Trends in nutrition education of children aged 5-11 y in the Fleurbaix-Laventie Ville-Santé study

Vincent Vigneron[1,2], Hichem Maaref[2], and A. Streri[3]

[1] CES SAMOS-MATISSE 8095, 90 rue de Tolbiac, 75634 Paris cedex 13
vigneron@univ-paris1.fr
[2] IBISC, CNRS FRE 2873, 91020 Evry-Courcouronnes, France
vincent.vigneron,hichem.maaref@ibisc.univ-evry.fr
[3] Laboratoire Cognition et Développement UMR-8605 (Univ. Paris V)
71, avenue Edouard Vaillant
92774 Boulogne-Billancourt Cedex, France
streri@psycho.univ-paris5.fr

Abstract. The “Fleurbaix-Laventie Ville Santé” survey is an epidemiological study on the relationships between nutrition and health in 2 small towns of northern France: Fleurbaix and Lavantie. The objective of this study was to understand and improve the relationship between children aged between 5-11 years and food. A second aim of this work is to evaluate the effectiveness of nutrition education programs through the drawing of the children, and whether it have a significant impact on physical and cognitive development. We outline some of the characteristics of this layout. A tool was conceived which is able to automatically extract low-level and high-level primitives from the drawing structuring which allow us to categorize main comportments. These descriptors could help to develop nutrition education programs to improve the eating habits of the children and reduce health risks.

1 The Fleurbaix-Laventie Ville-Santé study

Childhood obesity in America and recently in Europe is a growing problem. In Michigan, for instance, 1 on 4 children is obese [1]. These trends point on the need for understanding and improving the relationship between children and food. The Fleurbaix-Laventie Ville Santé study [2,3] studies the impact of physical activity, environmental and genetic factors on possible cardio-vascular risks as well as weight gain. It has been conducted on the parents, adolescents and children of 300 families in the North of France. Under the patronage of the Departement of Education and the Departement of Agriculture, Fisheries and Food, the survey promotes a better understanding of the links between nutrition and health and aims to establish efficient strategies of *nutritional prevention*. The purposes of this study were to improve the health status by identification, through group discussions, of healthy eating habits in young children. Study outcomes provided important data on the effects of a nutrition education program on attitudes, knowledge, practices, and interest in nutrition change. The survey demonstrates the effectiveness of nutrition education of children at school in order to appreciably correct the food disorders of the entire family, in particular a wiser concerning choice of the fats.

This observation of weight and food habits shall allow: (*i*) to better understand the mechanisms involved in excessive weight gain and in the distribution of the body fats, (*ii*) to follow the natural development in the choice of food and of sugars in particular, made by a population that is well informed on nutrition, (*iii*) to better understand the needs, the tastes and the food habits of the children, in order to offer and develop products that combine health and taste pleasures.

The instruments and measures used to collect data includes pre-existing instruments already used by Fleurbaix-Lavantie Ville-Santé and are obtained from *child’s drawings*. The record review includes the following: child health history, food preferences, height, weight, handwriting answers. Handwriting is only in this study a complementary way of expression for children, they are supposed to be familiarized with the use of the drawing tool.

The interest in children’s drawing is to make them able to focus their attention on the symbolic dimension of their production. Therefore the experimental protocol contains an exercise: “ *Draw a child who is eating well and a child who is not*”, for which nutrition apprenticeship, which begins early at the nursery and at home, and continues at primary school, interfere with personal preferences. However, no objectif tool allow to identify the key psychological factors influencing the conflicting images nor children's food choice.

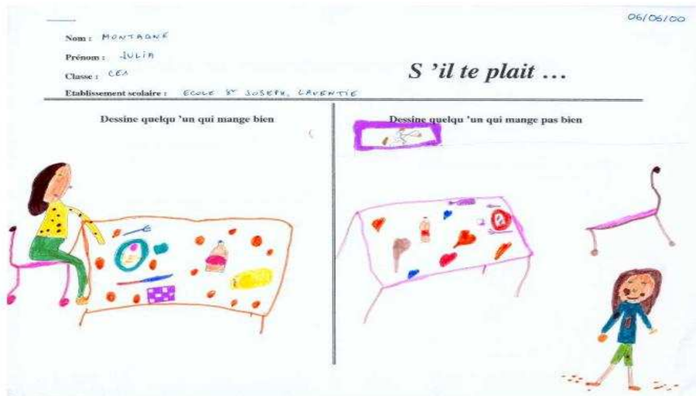

Fig. 1. Child draw on the thema "the child who does/does not eat well".

