MULTIDIMENSIONAL SCALING:

Euclidean distance models for exploring the complex structure of subjective perception, worldviews, and shared meaning

Lawrence Alfred Powell (PhD, MIT)
Honorary Research Fellow, COMPASS

la.powell@auckland.ac.nz *or*
lapowell.auckland@ymail.com

MDS offers a powerful 'pattern recognition' tool for the exploration and visualisation of structured patterns within complex numeric and textual observations, particularly those relating to human cognition, perception and contextualised 'meaning'.

This seminar briefly outlines the statistical basis, advantages/disadvantages, and available computer programmes for conducting MDS and perceptual mapping techniques. Examples are given of useful applications across the social sciences...

----

*I am grateful to the following scholars for their direct and indirect contributions to this presentation and their continued inspiration: Anthony Coxon, Alan Brier, Forrest Young, Angelina Anastasova, Natalia Jaworska, Joseph Kruskal, Louis Guttman, Warren Torgerson, Roger Shepard.

MDS: What Is It?

- Generally regarded as **exploratory data analysis,** but can also be **confirmatory** (i.e. test hypotheses re structure of cognition.).
- **Data/dimension reduction -** reduces large amounts of multivariate data into easier-to-visualise structures.
- Attempts to **find structure** (visual representation) in a set of **distance measures (**proximities - dis/similarities, between objects/cases.)
- **Globally/contextually maps** how objects/variables are inter-related perceptually, by assigning the objects to locations in a dimensional space.

MDS iteratively adjusts distances between points in the Euclidean space (the model) to match the matrix of dis/similarities (the data) as closely as possible. (Close points indicate similar objects; Far-apart points indicate dissimilar objects)

Origins & Development of MDS

– **Has origins in psychometrics advances of the 1920-'60s:**

--Scale construction, and dimensionality reduction

--Underwent a major burst of development in 1960s due to the "non-metric revolution"(Coombs), and emerging computing developments allowing for iterative estimation

– **Originally designed for analysis of similarities data**, taking a range of measures: "*anything which, by an act of faith, can be considered a similarity" (Shepard)*

--Extended rapidly to deal with a wide range of other types of data: Rectangular matrices, triads, pair-comparisons, free-sorting "stacks" of matrices (3-way scaling, INDSCAL)

--Originally referred to (by Guttman, Kruskal et al.) as **"smallest space analysis"**

A simple example: Constructing a map of U.S. cities . . .

--Ordinarily, you would start with the map, then measure the relative distances. MDS operates the other way round... Suppose you only had the distances between the cities, but didn't know what the map looked like . . .

-- **Given the data** [“distances”] **MDS attempts to find the original configuration** [location of points] **which generated the distances**

--This is “classic MDS”: developed in 1930s – but imperfect, not very robust, and works only if the data are ratio.

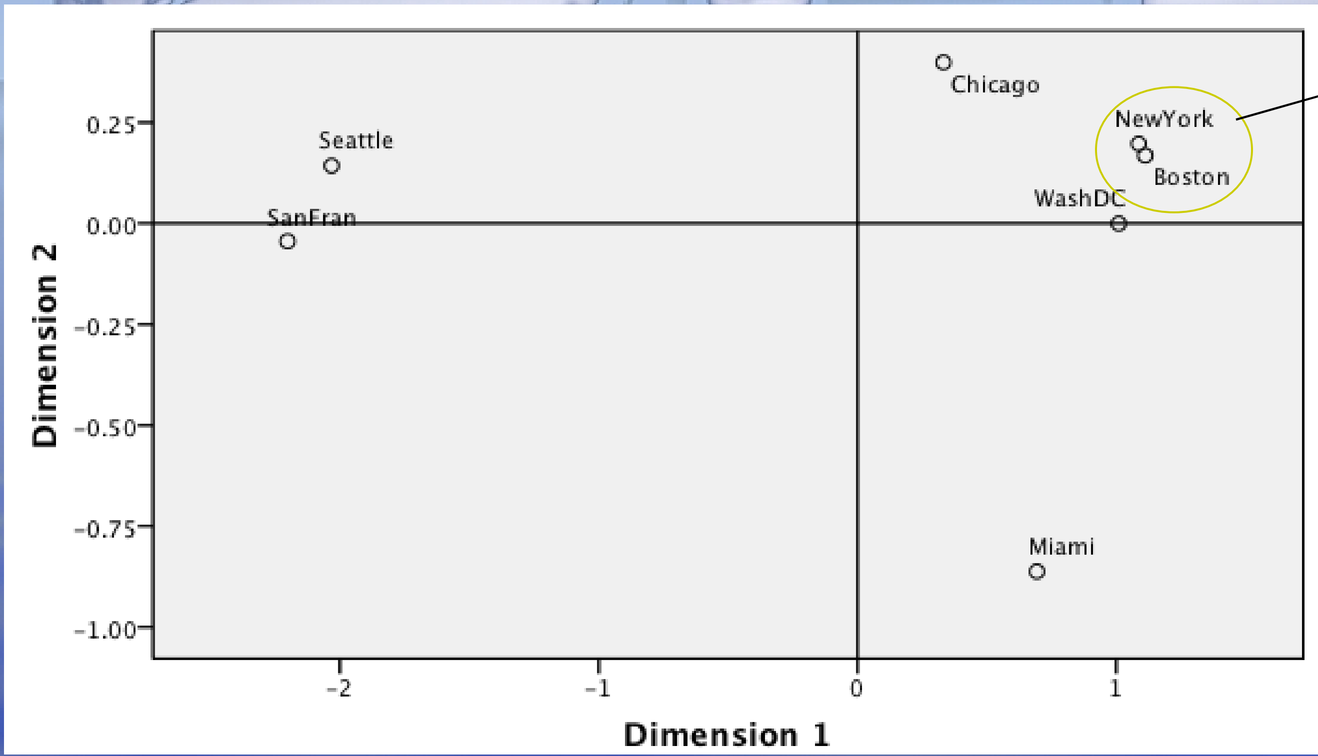
--Whereas more recent MDS can work when **just ordinal information** exists: relative rankings, ordinal, non-metric

What?? You can create an accurate map from knowing only the rank–order of the distances??? Yes, and it works . . .

MDS Example: City Distances

Distances Matrix: Symmetric

		1 BOST	2 NY	3 DC	4 MIAM	5 CHIC	6 SEAT	7 SF	8 LA	9 DENV
1	BOSTON	0	206	429	1504	963	2976	3095	2979	1949
2	NY	206	0	233	1308	802	2815	2934	2786	1771
3	DC	429	233	0	1075	671	2684	2799	2631	1616
4	MIAMI	1504	1308	1075	0	1329	3273	3053	2687	2037
5	CHICAGO	963	802	671	1329	0	2013	2142	2054	996
6	SEATTLE	2976	2815	2684	3273	2013	0	808	1131	1307
7	SF	3095	2934	2799	3053	2142	808	0	379	1235
8	LA	2979	2786	2631	2687	2054	1131	379	0	1059
9	DENVER	1949	1771	1616	2037	996	1307	1235	1059	0



Cluster

Spatial Map

Dimensions

1: North/South

2: East/West

Input data of MDS: a matrix of 'proximities' **similarities**, **dissimilarities**, **distances** (reflects amount of dis/similarity or distance between pairs of objects).

- Distinction between similarity and dissimilarity data dependent on type of scale used:
  - ⇒**Dissimilarity** scale: Low # = high similarity & High # = low dissimilarity.
  - ⇒**Similarity** scale: Opposite of dissimilarity.

  E.g. “On a scale of 1-9 (1 being the same and 9 completely different) how similar are political candidates A and B?”

Data Collection for MDS

Direct/raw data: Proximities' values are directly obtained from empirical, subjective scaling. E.g. pairwise comparison, grouping/sorting tasks, objective distance (e.g. city distances), direct ratings or rankings of dis/similarities of perceived stimuli/products/candidates.

Indirect/derived/inverted data: Computed from other measurements, Likert scales, semantic differential scales, or (inverted, transposed) correlations (any correlation matrix can be used with Gower conversion to Euclidean distances)

Types of MDS Models

MDS model classified according to . . .

1) . . . type of proximities:

– **Metric/quantitative**: Quantitative information, interval data about objects' proximities, e.g. city distance.

– **Non-metric/qualitative**: Qualitative information, nominal or ordinal data about proximities e.g. relative preference rankings of National, Labour, Greens, ACT

2) . . . number of proximity matrices (distance, dis/similarity)

– **Classical** MDS: One proximity matrix (metric, or non-metric).

– **Replicated** MDS: Several matrices.

– **Weighted MDS/Individual Difference Scaling**: Combines individual subject matrices (e.g. ratings of candidate attributes), to yield a common/averaged 'group space' as well as weighted individual subject spaces. (e.g. as implemented in INDSCAL, or ALSCAL within SPSS

– **Coombsian Unfolding**: Processes a joint matrix of objects x attributes.

Underlying Mathematical Model

- Classical MDS uses Euclidean principles to model data proximities in geometrical space, where distance (d*ij*) between points *i* and *j* is defined as:

$$d_{ij} = \sqrt{\sum (x_{ia} - x_{ja})^2}$$

x_i and x*j* specify coordinates of points i and j on dimension a, respectively.

- The modeled Euclidean distances are related to the observed proximities, δ*ij*, by some transformation/function (*f*).

$$= \sqrt{\sum (x_{ia} - x_{ja})^2}$$

- Most MDS models assume that the data have the form: δ*ij* = *f*(d*ij*)

All MDS algorithms are a variation of the above.

Output of MDS

Spatial Representation/Perceptual Map:

1) **Clusters:** Groupings in a MDS spatial representation. These may represent a domain/subdomain.

2) **Dimensions**: Hidden structures in data. Ordered groupings that explain similarity between items.

- Axes are meaningless, and orientation is arbitrary. (unlike, e.g. in factor analysis, PCA etc.)
- In theory, there is no limit to the number of derived dimensions.
- In reality, the number of dimensions that can be interpreted is limited (by human cognition)

Diagnostics of MDS

- MDS attempts to find a spatial configuration **X** such that the following is true: $f(\delta_{ij}) \approx d_{ij}(\mathbf{X})$
- **Stress (Kruskal's) function**: Measures degree of correspondence between distances among points on the MDS map and the matrix input.

  ⇒Proportion of variance of disparities *not* accounted for by the model:

$$\sqrt{\frac{\sum\sum (f(x_{ij}) - d_{ij})^2}{\sum\sum d_{ij}^2}}$$

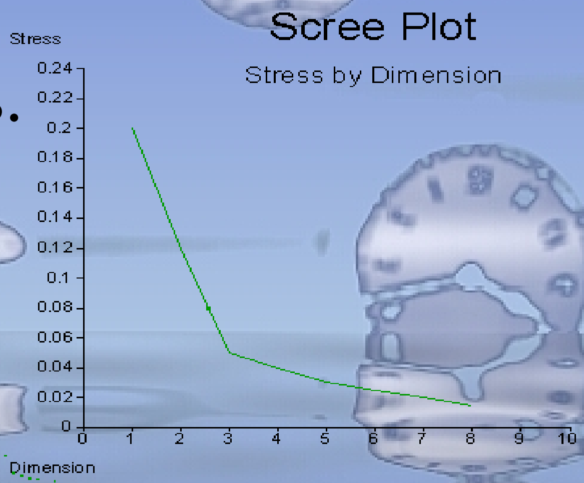
- Range 0-1: ***Smaller stress = better representation.***
- None-zero stress: Indicates some/all distances in the map are distortions of the input data.

