Universidade Federal do Rio Grande do Sul
Instituto de Matemática
Cadernos de Matemática e Estatística
Série F: Trabalho de Divulgação

Correspondence Analysis applied to Ethnographic Data: Case Examples

Jandyra M. G. Fachel
Ondina F. Leal
Ceres Victora
Cristina Freitag
Denise Jardim
Leandro Saraiva

Série F, nº 03, JAN/92
Porto Alegre, janeiro de 1992

INTRODUCTION

The purpose of this paper is to show how to combine anthropological analysis with statistical analysis using qualitative and multivariate data. The combination of ethnographic date with multiple correspondence analysis should enable anthropologists to refine the systematization of their descriptive material. Also, the use of non conventional investigation techniques, such as drawings of the body, or attitude scales structured from ethnographic and folklore material are innovative aproaches.

The first two cases which follow, were part of an extensive etnographic research about the social representations of the body and its sensations, and attitudes concerning disease and health in very low income groups in Southern Brazil. The third case is a methodological example. Although the specific percentage values are fictitious, they have been based on etnoghraphic material and the associations that came out of the graph correspond to what has been observed during field work.

CASE 1: Correspondence Analysis of the Children Body Map Data:

Variants of the Children's Perceptions of their own Body. Plane Diagram of 1st and 2nd axes of Inertia: Space of the Body Representations.

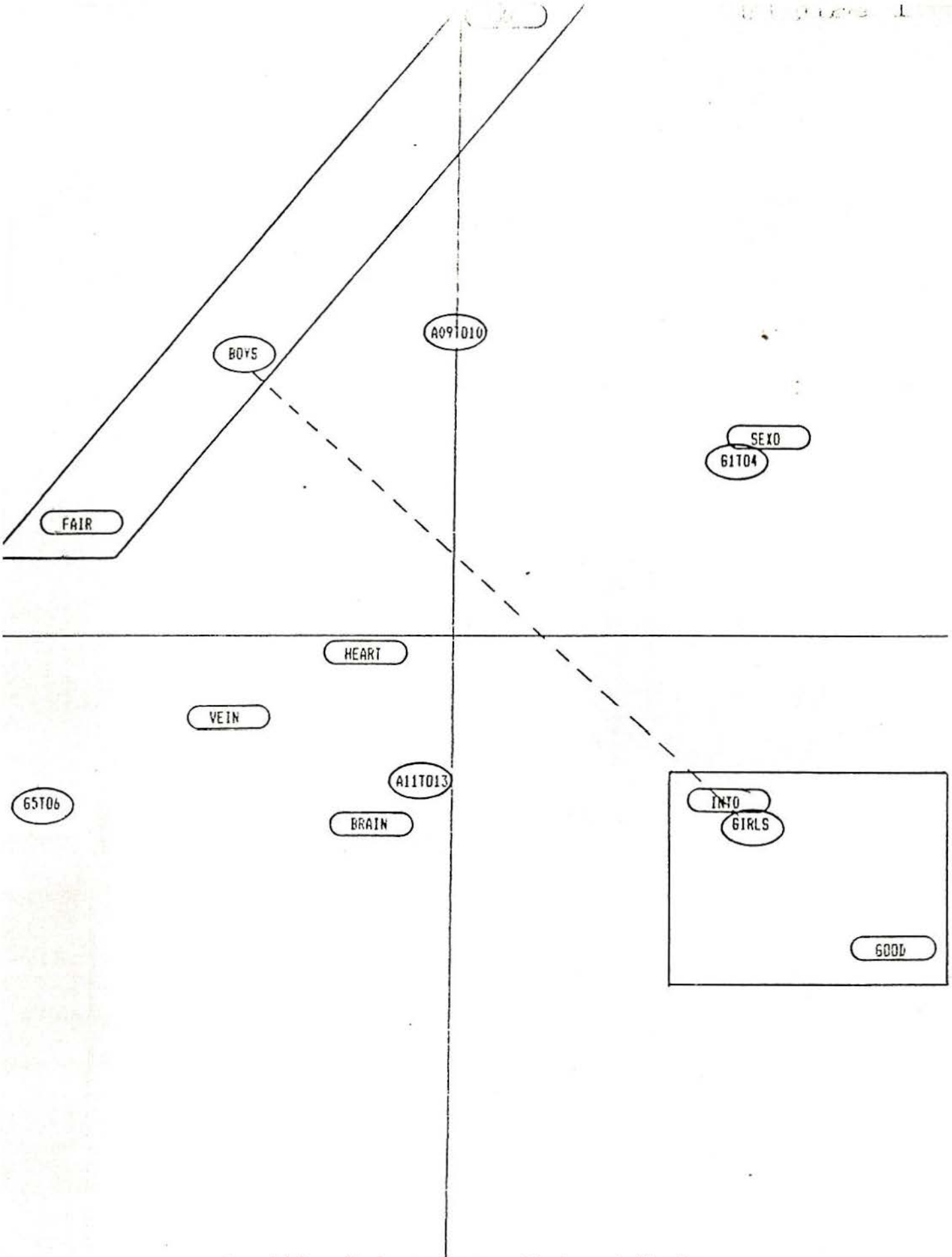
We asked children of an after-school program in a low-income community to draw their own body with its principal organs. The children had been taught an anatomy class by a medical doctor, a few weeks earlier. The drawings were classified by the researchers in "good" (a map where the main organs were indicated and shows good knowledge of the size and position of these organs); "fair" (not so good, fewer organs); and "poor" body maps, always having as parameter a standard anatomical figure.

