little evidence for more than two dimensions.

The difference between the fit values of the asymmetric and symmetric models was tested for statistical significance by the randomization-test procedure described earlier. Again, the improvement of fit obtained with the "true" (unscrambled) asymmetries was substantially higher than any of the improvements obtained in the 19 modified data sets, where the direction of the asymmetries was randomly reflected. (The range of fits obtained with the modified data sets is again indicated—for each dimensionality—by a bar in Figure 2.) Consequently, we conclude that there are systematic asymmetries in the car switching data, and that they are fit by the DEDICOM model at a better-than-chance level, i.e., at a level that would occur with probability of 0.05 or less on the basis of chance.

TABLE 3

Segment Composition of 1979 *Model Cars*[a]

1. Subcompact/Domestic
   - Bobcat
   - Chevette
   - Horizon
   - Monza
   - Omni
   - Pinto
   - Spirit
   - Sunbird

2. Subcompact/Captive Imports
   - Arrow
   - Champ
   - Colt
   - Colt Hatchback
   - Fiesta
   - Opel

3. Subcompact/Imports
   - Datsun 210
   - Datsun 310
   - Fiat 128
   - Honda Civic
   - Honda CVCC
   - Honda Accord
   - Renault LeCar
   - Toyota Corolla
   - Volkswagen Rabbit
     - —Gas
     - —Diesel

4. Small Specialty/Domestic
   - Camaro
   - Capri
   - Corvette
   - Firebird
   - Horizon TC-3
   - Mustang
   - Omni 024
   - Pacer
   - Skyhawk
   - Starfire

5. Small Specialty/Captive Imports
   - Challenger
   - Sapporo

6. Small Specialty/Imports
   - Datsun 280ZX
   - Fiat Spider 2000
   - Fiat X1/9
   - Toyota Celica
   - Volkswagen Scirocco

7. Low Price Compact
   - Concord
   - Fairmont
   - Nova
   - Volare

8. Medium Price Compact
   - Aspen
   - Omega
   - Phoenix
   - Skylark
   - Zephyr

9. Import Compact
   - Audi Fox
   - Datsun 510
   - Fiat Brava
   - Toyota Corona
   - Volkswagen Dasher

10. Midsize Domestic
    - Century
    - Cougar
    - Cutlass Salon
    - Diplomat
    - Granada
    - LeBaron
    - LeMans
    - LTD II
    - Malibu
    - Monarch

11. Midsize Imports
    - AUDI 5000
    - BMW 320I
    - Datsun 810
    - Toyota Cressida
    - Volvo 242/244
    - Volvo 245/265
    - Volvo 264

12. Midsize Specialty
    - Cordoba
    - Cougar XR-7
    - Cutlass Supreme
    - Grand Prix
    - Magnum XE
    - Monte Carlo
    - Regal
    - Thunderbird

13. Low Price Standard
    - Chevrolet
    - Ford

14. Medium Price Standard
    - Buick
    - Chrysler
    - Dodge
    - Mercury
    - Oldsmobile
    - Pontiac

15. Luxury Domestic
    - Cadillac
    - El Dorado
    - Lincoln Continental
    - Mark V
    - Riviera
    - Seville
    - Toronado
    - Versailles

16. Luxury Import
    - BMW 528I
    - BMW 733I
    - BMW 633CSI
    - Mercedes-Benz 240D
    - Mercedes-Benz 300D
    - Mercedes-Benz 300SD
    - Mercedes-Benz 280E/SE
    - Mercedes-Benz 450 SEL
    - Mercedes-Benz 450SL/SLC
    - Porsche 911
    - Porsche 924

[a]SOURCE: Rogers National Research, Toledo, Ohio.

TABLE 4

1979 *Car Switching Data*

From Segment i	To Segment j															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 SUBD	23272	1487	10501	18994	49	2319	12349	4061	545	12622	481	16329	4253	2370	949	127
2 SUBC	3254	1114	3014	2656	23	551	959	894	223	1672	223	2012	926	540	246	37
3 SUBI	11344	1214	25986	9803	47	5400	3262	1353	2257	5195	1307	8347	2308	1611	1071	288
4 SMAD	11740	1192	11149	38434	69	4880	6047	2335	931	8503	1177	23898	3238	4422	4114	410
5 SMAC	47	6	0	117	4	0	0	49	0	110	0	10	0	0	0	0
6 SMAI	1772	217	3622	3453	16	5249	1113	313	738	1631	1070	4937	338	901	1310	459
7 COML	18441	1866	12154	15237	65	1626	27137	6182	835	20909	566	15342	9728	3610	910	170
8 COMM	10359	693	5841	6368	40	610	6223	7469	564	9620	435	9731	3601	5498	764	85
9 COMI	2613	481	6981	1853	10	1023	1305	632	1536	2738	1005	990	454	991	543	127
10 MIDD	33012	2323	22029	29623	110	4193	20997	12155	2533	53002	2140	61350	28006	33913	9808	706
11 MIDI	1293	114	2844	1242	5	772	1507	452	565	3820	3059	2357	589	1052	871	595
12 MIDS	12981	981	8271	18908	97	3444	3693	1748	935	11551	1314	56025	10959	18688	12541	578
13 STDL	27816	1890	12980	15993	34	1323	18928	5836	1182	28324	938	37380	67964	28881	6585	300
14 STDM	17293	1291	11243	11457	41	1862	7731	6178	1288	20942	1048	30189	15318	81808	21974	548
15 LUXD	3733	430	4647	5913	6	622	1652	1044	476	3068	829	8571	2964	9187	63509	1585
16 LUXI	105	40	997	603	0	341	75	55	176	151	589	758	158	756	1234	3124

