Bridging Qualitative & Quantitative Methods: Lessons from Anthropology

Bryan Bayles, PhD
Department of Family & Community Medicine
University of Texas Health Science Center San Antonio
baylesb@uthscsa.edu’

Some Goals of Qualitative Research

- 1) identify themes and subthemes
- 2) build and apply codebooks
- 3) describe phenomena
- 4) make comparisons
- 5) build, display, test and validate models

Source: Gery Ryan, http://www.wjh.harvard.edu/nsfqual/Ryan%20Paper.pdf

Rigorous Qualitative Research: Some Definitions and Cautions

- Rigorous research
  - applies the appropriate tools to the stated objective
  - Is transparent and explicit
- Do not confuse research rigor with measurement precision, quantification, generalizability
- Be careful thinking some collection & analysis techniques are “more” rigorous than others. If techniques are just tools, don’t say “a saw is better than a hammer because it’s sharper”

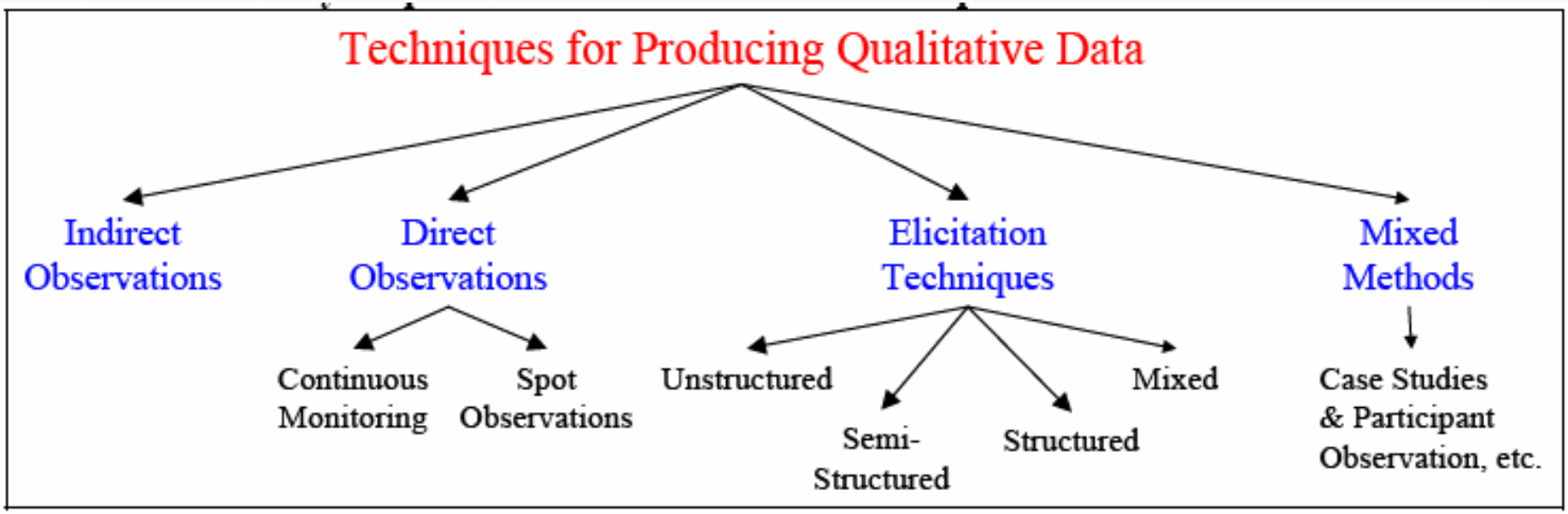
Source: Gery Ryan, http://www.wjh.harvard.edu/nsfqual/Ryan%20Paper.pdf

Key qualitative & quantitative distinctions

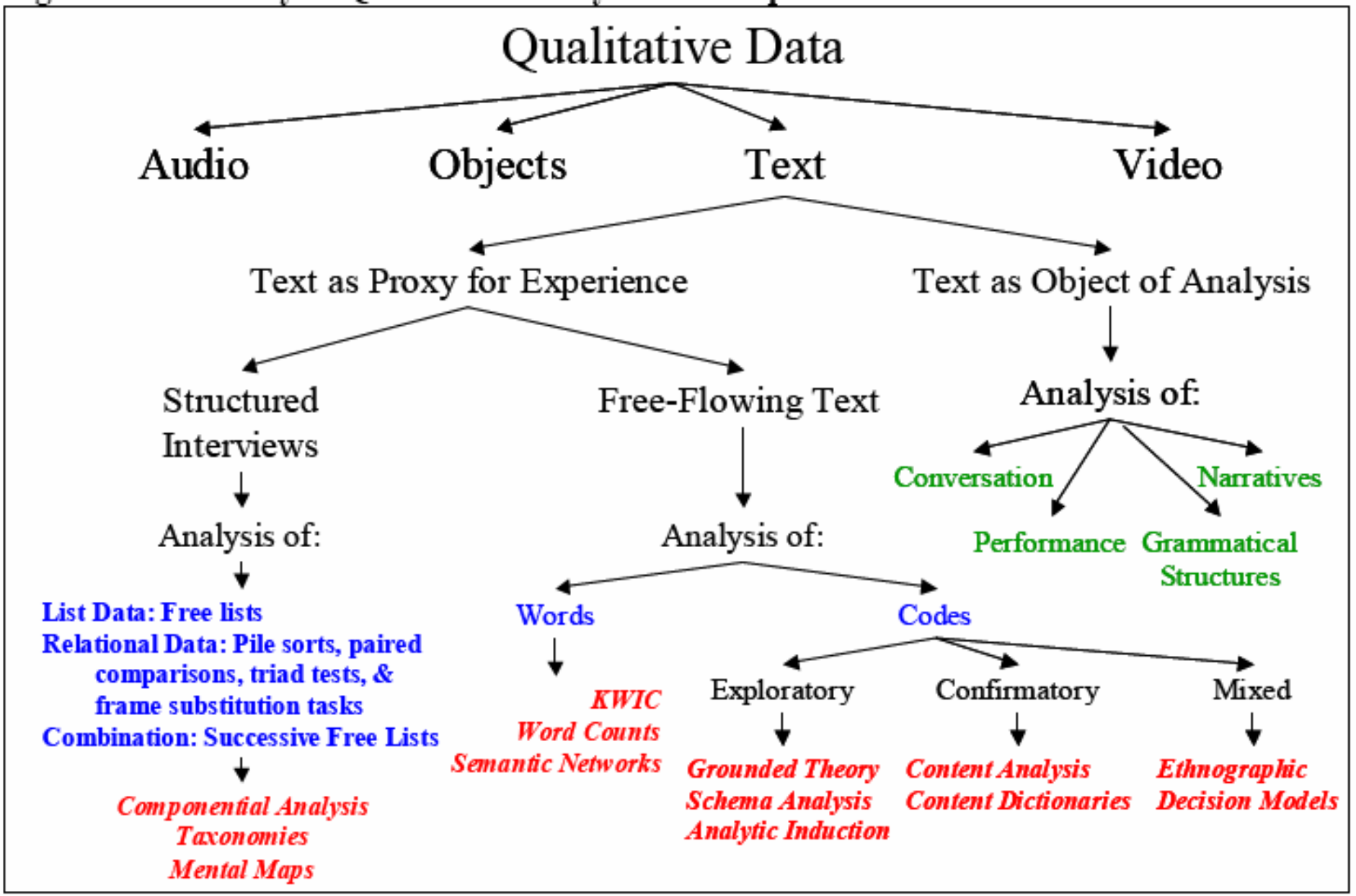
	Data	
Analysis	**Qualitative**	**Quantitative**
Qualitative	A Interpretive text studies. E.g., Hermeneutics, Grounded Theory, Phenomenology	B Search for and presentation of meaning in results of quantitative processing
Quantitative	C Turning words into numbers. E.g., Classic Content Analysis, Word Counts, Free Lists, Pile Sorts, etc.	D Statistical & mathematical analysis of numeric data