The correspondence analysis graphs (figure I) shows a strong association of female children with a reproduction of the body considered "good", and it indicates also that girls have an internal perception of their bodies, demonstrated by the specification of various internal organs. Male children drawings are associated to "poor" and "fair" maps. The other correspondence indicated by the multivariate analysis graphs shows that younger children and at lower school grades, independently of gender, made poorer body reproductions, with emphasis on sexual organs. Brain, blood vessels (vein), and heart appeared more frequently in the maps drawn by the children with higher levels of schooling and by children of 11 to 13 years old, than by the younger ones (9 to 10 years old); it was also observed a meaningful higher frequency of these organs in the girls' maps than in the boys' maps.

Although no part of the body was suggested by the researcher,

detailed blood vessels were represented in most drawings. This confirms other verbal representations of the body (among working class adults) which indicate the notion of body fluids as essential to their conceptions of body functioning.

From a statistical point of view we have that the two principal axes explain almost 100% of the total inertia. The first axis, the vertical one, explain 91% of the total inertia. We can observe that sex group separation take place along a diagonal axis.



variance totale expliquée par ces axes Dépendance significative.

Verticalement: Axe N° 1 (91 %)

Horizontalement: Axe N° 2 (8 %)

Dprt Stat. PORTO ALEGRE

	BOYS	GIRLS	A091010	A11T013	G1T04	G5T06
VEIN	57	100	40	100	50	100
INTO	14	80	20	57	33	50
HEART	57	100	60	88	50	83
SEXO	88	80	100	71	100	36
BRAIN	28	100	20	88	33	83
GOOD	0	60	0	43	17	33
FAIR	57	40	60	57	50	66
POUR	43	0	40	0	33	0

UFRGS Dprt Stat. PORTO ALEGRE

LIFIANCAS(Contributions relatives)

	AxeN° 1 (91%)	AxeN° 2 (8%)	CONTRIB ABSOLUES
VEIN	3%-	14%+	3%
INTO	6%-	11%-	7%
HEART	0%-	2%+	0%
SEXO	18%+	28%-	18%
BRAIN	12%-	1%+	11%
GOOD	16%-	17%-	15%
FAIR	4%+	27%+	5%
POUR	42%+	0%-	38%
BOYS	24%+	9%+	22%
GIRLS	18%-	29%-	18%
A091010	28%+	0%-	25%
A11T013	10%-	0%+	9%
G1T04	9%+	18%-	9%
G5T06	12%-	45%+	14%