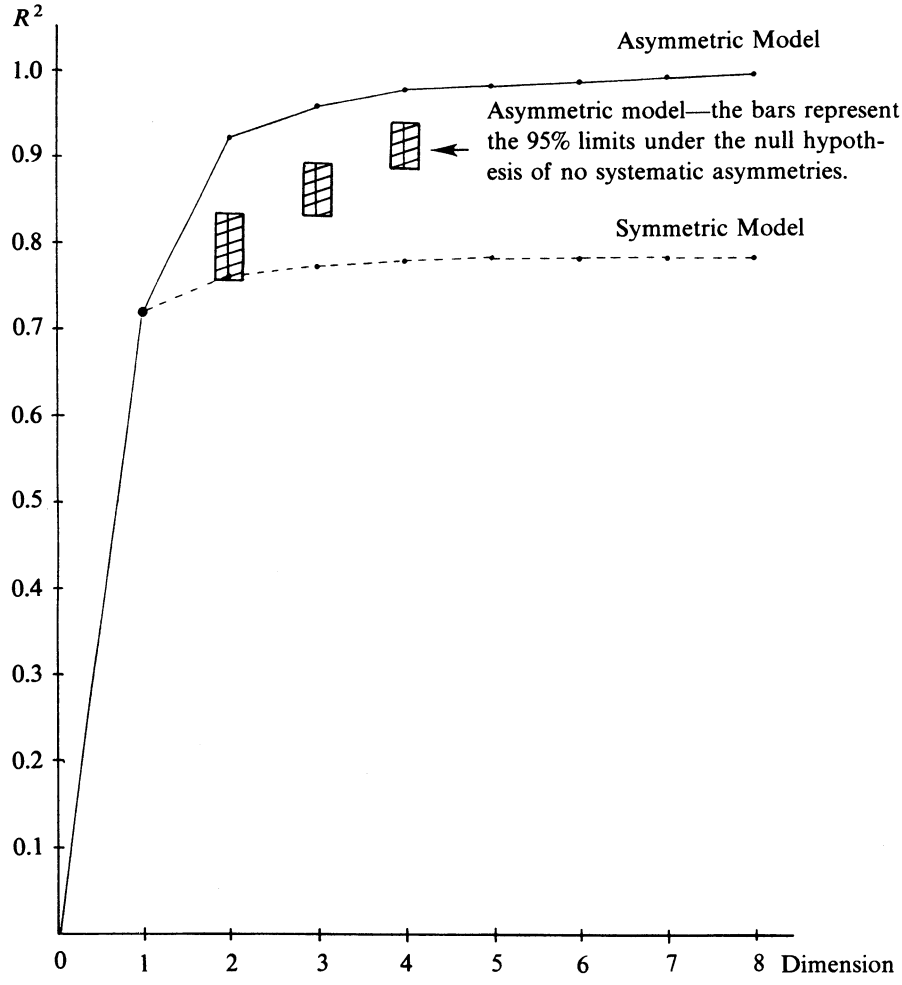

FIGURE 2. Fit Versus Dimensionality—Car Switching Data

Preliminary Interpretation of the Asymmetric Analysis

Examination of the **A** and **R** matrices for the one through five-dimensional solutions reveals that all of the asymmetric solutions are interpretable, up through four dimensions, further confirming our estimate of the "correct" dimensionality as four. The **A** and **R** matrices from the four-dimensional solution are shown in Table 5. This solution can be interpreted as breaking the automobile segments into four aspects: PLAIN LARGE-MIDSIZE, SPECIALTY, FANCY LARGE, and SMALL.

The Midsize Domestic, with a weight of 0.523, most closely resembles the PLAIN LARGE-MIDSIZE, the Midsize Specialty (0.898) appears to strongly resemble the SPECIALTY type, the Medium-Price Standard (0.699) appears

TABLE 5

A *and* **R** *Matrices from the Four-Dimensional Solution for the Car Switching Data*

	Dimension			
	1	2	3	4
A MATRIX	(PLAIN LARGE-MIDSIZE)	(SPECIALTY)	(FANCY LARGE)	(SMALL)
1 SUBD	−.026	−.027	.023	.318
2 SUBC	.012	−.017	−.002	.028
3 SUBI	−.058	.016	.032	.199
4 SMAD	.015	.221	−.078	.157
5 SMAC	−.000	.000	−.000	.001
6 SMAI	.006	046	−.023	.026
7 COML	.148	−.130	−.049	.177
8 COMM	.079	−.050	.015	.058
9 COMI	.012	−.035	.012	.031
10 MIDD	.523	.037	−.015	−.002
11 MIDI	.021	−.004	.000	.009
12 MIDS	.012	.898	.001	−.006
13 STDL	.336	−.015	.101	−.031
14 STDM	.010	−.002	.699	.002
15 LUXD	−.086	.058	.270	.031
16 LUXI	−.003	.004	.013	.002
R MATRIX				
1	161468	120680	89103	226400
2	22508	71778	31029	49225
3	51138	46755	87406	80910
4	83069	74248	10940	234614

to most strongly resemble the FANCY LARGE type, and finally the Subcompact Domestic (0.318) appears to most strongly resemble the SMALL type. But we have a number of puzzling features associated with this solution. Why are so many segments which one would expect to load on a particular aspect failing to do so? For example, why do the Subcompact Captive Import, the Small Specialty/Captive Import, and the Small Specialty Import fail to load on SMALL *or on any other dimension*? We see, in fact, that Small Specialty/Captive Import has a zero loading (to three decimal places) on three dimensions and only a 0.001 loading on the fourth. It appears that there is an *additional* source of variation in the size of the loadings which needs to be taken into account.

Dealing with Differences in Scale

The car switching data present a good example of a problem which is found in many data sets: large differences between data sources and in the size of entries. In the car switching data, for example, some of the 16 segments represent car types which have a large market share, and hence have large switch-from and swtich-to frequencies. Midsize Domestic is such a segment,

with a row sum representing 262,000 switches from this category, and a column sum representing 188,000 switches to this category. On the other hand, for segments with a small market share, there are few switches either to or from the category. The most extreme example is Small Specialty/Captive Import which has a switch-from row sum of only 339 and a switch-to column sum of only 612.

For some data, one might want to retain differences in overall row and column sums, because such differences are presumed to reflect the same kind of influences as those which show differences within each row and column. With the word-associations data, there were modest differences in the row-plus-column sums for different levels of i, but these were presumed to be meaningful because they arose from the differences in associative strength themselves. All the word-association stimuli had been presented an equal number of times, but some stimuli elicited many responses while other stimuli did not. Thus, differences in the number of associations involving each object might be interpreted as reflecting the strength of involvement of the object in the associative network under study.

In contrast, the large difference in level sums (row-plus-column sums) in the car switching data would seem to be due primarily to extraneous factors, such as market share. These overall differences are *not* of the same kind as the within-row differences that reflect factors of interchangeability and relative segment attractiveness to different kinds of buyers. Hence, we would like to remove the extraneous size differences in order to more clearly reveal the structural factors of interest.

The problem of unequal scales for different segments is analogous to problems which arise in factor analysis or cluster analysis because of unequal scales of the variables. For example, some of the variables might be measured in incommensurate units, such as milligrams of hormone per liter of blood vs. the number of correct responses on a 100-item test. In factor analysis, these differences in scale are often considered irrelevant, and are removed by standardization of the variances of the variables (e.g., by using correlation coefficients which implicitly set the variance of all variables to unity). We would like to have a similar way to remove irrelevant differences in scale from our two-way matrix of relationships.

There are two strategies which can be employed (and which are implemented as alternative DEDICOM options) in order to adjust the analysis for differences in overall row and column scale. One is to apply appropriate adjustments to the **A** matrix *after* analysis of the unadjusted data, and the other is to apply rescaling adjustments to the data *before* analysis. With error-free data that fit the model, both approaches have the same result. However, with fallible real data, the two approaches lead to different results because they differ in the implied weighting of errors for small vs. large-scale rows and columns. Here we focus on the method that adjusts the **A** matrix. Later, we briefly discuss the second method and the results obtained from it. However, before either of these methods is applied, a set of row or column scale factors is estimated by performing an iterative rescaling transformation

on a duplicate copy of the data (generated internally). This transformation is designed to remove scale inequalities among different objects. Since the single domain DEDICOM model assumes the same identity for row i as for column i (i.e. the row and the column correspond to the same entity), we seek a single scale adjustment that is applied to both the ith row and the ith column. While there are well-known techniques for equating the sums of rows and columns of a two-way table, when these methods are applied to asymmetric data, they will involve different scale multipliers applied to the rows than are applied to the columns. Techniques did not previously exist for equating the overall scale of the ith (row plus column) total across all n values of i, by means of scale adjustments applied equally (symmetrically) to the rows and columns.

Such a technique has been developed for DEDICOM and is discussed in more detail elsewhere (Harshman, 1982a). Suffice it to say here that an iterative procedure successively adjusts the rows and columns of the data matrix **X** until the mean entry (or equivalently, the sum of all entries) for each segment is equal (i.e., the sum for row i plus the sum for column i equals the sum for row j plus the sum for column j, for all i, j). As a coproduct of this process a single scale factor is determined for each segment. If the scale-adjusted **X** matrix is called $\mathring{\mathbf{X}}$ and the segment scale multipliers are assembled into a diagonal matrix **D**, then the original data are decomposed into a scale-free version $\mathring{\mathbf{X}}$ and a diagonal scale-factors matrix **D**, so that:

$$\mathbf{X} = \mathbf{D}\mathring{\mathbf{X}}\mathbf{D} \tag{6}$$

and it is the case that:

$$\bar{\mathring{x}}_i = \frac{1}{2n} \sum_{j=1}^{n} (\mathring{x}_{ij} + \mathring{x}_{ji}) = k \quad \text{(for all } i\text{)}. \tag{7}$$

(When data are being analyzed in which the diagonal is ignored, as in this article, the data rescaling is based on sums of off-diagonal cells only.) The value of k is usually set at the grand mean of the unadjusted data (or grand mean, ignoring the diagonal cells). Alternatively, similar adjustment procedures are available which equate sums of squares for segments; for the data considered here, however, equalization of sums was thought more appropriate.