Adapted from: Bernard, H. Russell. 1996. Qualitative data, quantitative analysis. *Cultural Anthropology Methods Journal* 8(1):9-11

Taxonomy of qualitative data collection techniques

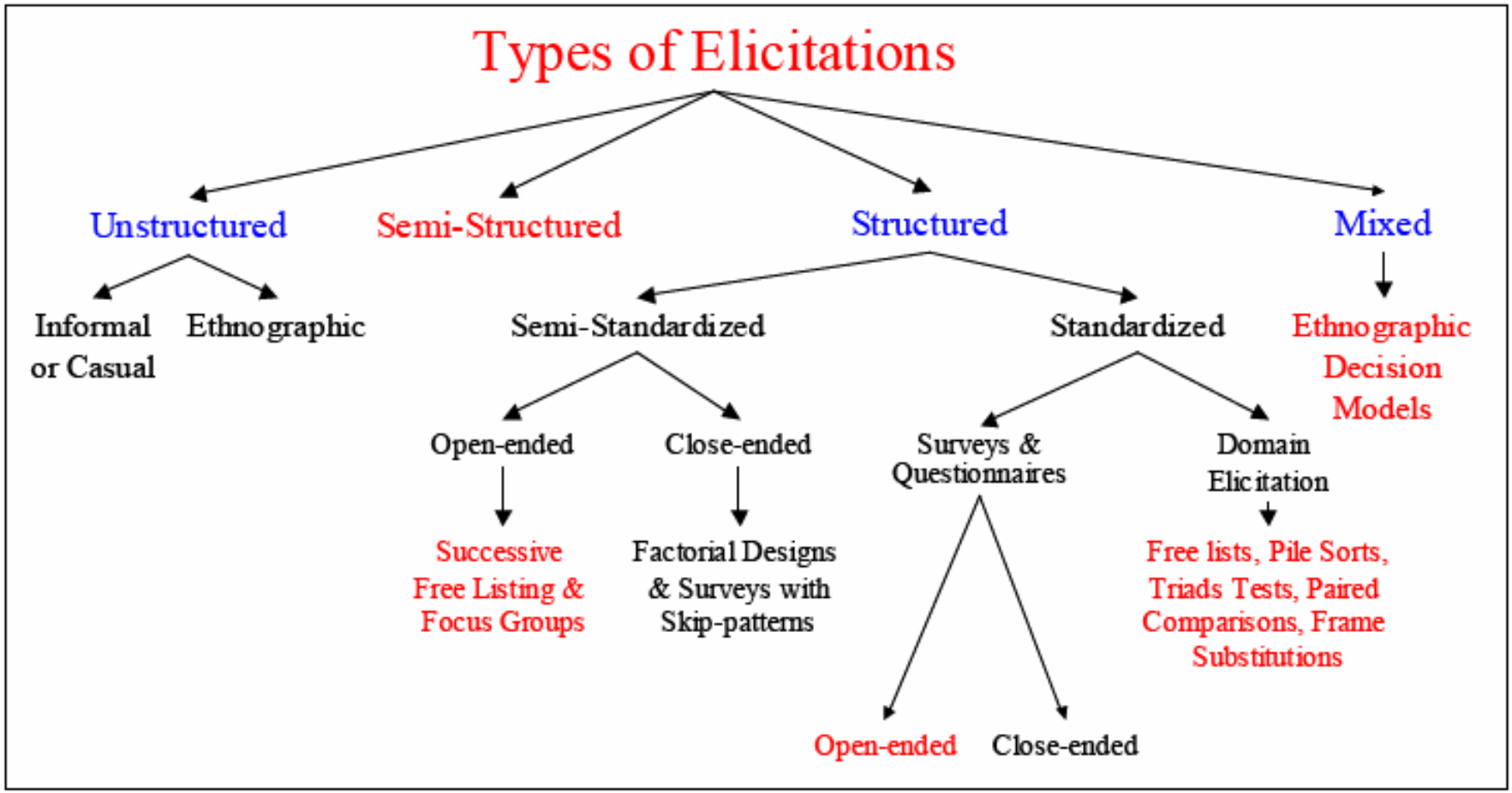


Source: Gery Ryan,
http://www.wjh.harvard.edu/nsfqual/Ryan%20Paper.pdf



Adapted from: Ryan & Bernard, 2000 Data Management and Analysis Methods. In: Handbook of Qualitative Research, 2nd ed., N. Denzin and Y. Lincoln eds., Thousand Oaks, CA: Sage Publications, 769-802..