Pourcentages arrondis à 1 entier

UFRGS Dprt Stat. PORTO ALEGRE

Lines Variables

VEIN: the child drew blood vessels
INTO: the child drew internal organs
HEART: the child drew the heart
SEXO: the child drew the sexual organs
BRAIN: the child drew the brain
GOOD: the child made a good representation of the body
FAIR: the child made a fair, reasonable representation of the body
POOR: the child made a poor representation of the body

Column Variables

BOYS: male children
GIRLS: female children
A09T010: age 9 to 10 years old
A11T013: age 11 to 13 years old
G1T04: 1st to 4th grade (Elementary School)
G5T06: 5th to 6th grade

CASE 2: Correspondence Analysis of the Women's Map of the Reproductive System Data:

Variants of the Women's Perceptions of their own Reproductive System. Plane Diagram of 1st and 2nd axes of inertia: Space of Female Body Representations.

The interviewer, a female anthropologist with previous long-term work with this specific group of women who live in an extremely poor area, asked them to draw their own reproductive system[1]. The reproductive system maps were classified by the researchers in relation to its similitude with the standard medical model of the reproductive system, as SIMMOD (similar to the medical model), DIFMOD (model different from the medical one), MOD+ (similar to the medical model plus other body organs, such as the heart), VAGY (the vagina is represented), VAGN (the vagina is not represented).

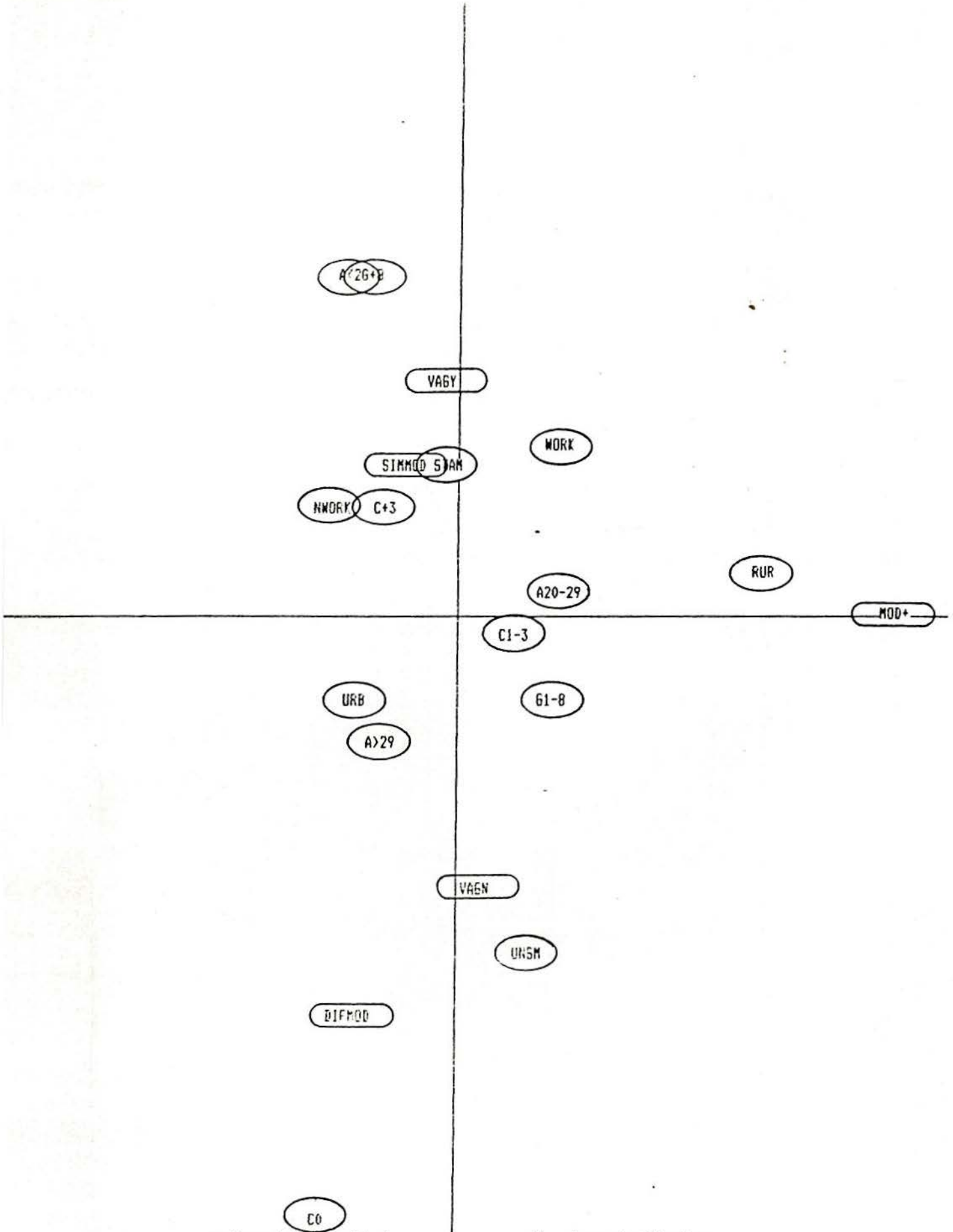
Previous ethnographic data indicated that the women with higher number of children tend to look for medical resources more frequently than the women with less children. The correspondence analysis graph (Figure II) confirms this association between "women with more than three children" (C+3) and higher knowledge of official medicine, when the graph aggregates in the same space: "women with more than three children" (C+3), a relatively stable marriage (STA-M) and the category SIMMOD ("body map more similar to the medical model").