Diagnostics of MDS (cont.)

- **R^2 (RSQ)**: Proportion of variance of the disparities accounted for by the MDS procedure.
  - $R^2 \geq 0.6$ is typically an acceptable fit.
- **Weirdness Index**: Used in Weighted Individual Differences Scaling (INDSCAL, ALSCAL) Indicates correspondence of subject's map and the aggregate map outlier identification.
  - Range 0-1: 0 indicates that subject's weights are proportional to the average subject's weights; as the subject's score becomes more extreme, index approaches 1.
- **Shepard Diagram of 'disparities'**: Scatterplot of input proximities (X-axis) against output distances (Y-axis) for every pair of items. (If plotted point distances fall on the step-line this indicates that input proximities are perfectly reproduced by the MDS model (the dimensional solution).

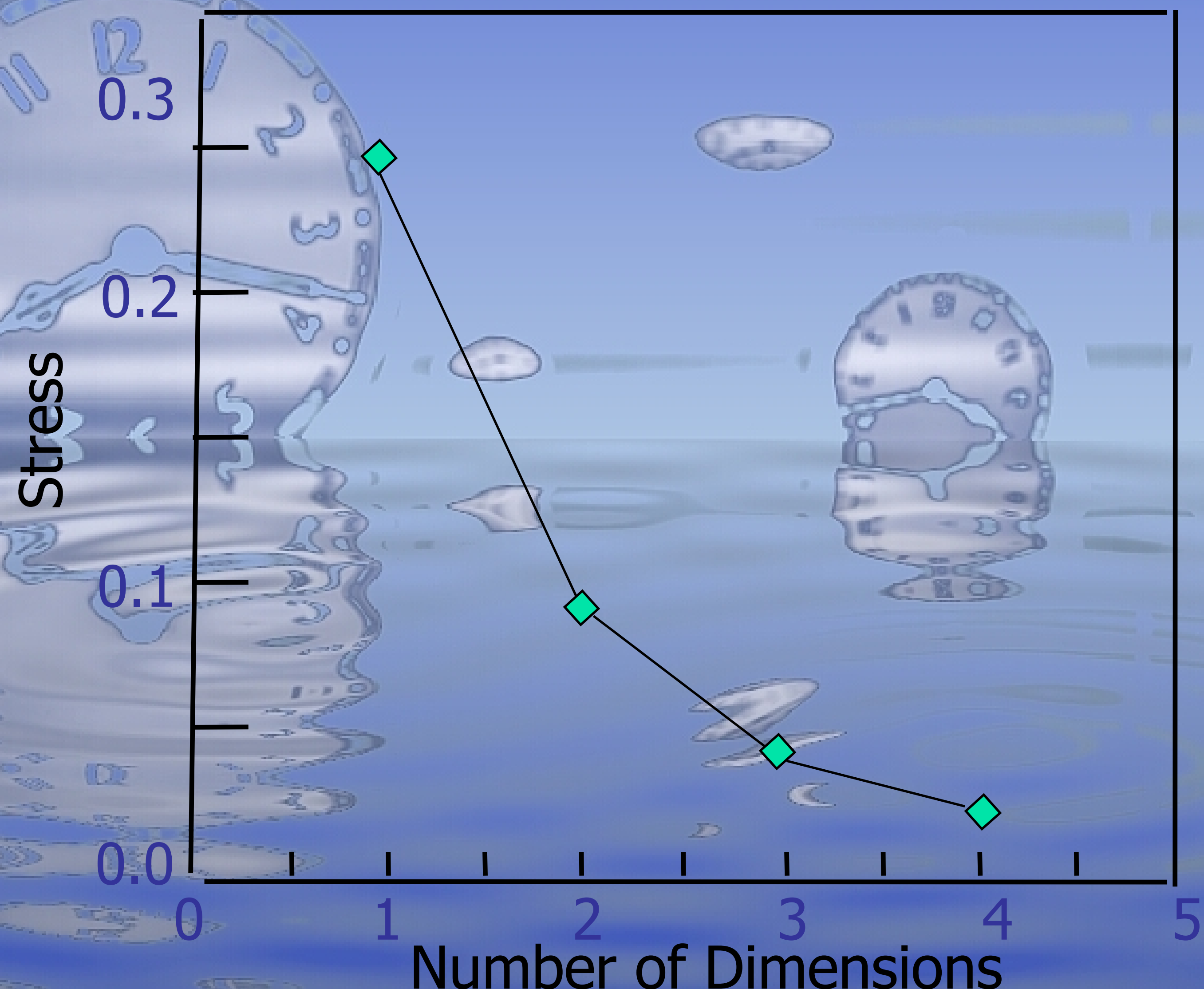
Interpretation of Dimensions

- Squeezing data into 2-D enables "readability" but may yield poor, distorted representation of the data (high stress); 3-D usually better.
- **Scree plot**: Stress vs. number of dimensions. (Similar function to scree plot in factor analysis.)



- Primary objective in dimension interpretation**: Obtain best fit with the smallest number of possible dimensions.**

Example: Stress reduction by # of dimensions



Meaning of Dimensions

- Label the dimensions by visual inspection, subjective interpretation, information & contextual clues from respondents.
- Externally validate dimensions by correlating with other related variables.

MDS Caveats

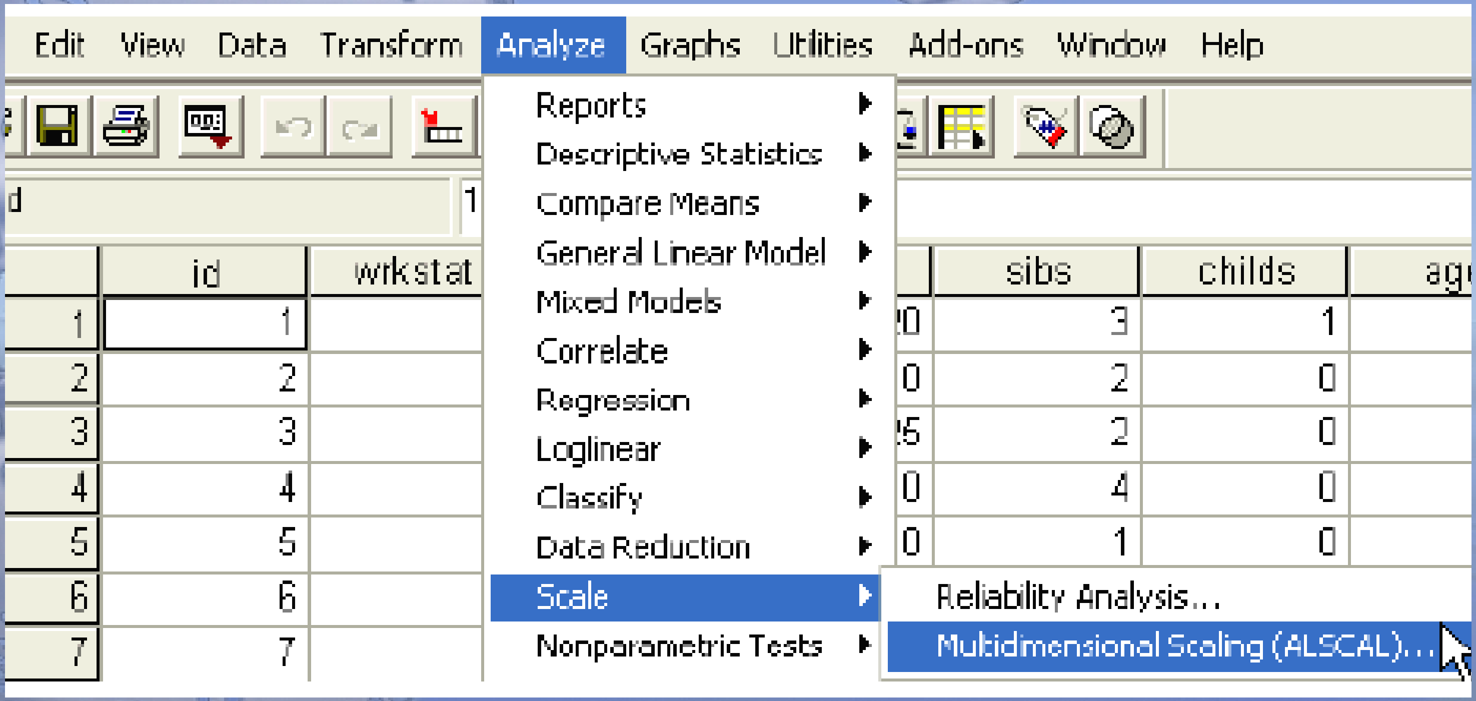
- Respondents may perceive stimuli differently. (i.e. you are comparing non-comparable responses)
- Respondents may attach different levels of importance to a dimension. (applies especially to INDSCAL, ALSCAL)
- Importance of a dimension may change over time.
- Interpretation of meaning of dimensions is subjective.
- Generally, at least four times as many objects as dimensions should be compared for the MDS model to be stable and avoid degenerate solutions.