The children drawings include for example issues like healthy eating, nutrition education, parenting and food, TV, presence of commercial marks, eating disorders, obesity, etc. Information from the focus groups will be used to develop nutrition education programs that empower families to improve their eating habits and reduce health risks, like growth deficiencies, for their children.
The remainder of this paper is organised as follows. Section 1 describes the material of this work. Section 2 outlines the characteristics of this layout, that we call *descriptors* encoding the graphic activity of the child. Section 4.2 presents the clustering process and results obtained in terms of food habits.

2 Material

The test of drawing The examination comprised measurements of height, weight and waist (BMI). The analysis was retricted to 434 children aged between 5 and 11 y (216 girls, 218 boys). In the test of drawing, the child has to reproduce two drawings labelled as "a child who eat well" and "a child who does not eat well" as illustrated in Fig. 1. The model has to be copied on a sheet whithout any reference mark and no realization time requirement. The production conditions are the same in the two drawings. Features describing the space organization (*e.g.* colors, height, width, surface of the drawings, etc.) are retained. The extraction of symbolic features related to the image (high level descriptors) is described in the next section.

The test of comments This test consists of the child oral or written opinion (depending of his age). Some characteristics of his expression would present significant variation with the drawing.

Coding Qualitative variables occupy a significant place in data ananlysis. Several questions arise when we try to deal with them: which coding of the variables to adopt? which distance to use for the variables and for the individuals ?
There are many aspects of categorical problems. We consider the situation in which N drawings "answer" to Q questions. Each question has a finite number of *modalities* and is answered by only one modality. The

purpose here is to *see* the relations between the modalities and to project them onto a *factorial subspace*. Qualitative variable are encoded the following way: when a feature appears in both drawings, the associated variable takes the modality ’1’, ’4’ when it is absent from both, ’2’ when present in the bad-eating representation and ’3’ when present in good-eating representation. ’6’ is the uncertainty indicator.

Statistics The Fleurbaix-Laventie Ville-Santé study concerned all the children enrolled in the last section of preschool and in primary school (all grades), which amounted to 827 children [4]. These children belonged to 579 different families and were asked to participate in clinical and biological examinations since 1992. The analysis was retricted to 434 children aged between 5 and 11 y (216 girls, 218 boys), as depicted in Tab. 1.

Table 1. Child age distribution.

	Group 1 < 6y	Group 2 ≥ 6y and < 9y	Group 3 ≥ 9y	Group 1+2+3
Nbr of childs	118	206	110	434
girls	52,5%	50,0%	45,5%	49,5%

The analysis was performed using R packages for data analysis of qualitative variables and graphics (R is a language for data analysis and graphics). The main biological characteristics of the children are collected in Tab. 2. Their mean age was 8 y. Boys and girls were similar. The children have been divided arbitrarilty in 3 groups, *i.e.* children less than 6 y., between 6-9 y., more than 9 y.

Table 2. Description of the population. (Means ± s.d. or 95% Cl.)

	Boys	Girls	P^a
number	251	223	
age	$7,9 \pm 1,7$	$7,7 \pm 1,5$	0,13
weight (kg)	$25,9 \pm 6,5$	$25,6 \pm 6,8$	0,50
Height (m)	$1,27 \pm 0,10$	$1,26 \pm 0,11$	$0,1$
waist to hip ratio	$0,865 \pm 0.04$	$0,831 \pm 0.049$	0.0001

[a] *P*-value for comparison between boys and girls with a mixed model taking into account familial relationships.

3 Data analysis

3.1 Material

The selection of an appropriate set of *low-level descriptors* which take into account the difficulties present in the extraction or selection process, and at the same time result in acceptable performance, is one of the most difficult tasks in the desig of pattern recognition systems. To facilitate the data analysis in our case, features will be classified into 3 categories: (*i*) *physical* features, (*ii*) *structural* features and (*iii*) *mathematical* features.

a)- Physical and structural features are detected by the eye. Shape, drawing composition and other geometrical properties of the drawing are considered as structural features. Colour is an example of physical feature. Although it could be argued that structural features are also physical features, the reader should keep in mind that the distinctions being made here are simply for the purpose of convenience and that they are in some sense, arbitrary (see *e.g.* Tou and Gonzalez [5]).

b)- Weight, height or other biological measures constitute our mathematical features.