Incorporating the Scale Adjustments in the **A** *Matrix*

In the basic DEDICOM single domain model of Equation (1), it is assumed that extraneous scale differences among rows and columns do not exist. Any differences in row and column sums for one stimulus or segment, compared to another, are assumed to be meaningful and to arise from the same structural processes as those which generate within-column and within-row variations (as noted earlier). When this is not the case, **A** must fulfill two functions: it must

express the associational relationships (i.e., the different degrees of relationship between a given cluster or aspect and the various segments) by means of different sized loadings in a given column and it must also reflect any extraneous differences in overall segment size by overall differences in the size of all loadings in a given row. When these two different sources of variation are present in the data, it is useful to split **A** into two component parts, one of which represents the scale factor differences for different segments and the other which represents structural relations among segments. We do this by letting:

$$\mathbf{A} = \mathbf{D}\bar{\bar{\mathbf{A}}} \tag{8}$$

where **D** is the diagonal matrix of scale factors for the segments which was estimated by the iterative process described earlier, and $\bar{\bar{\mathbf{A}}} = \mathbf{D}^{-1}\mathbf{A}$. By explicitly representing scale differences with the matrix **D**, we leave $\bar{\bar{\mathbf{A}}}$ free to represent the structural relationships present independently of scale. $\bar{\bar{\mathbf{A}}}$ can be thought of as giving the relationships among the segments in a data set from which differences in scale have been removed. If we let:

$$\mathbf{X} = \mathbf{A}\mathbf{R}\mathbf{A}' + \mathbf{E} \tag{9}$$

then, substituting $(\mathbf{D}\bar{\bar{\mathbf{A}}})$ for **A**, we obtain:

$$\mathbf{X} = \left(\mathbf{D}\bar{\bar{\mathbf{A}}}\right)\mathbf{R}\left(\mathbf{D}\bar{\bar{\mathbf{A}}}\right)' + \mathbf{E} \tag{10}$$

$$\mathbf{X} = \mathbf{D}\bar{\bar{\mathbf{A}}}\mathbf{R}\bar{\bar{\mathbf{A}}}'\mathbf{D} + \mathbf{E} \tag{11}$$

Ignoring, for the moment, the effects of our error matrix **E**, we can informally say that:

$$\mathbf{D}^{-1}\mathbf{X}\mathbf{D}^{-1} = \bar{\bar{\mathbf{A}}}\mathbf{R}\bar{\bar{\mathbf{A}}}'. \tag{12}$$

Since $\mathbf{D}^{-1}\mathbf{X}\mathbf{D}^{-1} = \mathring{\mathbf{X}}$, our adjusted **X** with the scale differences removed, we obtain (loosely speaking):

$$\mathring{\mathbf{X}} = \bar{\bar{\mathbf{A}}}\mathbf{R}\bar{\bar{\mathbf{A}}}'. \tag{13}$$

In other words, $\bar{\bar{\mathbf{A}}}$ provides a description of the structural relationships with scale differences removed. Now since $\mathbf{A} = \mathbf{D}\bar{\bar{\mathbf{A}}}$ and

$$\bar{\bar{\mathbf{A}}} = \mathbf{D}^{-1}\mathbf{A} \tag{14}$$

we simply remove the scale effects from **A** by premultiplying it by the inverse of the diagonal scale factor matrix.

But after this premultiplication, many desirable columnar properties of **A** are likely to be altered. If the columns had previously been orthogonal, it is unlikely that they would remain so, and if they had been rotated to simple structure, or its VARIMAX approximation, this would probably no longer be the case. Consequently, we might want to subject this row-rescaled **A** to column transformations which restore orthogonality and which would also standardize the sums (or sums of squares) of the column loadings as required by whichever scaling convention we had adopted. (This can be accomplished by performing Gram-Schmidt orthonormalization of the columns of $\bar{\bar{\mathbf{A}}}$, followed by VARIMAX rotation, and then a rescaling of the columns to the desired sum or sum-of-squares.) If these column operations are represented by the q by q transformation matrix **T**, then the new matrix of loadings for the structural part of the relationships can be written as $\bar{\bar{\mathbf{A}}}\mathbf{T}$. Let us call this matrix $\dot{\mathbf{A}}$. This is the matrix that *we wish to interpret* in order to gain insight into the different aspects or clusters underlying the switching patterns in **X**. The $\dot{\mathbf{A}}$ matrix is related to our original **A** by:

$$\dot{\mathbf{A}} = \mathbf{D}^{-1}\mathbf{A}\mathbf{T}. \tag{15}$$

To complete our enriched model for **X**, we need to obtain the appropriate **R** matrix to go with $\mathbf{D}\dot{\mathbf{A}}$. We saw earlier that $(\mathbf{D}\bar{\bar{\mathbf{A}}})$ could use our original **R** unaltered, since $(\mathbf{D}\bar{\bar{\mathbf{A}}}) = \mathbf{D}(\mathbf{D}^{-1}\mathbf{A}) = \mathbf{A}$. But $\dot{\mathbf{A}}$ has been subjected to the column adjustments represented by **T**. Hence the inverse adjustments need to be applied to **R**. A further complication arises because of our need to standardize the size of the columns of $(\mathbf{D}\dot{\mathbf{A}})$ in order for our new **R** matrix to have elements which sum to the same total as the elements in $\hat{\mathbf{X}}$ (or have some other desired property). To achieve this, we first set the columns of $(\mathbf{D}\dot{\mathbf{A}})$ to have sums of one. We do this with a q by q diagonal matrix **C** whose diagonal elements are reciprocals of the column sums of $(\mathbf{D}\dot{\mathbf{A}})$, or equivalently, the column sums of **AT**. We can then let:

$$\dot{\mathbf{R}} = \mathbf{C}^{-1}\mathbf{T}^{-1}\mathbf{R}\mathbf{T}'^{-1}\mathbf{C}^{-1}. \tag{16}$$

To complete this discussion, we can obtain an $\dot{\mathbf{A}}$ to go with $\dot{\mathbf{R}}$. We let $\dot{\mathbf{A}}$

represent the version of $\mathring{\mathbf{A}}$ which incorporates all scale effects, i.e., if we let:

$$\dot{\mathbf{A}} = \mathbf{D}\mathring{\mathbf{A}}\mathbf{C} = \mathbf{ATC}, \tag{17}$$

then we can write a simple expression for the original data in terms of $\dot{\mathbf{A}}$ and $\dot{\mathbf{R}}$, i.e.,

$$\mathbf{X} = \dot{\mathbf{A}}\dot{\mathbf{R}}\dot{\mathbf{A}}' + \mathbf{E}. \tag{18}$$

This representation differs only by "rotation" (or oblique transformation) from our original $\mathbf{ARA}'$. Whereas the original $\mathbf{A}$ had columns with maximal (VARIMAX) simple structure, just as they have appeared in $\mathbf{A}$, the $\dot{\mathbf{A}}$ matrix has columns oriented in a slightly different set of directions in the space spanned by $\mathbf{A}$. These revised directions are chosen so that the $\mathring{\mathbf{A}}$ *part* of $\dot{\mathbf{A}}$ will have maximal VARIMAX simple structure. This provides a maximally interpretable description of the clusters as defined in their "pure" associational structure (with scale differences removed). Since $\dot{\mathbf{A}} = \mathbf{D}\mathring{\mathbf{A}}\mathbf{C}$, we can expand the expression for $\mathbf{X}$, given above, into the interpretable components:

$$\mathbf{X} = (\mathbf{D}\mathring{\mathbf{A}}\mathbf{C})\dot{\mathbf{R}}(\mathbf{D}\mathring{\mathbf{A}}\mathbf{C})' + \mathbf{E} \tag{19}$$