Advantages of MDS

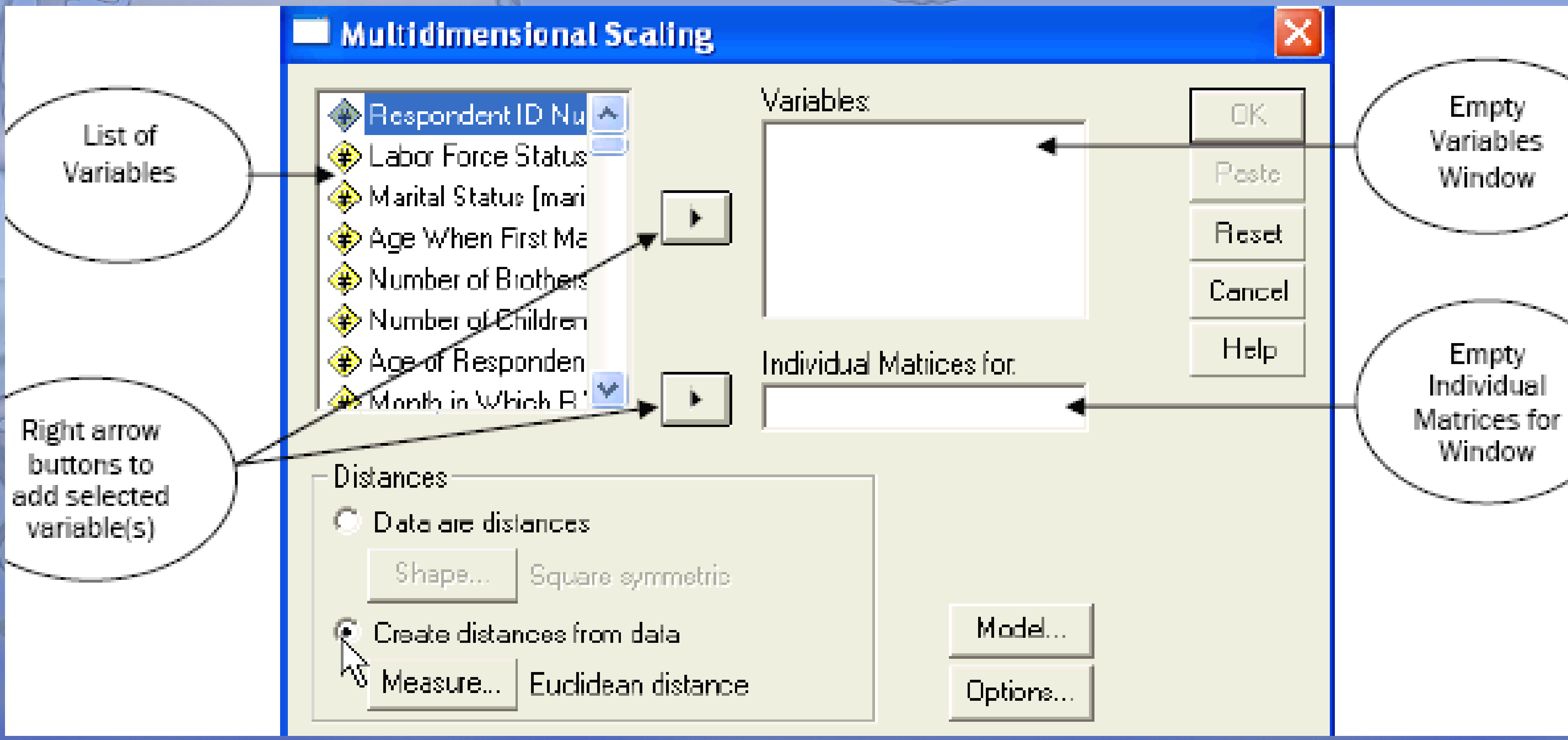
- An alternative to the GLM.
- Does not require assumptions of linearity, metricity, or multivariate normality.
- Can be used to model nonlinear relationships.
- Dimensionality "solution" can be obtained from individuals; gives insight into how individuals differ from aggregate data.
- Reveals dimensions without the need for pre-defined attributes. (i.e. Empirically-derived not ad hoc)
- Dimensions that emerge from MDS can be incorporated into regression analysis etc. to assess their relationship with other variables.

How to do MDS with SPSS

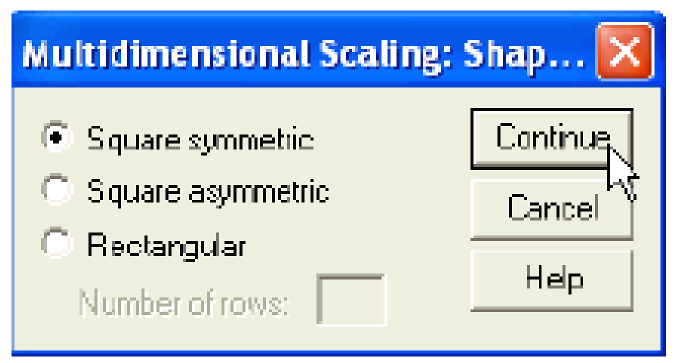
- In the SPSS Data Editor window, click: **Analyze** > **Scale** > **Multidimensional Scaling**



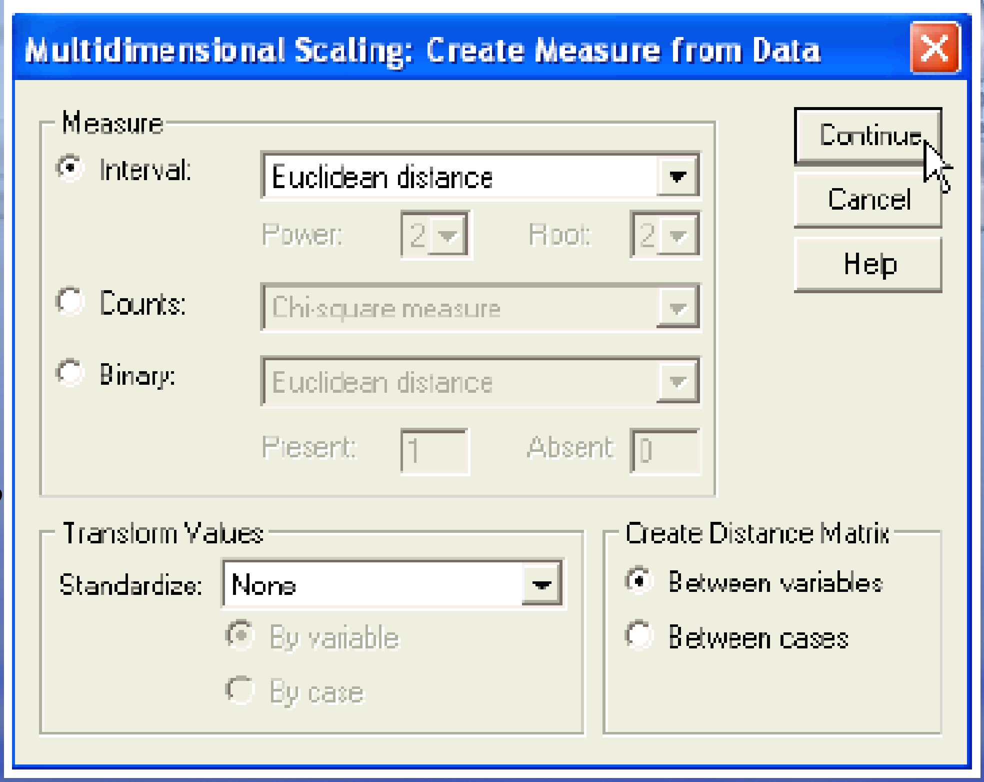
- Select four or more **Variables** that you want to test.
- You may select a single variable for the **Individual Matrices for** window (depending on the distances option selected).



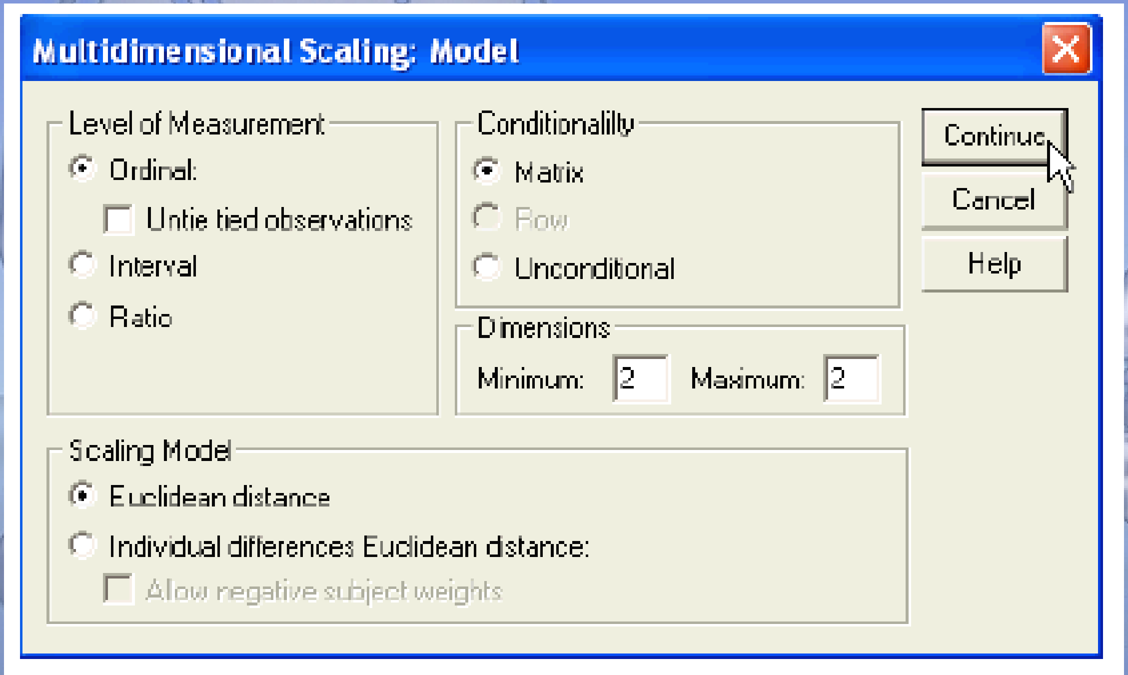
• If **Data are distances** (e.g. cities distances) option is selected, click on the **Shape** button to define characteristic of the dissimilarities/proximity matrices.



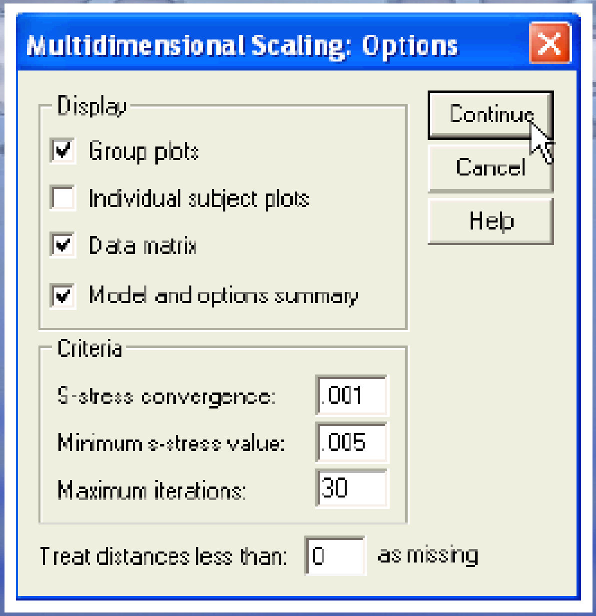
• If **Create distance from data** is selected, click on the **Measure** button to control the computation of dissimilarities, to transform values, and to compute distances.



- In the Multidimensional Scaling dialog box, click on the **Model** button to control the level of measurement, conditionality, dimensions, and the scaling model.



- Click on the **Options** button to control the display options, iteration criteria, and treatment of missing values.



Example of soc.sci. applications: Jacobowitz study of psycholinguistic structure of children's representations of body parts

The analysis located the points in the space, but did not draw the lines. The lines were drawn by Jacobowitz to interpret the psycholinguistic structure that people have for body-part words.

Jacobowitz theorized that the structure would be hierarchical. We can see that it is.

He further theorized that the structure would become more complex as the children become adults. This theory is also supported, since the adults' hierarchy also involves a classification of corresponding arm and leg terms.