Structural and physical features are strongly problem oriented in the sense that their use involves the development of specialized know-how. For instance, if one were trying to show disorder inherent to a drawing, abundance of objects on the table would be meaningful. The important point to keep in mind is that it is almost impossible to formulate general guidelines regarding the selection of physical and structural features.

3.2 Low-level descriptors

Attention is focused on the selection and extraction of the structural and physical features to determine certain invariant attributes summerized in the following table:

Drawing composition		**People**		**Food**		**Comments**	
table	TABLE	disorder	DISORD	food P/NP	FOOPNP[a]	intellectual perf.	INTP
food only	FOONLY	P/A child	CHLDPNP	food symbol	SYMB	physical perf.	PHYP
colours	COLOR	dimorphism	DIMORP	food variety	VARIET	vital value	VITAL
				milk products;	MIPRO	health state	HEST
				meat,fish/eggs	MEFIEG	perception	FEEL
				feculents	FECUL		
				drinks	NADRNK		
				oil,greases	OIL		
				vegetables,fruits	VEGFRU		

[a] P/NP=present or not present.

Table 3. 23 items grouped in 4 main set of variables.

An overall number of 23 qualitative features are extracted:

A. composition of the drawing
- presence of the *table*
- presence of *food* when child, table or other elements are not
- use of *colours* by the child

B. people
- *disordered*: disorder being present in the drawing
- *child present or not* in the drawing
- *dimorphism* : the fact of having a distinct pace from normal child

C. food
- Food representation
  - food *being present or not*
  - food materialized with *symbols* (eventually commented)
  - food *variaty* : diversity of food present in the drawing (at minus 3)
- elements of the meal
  - *milk based* products (cheese, milk,. . .)
  - *meat,fish or eggs*
  - *feculents*: potatoes, chips, pasta chute,. . .
  - *drinks* : non alcoolic drinks
  - *sugar* : cakes, chocolate, sweets,. . .
  - *alimentary greases* : oil, mayonnaise,. . .
  - *vegetables and fruits*

D. mention on
- *intellectual peformances*: intelligence,
- *physical performances*
- *vital value*: life and death
- *health state*: dynamism, health,...
- pleasant and displeasant (coldness, thirst, hunger,...) *feelings*

4 Methods

4.1 Classical contingency analysis

Qualitative data are collected in a *contingency table* with Q columns, where Q is the number of discrete variables (see Table 4). Each row is a schematic description of a drawing.

$$\begin{array}{|c|ccccc|}\hline & 1 & & q & & Q \\ \hline 1 & x_{11} & \dots & x_{1q} & \dots & x_{1Q} \\ & \vdots & & \vdots & & \vdots \\ n & x_{n1} & \dots & x_{nq} & \dots & x_{nQ} \\ & \vdots & & \vdots & \ddots & \vdots \\ N & x_{N1} & \dots & x_{Nq} & \dots & x_{NQ} \\ \hline \end{array} \longrightarrow B_{(M\times M)} = \begin{array}{|c|ccc|}\hline & 1 & & M \\ \hline 1 & n_{11} & \dots & n_{1M} \\ & \vdots & & \vdots \\ m & n_{m1} & \dots & n_{mM} \\ & \vdots & & \vdots \\ M & n_{M1} & \dots & n_{MM} \\ \hline \end{array}$$

Table 4. Two-way contingency collected from N drawings.

Let us give some notations. Let us denote by M the total number of all the modalities: $M = \sum_{q=1}^{Q} m_q$. We form a M-way contingency table $K_{(N\times M)}$. This matrix K gives the complete data and is called the *complete disjunctive table*: it is essential if we want to remember who answered what. Let us denote by $K_{(N\times M)}$ the matrix with N rows and M columns which corresponds to the complete disjonctive table :

$$K = \{k_{ij}\} \text{ where } k_{ij} = \begin{cases} 1 & \text{if the individual } i \text{ chooses the modality } j, \\ 0 & \text{otherwise.} \end{cases}$$

The marginals of the rows of K are constant equal to the number Q of questions :

$$k_{i\cdot} = \sum_{j=1}^{M} k_{ij} = Q.$$

If we only have to study the *relations between the Q variables* (or questions), we can sum up the data in a *Burt matrix*, defined by $B = K^T K$, where K^T is the transposed matrix of K (Tab. 4).