This provides a representation in terms of four different matrices: $\mathbf{D}$ gives the scale factors for the different segments in $\mathbf{X}$ (i.e., the overall size differences for the different levels of $\mathbf{X}$) and $\mathring{\mathbf{A}}$ gives the relations between different segments in terms of shared aspects or cluster membership, independent of size of the segments. $\mathbf{C}$ gives an adjustment in the overall size of loadings for each cluster or aspect, so that their loadings have the proper sum (or sum of squares) after the various cluster members have been shrunk or expanded according to their scale factors. (One might say that $\mathbf{C}^{-1}$ represents the effect of segment scale factors on the relative "impact" of the clusters.) Finally, if the earlier-selected scaling convention is adopted which makes the columns of $\dot{\mathbf{A}}$ sum to 1.0, the elements of $\dot{\mathbf{R}}$ represent the number of switches among the underlying clusters or aspects, i.e., the switches among the clusters which are presented in "pure" form by $\mathring{\mathbf{A}}$, and in a form with scale differences incorporated by $\mathbf{D}\mathring{\mathbf{A}}\mathbf{C}$ or $\dot{\mathbf{A}}$.

Application of the After-Analysis Approach

The iterative method for estimation of scale differences was applied to the car switching data of Table 4; the segment scale factors d_i and the rescaling coefficients $(1/d_i)$ are given in Table 6. These are the diagonal values of $\mathbf{D}$ and $\mathbf{D}^{-1}$, and if each entry in the original data matrix is multiplied by its corresponding row and column rescaling value (i.e., $\mathring{x}_{ij} = x_{ij}1/d_i1/d_j$), then the rescaled data would show the same total number of switches (row sum

TABLE 6

Segment Scale Factors (The Diagonals of **D**) *and Rescaling Coefficients (The Diagonals of* $\mathbf{D}^{-1}$) *for the Car Switching Data*

Segment	Scale Factors	Rescaling Coefficients
1	1.504	0.665
2	0.213	4.686
3	1.393	0.718
4	1.599	0.626
5	0.005	182.030
6	0.377	2.650
7	1.119	0.894
8	0.635	1.575
9	0.258	3.875
10	2.608	0.383
11	0.277	3.613
12	1.998	0.500
13	1.386	0.722
14	1.526	0.655
15	0.699	1.431
16	0.106	9.440

plus column sum) for each segment. Because our analysis of Table 4 will ignore the diagonal values (we suspect that special considerations of "loyalty" make these non-comparable to the rest of the data), the rescaling of the data is based on sums of off-diagonal cells only.

A rescaled $\mathring{\mathbf{A}}$ matrix was produced by multiplying each row of the original **A** matrix by the corresponding rescaling coefficient, then orthonormalizing and re-rotating the resulting matrix to VARIMAX simple structure in order to adjust for any deviations from the orthonormal simple structure produced by the row adjustments. The resulting $\mathring{\mathbf{A}}$ matrix is shown in Table 7. Although the same basic four dimensions are apparent, their interpretation is much clearer. Note, for example, that Small Specialty/Domestic now loads on the SPECIALTY dimension (0.225). The segment Small Specialty/Captive Import, which loaded on nothing in the non-rescaled **A** matrix, now is found to load 0.143 primarily on the SMALL dimension (although there is a moderate loading of 0.101 on SPECIALTY). At the same time, some segments which previously had inflated loadings because of their very large overall scale, or market share, are placed in better perspective in this rescaled solution. The Midsize Domestic, which previously dominated the PLAIN LARGE-MIDSIZE cluster or dimension, now is reduced to being the second largest segment, with a loading of 0.312. The largest loading for this dimension/cluster is now 0.398 for Low-Priced Standard, which seems appropriate given the overall interpretation of this dimension.

Interpretation of DEDICOM *Dimensions as Clusters*

It is useful to consider more carefully how the columns of $\mathring{\mathbf{A}}$ might be thought of as identifying clusters of related things, much in the manner of

TABLE 7

$\overset{*}{\mathbf{A}}$ (*Rescaled* **A**) *and Its Associated* $\dot{\mathbf{R}}$ *Matrix from the Four-Dimensional Solution for the Car Switching Data*

	Dimension			
	1	2	3	4
$\overset{*}{\mathbf{A}}$ MATRIX	(PLAIN LARGE-MIDSIZE)	(SPECIALTY)	(FANCY LARGE)	(SMALL)
1 SUBD	−.074	−.004	.035	.172
2 SUBC	.053	−.079	−.009	.109
3 SUBI	−.092	.023	.041	.115
4 SMAD	−.013	.225	−.040	.081
5 SMAC	−.067	.101	−.039	.143
6 SMAI	.003	.199	−.055	.057
7 COML	.156	−.108	−.049	.131
8 COMM	.175	−.069	.017	.075
9 COMI	.046	−.176	.050	.098
10 MIDD	.312	.082	−.028	.000
11 MIDI	.111	.005	−.005	.025
12 MIDS	.028	.669	.011	−.004
13 STDL	.398	.050	.047	−.020
14 STDM	.103	−.031	.471	−.014
15 LUXD	−.120	.067	.471	.022
16 LUXI	−.017	.047	.136	.011
$\dot{\mathbf{R}}$ MATRIX				
1	150676	133409	87244	220712
2	9487	105856	38157	42550
3	35892	45205	76584	57338
4	84511	95245	18064	240341

cluster analysis. Although DEDICOM decomposes relationships into latent aspects and hence has its strongest kinship with factor analysis and MDS, it frequently happens that the aspects identified by an analysis can be taken to define a rough partition of the original stimuli or objects (in this case, the original segments of the automobile market) into a few clusters. By choosing a simple structure rotation with orthogonal columns of **A**, we encourage the model to provide a solution in this form (see Harris and Kaiser, 1964). When a "rotation" of the solution can be found that shows each stimulus or object as loading strongly on only one aspect, then we can interpret the DEDICOM solution as providing a "fuzzy" disjoint clustering of the stimuli or objects whose relations are being analyzed. The size of the loadings gives the strength of cluster membership. And even when there are stimuli which load substantially on several columns of $\overset{*}{\mathbf{A}}$, this might be interpreted as an additive-cluster representation, where objects can participate in several clusters simultaneously. Under the cluster interpretation, the $\dot{\mathbf{R}}$ matrix gives the asymmetric relationships among the several fuzzy clusters identified by the analysis.

Pseudo-Hierarchical Clustering of the Segments

By examining the way in which DEDICOM partitions the segments into clusters at each different dimensionality, from one to four, we can establish a loose analog to a hierarchical clustering of the segments. Table 8 gives the one, two, and three-dimensional $\mathbf{\dot{A}}$ and $\mathbf{\dot{R}}$ matrices for the car switching data. In the one-dimensional solution, all the car makes load to roughly the same degree (perhaps larger cars having slightly larger loadings), except that the import segments (rows 3, 6, 9, 11, and 16) all seem to be smaller than their neighboring loadings. (This suggests that there may be some special property of switching involving imports which is not being captured in these analyses.)