2-dimensional comparison of children's & adults' similarity judgements about body parts

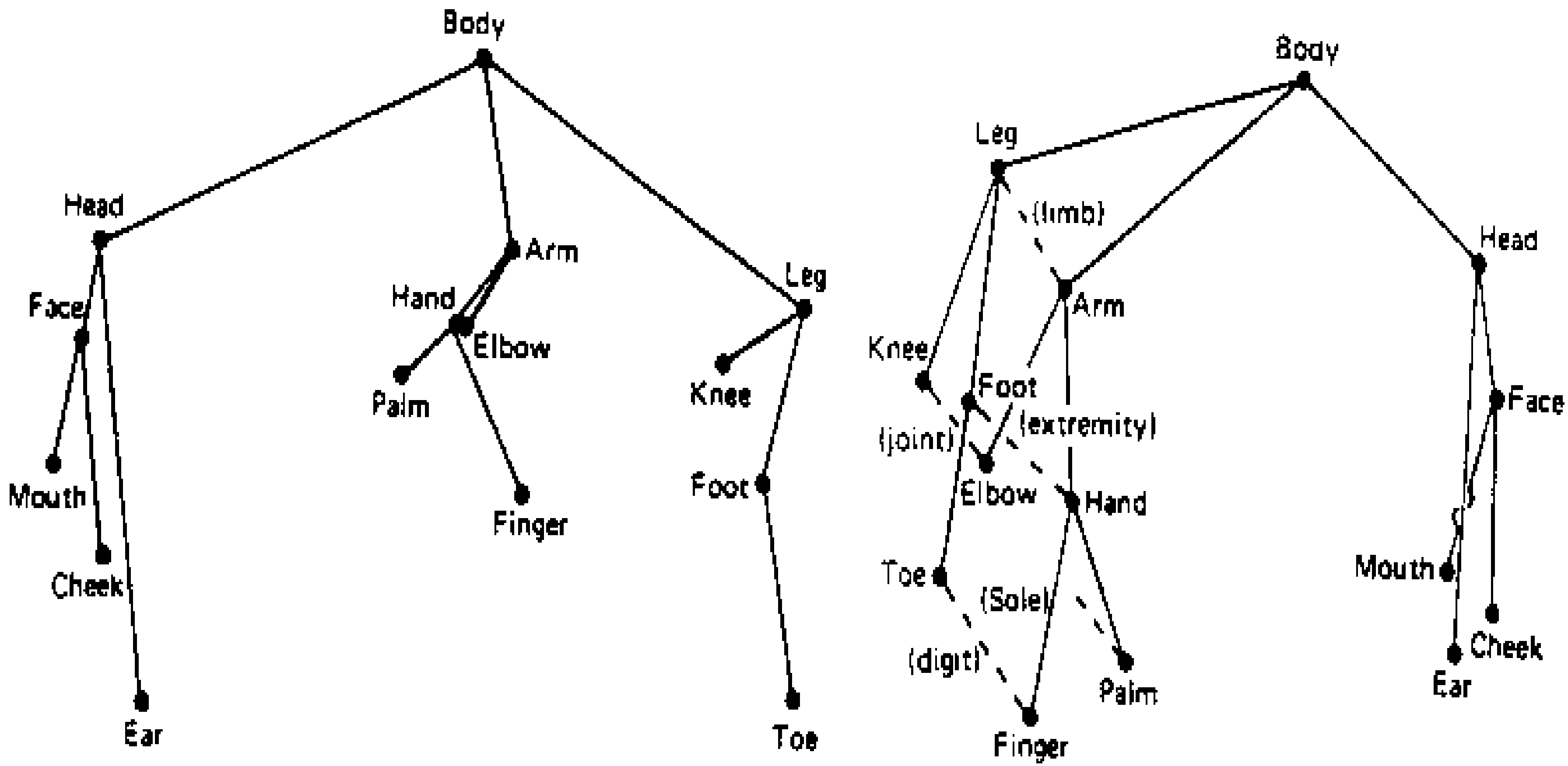
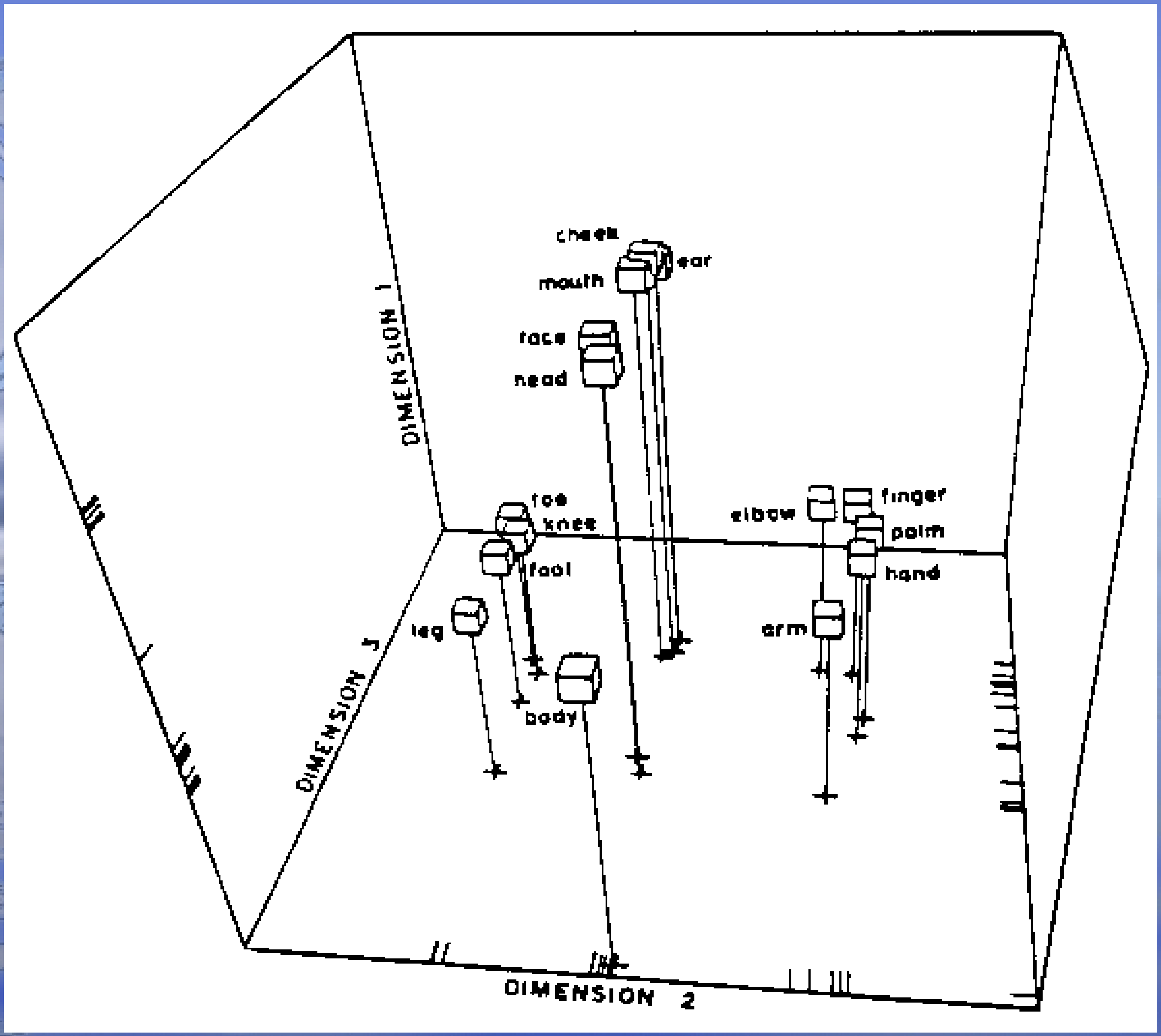
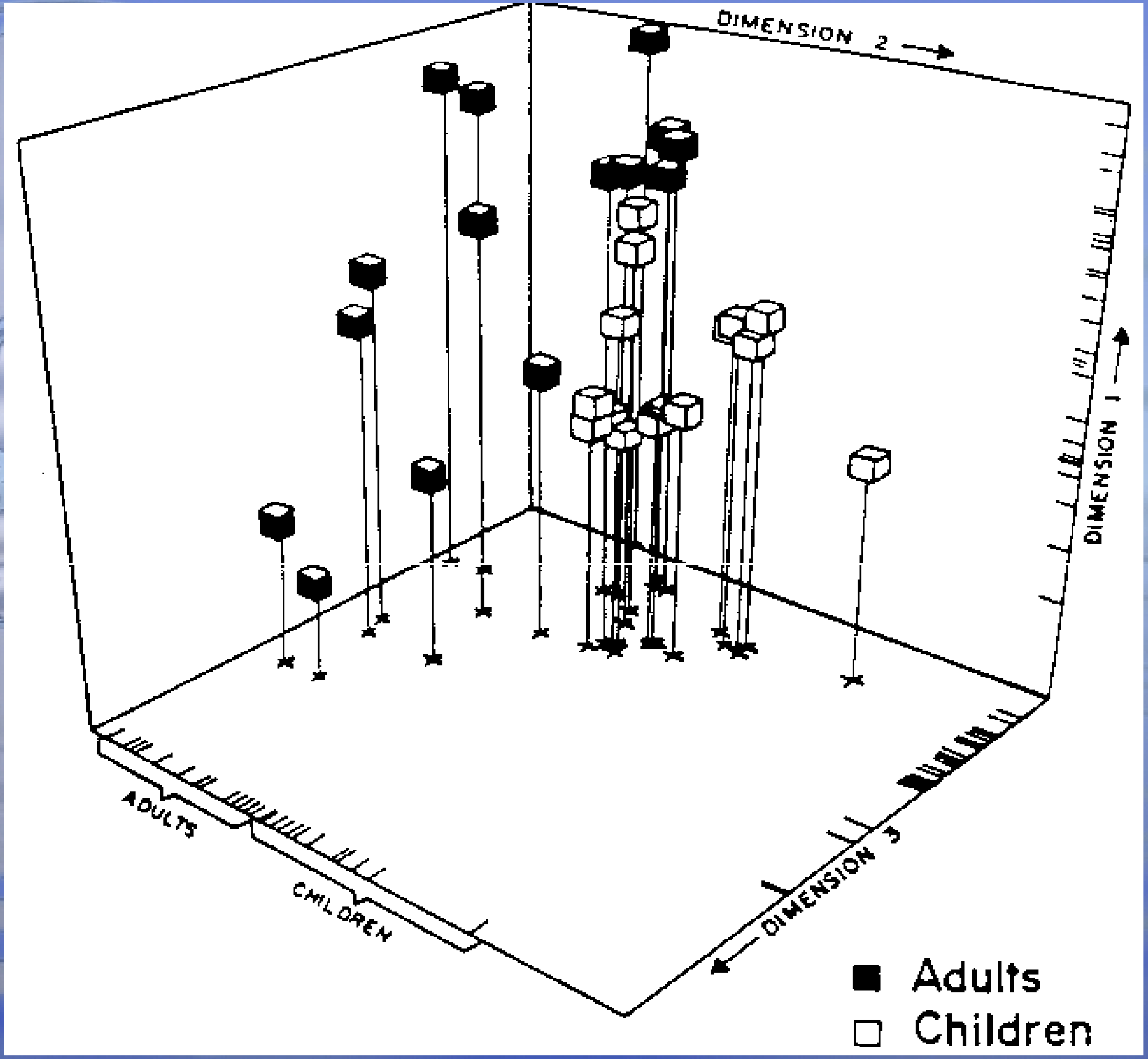


Figure 2 (a) RMDS of children's similarity judgments about 15 body parts; (b) RMDS of adults' similarity judgments about 15 body parts.