Example 1. Let $A, B, C, \dots$ be a N-sample of Q discrete features. Each feature has m_q possible modalities. Let A have m_A levels, B have m_B levels, C have m_C levels,.... If we assume that $Q = 3$ and $m_A = 3$, $m_B = 2$ and $m_C = 3$, then an answer of an individual could be $(0, 1, 0|0, 1|1, 0, 0)$, where 1 corresponds to the chosen modality for each question. □

$$
K_{(12\times 3)} = \begin{array}{|ccc|cc|cccc|}
\hline
 & m_A & & m_B & & & m_C & & \\
\hline
0&1&0&0&1&0&0&0&1\\
0&1&0&1&0&0&0&1&0\\
0&0&1&1&0&0&1&0&0\\
1&0&0&0&1&0&0&0&1\\
1&0&0&0&1&0&0&1&0\\
0&1&0&0&1&0&0&1&0\\
0&0&1&1&0&1&0&0&0\\
1&0&0&1&0&1&0&0&0\\
0&1&0&1&0&0&1&0&0\\
0&1&0&0&1&0&0&1&0\\
0&0&1&0&1&0&1&0&0\\
1&0&0&1&0&0&0&0&1\\
\hline
\end{array}
\longrightarrow B_{(9\times 9)} = \begin{array}{|ccc|cc|cccc|}
\hline
4&0&1&2&2&1&0&1&2\\
0&5&0&2&3&0&1&3&1\\
0&0&3&2&1&1&2&0&0\\
\hline
2&2&2&6&0&1&2&1&1\\
2&3&1&0&6&0&1&3&2\\
\hline
1&0&1&2&0&2&0&1&0\\
0&1&2&2&1&0&3&0&0\\
1&3&0&1&3&0&0&4&0\\
2&1&0&1&2&0&0&0&3\\
\hline
\end{array}
$$

Table 5. Example of a disjunctive table $K_{(12\times 3)}$ and associated Burt table $B_{(9\times 9)}$.

The first stage consists in performing a χ^2-analysis to measure the level of dependance betweens the features (of B in Tab. 4) to make sure that every visualisation of pair of features $(x_i, x_j), 1 \leq i, j \leq M$ has a sense. The question whether the relationship between two variables is significant or not is formulated by the following problem of hypothesis testing:

$$H_0 : \chi^2_{Q-1,Q-1} = 0, \quad \text{vs} \quad H_1 : \chi^2_{Q-1,Q-1} > 0. \tag{1}$$

The null hypothesis H_0 is rejected or not at a given significance level p on the basis of the outcome of this test compared with the critical value of the theoretical distribution.
In practice, χ^2 is computed as follows. Let n_{ij} denote the joint effective associated to the couple (i, j), $n = \sum_{i,j} n_{ij}$ denote the complete size of the two-way contingency table crossing the variables i and j, $n_{i.} = \sum_j n_{ij}, n_{.j} = \sum_i n_{ij}$ the row i and the column j marginals respectively. The density of the deviation from independence given by

$$d_{Q-1,Q-1} = n\left(\sum_{i,j} \frac{n_{ij}^2}{n_{i.}n_{.j}} - 1\right), \tag{2}$$

is the chi-square value with $(Q-1)\times(Q-1)$ degrees of freedom.

Results for Tab. 3.2 We need to check a lot of different combinations of factors to see what's going on. However, doing so in a haphazard manner can be dangerous. Remember that each comparison we make (assuming we are using the standard hypothesis testing model) entails a type I error risk equal to our predefined α. We might assign the conventional value $\alpha = 0,10$. Each comparison we make has a $(1 - 0,10) = 0,95$ probability of avoiding a Type I error. We have here 23 process variables, and we want to see what the relationships are among them. We might temped to calculate the 253 possible correlations, and see which ones turn out to be statistically significant. But in the best-case scenario where the comparisons are independent, the probability of getting all the comparisons right is the product of the probabilities for getting each comparison right. In this case that would be $0,90^{253} \approx 9,55.10^{-4}$. Thus our chance of getting all 253 comparisons right is almost zero.
In fact, if you were to take 23 *uncorrelated* variables and calculate the set of 253 correlations, you should expect to see about 12 spurious correlations in the set. Among the 253 relationships that have been tested,

20 of them seem to be significant. We 'll have a tough time sorting pout which ones are real and which are spurious. Several strategies can be used to overcome this problem. The easiest, but probably the least acceptable, is to make α smaller, but it is impractical for large numbers of comparisons: as α becomes smaller, the power of the test is reduced to almost zero.The best strategy is *replication* – rerun of the experiment and see which comparisons show differences in both group: it should give an idea which effects are real abd which are not. As we can't replicate in our case, the next best thing is *cross-validation*, which involes setting aside part of the sample as a validation sample. The statistics of interest is computed on the main sample, and then checked against the validation sample to verify that the effects are real. Results that are spurious are usually revealed ny the validation sample.