The two-dimensional rescaled solution in Table 8 separates the car segments into two groups, which can loosely be characterized by the "winners" and "losers" in the competition for market share. The winners include all of the small cars (the first six rows of $\mathbf{\dot{A}}$) with the interesting exception of Subcompact Captive Import. The winners also include a large loading for Midsize Specialty and Luxury Domestic. In essence, the winners include the three groups which had advantageous switching patterns in the four-dimensional

TABLE 8

$\mathbf{\dot{A}}$ *and* $\mathbf{\dot{R}}$ *Matrices from One- Through Three-Dimensional Solutions for the Car Switching Data*

	1-Dimensional Solution	2-Dimensional Solution		3-Dimensional Solution		
$\mathbf{\dot{A}}$	Dimension	Dimensions		Dimensions		
Matrix	1	1	2	1	2	3
1 SUBD	.071	.126	−.003	.157	−.015	.006
2 SUBC	.054	.040	.073	.089	.071	−.089
3 SUBI	.046	.101	−.027	.113	−.043	.034
4 SMAD	.060	.106	−.002	.116	−.022	.063
5 SMAC	.063	.125	−.020	.160	−.040	−.014
6 SMAI	.046	.080	.001	.092	−.014	.033
7 COML	.076	.038	.125	.103	.139	−.134
8 COMM	.071	.020	.138	.053	.151	−.052
9 COMI	.044	.017	.080	.056	.085	−.079
10 MIDD	.077	−.008	.192	.006	.213	−.002
11 MIDI	.038	.006	.081	.020	.088	−.017
12 MIDS	.082	.169	−.033	.105	−.052	.299
13 STDL	.096	−.027	.260	−.031	.292	.049
14 STDM	.081	.056	.113	−.033	.139	.325
15 LUXD	.063	.107	.005	−.007	−.003	.414
16 LUXI	.032	.045	.014	.000	.011	.166
$\mathbf{\dot{R}}$ Matrix						
1	1348106	395745	164137	316678	109668	51604
2		493066	300131	324085	219939	156065
3				42788	44667	123642

solution of Table 7. The loser cluster picks out the Compact Non-Imports, the Midsize Non-Imports, and the Low-Priced Standard. The Medium-Priced Standard seems to straddle the two clusters, which is plausible since it has some of the glitter of the more luxurious cars.

A glance at the $\dot{\mathbf{R}}$ matrix quickly reveals why the second dimension was labeled "losers." There is a very large asymmetry in the estimated number of switches, with the number of switches *to* the second cluster being roughly one-third of the estimated switches *from* this cluster.

The three-dimensional solution splits the winners cluster into two meaningful subsets: SMALL and FANCY LARGE-MIDSIZE. However, the Midsize Specialty loads on both dimensions. Presumably, it has some of the appeal of small cars, and yet also some of the appeal of luxury cars. The $\dot{\mathbf{R}}$ matrix for this solution reveals that both the SMALL and FANCY LARGE-MIDSIZE clusters gain at the expense of PLAIN LARGE-MIDSIZE. There is relatively little asymmetry in switching between these two winners.

It is possible to take the DEDICOM solutions, interpreted as representing fuzzy additive clusters, and plot the evolution of progressively finer distinctions which emerge at successive dimensionalities, in order to obtain a pseudo-hierarchical clustering of the data. Cluster distinctions which occur at lower dimensionalities are more important in the sense of accounting for a higher proportion of the variance than the distinctions which occur as we progress to higher and higher-dimensional solutions. For example, the basic division into winners and losers improved the R^2 from 0.72 to 0.92. The next division, which splits up the SMALL from the FANCY LARGE, raises the fit from 0.92 to 0.96, and the final isolation of the SPECIALTY cluster raises the fit to only 0.98. Thus it might be justified to draw a hierarchical tree (or an embedded clustering) such as that shown in Figure 3.

Analysis of the Car Switching Data Via Matrix Rescaling Only

As briefly mentioned earlier, there is an alternative approach to the problem of unequal scale of the observations for different rows and columns of our data. We can simply *adjust the data themselves* and then analyze the matrix in the usual manner. Suppose we have determined the diagonal rescaling matrix **D** which allows **DXD** to have equal sums of observations (i.e., equal row-plus-columns sums, ignoring the diagonal) for all the segments. We let:

$$\mathring{\mathbf{X}} = \mathbf{DXD} \tag{20}$$

and then decompose these data according to the DEDICOM model:

$$\mathring{\mathbf{X}} = \mathring{\mathbf{A}}\mathring{\mathbf{R}}\mathring{\mathbf{A}}' + \mathring{\mathbf{E}}. \tag{21}$$

In other words, the second approach (data rescaling) simply uses the prelimi-

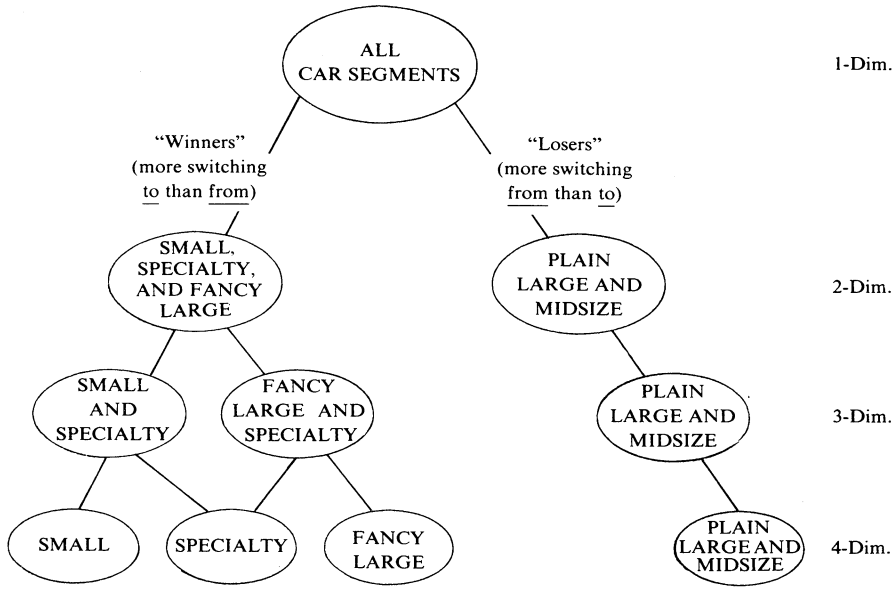

FIGURE 3. Implied Pseudo-Hierarchical Additive Clustering of Automobile Segments Obtained By Comparing DEDICOM Solutions At Successive Dimensionalities.

nary diagonal transformation to find $\mathring{\mathbf{X}}$—as already described in Equation (6)—and then directly analyzes this matrix in the usual manner.

The interesting thing about this approach (as described earlier) is that it reweights and equalizes the relative importance of the different rows and columns. For example, originally row 5 and column 5 of **X** contained very small entries. Consequently, the estimates for this row and column were also small, and (what is crucial for this argument) the errors were small, compared to the errors made in estimating large rows and columns such as those for segment 10. One consequence of this inequality of size of errors is that the least-squares fitting procedure pays little or no attention to row 5. Differences in the overall error of a given solution arising from row and column 5 are so small as to have virtually no effect on the choice of parameter values. Instead, large rows and columns such as those of segment 10 tend to dominate the solution. Thus, the parameter values are adjusted so as to fit these large numbers as closely as possible.

Sometimes, however, we are interested in the switching patterns of the small segments, even though they may not be as accurately estimated, and even though they may not have the same economic importance. In such a case, the direct analysis of the rescaled data might be appropriate.

To demonstrate this approach, rescaled car switching data were generated by multiplying each element of **X** (Table 4) by the corresponding row and column rescaling coefficients (Table 6). While not discussed in detail here, the resulting matrix was analyzed directly by DEDICOM in one through eight

dimensions. Unfortunately, two of the smallest segments—Small Specialty/Captive Import and Luxury Import—produced two extra dimensions on which each was the only segment loading substantially.