3-dimensional MDS solution



3-dimensional comparison of children's & adults' judgements



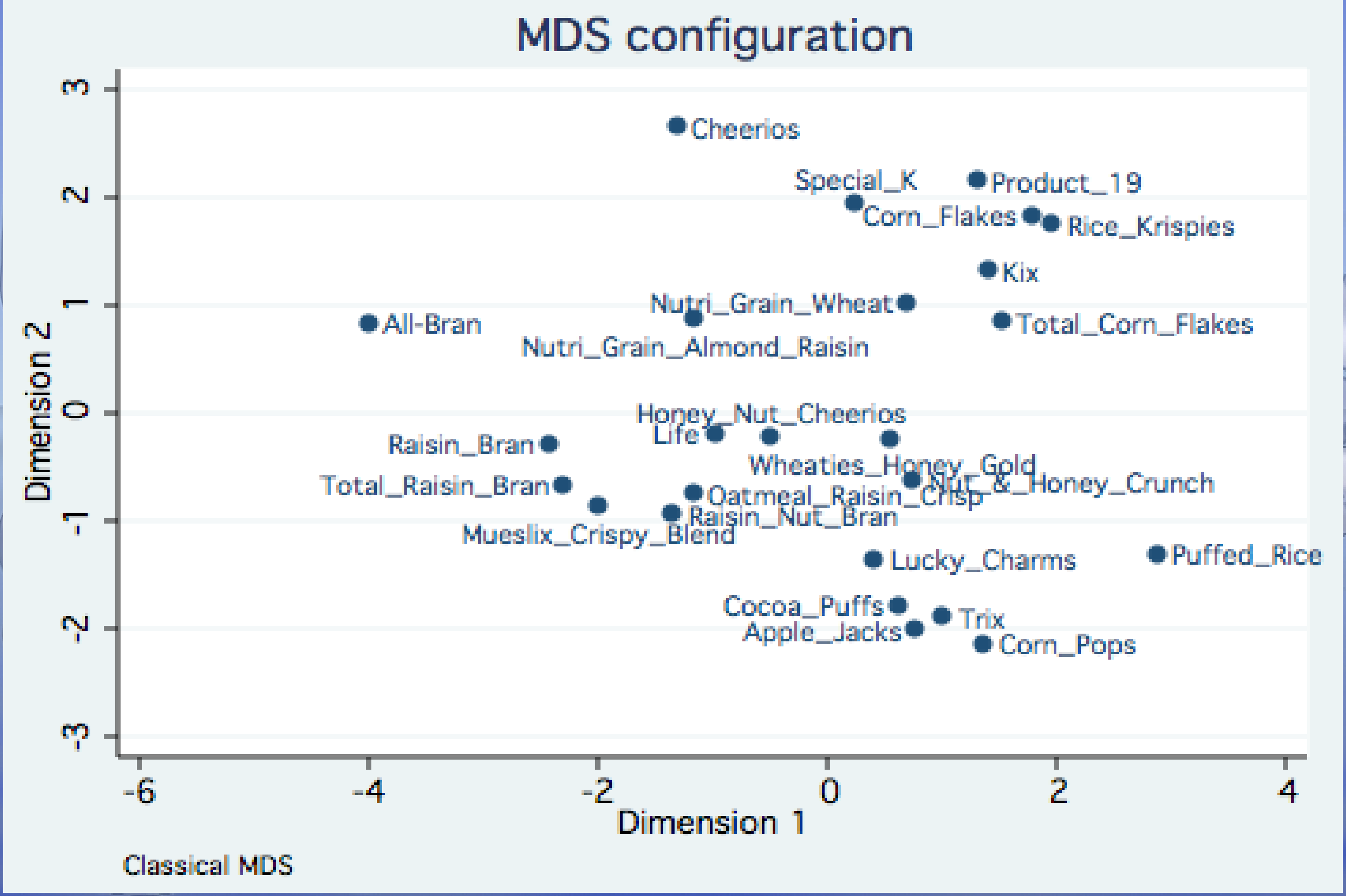
Examples of soc.sci. applications: Perceptions of breakfast cereal brands, in 2 dimensions

	brand
1.	Cheerios
2.	Cocoa_Puffs
3.	Honey_Nut_Cheerios
4.	Kix
5.	Lucky_Charms
6.	Oatmeal_Raisin_Crisp
7.	Raisin_Nut_Bran
8.	Total_Corn_Flakes
9.	Total_Raisin_Bran
10.	Trix
11.	Wheaties_Honey_Gold
12.	All-Bran
13.	Apple_Jacks
14.	Corn_Flakes
15.	Corn_Pops
16.	Mueslix_Crispy_Blend
17.	Nut_&_Honey_Crunch
18.	Nutri_Grain_Almond_Raisin
19.	Nutri_Grain_Wheat
20.	Product_19
21.	Raisin_Bran
22.	Rice_Krispies
23.	Special_K
24.	Life
25.	Puffed_Rice

Approximating configuration in 2-dimensional Euclidean space

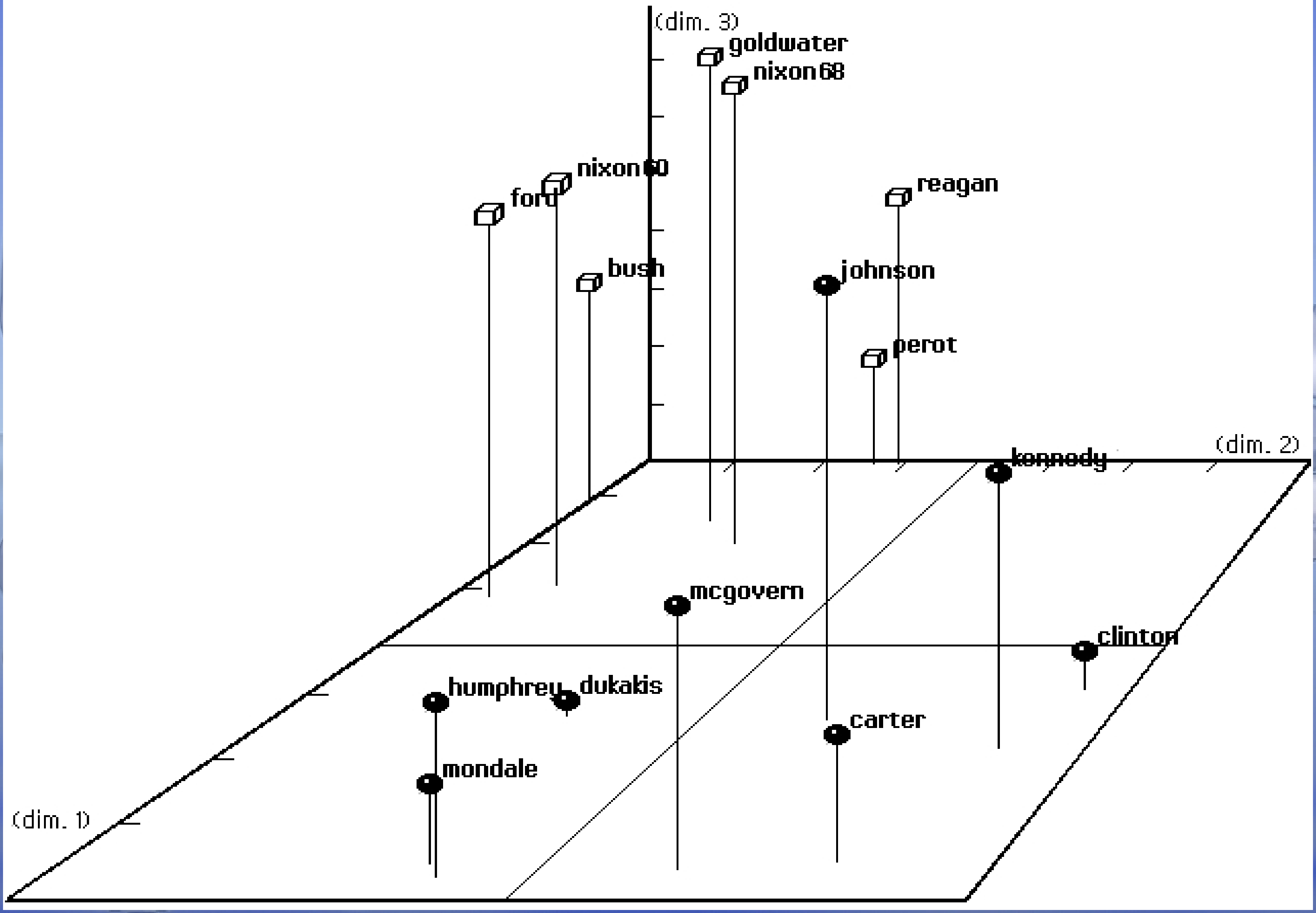
brand	dim1	dim2
Cheerios	-1.3080	2.6638
Cocoa_Puffs	0.6296	-1.7910
Honey_Nut_~s	-0.5050	-0.2227
Kix	1.4003	1.3242
Lucky_Charms	0.4178	-1.3534
Oatmeal_Ra~p	-1.1762	-0.7533
Raisin_Nut~n	-1.3523	-0.9414
Total_Corn~s	1.5175	0.8541
Total_Rais~n	-2.3049	-0.6710
Trix	1.0107	-1.8899
Wheaties_H~d	0.5404	-0.2336
All-Bran	-4.0119	0.8411
Apple_Jacks	0.7712	-2.0103
Corn_Flakes	1.7864	1.8346
Corn_Pops	1.3661	-2.1499
Mueslix_Cr~d	-2.0077	-0.8722
Nut_&_Hone~h	0.7470	-0.6259
Nutri_Grai~n	-1.1706	0.8679
Nutri_Grai~t	0.6929	1.0345
Product_19	1.3073	2.1645
Raisin_Bran	-2.4414	-0.2820
Rice_Krisp~s	1.9619	1.7543
Special_K	0.2362	1.9531
Life	-0.9843	-0.1881
Puffed_Rice	2.8769	-1.3072

Perceptions of breakfast cereal brands, in 2 dimensions (plotted coordinates)



Examples of soc.sci. applications: Perceptions of U.S. presidential candidates, 1960-1996 (2-dim. map)

Examples of soc.sci. applications: Perceptions of U.S. presidential candidates, 1960-1996 (3-dim. map)

