A new approach for analyzing hierarchical sorting task data

Marine Cadoret, Sébastien Lê, Jérôme Pagès
AGROCAMPUS OUEST, France

10th Sensometrics conference

July 25-28 2010, Rotterdam

Introduction

- Sensory analysis:
  - Egoroff, 2005
  - Blancher *et al.*, 2008
  - Santosa *et al.*, 2010
- Hierarchical sorting task consists in asking subjects to perform successively several nested sorting tasks
- Used to understand the process during a sorting task

Data

- Binary hierarchical sorting
- 22 subjects
- 16 advertisements concerning an orange juice
- Construction according to a 2^{5-1} fractional factorial design

Data: factors of the design

- Background color

Green

White

Data: factors of the design

- Figurative

With

Without

Data: factors of the design

- Catchword: “It’s winter… Are you feeling weak? Give your body vitamins!”

With

Without

Data: factors of the design

- Allegation: “Rich in vitamins A, C, B1 and B6”

With

Without

Data: factors of the design

- Realization of a performance' test: "90% of women who tested it, felt the difference in 2 weeks"

With

Without

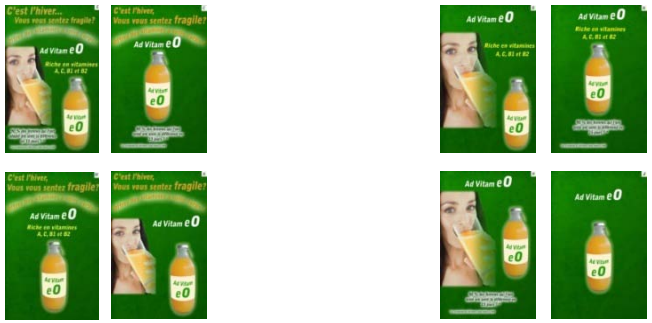
Hierarchical sorting of subject 3

L1

Color

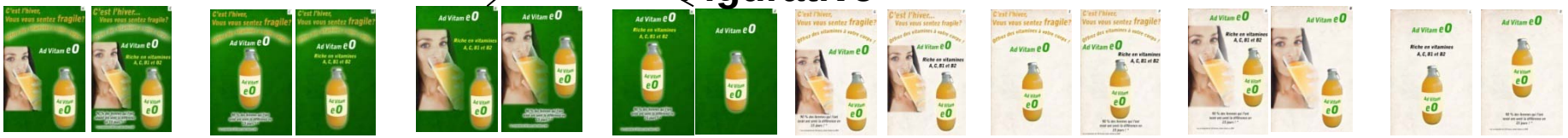
L2

Catchword



L3

Figurative

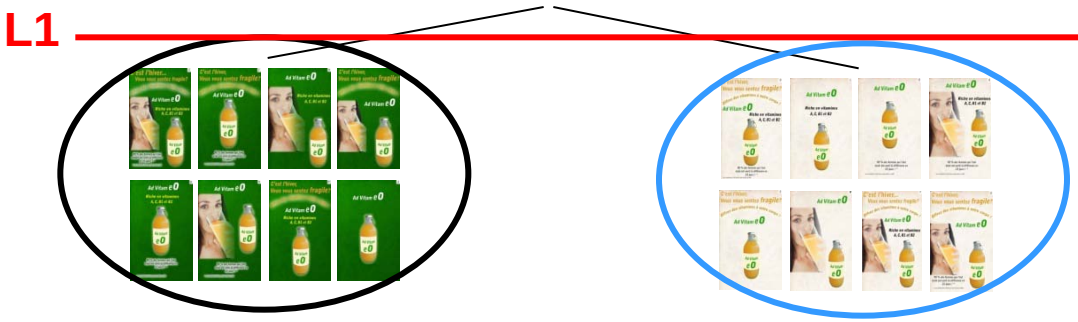


The data gathering

How the data can be gathered in a data table?