Taxonomy of elicitation techniques



Source: Gery Ryan, http://www.wjh.harvard.edu/nsfqual/Ryan%20Paper.pdf

One Strategy for Successive Qualitative – Quantitative Research

Collection	Analysis
Free Lists (Illnesses, causes, symptoms, conventional/alternative treatments)	-Salience Indices -Use-Value Index for medicinal plants
Going Deeper: Semi-structured Interviews Instrument (Spradley 1979 as guide)	-Grounded Theory, KWIC, Identifying themes and metaphors -Comparison of contextualized passages
Structured Frame Questionnaire (Illness-by-Frame Instrument)	-Multi-dimensional Scaling -Cluster Analysis

Bayles B (2002) The Belly Wants its Heat: Cultural Models of Health & Fertility among Tojolab'al Maya Midwives, Dissertation, Univ. Missouri Columbia.

Sample Size

- Depends on the degree of consensus
- Complete relative frequencies after 20, then again after 30. If no change, no need for more info
  - BUT the informants must be randomly selected.

Universal Attributes in Cultural Domains

- Similarity (X *is similar to* Y)
- Co-occurrence (X *goes with* Y)
- Others:
  - Cause & effect (X *causes* Y)
  - Inclusion (X *is a kind of* Y)
  - Rationale (X *is a reason for doing* Y)
  - Means-end (X *is a way to accomplish* Y)
  - Sequence (X *follows* Y)

Salience Indices

- Items recalled first are assumed to be more salient than items recalled last
- Smith's s
  - Rank of term in list * (1/(list length -1))
- Sutrop's index:
  - Salience of a term= F/(NmP)
    - Frequency (F) with which a term is mentioned, N is number of respondents, mP is mean position of that term
- Can use to compare groups, male-female, experts-laypersons, old-young, regional, etc

Simple tricks you can do with a word processor

- ID009. F1030. Boy. Fthr. Loving. Obedient. Maintains own identity. Likes being home. Independent. Anxious to go to California to school.
- ID016. F1130. Boy. Fthr. Smart. Energetic. Arrogant. Dependent. Slick. Passive. Lack of imagination. Attraction to inner-city lifestyle.
- ID124. F1130. Girl. Mthr. Great kid. Willing to communicate with parents. Listens. Motivated in school. Helpful around the house. Healthy. Active. Lots of friends. She tends to play it safe.

Source: Gery Ryan and Thomas Weisner
http://www.analytictech.com/borgatti/ryantext.htm

Simple Text Statistics

	Mothers	Fathers	Total
Characters	9748	7625	17373
Word Count	1692	1346	3038
Sentence Count	528	411	939
Average Word Length	5.76	5.66	5.72
Average Words per Sentence	3.20	3.27	3.24
Maximum Words per Sentence	14	17	17

Source: Gery Ryan and Thomas Weisner
http://www.analytictech.com/borgatti/ryantext.htm

Word frequencies sorted by standardized frequency difference in gender

TERM	Both	Mother	Father	Expected Father	Standardized difference
school	26	10	16	20.32	-10.32
good	45	22	23	29.21	-7.21
lack	9	2	7	8.89	-6.89
student	9	2	7	8.89	-6.89
enjoys	6	1	5	6.35	-5.35
independent	13	5	8	10.16	-5.16
extremely	4	0	4	5.08	-5.08
like	4	0	4	5.08	-5.08
ability	7	2	5	6.35	-4.35
own	7	2	5	6.35	-4.35
wants	7	2	5	6.35	-4.35
high	5	1	4	5.08	-4.08
interested	5	1	4	5.08	-4.08
great	11	6	5	6.35	-0.35
mature	11	6	5	6.35	-0.35

Source: Gery Ryan and Thomas Weisner
http://www.analytictech.com/borgatti/ryantext.htm

Mothers' Descriptions

Total number of running words in file: 1,721

Number of unique word forms in file: 734

The following counts exclude 542 occurrences of 125 common word forms.

#	Count/Word
1	22 good
2	12 friends
3	11 loving
4	11 out
5	11 people
6	10 doesn't
7	10 hard
8	10 school
9	9 responsible
10	9 sense
11	8 caring
12	8 intelligent
13	8 lacks
14	8 sensitive
15	7 bright
16	7 honest
17	7 others

Fathers' Descriptions

Total number of running words in file: 1,355

Number of unique word forms in file: 607

The following counts exclude 419 occurrences of 125 common word forms.

#	Count/Word
1	23 good
2	16 school
3	11 hard
4	9 intelligent
5	8 bright
6	8 independent
7	8 out
8	8 well
9	7 doesn't
10	7 lack
11	7 loving
12	7 people
13	7 sensitive
14	7 sports
15	7 student
16	6 caring
17	6 does
18	7 self
19	7 time
20	7 well
21	7 work
22	6 creative
23	6 does
24	6 great
25	6 mature
26	6 sports
27	5 academically
28	5 artistic
29	5 cares
30	5 concerned
31	5 goals
32	5 going
33	5 humor
34	5 independent
35	5 other
36	5 social
37	5 times
...	...
##	1 zest
18	6 life
19	6 others
20	6 work
21	5 ability
22	5 enjoys
23	5 great
24	5 lacks
25	5 likes
26	5 mature
27	5 own
28	5 sense
29	5 social
30	5 wants
...	...
54	1 zero
8	

Respondent-by-item Matrix

(1=mentioned; 0=did not mention;
Columns could be themes, metaphors, items from lists, etc)

Respondent	Theme 1	Theme 2	Theme 3	Theme 4	Theme 5
1	1	0	1	0	1
2	1	0	1	1	1
3	1	0	1	0	0
4	0	1	1	0	1
5	0	0	1	0	1
6	1	1	1	1	0
7	0	0	0	1	0
8	1	0	1	1	0
9	1	1	0	1	1
10	1	1	1	1	0