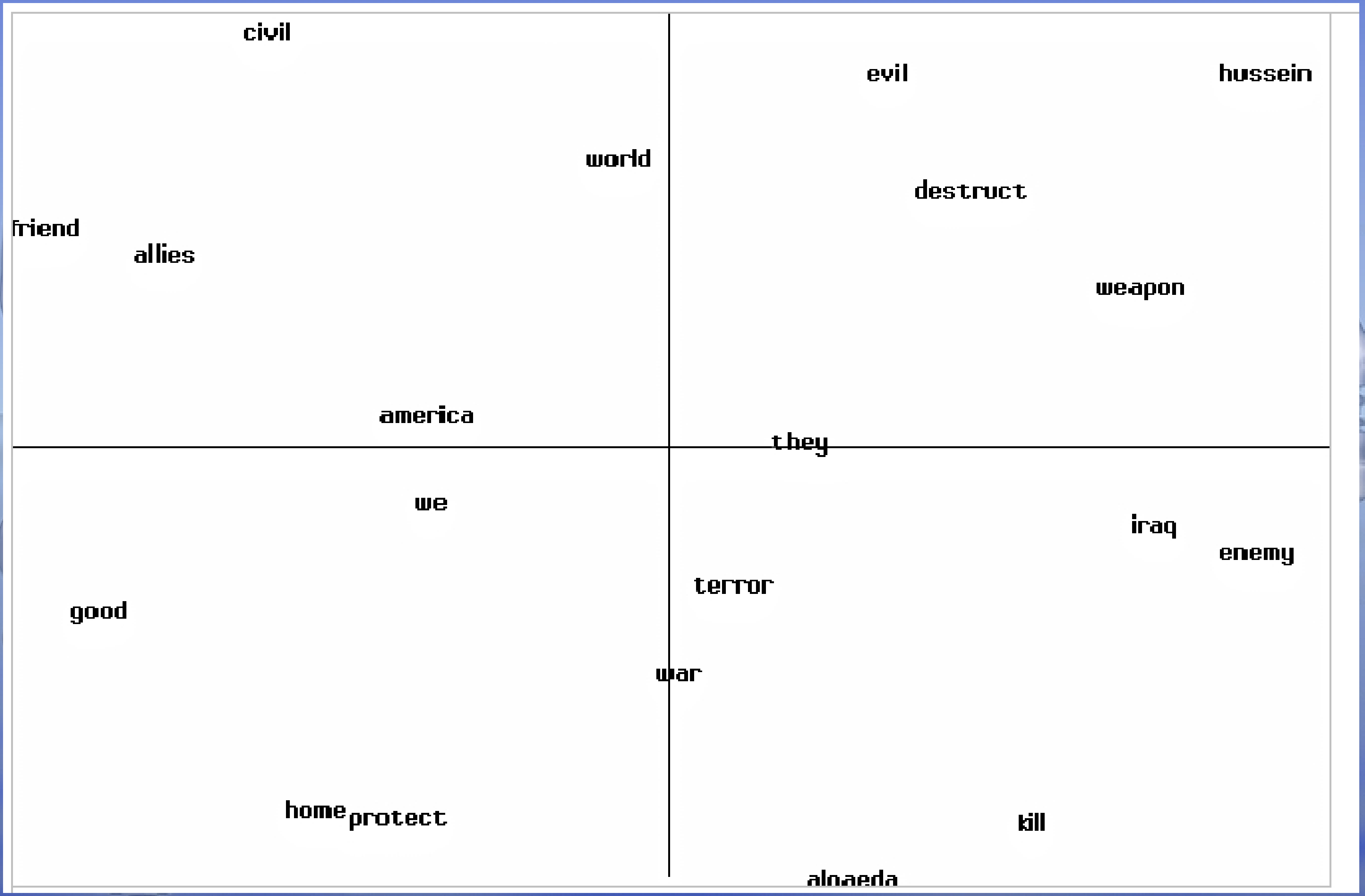


Examples of soc.sci. applications: MDS of keywords in post-9/11 George Bush speeches (word co-occurrences matrix)

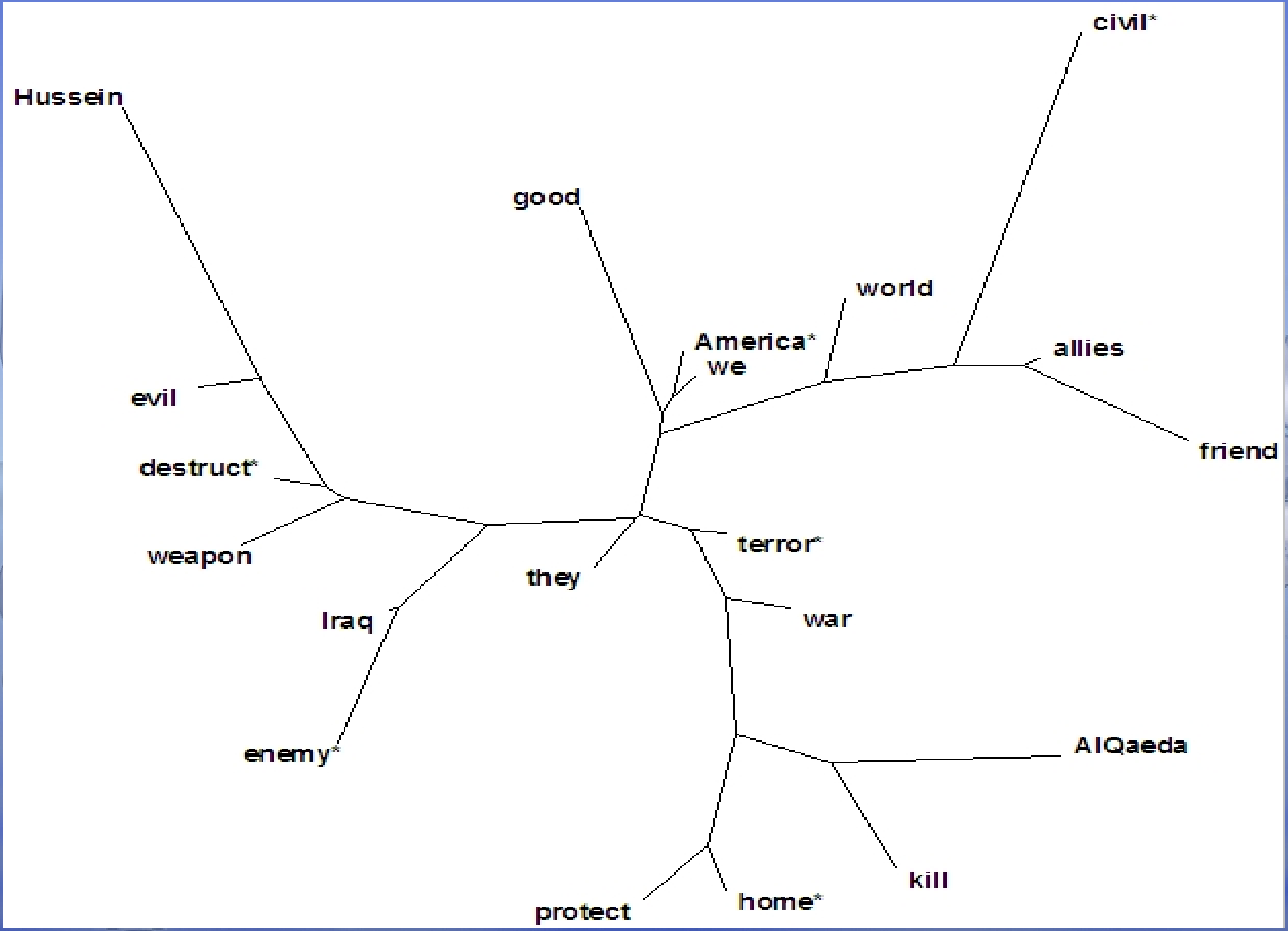
Table 2b: 'Raw' Joint Frequencies (Co-occurrences) for 20 Keywords, Counted Across 4 Bush Congressional Addresses. (Co-occurrences were counted within fixed context units of 9 words.)

Keyword		*Joint Frequencies (Co-occurrences) Matrix*																			
allie*	1																				
alq*	2	0																			
america*	3	5	0																		
civil*	4	0	0	2																	
destr*	5	0	0	1	0																
enem*	6	0	0	2	0	0															
evil*	7	0	0	1	0	1	0														
friend*	8	5	0	4	0	0	0	0													
good*	9	1	1	5	0	0	0	1	1												
home*	10	0	0	6	0	0	0	0	0	1											
hussein	11	0	0	0	0	0	0	1	0	0	0										
iraq*	12	0	0	1	0	0	2	1	0	0	1	1									
kill*	13	0	2	2	0	0	0	0	0	0	0	0	1								
protect*	14	0	1	4	0	0	0	0	0	0	5	0	0	0							
terror*	15	7	3	7	0	1	3	1	0	1	1	0	1	2	2						
they	16	3	2	14	0	4	1	1	2	2	3	1	5	5	2	19					
war*	17	0	0	5	0	2	0	0	0	1	3	0	3	1	3	12	1				
we	18	17	2	69	2	4	13	1	16	30	18	2	10	3	12	27	50	22			
weapon*	19	1	0	4	0	11	1	0	0	0	0	6	1	1	0	5	3	2	7		
world*	20	1	0	11	6	3	0	2	2	2	0	1	3	0	0	1	4	2	21	5	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

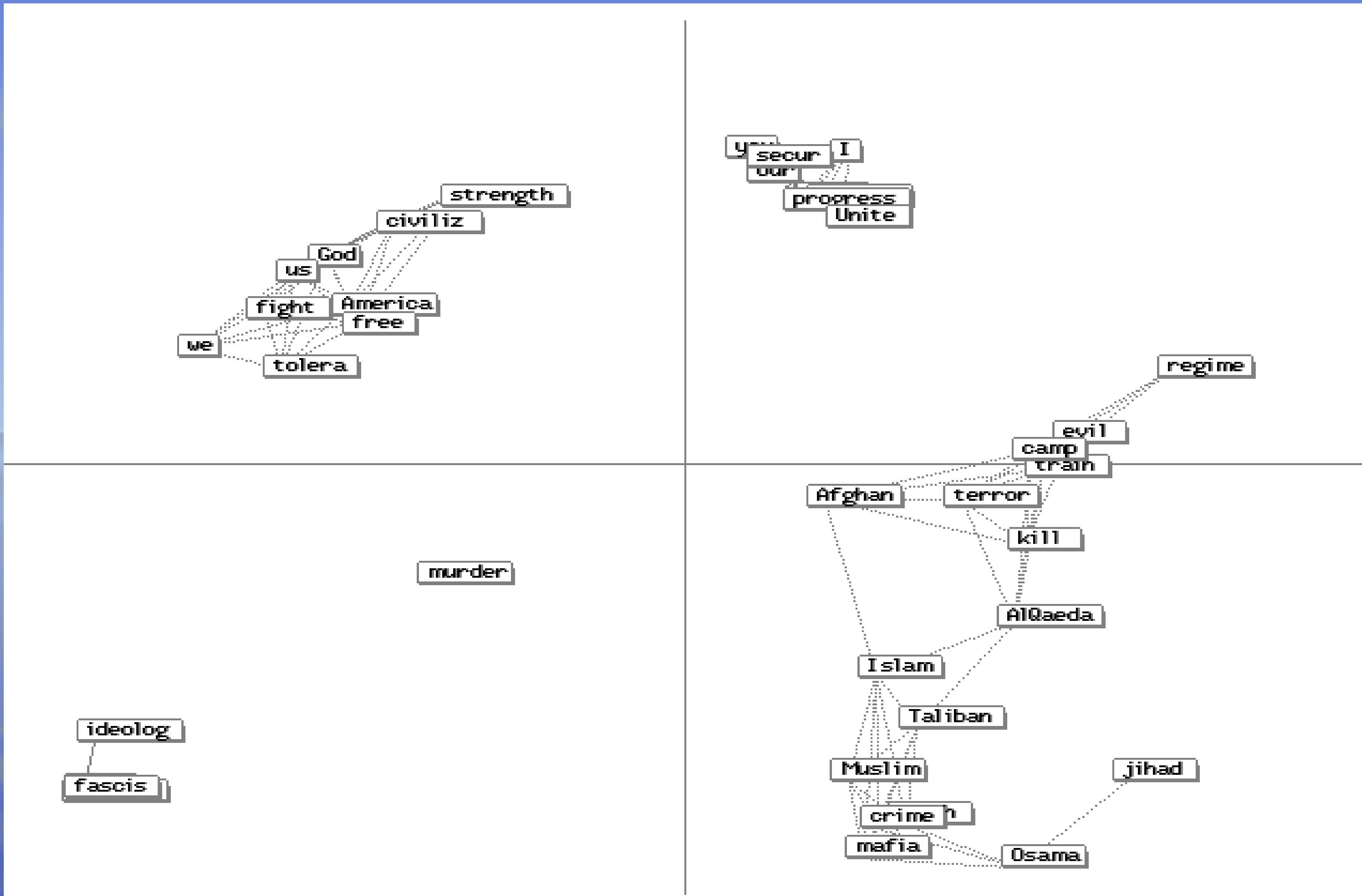
2-dim plot of MDS of keywords in post-9/11 George Bush speeches (note “us” words on left, “them words on right)