4.2 Kohonen contingency table analysis

When we deal with qualitative variables, we seek to highlights the typology of the modalities and we try to emphasize the relations existing betwen the modalities of the variables. A remark has to be made at this stage: two modalities will be "close" if there is a large proportion of individuals that choose them simultaneously. For example, if a considerable amount of individuals choose the modality '1' of the question 'A' and the modality '2' of the question 'B', then we will say that the modalities $[A-1]$ and $[B-2]$ are close and that they *attract* each other. We would like to get these individuals grouped in the same region. Conversely, we like to observe distant representations in the case of a large amount of individuals choose $[A-1]$ and reject $[B-2]$.
It is clear that the Kohonen algorithm which organizes the units by respecting the proximities in the input space is appropriate for this type of treatment. In a previous paper, Cottrell *et al.* [6] defined a new algorithm (KOUPLET) which allows to handle two qualitative variables. The algorithm used here is inspired from this algorithm and analyzes the relations between Q (≥ 2) qualitative variables.
Recall that the matrix B is a $(M \times M)$ symetric matrix and is composed of $Q \times Q$ blocks, such that the $(q \times r)$ block B_{qr} $(for\ \ 1 \leq q, r \leq Q)$ contains the N answers to the question r. The block B_{qq} is a diagonal matrix, whose diagonal entries are the numbers of individuals who have respectively chosen the modalities $1, ..., m_q$, for the question q. The Burt table $B_{(M \times M)}$ has to be seen as a generalized contingency table, when more than 2 kinds of variables are to be studied simultaneously. In this case we loose a part of the information about the individuals answers but we keep the information regarding the relations between the modalities of the qualitative variables. Each row of the matrix B characterizes *a modality of a question* (or variable). Let us denote by b_{ij} the entries of the matrix B, whatever are the questions which contain the modalities i or j. The total sum of all the entries of B is $b = \sum_{i,j} b_{ij} = Q^2 N$. One defines successively

- the table F of the relative frequencies, with entry $f_{ij} = \frac{n_{ij}}{b}$,
- the margins with entry $f_{i.} = \sum_j f_{ij}$ or $f_{.j} = \sum_i f_{ij}$,
- the table P of the profiles which sum to 1, with entry $P_{ij} = \frac{f_{ij}}{f_{i.}}$.

Classical correspondence analysis (CA) is a *weighted* principal component analysis (PCA) performed on the row profiles or column-profiles of the matrix P, obtained with a χ^2 metric (see section 4) ([7]), each row being weighted by $f_{i.}$. Let $r(i)$ and $r(i')$ be two row-profiles of the matrix P. One has :

$$\chi^2(r(i), r(i'))r = \sum_j \frac{1}{f_{.j}} (P_{ij} - P_{i'j})^2 = \sum_j \left(\frac{f_{ij}}{\sqrt{f_{.j}} f_{i.}} - \frac{f_{i'j}}{\sqrt{f_{.j}} f_{i'.}} \right)^2 . \tag{3}$$

So it is equivalent to compute profile matrix C whose entry is $c_{ij} = \frac{f_{ij}}{\sqrt{f_{.j}} f_{i.}}$, to consider the Euclidean distance between its rows and to weighten each row by $f_{i.}$ (for more details see also [8]). CA provides a simultaneous representation of the M vectors on a low dimensional space which gives some information about the relations between the M variables.
It is possible at this stage to use a Kohonen algorithm to get a non-linear representation of the contingency table, as it has been already proposed by [8].
The main tool is a Kohonen network, generally a two-dimensional grid with $n \times n$ units, called *code-vectors*,

where a topological neighborhood is defined in a homogeneous way around each unit (but the method can be used with any topological organization of the Kohonen network). Note that it is an iterative algorithm. We propose to study the resulting map to extract the relevant information about the relations between the Q variables.
The general principal in SOM algorithm is given below:

a). by means of parallel computation, find the cell in the network that gives the “best” response to the present input, in the sense of some criterion .
b). modify this cell and its neighbors in the network to enhance their responses to the present input.