These two segments were dropped from the raw data matrix and the rescaled data were again analyzed—this time in one to six dimensions. At four dimensions the solution also revealed the presence of the PLAIN LARGE-MIDSIZE, FANCY LARGE, and SMALL dimensions, already noted in Tables 5 and 7, but the SPECIALTY dimension did not emerge. Instead, an IMPORT dimension appeared. Attempts to find a five-dimensional solution that contained all four original aspects *and* the IMPORT aspect were not successful.

Finally, symmetric solutions were also obtained, using *both* methods of scale adjustment (i.e. solving for $\overset{*}{\mathbf{A}}$ and $\dot{\mathbf{R}}$ versus direct analysis of the rescaled data $\mathring{\mathbf{X}}$). As was the case in the word associations data, the symmetric analyses were considerably more difficult to interpret and less informative than their asymmetric counterparts. As an illustration of the difficulty of interpretation of the symmetric solutions, Table 9 shows the one, two, and three-dimensional

TABLE 9

$\overset{*}{\mathbf{A}}$ *and* $\dot{\mathbf{R}}$ *Matrices from One- Through Three-Dimensional Solutions for a Symmetric Analysis of the Car Switching Data*

	1-Dimensional Solution	2-Dimensional Solution		3-Dimensional Solution		
$\overset{*}{\mathbf{A}}$	Dimension	Dimensions		Dimensions		
Matrix	1	1	2	1	2	3
1 SUBD	.069	.104	−.017	.142	−.025	.047
2 SUBC	.053	.091	−.047	.082	−.048	.092
3 SUBI	.045	.065	−.006	.060	−.007	.065
4 SMAD	.060	.075	.019	−.009	.011	.206
5 SMAC	.069	.110	−.045	.043	−.039	.184
6 SMAI	.047	.066	−.005	−.045	.015	.195
7 COML	.075	.124	−.048	.223	−.080	.003
8 COMM	.069	.087	.021	.159	.007	−.007
9 COMI	.044	.075	−.038	.061	−.035	.080
10 MIDD	.076	.068	.091	.118	.086	.004
11 MIDI	.040	.045	.023	.032	.029	.056
12 MIDS	.080	.051	.153	.048	.164	.058
13 STDL	.095	.070	.155	.174	.146	−.055
14 STDM	.080	.003	.304	.069	.296	−.061
15 LUXD	.065	−.028	.313	−.066	.332	.032
16 LUXI	.032	−.006	.129	−.089	.149	.100
$\dot{\mathbf{R}}$ Matrix						
1	1354037	741703	185997	492893	123761	123140
2		185997	274164	123761	232912	51651
3				123140	51651	88868

TABLE 10

$\overset{*}{\mathbf{A}}$ *and* $\dot{\mathbf{R}}$ *Matrices from the Four-Dimensional Solution for a Symmetric Analysis of the Car Switching Data*

	Dimension			
$\overset{*}{\mathbf{A}}$ MATRIX	1	2	3	4
1 SUBD	.162	−.002	−.005	.031
2 SUBC	.093	−.022	−.069	.107
3 SUBI	.069	.006	−.053	.091
4 SMAD	.048	−.061	.195	.082
5 SMAC	.099	−.104	.134	.093
6 SMAI	−.041	−.077	.132	.195
7 COML	.243	−.020	−.062	−.013
8 COMM	.161	.065	−.046	.002
9 COMI	.044	.031	−.219	.186
10 MIDD	.110	.068	.141	−.013
11 MIDI	.023	.009	.027	.082
12 MIDS	−.007	−.003	.624	−.001
13 STDL	.157	.156	.144	−.061
14 STDM	.025	.506	−.035	−.018
15 LUXD	−.085	.332	.093	.077
16 LUXI	−.098	.116	−.001	.161
$\dot{\mathbf{R}}$ Matrix				
1	389421	68118	113947	62874
2	68118	136738	72164	26696
3	113947	72164	138002	31608
4	62874	26696	31608	50098

results and Table 10 the four-dimensional solution for the $\overset{*}{\mathbf{A}}$ and $\dot{\mathbf{R}}$ approach. These can be compared with the earlier-described results of the asymmetric case, appearing in Table 7.

In contrast to Table 7's relatively clear interpretation, the symmetric solution appears to be less informative. In Table 9 the two-dimensional solution appears to distinguish the lower priced makes from expensive makes; the three-dimensional solution splits the lower priced cars into Small Specialty cars versus a rather heterogeneous mixture of Compacts, a Low-Priced Standard, and a Subcompact Domestic. The four-dimensional solution (Table 10) also appears difficult to understand. While dimensions two and three appear to be FANCY LARGE and SPECIALTY, respectively, dimension four now looks like an IMPORT dimension while dimension one contains a heterogeneous mixture of Subcompacts, Compacts, and even a Low-Priced Standard.

All in all, we find the $\overset{*}{\mathbf{A}}$ and $\dot{\mathbf{R}}$ solutions appearing in Tables 7 and 8 (and portrayed as a cluster structure in Figure 3) to be the most direct and appealing. Still, the availability of alternative approaches provides the researcher with complementary ways to analyze the same data set. The availability of complementary approaches provides the potential for obtaining richer

information about a data structure; whether that potential is realized or not, of course, depends on the specific data set.[2]

Discussion

It seems fair to conclude from the two preceding field-level applications that diverse sets of data are amenable to DEDICOM analysis, and that DEDICOM can sometimes provide significantly better fit values and more interpretable solutions than are obtained by symmetric models (that are analogous to factor analysis and multidimensional scaling).

It would seem that there are at least five distinct advantages to performing a DEDICOM analysis of data sets with real and potentially important asymmetries: (1) The data are better represented by the asymmetric DEDICOM model, in the sense that the fit is better, and thus more of the structure of the data is incorporated into the definition of the dimensions, resulting in dimensions that are usually more meaningful; (2) the **R** matrix which is obtained as part of the solution provides a source of information on patterns of asymmetric (and symmetric) flows among clusters, or aspects, of the data, and this

[2]We have been applying DEDICOM to transition matrices **X** and symmetrically rescaled transition matrices $\mathring{\mathbf{X}}$. Suppose, instead, we preferred to think in terms of matrices of conditional probabilities, where each row summed to 1.0 and each entry gave the probability of switching to car make j given that one had previously owned car make i. How should we proceed with the analysis? Unfortunately, it seems that if the original data matrix **X** were appropriate for the single domain DEDICOM model, then the matrix of conditional probabilities obtained from **X** would not be appropriate for that model. The problem arises because in order to obtain the conditional probability matrix, we apply a *nonsymmetrical* rescaling to the data. Each row is multiplied by a constant so that the row sum is equal to 1.0. If we let **W** represent an n by n diagonal matrix of rescaling weights, with the diagonal entries $w_{ii} = (\sum_j x_{ij})^{-1}$, then our original data matrix with DEDICOM structure

$$\mathbf{X} = \mathbf{ARA}' + \mathbf{E}$$

is transformed into the matrix of conditional probabilities $\mathbf{P}_c$ as follows

$$\mathbf{P}_c = \mathbf{WX} + \mathbf{WARA}' + \mathbf{WE}.$$

The structure of $\mathbf{P}_c$ is not suitable for the single domain DEDICOM model, since the row dimensions have the form **WA** and the column dimensions have the form **A**.

While dual domain analysis could be carried out on this data, the presence of error in the data would prevent the obtained solution from exactly fulfilling the constraint that $\mathbf{WA}_{\text{left}} = \mathbf{A}_{\text{right}}$, even if the underlying structure of **X** did exactly fulfill the single domain model. It seems more straightforward, therefore, to simply analyse **X** and then (if desired) rescale the rows of **A** by **W** to obtain loadings appropriate for a "constrained" dual-domain solution.