MDS of keywords in post-9/11 George Bush speeches (superimposing results of cluster analysis to define branchings)



2-dim. plot of MDS of keywords in post-9/11 George Bush speeches (additional keywords / using WORDPROX program)



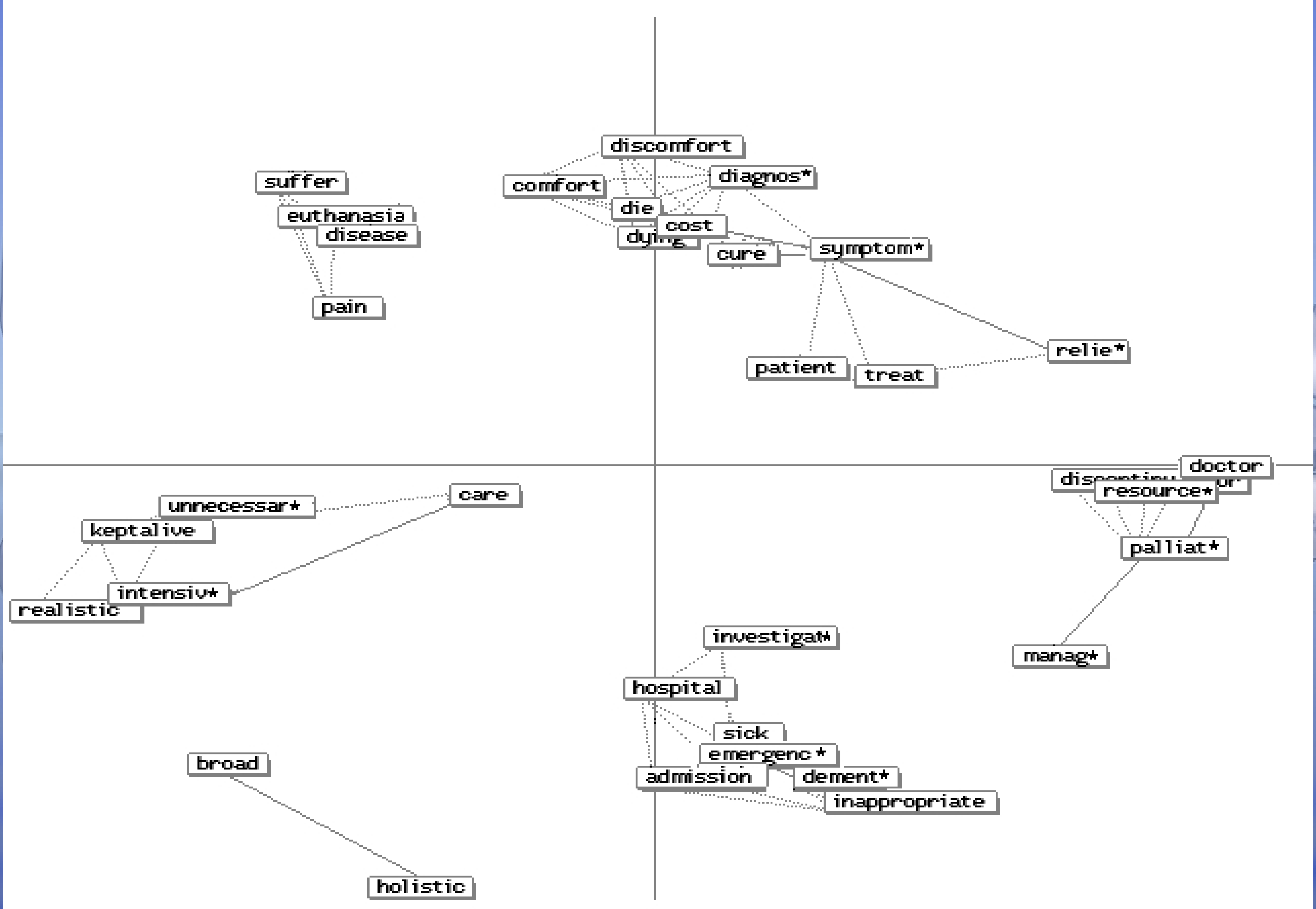
Examples of soc.sci. Applications: Study of 'Biomedical' vs 'Holistic' worldviews among Auckland physicians

Figure 2 Derived spatial configuration for a Biomedical Worldview participant text (interview no.5), in two dimensions. (35-word sample, WORDPROX Kruskal-Guttman-Lingoes-Roskam smallest space coordinates)

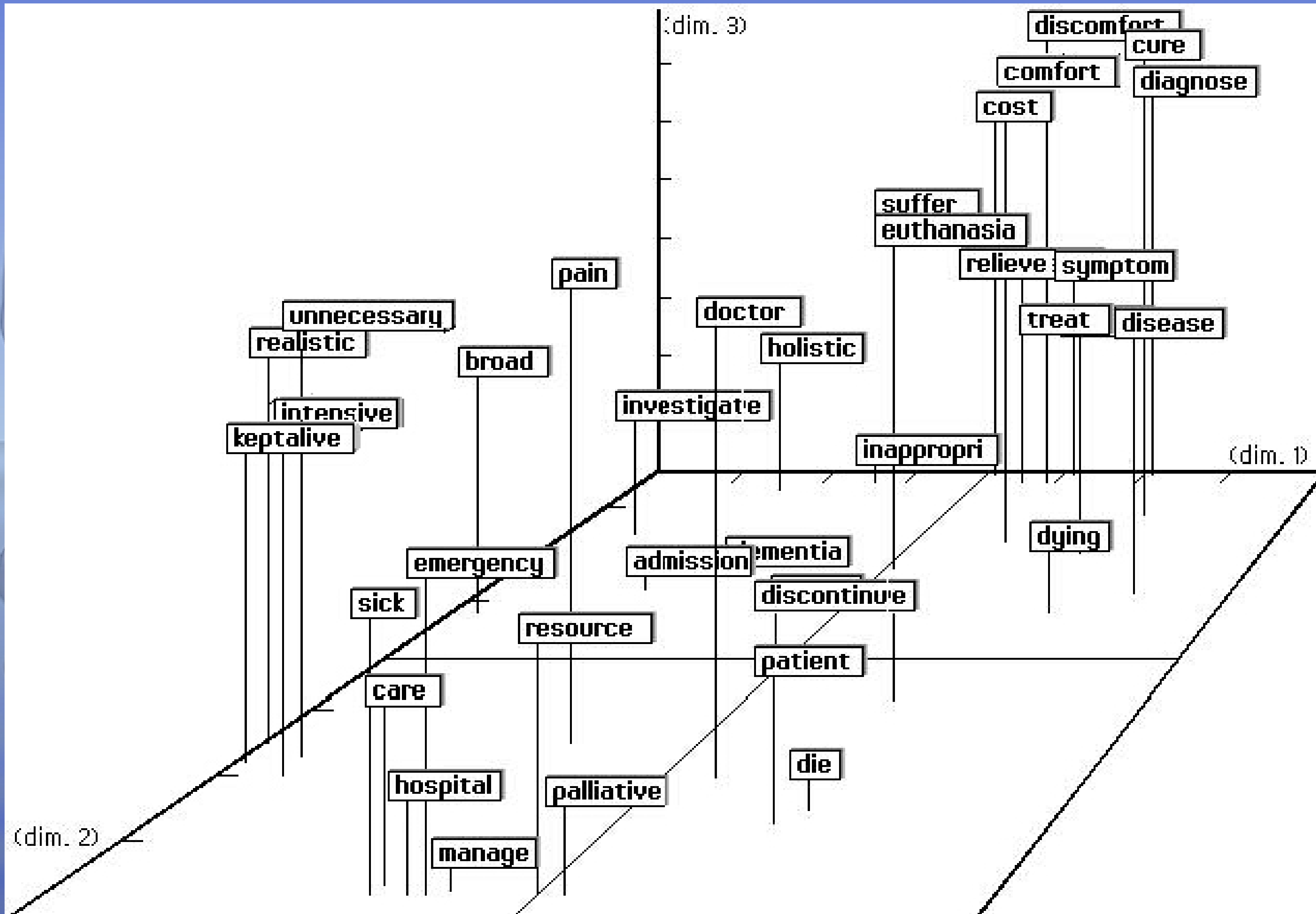
Word	*Dimension*		*Word*	*Dimension*	
	I	II		I	II
holistic	-.7977	-1.8642	emergency	.2899	-1.3807
broad	-1.3478	-1.3132	sick	.2584	-1.1773
comfort	-.3813	1.2674	euthanasia	-1.0704	1.1328
pain	-.9668	.7248	symptom	.5551	.9901
relieve/relief	1.2781	.527	die	-.0536	1.1648
suffer	-1.1382	1.2865	dying	-.0315	1.0398
discomfort	-.0863	1.4461	demented	.501	-1.3732
realistic	-1.8881	-.6337	disease	-.9532	1.0491
diagnos*	.2442	1.3114	manag*	1.1748	-.8371
cure	.1987	1.0595	unnecessary	-1.4362	-.1624
patient	.4835	.4066	waste	1.5285	-.0759
doctor	1.6169	-.0555	discontinue	1.2919	-.0486
hospital	-.0184	-.9808	keptalive	-1.6708	-.2785
care	-.5445	-.1158	cost	.0858	1.0883
palliative	1.5054	-.3512	resource	1.4224	-.0956
treat	.692	.4195	inappropriate	.5968	-1.4838
investigate	.2274	-.7527	admission	.0245	-1.38
intensive	-1.5906	-.5585			