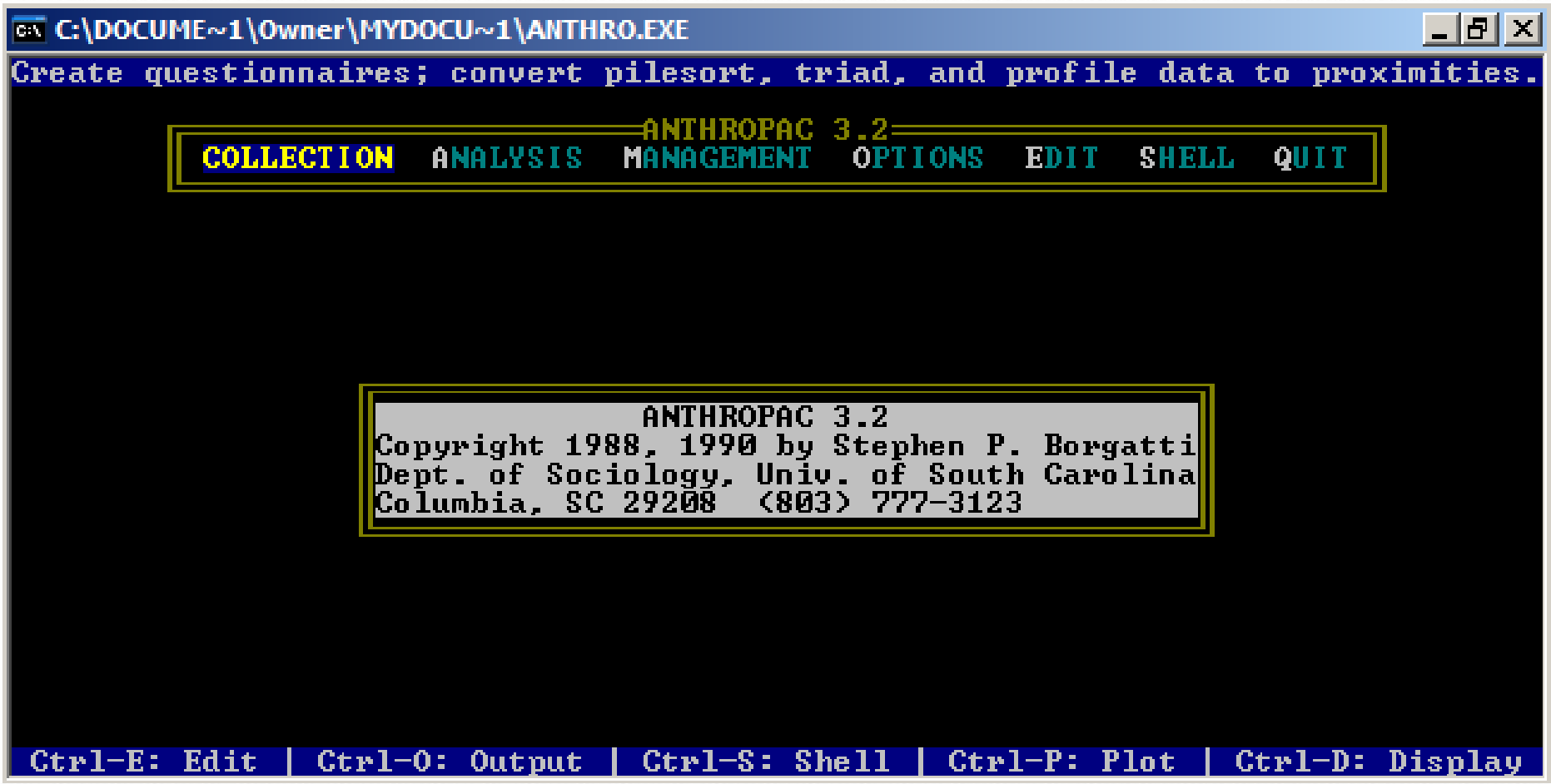
Can also compute co-occurrence then submit to MDS

	A	B	C	D
A	4	3	2	2
B	3	4	1	1
C	2	1	3	2
D	2	1	2	2

Item-by-item matrix of co-occurrences

(For a 0,1 matrix) in SPSS=
Analyze->Scale->
Multidimensional Scaling->
Create Distances from Data

Anthropac


C:\DOCUME~1\Owner\MYDOCU~1\ANTHRO.EXE
Create questionnaires; convert pilesort, triad, and profile data to proximities.
ANTHROPAC 3.2
COLLECTION ANALYSIS MANAGEMENT OPTIONS EDIT SHELL QUIT
ANTHROPAC 3.2
Copyright 1988, 1990 by Stephen P. Borgatti
Dept. of Sociology, Univ. of South Carolina
Columbia, SC 29208 (803) 777-3123
Ctrl-E: Edit | Ctrl-O: Output | Ctrl-S: Shell | Ctrl-P: Plot | Ctrl-D: Display

Step-by-Step

- 1. Collect data for a person-by-method matrix which contains a 1 if a given person has used a given statistical method, and 0 otherwise. Here is a hypothetical example of such a matrix:

	Correlation	Regression	ANOVA	MDS	FACTOR	Chi-square	Log-Linear
Bill	1	1	1	0	1	0	0
John	1	0	0	1	1	0	0
Mary	0	0	1	0	0	1	1
Don	0	0	1	1	0	1	0
Jan	1	1	0	0	0	1	0
Sally	0	1	1	0	0	1	1

Source: Borgatti, S: http://www.analytictech.com/borgatti/proximit.htm

2. Enter the data into an ascii file called STATMETH.DAT using the following format:

```
dl nr 6 nc 7
Title='Experience with statistical methods'
col labels:
Correlation Regression ANOVA MDS FACTOR Chi-square Log-Linear
row labels embedded:
data:
Bill  1 1 1 0 1 0 0
John  1 0 0 1 1 0 0
Mary  0 0 1 0 0 1 1
Don   0 0 1 1 0 1 0
Jan   1 1 0 0 0 1 0
Sally 0 1 1 0 0 1 1
```