The graph (Figure II) indicates a line that describes the

[1] The anthropologist employed other words.

increasing numbers of children, which takes the same direction of the line that describes the proximity to the medical knowledge, that is, the number of children has a direct relationship with the interaction of these women with medical services.

At this point, having the ethnographic experience as background, and having this correspondence analysis graph as a model, we are able to formulate a few hypotheses: 1. There is a specific working-class logic which assimilates general knowledge about health from the official medicine, but maintains a relatively high rate of natality; 2. "Stable marriage" in this group actually means a current stable marriage (defined by the male presence as a "provider" and by having a child together) not a long term marriage. If so, a variable such as "number of years with the present partner" or "number of children with the current husband" has to be added to a further correspondence analysis to verify this kind of observational information. At this point we will be able to consider one of the main hypothesis of this research project, that is, the descendent kin is a defining element to an actual marital alliance and/or a female strategy to engage man in structuring a family.



variance totale expliquée par ces axes Dépendance significative.

Verticalement: Axe N° 1 (66 %)

Horizontalement: Axe N° 2 (21 %)

	RUR	URB	A<20	A20-29	A>29	STAM	UNSM	G1-8	G+8	C0	C1-3	C+3	WORK	NWORK
SIMMOD	50	66	100	43	80	75	40	50	100	0	60	100	57	66
DIFMOD	0	33	0	28	20	12	40	25	0	100	20	0	14	33
MOD+	50	0	0	28	0	12	20	25	0	0	20	0	28	0
VAGY	50	44	100	71	20	75	0	41	100	0	50	50	86	83
VAGN	50	55	0	28	80	25	100	59	0	100	50	50	14	17

UFRGS Dprt Stat. PORTO ALEGRE

FEMALE REPRODUCTIVE SYSTEM MAP(Contributions relatives)