AGRO CAMPUS OUEST

Data coding: subject 3



	L1
A	G2
B	G1
C	G1
D	G2
E	G1
F	G2
G	G1
H	G2
I	G2
J	G1
K	G1
L	G2
M	G1
N	G1
O	G2
P	G2

Data coding: subject 3

	L1	L2
A	G2	G4
B	G1	G2
C	G1	G1
D	G2	G4
E	G1	G1
F	G2	G3
G	G1	G3
H	G2	G2
I	G2	G3
J	G1	G2
K	G1	G1
L	G2	G4
M	G1	G1
N	G1	G2
O	G2	G4
P	G2	G3

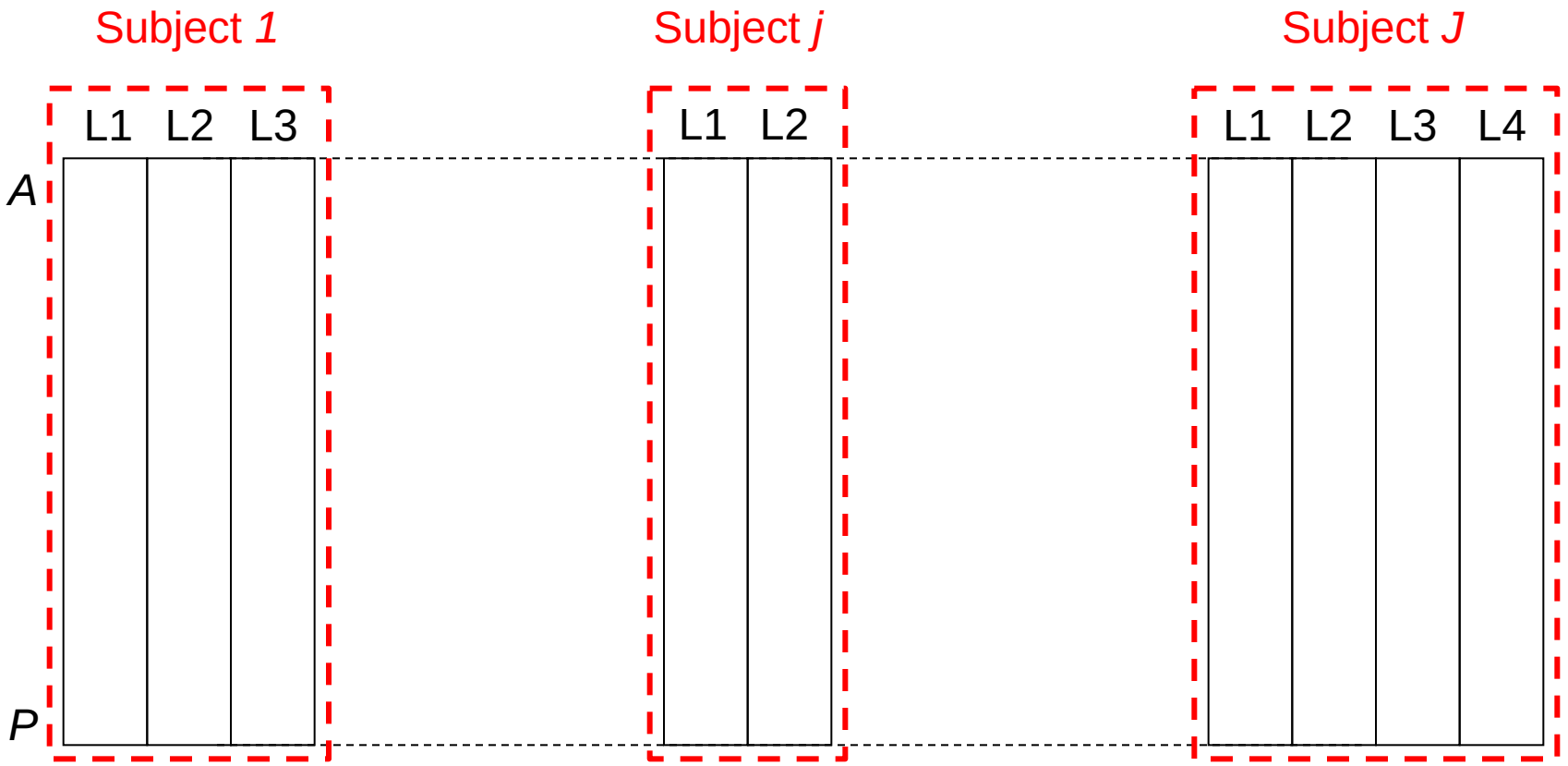
Data coding: subject 3

	L1	L2	L3
A	G2	G4	G8
B	G1	G2	G3
C	G1	G1	G2
D	G2	G4	G8
E	G1	G1	G1
F	G2	G3	G5
G	G1	G3	G5
H	G2	G2	G3
I	G2	G3	G6
J	G1	G2	G4
K	G1	G1	G1
L	G2	G4	G7
M	G1	G1	G2
N	G1	G2	G4
O	G2	G4	G7
P	G2	G3	G6