It is interesting to note in passing, however, that if the data are globally rescaled, so that each cell represents an *unconditional* transition probability (i.e. if $\overset{*}{x}_{ij} = x_{ij}/\sum_i\sum_j x_{ij}$), then the single-domain analysis is appropriate and the parameters have a particularly nice interpretation in terms of conditional and unconditional probability components which multiply together to give the unconditional transition probabilities (see Harshman 1982a for details).

information can often have useful marketing implications; (3) the **A** and **R** matrix scaling conventions adopted for DEDICOM can provide particularly direct and concrete relationships between the data and the inferences that can be drawn from the solution (e.g., one can obtain "actual" switching frequencies between latent clusters, or "actual" associative strengths between latent aspects of phrases); (4) DEDICOM's method for rescaling the rows of **A**, and/or removing the scale-factor inequalities of overall scale from the rows and columns of **X** (by means of the same multiplier applied to rows as to columns) provides a novel solution to problems of scale inequalities that sometimes arise in marketing and other social science data; and (5) significance tests for the asymmetries provide a method for evaluating the "reality" of the asymmetrical structure in a set of data.

Limitations of the DEDICOM *Approach*

Now that the DEDICOM model has been described and illustrated in terms of two actual data sets, it seems useful to discuss some current limitations of the approach and planned methodological developments for the future. Following this, we list some types of marketing and behavioral research data that might be fruitfully analyzed by the technique.

The solutions presented in this article were obtained by a closed-form algorithm which provides only an approximation to a least-squares solution. A true least-squares solution can be obtained by an iterative method, but the potentially most efficient iterative approach (involving alternating least squares) has not yet been programmed. Thus, current solutions are only approximations to the parameter values that will be obtained when the alternating least squares method becomes available. Tests comparing the solutions obtained by the approximate least squares methods with current (inefficient but true) least squares methods indicate that interpretation of the solution would not be changed by the modest differences obtained.

As currently programmed, DEDICOM expects ratio-scale data. However, this does not usually seem to seriously impair the ability of the model to obtain good fits and meaningful solutions, as demonstrated by the examples above. Data such as brand switching or association frequency matrices possess a natural origin or zero, and so the problem of an additive constant is not a problem. However, many data sets are really only interval-scaled, at best. These kinds of data should be adjusted by estimation of an "additive constant" as in MDS, or by double centering; alternatively, the issue could be handled by a parameter within the model for estimating the additive constant required to convert the data to ratio-scale. Such an approach is under investigation.

Recent and Future Research for DEDICOM

Other extensions of the DEDICOM model have already been envisioned and, for some of these, the program algorithms have been specified mathematically. One example is the estimation of row and/or column bias terms. In some data, "uninteresting" asymmetries are introduced as a consequence of

such biases, and it would be useful to remove these effects as part of fitting the model to the data. Approaches to this are discussed in Harshman (1982a), along with other extensions of the model.

It has been shown (Harshman, 1981; Harshman and Lundy, 1982) that DEDICOM provides a new way to analyze dominance matrices of the type often associated with paired comparisons preference data. Typically, dominance matrices are analyzed by some model such as the Thurstonian Case V (Green and Tull, 1978), or the Bradley, Terry, Luce (Restle and Greeno, 1970) procedures.

DEDICOM contains a procedure for analyzing such skew-symmetric matrices. Data that are strictly scalable by a method such as Case V appear in the DEDICOM analyses as a straight line in two dimensions; more complex scales show up as curved lines or other configurations in the plane. Multiple scales show up as lines in sets of planes (one plane for each scale) while nonscalable points plot near the origin.

A version of DEDICOM has also been proposed for three-way matrices (e.g., car switching by successive model years). In this extension the model of Equation (1) becomes:

$$\mathbf{X}_i = \mathbf{A}\mathbf{D}_i\mathbf{R}\mathbf{D}_i\mathbf{A}' + \mathbf{E}_i \tag{22}$$

where each $\mathbf{D}_i$ matrix is a diagonal matrix giving the weights for the underlying aspects that are appropriate for the ith model year.[3]

Other Application Areas

In addition to the kinds of data illustrated here, DEDICOM is applicable to other classes of marketing and behavioral research data. Illustrative of these classes of data are:

- Attributions data
- Compatibility data
- Communications flows
- Second choice data
- Like/dislike data
- Confusions data
- Subjective judgments on causality
- Temporal (precedence) data.

No attempt will be made to illustrate these many possibilities; moreover, as additional experience with DEDICOM is gained, other kinds of data will surely be suggested.

One way to distinguish these kinds of analyses from the more traditional analysis of symmetric data is to say that the latter emphasizes relations that pertain to the *shared* attribute levels that two objects exhibit. For example, two

[3] A related approach to the problem of representing three-way asymmetric data has been described by Carroll (1977).

brands are "similar" to the extent that they exhibit common attribute levels. In contrast, in the DEDICOM analysis of asymmetric data, emphasis may be placed on the complementarity of the object pairs; for example, the degree of liking for some second object, given that one has already received his/her first choice. In this case, the two objects may exhibit relatively little in the way of common attribute levels; indeed, the second object may be highly liked precisely because it is different from (but complementary to) the first-choice object. It seems to us that relations of this type are both prevalent in marketing and important to model.[4]

Technical Appendix

Only the essentials of the single domain DEDICOM model were outlined in the body of the paper. Here, we describe a related model in the DEDICOM family as well as some features of the algorithm used to find numerical solutions.

The Dual Domain Model

The primary model of Equation (1) can be described as a *single domain* model in the sense that the solution technique requires the row space to be the same as the column space (i.e., where one is a linear transformation of the other). However, suppose we reconsider the situation involving the car switching data. It may be that certain characteristics of cars (e.g., repair history) are more salient when one is disposing of a car than when one is considering the switched-to car; in the latter case, body style and appearance may be more salient.

In any case, DEDICOM has an option that permits the user to fit a model in which the column space does not have to equal the row space. The *dual domain* model (Harshman, 1982a) can be written as:

$$\mathbf{X} = \check{\mathbf{A}}\tilde{\mathbf{R}}\hat{\mathbf{A}}' + \mathbf{E} \tag{23}$$

where $\check{\mathbf{A}}$ and $\hat{\mathbf{A}}$ are $n \times q$; $\tilde{\mathbf{R}}$ is $q \times q$; and $\mathbf{E}$, as usual, is $n \times n$.

In contrast to Equation (1), the present model provides two sets of latent entities. $\check{\mathbf{A}}$ gives the weights of the observed objects in their row positions, while $\hat{\mathbf{A}}$ gives the counterpart weights for the column positions. The $\tilde{\mathbf{R}}$ matrix gives the directed relationships from the $\check{\mathbf{A}}$ to $\hat{\mathbf{A}}$ aspects.

Assuming that $\check{\mathbf{A}}$ and $\hat{\mathbf{A}}$ are not linear transformations of each other, there generally is no q-dimensional solution of the form given by Equation (1) for data represented by Equation (23). In this sense, the assumptions underlying

[4]Finally, an important area for future research involves the possibility of extending the DEDICOM methodology to "comfirmatory" analyses, in the spirit of recent developments in structural modeling (Bagozzi, 1980; Jöreskog, 1978).

the model of Equation (23) are weaker than those of Equation (1). Hence, if data are fitted about as well with Equation (1) as with Equation (23), we have greater confidence in the existence of a *common process* underlying the row and column interaction patterns. (Randomization tests for comparing the two models are under development; Harshman, 1982a.)

Model Solutions

The more general ("weak") model of Equation (23) is solved as follows:

1. Find the singular value decomposition of $\mathbf{X}$

$$\mathbf{X} = \mathbf{PDQ}' \tag{24}$$

where $\mathbf{D}$ is diagonal, and $\mathbf{P}$ and $\mathbf{Q}$ are orthonormal sections.