3. Import the data as a dataset called STATMETH.

Source: Borgatti, S: http://www.analytictech.com/borgatti/proximit.htm

Proximities Among Persons

1. Choose TOOLS>SIMILARITIES from the menu. Fill in the input form as shown below:

```
Input dataset:
STATMETH
Measure of profile similarity:                  MATCHES
Compute similarities among ROWS or COLUMNS?     ROWS
(For square matrices) Diagonal valid?:          YES
Output dataset:                                 MATCHES
```

To run the program, press F10. The result should be the following matrix:

		1 Bill	2 John	3 Mary	4 Don	5 Jan	6 Sally
1	Bill	1.00	0.57	0.29	0.29	0.57	0.43
2	John	0.57	1.00	0.14	0.43	0.43	0.00
3	Mary	0.29	0.14	1.00	0.71	0.43	0.86
4	Don	0.29	0.43	0.71	1.00	0.43	0.57
5	Jan	0.57	0.43	0.43	0.43	1.00	0.57
6	Sally	0.43	0.00	0.86	0.57	0.57	1.00

Source: Borgatti, S: http://www.analytictech.com/borgatti/proximit.htm

2. Choose TOOLS>SIMILARITIES from the menu. Fill in the input form as shown below (note change of measure to CORRELATION):

```
Input dataset:
STATMETH
Measure of profile dissimilarity:
EUCLIDEAN
Compute similarities among ROWS or COLUMNS?  ROWS
(For square matrices) Diagonal valid?:       YES
Output dataset:
EUCLIDEAN
```

To run the program, press F10. The result should be the following matrix:

```
                    1      2      3      4      5
6
                 Bill   John   Mary    Don    Jan
Sally
               ------ ------ ------ ------ ------
------
   1   Bill    1.00   0.17  -0.42  -0.42   0.17
-0.17
   2   John    0.17   1.00  -0.75  -0.17  -0.17
-1.00
   3   Mary   -0.42  -0.75   1.00   0.42  -0.17
0.75
   4    Don   -0.42  -0.17   0.42   1.00  -0.17
0.17
   5    Jan    0.17  -0.17  -0.17  -0.17   1.00
0.17
   6  Sally   -0.17  -1.00   0.75   0.17   0.17
1.00
```

Source: Borgatti, S: http://www.analytictech.com/borgatti/proximit.htm

3. Choose TOOLS>DISSIMILARITIES from the menu (note change to dissimilarities). Fill in the input form as shown below.

```
Input dataset:
STATMETH
Measure of profile dissimilarity:
EUCLIDEAN
Compute similarities among ROWS or COLUMNS?  ROWS
(For square matrices) Diagonal valid?:       YES
Output dataset:
EUCLIDEAN
```

To run the program, press F10. The result should be the following matrix:

		1 Bill	2 John	3 Mary	4 Don	5 Jan	6 Sally
1	Bill	0.00	1.73	2.24	2.24	1.73	2.00
2	John	1.73	0.00	2.45	2.00	2.00	2.65
3	Mary	2.24	2.45	0.00	1.41	2.00	1.00
4	Don	2.24	2.00	1.41	0.00	2.00	1.73
5	Jan	1.73	2.00	2.00	2.00	0.00	1.73
6	Sally	2.00	2.65	1.00	1.73	1.73	0.00

To compute proximities among each pair of methods, just repeat the process above, but change "ROWS" to "COLUMNS" in every case.

Source: Borgatti, S: http://www.analytictech.com/borgatti/proximit.htm

Free List Example: Data Format

- For each informant, the data should begin with a pound sign (#) and an optional identifier (such as the informant's name or id number), followed by the items he or she mentioned. Each item should be typed on a line by itself. For example, the following file contains hypothetical free lists of the birth control domain from 3 respondents, identified as "Father Mulcahey", "Tammy Faye Bakker", and "Jim Bakker".

Contents of file BCONTROL.FL:

```
# Father Mulcahey
abstinence
# Tammy Faye Bakker
douche
make-up
Jimmy Bakker
Jessica Hahn
# Jim Bakker
abortion
morning after pill
```

Free-Lists: Outputs

- SORTED FREQUENCIES. A list of items mentioned, sorted in descending order by frequency, is output to the screen (unless redirected via the OUTPUT procedure).

- CLEANED UP FILE. A copy of the data file in which spelling differences have been resolved is written to disk.

- RESPONDENT-ITEM MATRIX. A binary respondent-by-item matrix is also written to disk. The Xij cell of the matrix is 1 if respondent i mentions item j and 0 otherwise. For example, in the case of the file BCONTROL.FL, the following output file is created:

Outputs cont.

Contents of file FLMATRIX.DAT:

DL NR = 3 NC = 7
ROW LABELS:
"FATHER MUL“
"TAMMY FAYE“
"JIM BAKKER“
COLUMN LABELS:;
"ABSTINENCE“
"DOUCHE“
"LOTS OF MAKE-UP“
"JIMMY BAKKER“
"JESSICA HAHN“
"ABORTION“
"MORNING AFTER PILL“
DATA:
1 0 0 0 0 0 0
0 1 1 1 1 0 0
0 0 0 0 0 1 1

Input for Multidimensional Scaling Procedure

TRIADS

STEP 1. Obtain a list of items in a domain.
Examples of items are diseases, treatments, clinics, etc.

STEP 2. Second, compose a triads questionnaire. In such a questionnaire, the informant is (repeatedly) shown three items and asked to choose the item most DIFFERENT=