Data coding: subject 3

	L1	L2	L3
A	white	G4	G8
B	green	G2	G3
C	green	G1	G2
D	white	G4	G8
E	green	G1	G1
F	white	G3	G5
G	green	G3	G5
H	white	G2	G3
I	white	G3	G6
J	green	G2	G4
K	green	G1	G1
L	white	G4	G7
M	green	G1	G2
N	green	G2	G4
O	white	G4	G7
P	white	G3	G6

Data coding: *J* subjects



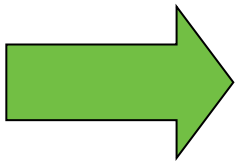
The data analysis

How the data table can be analyzed?

AGRO CAMPUS OUEST

Data analysis

- With these data, we are interested in different representations
- To obtain these representations, we want to:
  - Balance the influence of each subject
  - Keep the information provided by each subject



Multiple Factor Analysis (MFA) in which
1 subject = 1 group

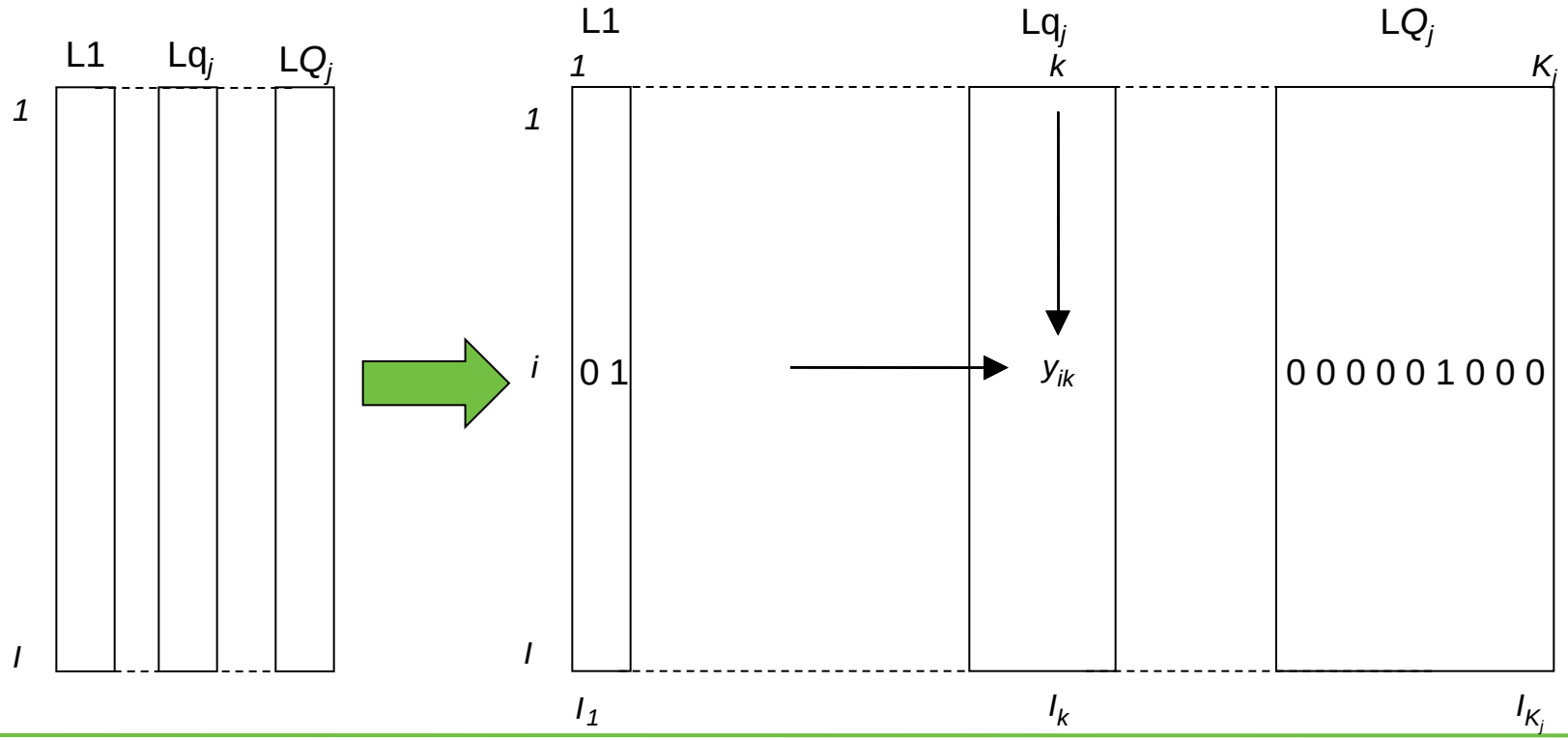
Data analysis

- MFA is looking for:
  - Objects oppositions provided by several subjects
  - Objects oppositions provided at upper levels