2. For a q-dimensional solution, take the first q singular vectors and rotate them to the desired pattern (e.g., a simple structure approximation) via $\check{\mathbf{T}}$ and $\hat{\mathbf{T}}$:

$$\check{\mathbf{A}} = \check{\mathbf{P}}_q\check{\mathbf{T}}; \qquad \hat{\mathbf{A}} = \mathbf{Q}_q\hat{\mathbf{T}}. \tag{25}$$

3. Estimate $\tilde{\mathbf{R}}$ by least squares, given $\check{\mathbf{A}}$ and $\hat{\mathbf{A}}$:

$$\tilde{\mathbf{R}} = \check{\mathbf{A}}^+\mathbf{X}(\hat{\mathbf{A}}^+)' \tag{26}$$

where

$$\check{\mathbf{A}}^+ = (\check{\mathbf{A}}'\check{\mathbf{A}})^{-1}\check{\mathbf{A}}'; \qquad \hat{\mathbf{A}}^+ = (\hat{\mathbf{A}}'\hat{\mathbf{A}})^{-1}\hat{\mathbf{A}}'. \tag{27}$$

That is, $\check{\mathbf{A}}^+$ and $\hat{\mathbf{A}}^+$ are pseudoinverses of $\check{A}$ and $\hat{A}$, respectively. This provides a true least-squares solution.

One method of obtaining an approximate least-squares solution for the "strong" model of Equation (1) is as follows:

1. Find the singular value decomposition of $\mathbf{X}$ as in Equation (24).
2. Define the symmetrizing rotation $\mathbf{L}$

$$\mathbf{L} = \mathbf{QP}'. \tag{28}$$

3. Find $\mathbf{V}$, the common-space transformation of $\mathbf{X}$:

$$\mathbf{V} = \mathbf{XL} + \mathbf{LX} = \mathbf{PDP}' + \mathbf{QDQ}'. \tag{29}$$

4. Find the singular value decomposition of **V**:

$$\mathbf{V} = \overset{*}{\mathbf{P}}\overset{*}{\mathbf{D}}\overset{*}{\mathbf{Q}}' \tag{30}$$

and take the first q singular vectors.

5. Rotate the first q singular vectors to the desired pattern (e.g., a simple structure approximation) to obtain **A**:

$$\mathbf{A} = \overset{*}{\mathbf{P}}_q\mathbf{T}. \tag{31}$$

6. Estimate **R** by regression methods, similar to those already shown in Equations (26) and (27).

Iterative solutions also exist for DEDICOM, and research is underway in the development of efficient exact least squares solutions employing alternating least squares.

User Options

The DEDICOM computer program contains a number of features for facilitating different kinds of analyses. In addition to fitting either the strong or weak forms of the model, the user can specify a number of other options:

1. The original main diagonal entries in **X** can be considered as fixed or as starting values for successive estimation via an iterative procedure.
2. Various preliminary row and column normalizations of **X** can be carried out prior to model fitting (as illustrated in the main text).
3. Various rotation matrices **T**, including orthogonal, oblique, or fixed target, can be applied to the **A** matrix.
4. Alternative scalings of **A** and **R** can be selected.
5. The input matrix **X** can be decomposed into its symmetric and skew symmetric parts and each matrix analyzed separately.

The various options described above can be called upon in a flexible manner so that several different analyses can be conducted during a single run of the program.

A number of programming refinements are currently being added to DEDICOM. An exportable version of DEDICOM should be available for leasing by commercial or academic researchers in late 1982 or early 1983. Details can be obtained from Scientific Software Associates, 48 Wilson Avenue, London, Ontario, Canada, N6H 1X3.

References

Bagozzi, Richard P. (1980), "Performance and Satisfaction in an Industrial Sales Force: An Examination of Their Antecedents and Simultaneity," *Journal of Marketing*, 44 (Spring), 65–77.

Carroll, J. Douglas (1976), "Spatial, Non-Spatial, and Hybrid Models for Scaling," *Psychometrika*, 41, 439–63.

——— (1977), "Impact Scaling: Theory, Mathematical Model, and Estimation Procedures," *Proceedings of the Human Factors Society*, 21, 513–17.

——— and Phipps Arabie (1980), "Multidimensional Scaling," in *Annual Review of Psychology*, M. R. Rosenzweig and L. W. Porter (eds.), Palo Alto, California: Annual Reviews.

Chino, Naohito (1978), "A Graphical Technique in Representing the Asymmetric Relationships Between *n* Objects," *Behaviormetrika*, 5, 23–40.

Constantine, A. G. and John C. Gower (1978), "Graphical Representation of Asymmetric Matrices," *Applied Statistics*, 27, 294–304.

Gower, John C. (1977), "The Analysis of Asymmetry and Orthogonality," in *Recent Developments in Statistics*, J. R. Barra et al. (eds.), Netherlands: North Holland Publishing Co.

Green, Paul E. (1975), "Marketing Applications of MDS: Assessment and Outlook," *Journal of Marketing*, 39 (January), 24–31.

——— and Donald S. Tull (1978), *Research for Marketing Decisions*, Fourth Edition, Englewood Cliffs, N.J.: Prentice Hall, Inc., Chapter 6.

———, Yoram Wind, and Arun K. Jain (1973), "Analyzing Free-Response Data in Marketing Research," *Journal of Marketing Research*, 10 (February), 45–52.

Harris, Chester W. and Henry F. Kaiser (1964), "Oblique Factor Analytic Solutions by Orthogonal Transformations," *Psychometrika*, 29 (December), 347–62.

Harshman, Richard A. (1978), "Models for Analysis of Asymmetrical Relationships Among *N* Objects or Stimuli," Paper presented at the First Joint Meeting of the Psychometric Society and the Society for Mathematical Psychology, McMaster University, Hamilton, Ontario, August.

——— (1981), "DEDICOM Multidimensional Analysis of Skew Symmetric Data. Part 1: Theory," Unpublished technical memorandum, Bell Laboratories, Murray Hill, N.J.

——— (1982a), "DEDICOM: A Family of Models Generalizing Factor Analysis and Multidimensional Scaling for Decomposition of Asymmetric Relationships," Unpublished manuscript, University of Western Ontario.

——— (1982b), "Scaling and Rotation of DEDICOM Solutions," Unpublished manuscript, University of Western Ontario.

——— and Margaret E. Lundy (1982), "DEDICOM Multidimensional Analysis of Skew Symmetric Data. Part 2: Applications," Unpublished manuscript, University of Western Ontario.

Jöreskog, K. G. and D. Sörbum (1978), "LISREL: Analysis of Linear Structural Relationships by the Method of Maximum Likelihood," Version IV, Release 2, Chicago: National Educational Resources, Inc.

Levin, Joseph and Morton Brown (1979), "Scaling a Conditional Proximity Matrix to Symmetry," *Psychometrika*, 44 (June), 239–44.

Restle, Frank and James G. Greeno (1970), *Introduction to Mathematical Psychology*, Reading, Mass.: Addison-Wesley Publishing Co., Chapter 6.

Shepard, Roger N. (1962), "Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function," *Psychometrika*, I, 27, 125–40: II, 27, 219–46.

——— (1974), "Representation of Structure in Data: Problems and Prospects," *Psychometrika*, 39, 373–421.

Tobler, Waldo (1979), "Estimation of Attractivities from Interactions," *Environment and Planning*, Series A, 11, 121–27.

Young, Forrest W. (1974), "An Asymmetric Euclidean Model for Multiprocess Asymmetric Data," in *Theory, Methods, and Applications of Multidimensional Scaling and Related Techniques*, Proceedings of the U.S.-Japan Joint Seminar, University of California at San Diego, August 20–24.