	AxeN° 1 (66%)	AxeN° 2 (21%)	CONTRIB ABSOLUES
RUR	0%+	44%-	9%
URB	1%-	5%+	2%
A<20	14%+	5%+	10%
A20-29	0%+	5%-	1%
A>29	2%-	3%+	2%
STAM	3%+	0%+	2%
UNSM	14%-	2%-	10%
G1-8	1%-	5%-	2%
G+8	14%+	5%+	10%
C0	46%-	8%+	32%
C1-3	0%-	2%-	0%
C+3	1%+	3%+	2%
WORK	3%+	5%-	3%
NWORK	1%+	7%+	3%
SIMMOD	12%+	6%+	9%
DIFMOD	33%-	9%+	23%
MOD+	0%-	84%-	18%
VAGY	25%+	1%+	16%
VAGN	30%-	1%-	20%

Pourcentages arrondis à l'entier

UFRGS Dprt Stat. PORTO ALEGRE

Line Variables

SIMMOD: similar to the medical model
DIFMOD: model different from the medical one
MOD+: similar to the medical model plus other body organs, such as the heart
VAGY: vagina yes, the vagina is represented
VAGN: vagina no, the vagina is not represented.

Column Variables

RUR: from rural origin
URB: from urban origin

A<20: age group
A20-29
A>29

STAM: stable marriage
UNSM: unstable marriage

G1-8: last school grade attended
G+8

C0: has no children
C1-3: has from 1 to 3 children
C+3: has more than 3 children

WORK: she works out
NWORK: she does not works out

CASE 3: ; Correspondence Analysis of Pain, Illness and Cure Resources Data:

Variants of a population perception of pain and illness and the medical diagnosis. Plane Diagram of 1st and 2nd axes of inertia: Space of Pain, illness and Diagnosis.

This example intends to be just a methodological example of how the anthropological information can be combined with statistical correspondence analysis. Guided by the ethnographic material and field experience we have about a low-income population, where preventive community medical services are available, using also the medical records as a source of information, we projected fictitious percentage values to be able to built a table of contingencies and the correspondence analysis graph. At this point, our figures can only indicate hypotheses to further research and hints to what is crucial to be systematized in the research procedures. Our intention with this example is to indicate that, having the relevant specific data, methodologically we are able to associate the people's "world view" with their decisions about health resources and reproduction.

The correspondence analysis graph (Figure III) shows gender as the variable that ordinates the vertical axis of inertia, male and female being localized in opposite poles along this axis, as main signifiers for the entire space of significance of pain, illness and cure resources.

In this graph, the variables being considered are: Frequency to medical consultation by year F>5X (more than 5 times/year) and F<5X (less than 5 times by year); the "domain" which is identified by the patient as the cause of his illness, categorized in "home",

Taking into consideration the category age over 60 years old (A>60) and less than 15 years old (A<15), we observe a spatial disposition of points that concentrates the multilocalized pain (MULTILOC), high frequency of medical appointments (F>5X), and the patient's disease rationalia related to the street (STREET). Home, for young people and old persons seems to be a safe place. The point here is that the representation of the sickness and its cure are close to each other.

The variable medical diagnosis with the categories: Hypertension (HYPE), Alcoholism (ALCO), Cardiopathy (CARDIO), Anxiety (ANSI) and Acute Situation (ACUTE), was also cross tabulated with the frequent medical appointment, cause of the illness, kind of pain and cure resources. The analysis shows that as the male-female define a polarity in the first axis, the categories acute condition (ACUTE) and anxiety (ANSI) reproduce this polarity, having the same profile as the gender variables: there is a clear homology between women/anxiety and men/acute cases. This let us interprete the relationship of these categories with the line variables, in the same way people with anxiety problems have a high frequency of medical appointments and multilocalized pain.