1.	DOG	SEAL	MOSQUITO
2.	BEAR	SHARK	DOLPHIN
....			

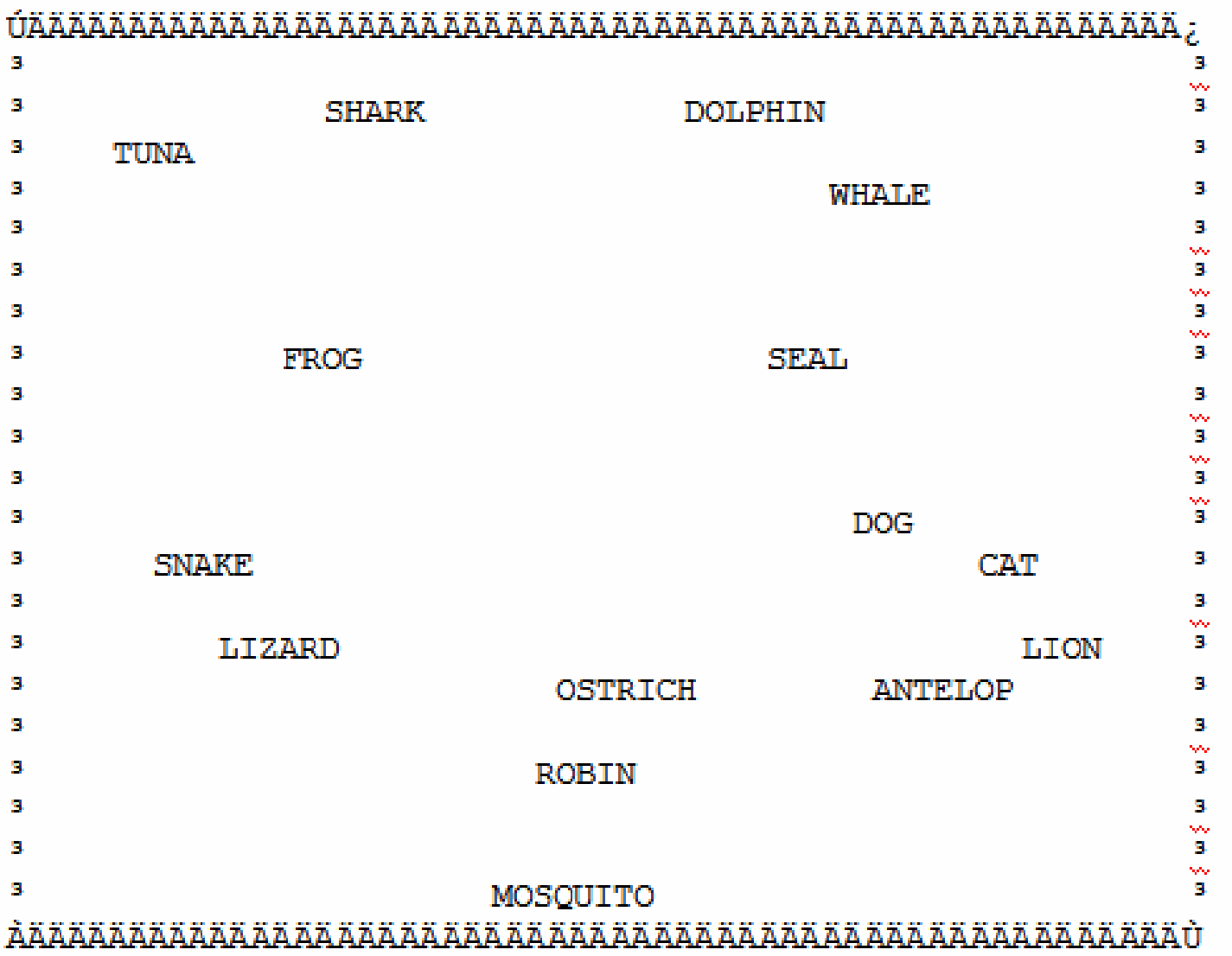
STEP 3. Administer the questionnaire. Typically, respondents are asked to circle the item most different, although logically they could be asked to pick the pair most similar. However, doing that makes recording the data more difficult. Then the researcher types the data into the computer in matrix form where rows represent respondents and columns represent triads.

A data matrix for 5 subjects on a lambda-2 design for 9 items (which makes 24 triads)=

```
3 1 1 2 1 3 2 1 2 3 1 2 1 2 1 2 1 3 2 1 3 3 1 2
3 1 2 3 3 1 2 2 3 3 1 3 3 2 3 2 2 2 3 1 1 2 3 2
3 2 1 2 1 3 2 1 2 3 1 2 1 2 1 2 1 3 2 1 3 3 1 2
2 1 1 3 3 1 2 2 3 3 1 3 3 2 3 2 2 2 3 1 1 2 3 2
3 2 2 3 3 1 2 2 3 3 1 3 3 2 3 2 2 2 3 1 1 2 3 2
```

STEP 4. Create an item-by-item similarity matrix by counting up the number (or proportion) of times that each pair of items was judged similar. For example, in the first triad above, "dog" and "seal" get one vote towards their similarity if the respondent chooses "mosquito" as most different. Counting is performed across all triads and all respondents to get a single similarity matrix representing the total number of times any person judged a given pair of items the same relative to any third item.

STEP 5. Submit the similarity matrix to a clustering or multidimensional scaling routine in order to reveal clusters and/or underlying dimensions organizing the domain.