- MFA provides different representations:
  - An objects representation
  - A subjects representation
  - A levels representation

Disjunctive data table associated with subject *j*

- Each level is represented by a set of dummy variables



Objects representation

- Distance between 2 objects *i* and *l*:

$$d^2(i,l) = \sum_j \sum_{k \in K_j} \frac{1}{Q_j} \frac{I}{I_k} \left(y_{ik} - y_{lk} \right)^2,$$

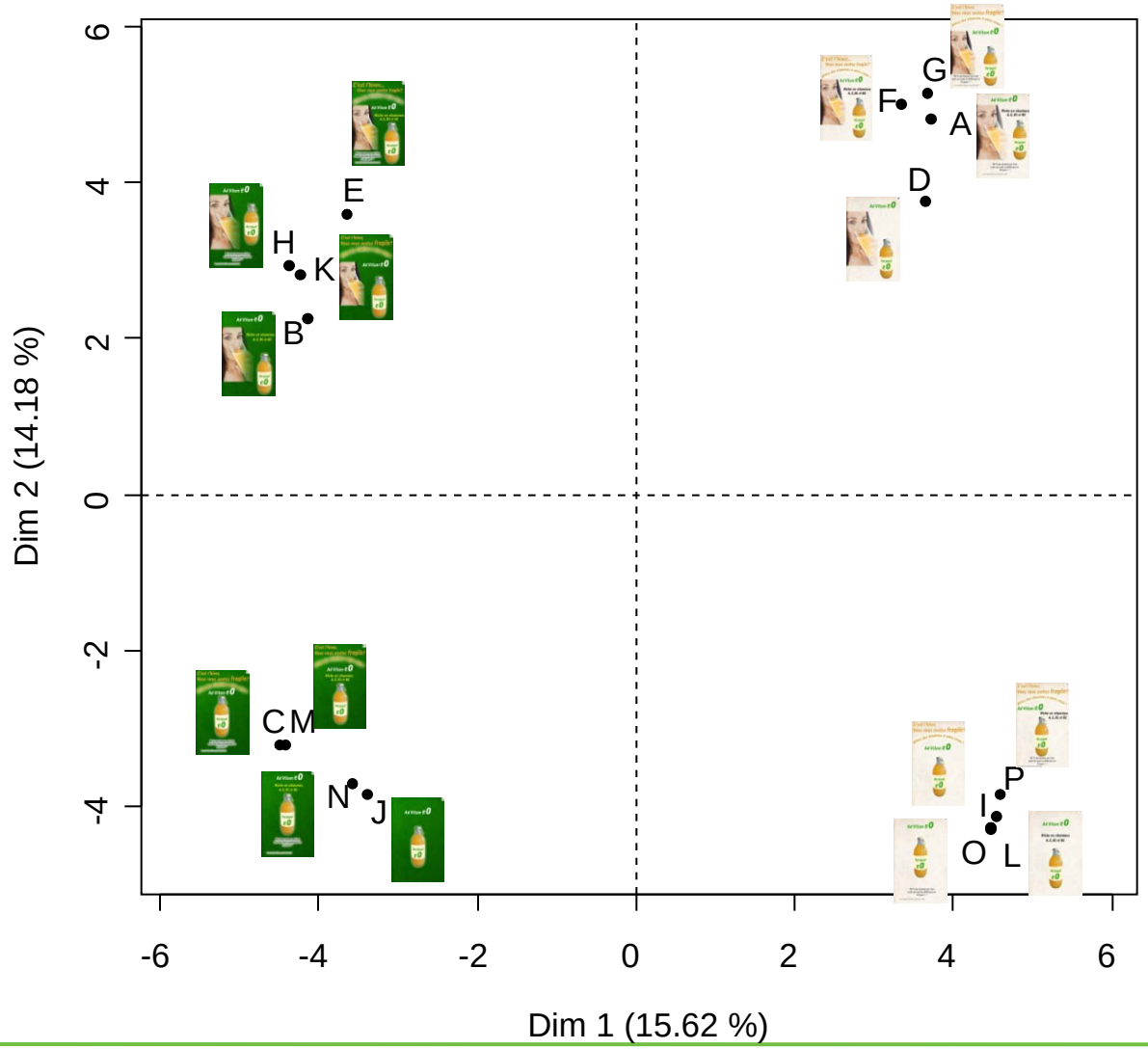
with Q_j the number of level of subject *j*