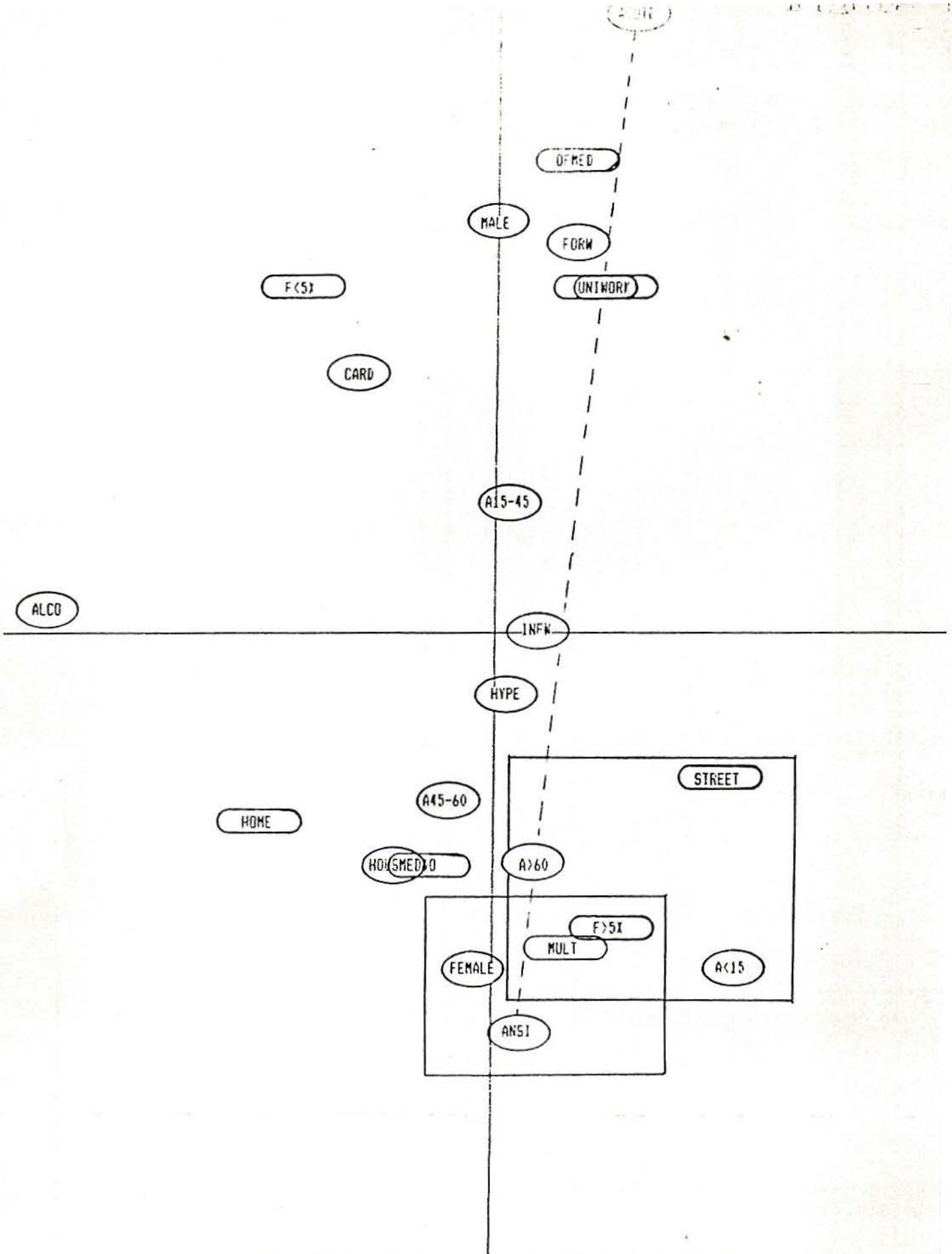
The use of multivariate correspondence analysis that conveys this kind of data allows us to understand the representations of diseases, diagnosis and their cure, associating it to gender, age and other demographic variables. Through the use of the ethnographic material combined with the multivariate analysis we are able to draw cognitive maps of working-class world view.