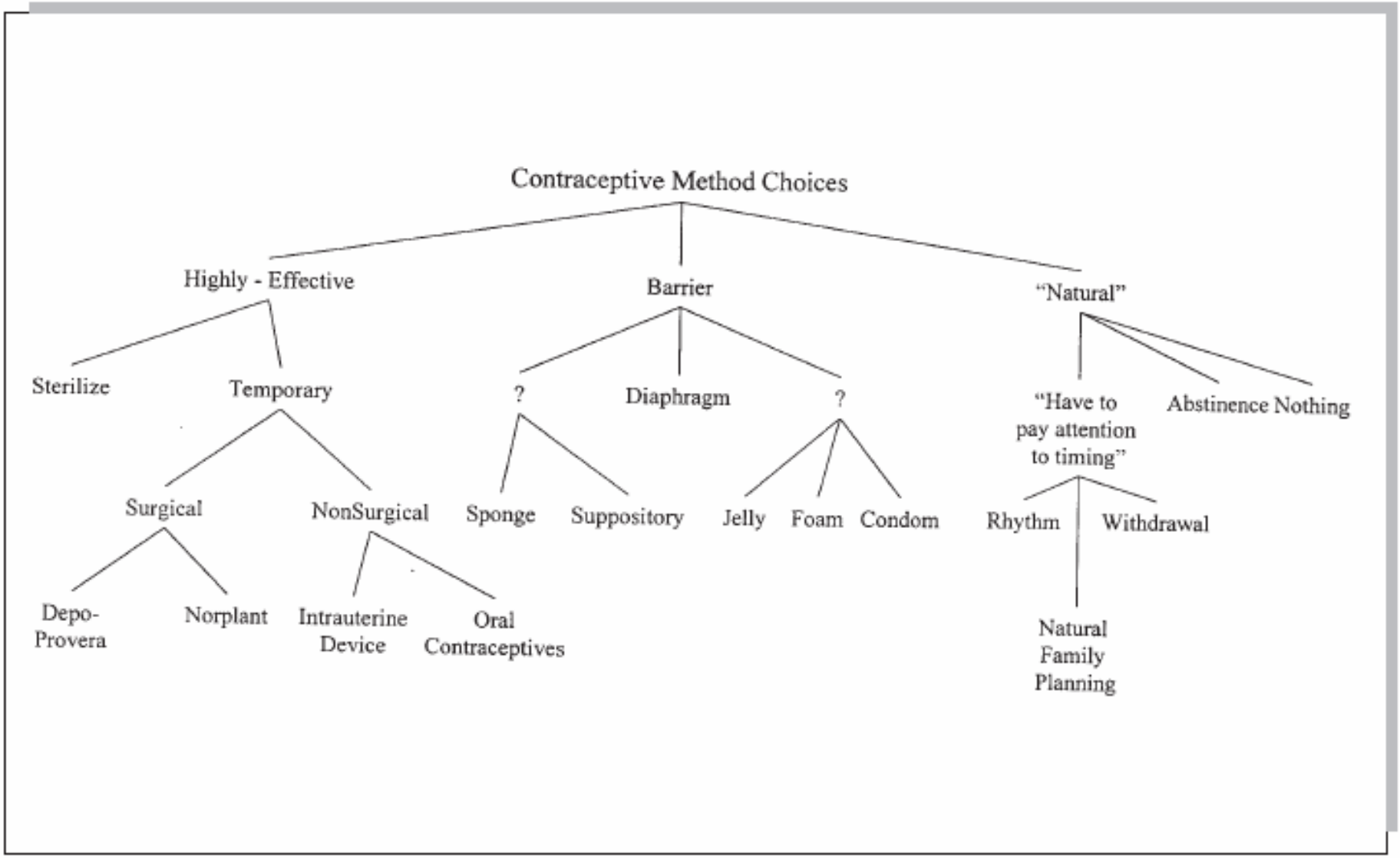
PILE SORTS

Collecting Pile Sorts

- List Items on Index Cards
- Pull out any informant is not familiar with
- Sort into piles according to ‘how similar’ they are
  - Definition of ‘similar’ is from the viewpoint of respondent
- May duplicate & put item in more than one pile