At the end of the iterative scheme, Each observation x belongs to classe i if and only if the associated code vector $\boldsymbol{u}_i$ is the closest among all the code vectors. The main characteristics of Kohonen classification is the conservation of the topology: after learning “close” observations are associated to the same class or to some “close” classes according to the definition of neighborhood in the Kohonen network. This feature allows to consider the resulting classification as a good starting point for further developments as stated below.
As mentionned above, the first and raw result we get after learning is a classification of the N observations into n classes[4].

5 Results

Among the 253 relationships between variables that have been tested, 20 of them seem to be significant, *i.e.* there is an interaction involving these indicators that can be visualized, such as in Fig. 2.a. There are 8 units in a grid. In each cell, the final code vector is drawn. In this representation, neighboring cells correspond to similar vectors. For sake of simplicity, only paired relationships are analysed, but more than 2variables relationships can be studied as well, the equivalent of multicomponent analysis (MCA).

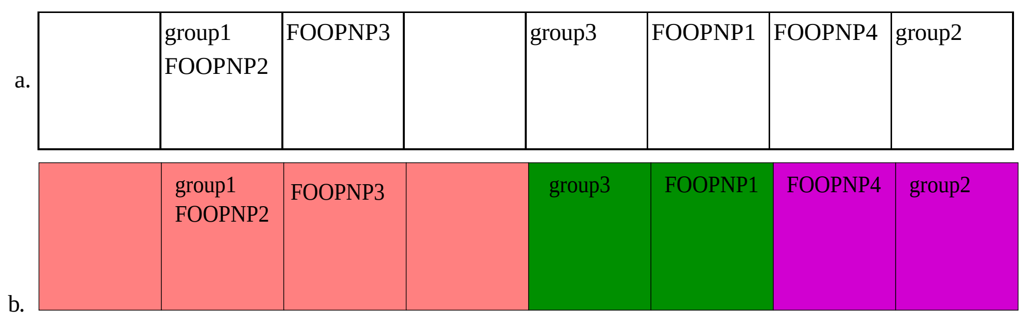

Fig. 2. a)String class representation of variables ’`group`’ and ’`FOOPNP`’ b) Hierarchical classification in 3 macro-classes.

We propose to reduce the number of classes by means of a hierarchical classification of the 8 code vectors using the *Ward distance*. We define 3 embedded classes and can distinguish the “micro-classes” (individual code vector) and the “macro-classes” which group together some of the micro-classes. To make visible this two-way classification, we affect to each “macro-class” some colour. See in Fig. 3 a representation of the “micro-classes” grouped together to constitute 3 “macro-classes”(see Fig. 2.b).
The advantage of this double classification is the possibility to analyse the data set at a “macro-level” where general features emerge and at a “micro-level” to determine the characteristics of more precise phenomena and especially the paths to go from one class to another one.

[4] The choice of the number n of units is arbitrary, and there does not exist any method to better choose the size of the network. We can only guess that the “relevant” number of classes could be smaller that n.

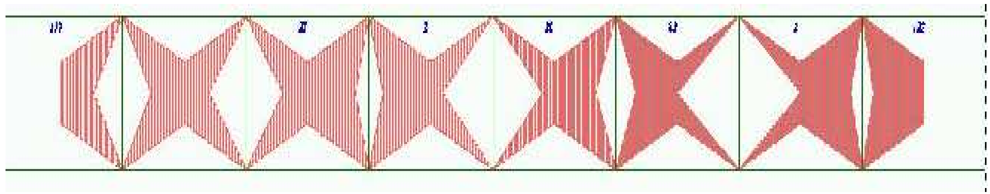

Fig. 3. Interclass Ward distances.