"work" and "street" (meaning by "street", everything that is not related with none of the others two, e.g. the children got a cold because they went out of the house, sexual transmitted diseases are also identified as being caught on the "street"); kind of pain ("multilocalized" and "unilocalized": these categories try to indicate if the patient pain sensation is specific or if it is general, difficult to define, "a pain that moves"); cure resources the patient usually searches for, that he(she) uses excluqively the official medicine (OFMED) or he(she) also searches for traditional and/or religious healers (MED&O).

Associated to the male-gender pole in the graph there is a concentration of the variables low medical consultation (F<5X); WORK as the domain to which the cause of the illness is associated; the kind of pain the patient relates is specific and has a central focus (UNILOC); and there is an emphasis in searching only official medicine recourse (OFMED).

On the other hand, around the female-gender pole in the graph are concentrated variables which indicate a high frequency of medical consultations (F>5X), among the most common causes of the illness, the patient indicates the domestic domain (HOME); the kind of pain is difficult to define, the patient has multi-complains (MULTILOC) and with this kind of patient we also observe that he/she searches for cure resources of various kinds (MED&O).

Having in mind this correspondence analysis, it is possible to say that the male and female behavior are not only differentiated, but opposite concerning representation and practices related to disease and its cure.



variance totale expliquée par ces axes Dépendance significative.

Verticalement: Axe N° 1 (42 %)

Horizontalement: Axe N° 2 (24 %)

UFRGS Dprt Stat. PORTO ALEGRE

	Axe N° 1 (42%)	Axe N° 2 (24%)	CONTRIB ABSOLUES
MALE	12%+	0%+	5%
FEMALE	9%-	0%+	4%
A<15	9%-	22%-	9%
A15-45	1%+	0%-	0%
A45-60	2%-	1%+	1%
A>60	4%-	1%-	2%
FORW	11%+	3%-	6%
IUFW	0%+	1%-	0%
HOUSEWIFE	5%-	3%+	3%
ALCO	0%+	55%+	13%
ANSI	12%-	0%-	5%
HYPE	0%-	0%-	0%
CARD	5%+	7%+	4%
ACUTE	29%+	7%-	14%
F<5X	15%+	23%+	11%
F>5X	13%-	10%-	7%
HOME	5%-	32%+	10%
WORK	11%+	4%-	6%
STREET	1%-	13%-	4%
MULT	13%-	4%-	7%
UNIL	14%+	6%-	7%
OFMED	19%+	2%-	9%
MED&O	9%-	5%+	5%

Pourcentages arrondis à l'entier

FRGS Dprt Stat. PORTO ALEGRE

	MALE	FEMA	A<15	A15-45	A45-60	A>60	FORK	INFW	MMUS	ALCO	ANSI	HYPE	CARD	ACUT
F<5I	90	20	10	50	40	30	60	40	40	95	20	20	80	70
F>5I	10	80	90	50	60	70	40	60	60	5	80	80	20	30
HOME	10	65	15	25	55	40	20	30	80	80	50	80	80	10
WORK	65	20	5	70	35	40	70	60	0	10	30	10	10	60
STREET	25	15	80	5	10	20	10	10	20	10	20	10	10	30
MULT	20	70	60	30	60	80	40	50	60	0	90	60	40	10
UNIL	80	30	40	70	40	20	60	50	40	0	10	40	60	90
OFMED	30	10	20	10	10	20	70	20	20	10	10	70	70	90
MED&O	70	90	80	90	90	80	30	80	80	90	90	30	30	10

UFRGS Dprt Stat. PORTO ALEGRE

REFERÊNCIAS BIBLIOGRÁFICAS

1. BENZÉCRI, J. P. (1973). Analyse des donées, tome II: Analyse des correspondendes. Dunod, Paris.

2. CALLEGARI-JAQUES, S. M. (1992). Análise de Correspondência: Aplicação. Cadernos de Matemática e Estatística , Série F, número 02, dezembro 1991, UFRGS, Porto Alegre.