Kruskal's Stress = 0.317

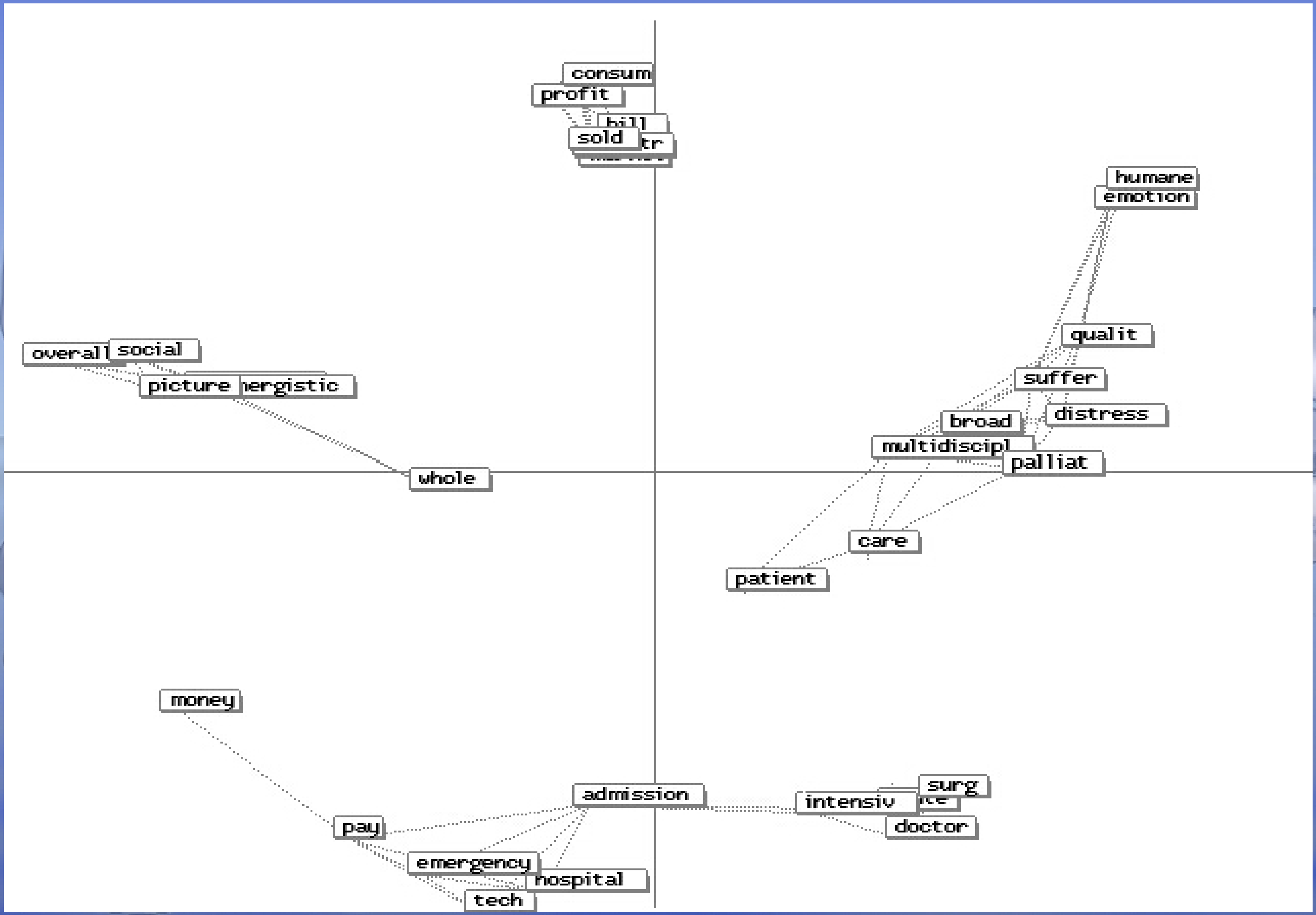
'Biomedical' vs 'Holistic' worldviews among Auckland physicians (2-dim. 'biomedical' example, using WORDPROX)



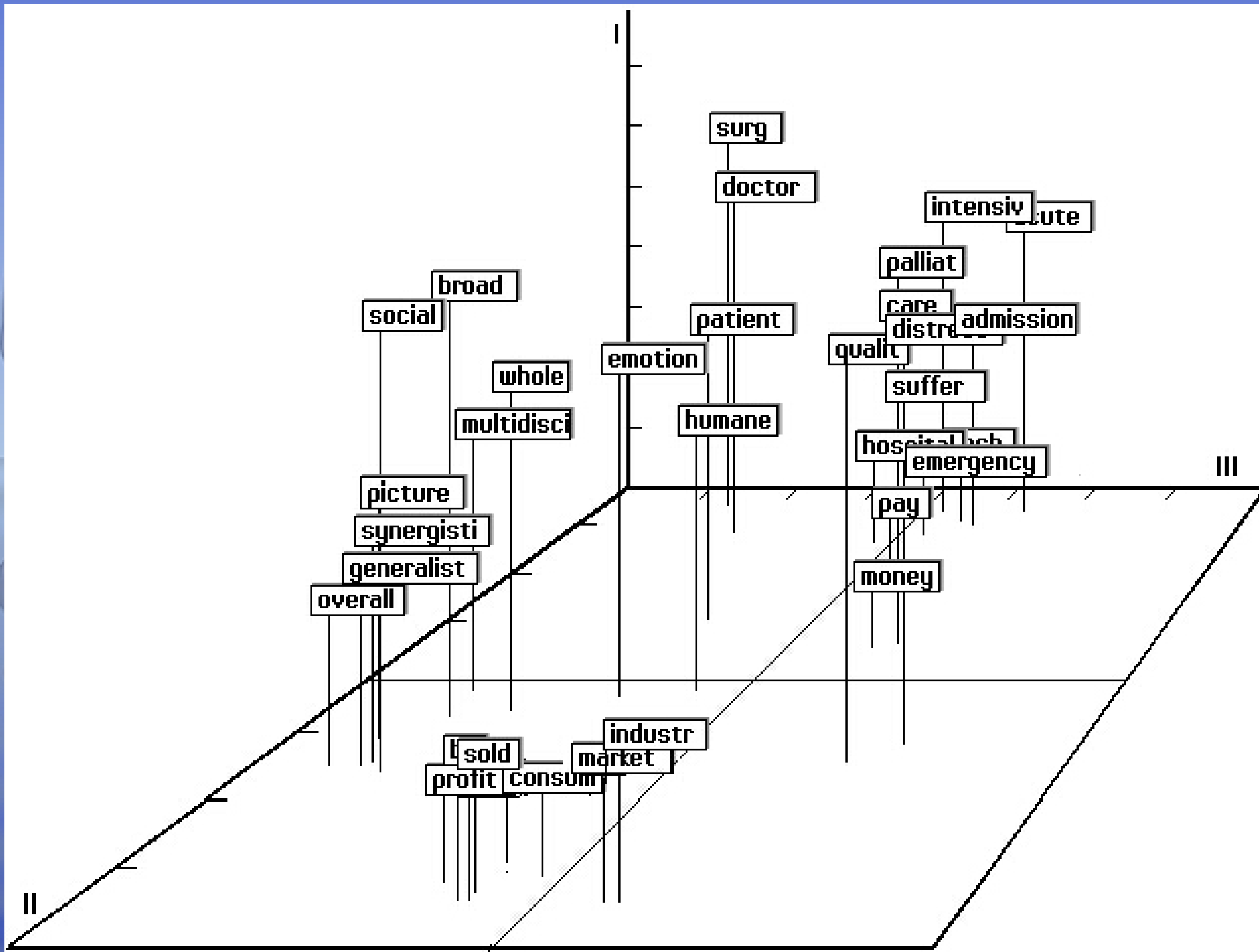
'Biomedical' vs 'Holistic' worldviews among Auckland physicians (3-dim. 'biomedical' example, using WORDPROX)



'Biomedical' vs 'Holistic' worldviews among Auckland physicians (2-dim. 'holistic' example, using WORDPROX)



'Biomedical' vs 'Holistic' worldviews among Auckland physicians (3-dim. 'holistic' example, using WORDPROX)



MDS computer programmes/packages:

--Summary versions of the major MDS procedures are available (but with limited options) in . . . SPSS, SAS, STATA and SYSTAT

--The full, original MDS programmes (many of which can now be obtained in separate or combined form via the NEW MDSX website (Coxon, Brier) include...

MINISSA (classical metric & nonmetric MDS, limited to square matrix)

KYST (similar to MINISSA, but can also do replicated MDS)

MINIRSA (can handle a rectangular matrix, 2-way Coombsian data)

INDSCAL, SINDSCAL (individual differences scaling)

PINDSCAL (does Procrustean individual differences scaling)

ALSCAL (F. Young, original version, handles many different models)

VISTA (F. Young, generalised teaching version of MDS with many options)

PERMAP (allows interactive, in-process control of the MDS solutions)

HAMLET (A.Brier, does MDS of word co-occurrences in a text, & word maps, can do Procrustean and some other MDS models)

WORDPROX (L. Powell, uses both “word co-occurrences” & “word proximities” to triangulate on the word patterns in a text, generally yielding tighter clusters in 2- and 3-dim MDS perceptual maps)

GRID2MDS (L. Powell, does MDS [Coombsian unfolding] on repertory grids, and plots the results)

Useful sources on MDS:

Ashby, F. G. (1992). *Multidimensional models of perception and cognition*. Mahwah, NJ: Erlbaum.

Borg, I., & Groenen, P. (1997). *Modern multidimensional scaling. Theory and applications.* New York: Springer.

Brier, A. (2003). *Analysis of joint frequencies of words in a text: User notes for HAMLET for Windows.* Southampton University: University Computing Service.

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

Coombs, C. H. (1964). *A theory of data*. New York: Wiley.

Ding, C. S. (2006). Multidimensional scaling modelling approach to latent profile analysis in psychological research. *International Journal of Psychology* 41 (3), 226-238.

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

Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a non-metric hypothesis. *Psychometrika, 29*, 1-27.

Useful sources on MDS (cont.):

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

Kruskal, J.B. & Wish M. (1978). *Multidimensional Scaling*. Sage.

Rapoport, A. and Fillenbaum, S. (1972). An experimental study of semantic structures. In K. Romney, R. Shepard, and S. Nerlove (Eds.), *Multidimensional scaling: theory and applications in the social sciences, Volume II: Applications* (pp. 93-131). New York: Seminar Press.

Roskam, E.E. & Lingoes, J.C (1970). MINISSA-1: A FORTRAN IV (G) program for the smallest space analysis of square symmetric matrices, *Behavioral. Science, 15* , 204-205.

Shepard, R. N. (1962a). The analysis of proximities: multidimensional scaling with unknown distance function. Parts I, II. Psychometrika, 27, 125-246.

Takane, Y., Young, F.W., & de Leeuw, J. (1977). Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features, *Psychometrika* 42 (1), 7–67.

Wish, M., Deutsch, M. & Biener, L. (1970). Differences in conceptual structures of nations: An exploratory study. *Journal of Personality and Social Psychology, 16*, 361-373.

Young, F.W., Takane, Y., & Lewyckyj, R. (1978). Three notes on ALSCAL, *Psychometrika* 43 (3), 433–435.

Young, F. W. (1987). *Multidimensional scaling: History, theory and applications.* Hillsdale, NJ: Lawrence Erlbaum.