Once the classes are defined, one can compute the intra/inter class variances. As a complement and in order to have a better representation of the map geometry, we propose to vizualise the distances between classes [9]. This visualization avoids misleading interpretation and gives an idea of the discrimination between classes. We use the method proposed by Cottrel and De Bodt [10]: each unit is represented as an octogon. The bigger it is, the closer the unit is to its neighbors. So the clusters appear to be regions in which octogons tend to be big and frontiers are regions largely unshaded (Fig.3).
We observe generally (and in our application in particular) that the "macro-classes" boundaries coincide with the most important distances between the classes. This confirm the pertinence of the second-level classification. On the contrary, if a boundary occurs between two classes with small distance, that means that the second level classification splits a large group into two groups and that the path from one to the other is continuous. A hierarchical classification with fewer classes should then be considered.
Another question is how to put in evidence the intra-classes dispersion. This is related to the problem of the outliers, of the existence of a small typical group different from the rest of the data. We try to present a visual tool to decide which observation could be deleted in the learning phase and also how it would have been classified after learning. We can immediately see in which units the dispersion is large with respect to the others, which observation could be deleted or examined separately.
The main advantage of this topology preserving mapping is that there is no arbitrary choice of the representation. The *M* vectors which correspond to the modalities are correctly classified by the network and the map is realized in a very natural way. However, it is well-known that the classical representations uses a strong approximation which can make an interpretation of the relations very difficult.

Study outcomes from Kohonen maps provided important remarks:

1. sex is related to disorder and to the symbolisation of food. Girls resort to symbolisation of aliments because they draw better.
2. when less than 9 y, the child stress on the table and on the food, on the contrary, when more than 9 y, he represents the meal with a child.
3. the perception of the table changes with age. The child less than 6 y associates the table with bad-eating: the table means "constraints" for him while it is associated with comfort for the child more than 6 y (integration of norms).
4. the more than 9 y child stresses on the physical appearance and represents easily.
5. the more than 9 y child associates bad eating , with variables `DIMORP, NADRNK, OIL`. When `DIMORP` is used, the drawn child is thin, or fat or ill.
6. the disorder in the less than 6 y child indicates the quatity?
7. girls were overweight with respect to boys

One of our primary goal was to identify the most influencing variables in the genesis of *food perception*.. . .
Involvement of the family in nutrition education programs could promotes positive outcomes in knowledge, attitudes, parent-child communication. This observation of weight and food Habits shall allow the healt-department to: - better understand the mechanisms involved in excessive weight gain and in the distribution of the body fats, - follow the natural development in the choice of food and of oils in particular, made by a population that is well informed on nutrition, - better understand the needs, the tastes and the food habits of the consumer, in order to offer and develop products that combine health and taste pleasures.

References

1. J.E. Blundell and N.A. King. *The origins and consequences of obesity*, chapter Overconsumption as a cause of weight gain : behavioral-physiological interactions in the control of food intake (appetite), pages 138–158. Wiley, Chichester, 1993.
2. S. Plancoulaine, M.A. Charles, L. Lafay, M. Tauber, J.M. Thibult, and E. Eschwège. Infant feeding patterns are related to blood cholesterol concentration in prepubertal children aged 5-11 y the Fleurbaix-Lavantie ville santé study. *European Journal of Clinical nutrition*, 54:114–119, 2000.
3. L. Lafay and M.-A. Charles. Anthropométrie et alimentation des enfants prépubères de l'étude Fleurbaix-Laventie ville-santé (flvs). rapport technique, INSERM unité 258, 1997.
4. M.A. Charles, F. Thomas, A. Fontbonne, J.M. Borys, D. Boute, C. Toursel, P. Fossati, and E. Eschwege. *Atherosclerosis X*, chapter Familial approach of cardiovascular risk factors : the "FLeurbaix Laventie Ville Santé" study, pages 73–78. Elsevier Sciences, 1995.
5. J.T. Tou and R.C. Gonzalez. *Pattern recognition principles*. Applied mathematics and Computation. Addison-Weslay, Reading, MA, 1974.
6. M. Cottrell, P. Letremy, and E. Roy. Analysing a contingency table with kohonen maps: a factorial correspondence analysis. In J. Cabestany, J. Mary, and A. Prieto, editors, *IWANN'93*, pages 305–311, 1993.
7. L. Lebart, A. Morineau, and M. Piron. *Statistique exploratoire multidimensionnelle*. Dunod, 1995.
8. S. Ibbou and M. Cottrell. Multiple correspondence analysis of a crosstabulations matrix using the kohonen algorithm. In M. Verleysen, editor, *ESANN'99*. Springer, 1999.
9. M. Cottrell and P. Rousset. The kohonen algorithm: a powerful tool for analyzing and representing multidimensional quantitative and qualitative variables. In *IWANN'97*, pages 861–871, 1997.
10. M. Cottrell and E. De Bodt. A kohonen map representation to avoid misleading interpretations. In M. Verleysen, editor, *ESANN'96*, pages 103–110, Bruxelles, 1996.