3. FISHER,R.A. (1940). The precision of discriminant functions. Ann Eugen. Lond.10:422-429.

4. GIFI, A. (1981). Nonlinear multivariate analysis. University of Leiden, Afdeling Datatheorie, Leiden, The Nederlands.

5. GREENACRE, M.J. (1981). Theory and applications of correspondence analysis. Academic Press, New York.

6. HIRSCHFELD,D.O. (1935). A connection between correlation and contingency. Proc. Camb..Phil.Soc. 31: 520-524.

7. HILL,M.O. (1974). Correspondence analysis: a neglected multivariate method. Appl.Statist. 23: 340-354.

8. JAMES,F.C. & McCULLOCH,C.E. (1990). Multyvariate analysis in Ecology and Systematics: panacea or Pandora's box. Annual Rev.Ecol. Syst. 21: 129-166.

9. LEBART,L. & FENELON,J.P. (1971). Statistique et informatique appliquées. Dunod, Paris.

10. LEBART,L.;MORINEAU,A. & TABARD,W. (1977). Techniques de la description statistique. Dunod, Paris.

11. LEBART,L.,MORINEAU,A. & WARWICK, K.M. (1984). Multivariate descriptive analysis: correspondence analysis and related techinques for large matrices. Wiley-Intercience, New york.

12. MOSER,E.B. (1989). Exploring contingency tables with correspondence analysis. Computer Applications in Biosciences, 5: 183-189.

13. NISHISATOP,S. (1980). Analysis of categorical data: dual scalling and its aplications. University of Toronto Press, Toronto.

14. SOUZA,A.M.R. (1982). Análise de correspondência. Dissertação de Mestrado, Instituto de Matemática e Estatística, Universidade de São Paulo.

15. SOUZA,N.M. (1990). Análise de correspondência. Monografia, Bacharelado em Estatística, Universidade Federal do Ro Grande do Sul.

16. VERDINELLI,M.A. (1980). Análise inercial em Ecologia. Tese Dout., Instituto Oceanográfico, Universidade de São Paulo, São Paulo.

Publicações do Instituto de Matemática da UFRGS
Cadernos de Matemática e Estatísitca

Série F: Trabalho de Divulgação

1. Nubem A. C. Medeiros e Jaime B. Ripoll - Superfícies Invariantes - MAI/90.
2. Sídia Maria Calegari Jacques - Análise de Correspondência - aplicações - DEZ/91
3. Jandyra M. G. Fachel, Ondina F. Leal, Ceres Victora, Cristina Freitag, Denise Jardim e Leandro Silva - Correspondence Analysis applied to ethnografic data: Case Examples. - JAN/92

Universidade Federal do Rio Grande Sul
Reitor: Professor Tuiskon Dick

Instituto de Matemática
Diretor: Professor Aron Taitelbaum
Núcleo de Atividades Extra Curriculares
Coordenador: Professor Jaime Bruck Ripoll
Secretária: Faraildes Beatriz da Silva

Os Cadernos de Matemática e Estatística publicam as seguintes séries:
Série A: Trabalho de Pesquisa
Série B: Trabalho de Apoio Didático
Série C: Colóquio de Matemática SBM/UFRGS
Série D: Trabalho de Graduação
Série E: Dissertações de Mestrado
Série F: Trabalho de Divulgação
Série G: Textos para Discussão

Toda correspondência com solicitação de números publicados e demais informações deverá ser enviada para:
NAEC - Núcleo de Atividades Extra Curriculares
Instituto de Matemática - UFRGS
Av. Bento Gonçalves, 9500
91.500 - Agronomia - POA/RS
Telefone: 336.98.22 ou 339.13.55 Ramal: 6176