I the number of objects

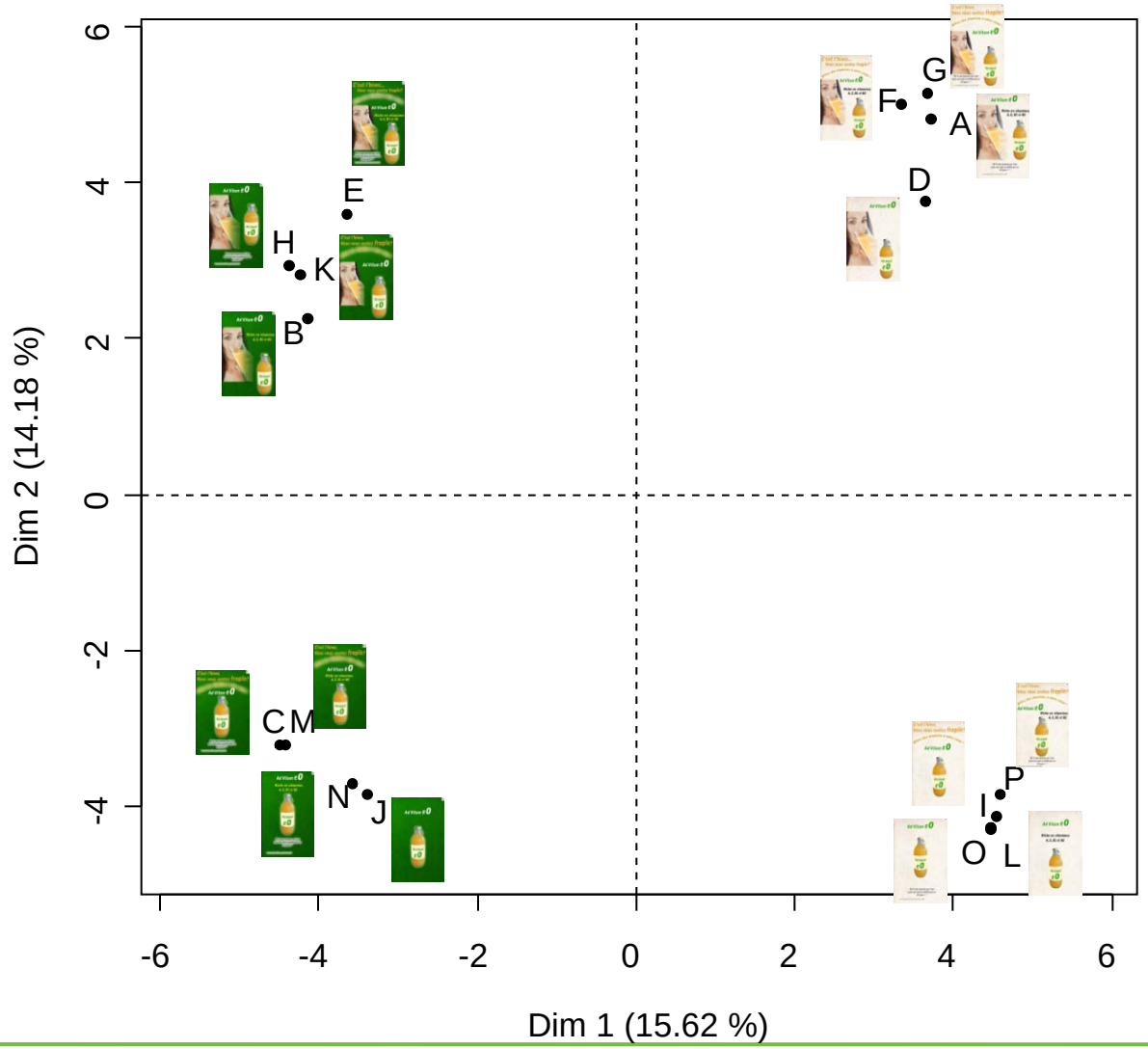
I_k the number of objects into the group *k*

y_{ik} the element of the disjunctive data table which is equal to *1* if the object *i* belong to group *k* and *0* in the opposite case