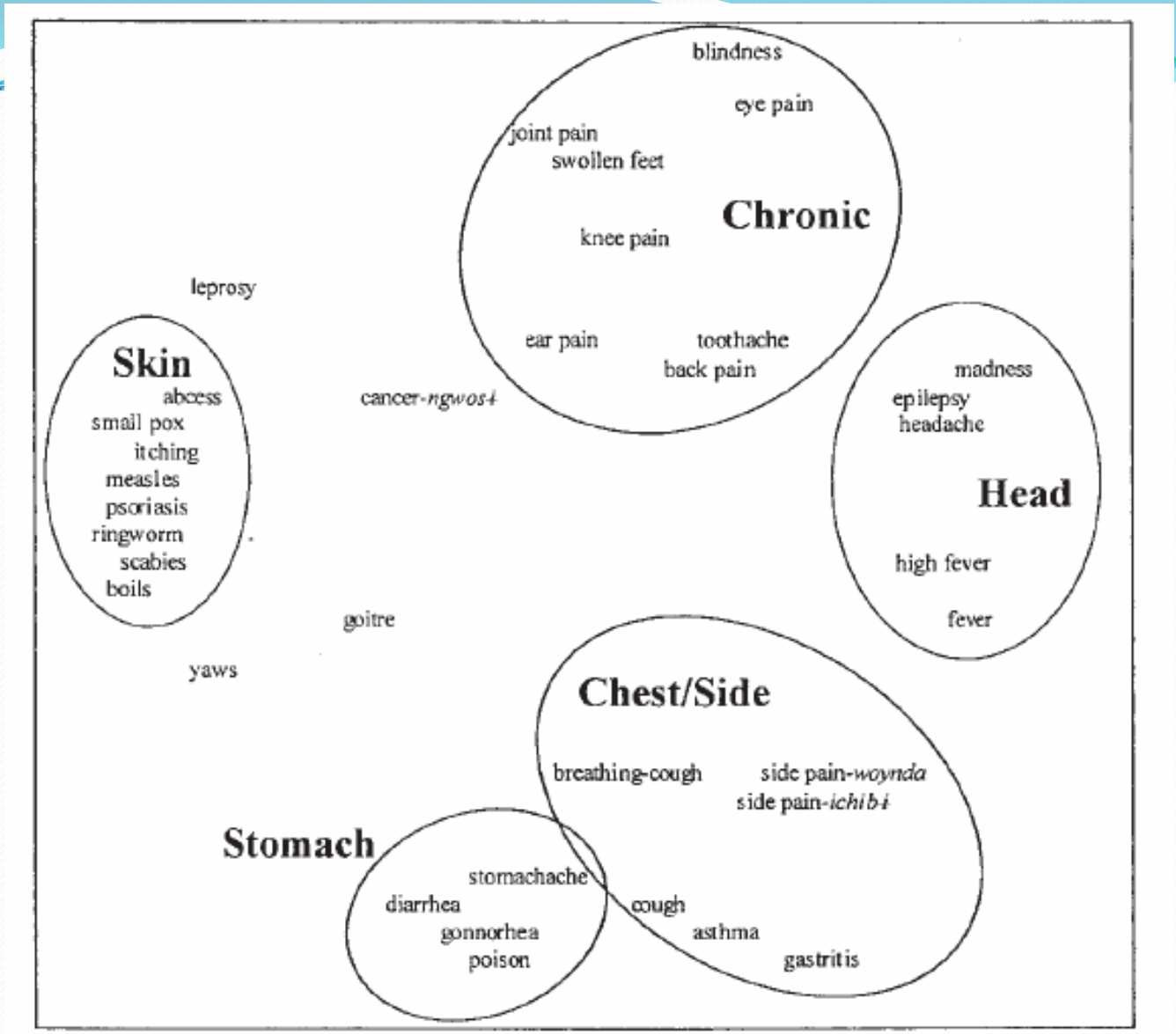
Analyzing Pile Sorts

	Frog	Salamander	Beaver	Raccoon	Rabbit	Mouse	Coyote	Deer	Moose
Frog	100	96	6	2	2	0	0	2	2
Salamander	96	100	4	0	0	2	0	0	0
Beaver	6	4	100	62	65	56	17	25	13
Raccoon	2	0	62	100	71	58	23	29	15
Rabbit	2	0	65	71	100	75	17	27	15
Mouse	0	2	56	58	75	100	17	15	10
Coyote	0	0	17	23	17	17	100	21	15
Deer	2	0	25	29	27	15	21	100	77
Moose	2	0	13	15	15	10	15	17	100



Source: Fig 29.2, Ryan & Bernard, Handbook of Qualitative Research, 2nd ed. Denzin & Lincoln eds. Sage Publications

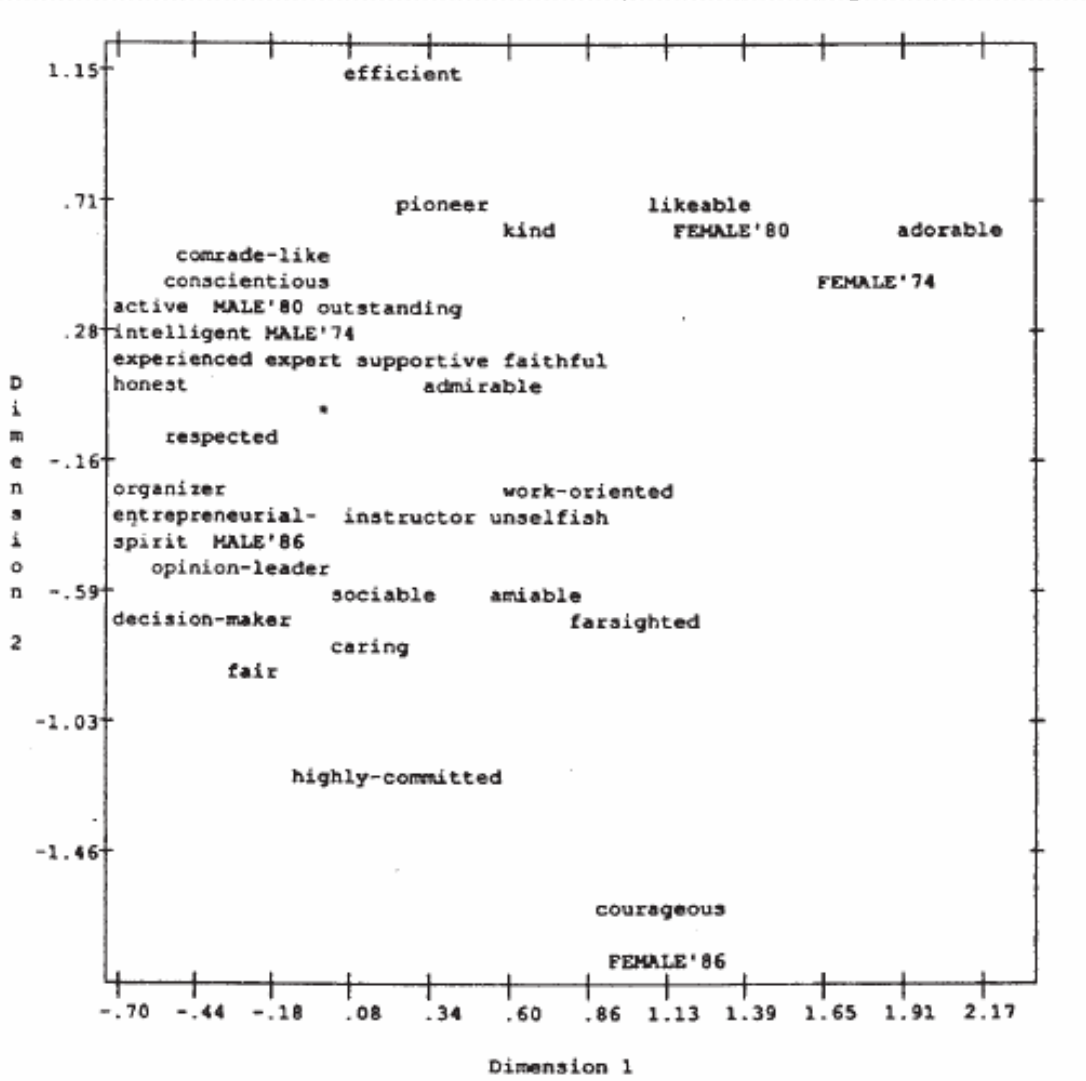
Sample MDS & Correspondence Analysis Outputs



Source: Fig 29.3, Ryan & Bernard, Handbook of Qualitative Research, 2nd ed. Denzin & Lincoln eds. Sage Publications

Correspondence Analysis of Frequencies of 31 Disruptive Obituary Categories by Gender & Year of Publication

(©Kirchler, E. Eur J Social Psych 22 (1992), p 371)



Source: Fig 29.4, Ryan & Bernard, Handbook of Qualitative Research, 2nd ed. Denzin & Lincoln eds. Sage Publications

Multidimensional Scaling of Informants based on words used in descriptions of Horror Films

Source: Fig 29.5, Ryan & Bernard, Handbook of Qualitative Research, 2nd ed. Denzin & Lincoln eds. Sage Publications; Based on data in Nolan & Ryan 1999.

Some Resources for Cultural Domain Analysis

- Spradley J. 1979 The ethnographic interview. NY; Holt, Rinehart & Winston
- Weller SC & AK Romney 1988 Systematic Data Collection. Newbury Park, Sage Publications
- FIELD METHODS (formerly Cultural Anthropology Methods)
  - http://www.qualquant.net/FM/
- ANTHROPAC
  - http://www.analytictech.com/anthropac/apacdesc.htm