Advertisements representation

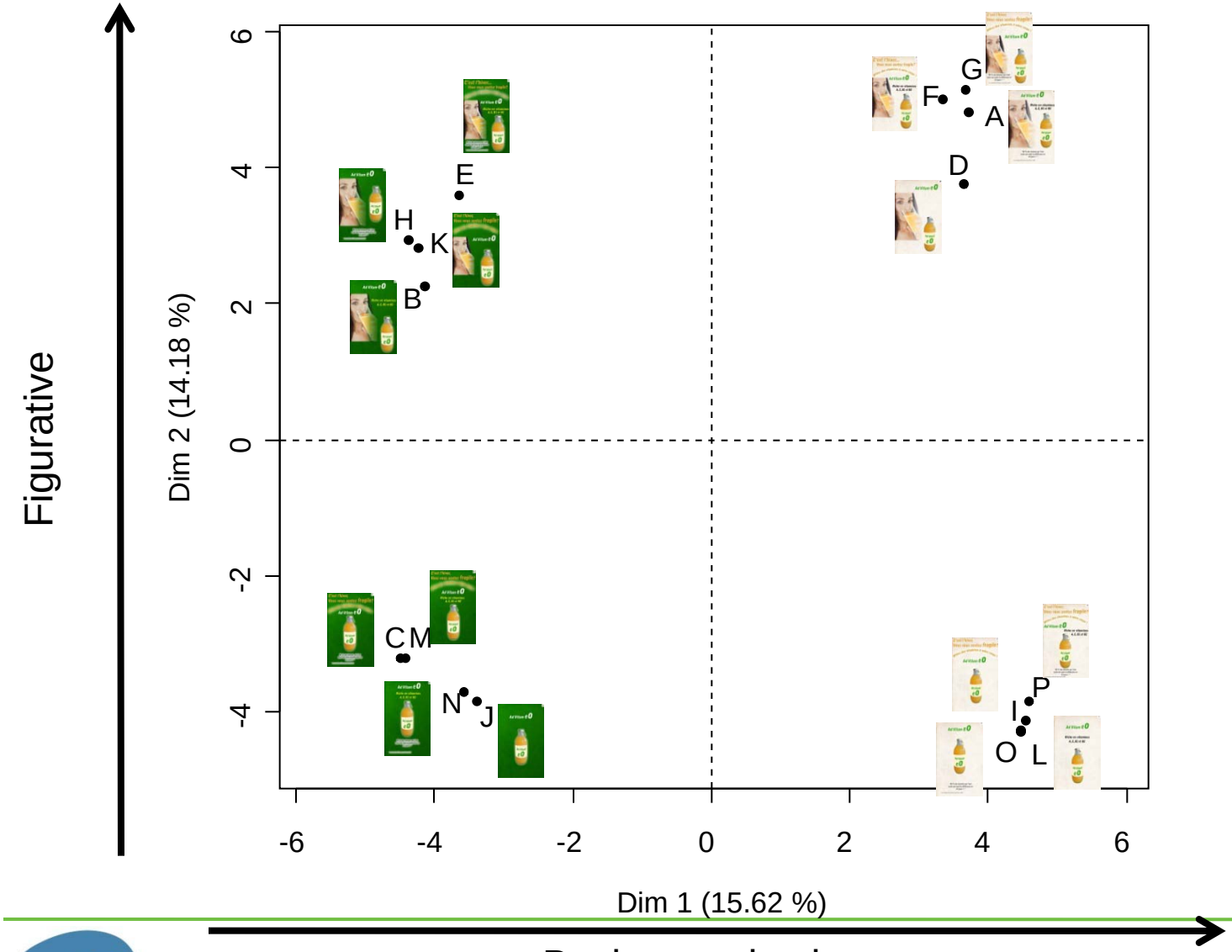


Advertisements representation



Background color

Advertisements representation



Subjects representation

- Coordinate of subject *j* on axis *s*:

$$\frac{1}{Q_j}\sum_{k\in Q_j}\eta^2(z_s, V_k),$$

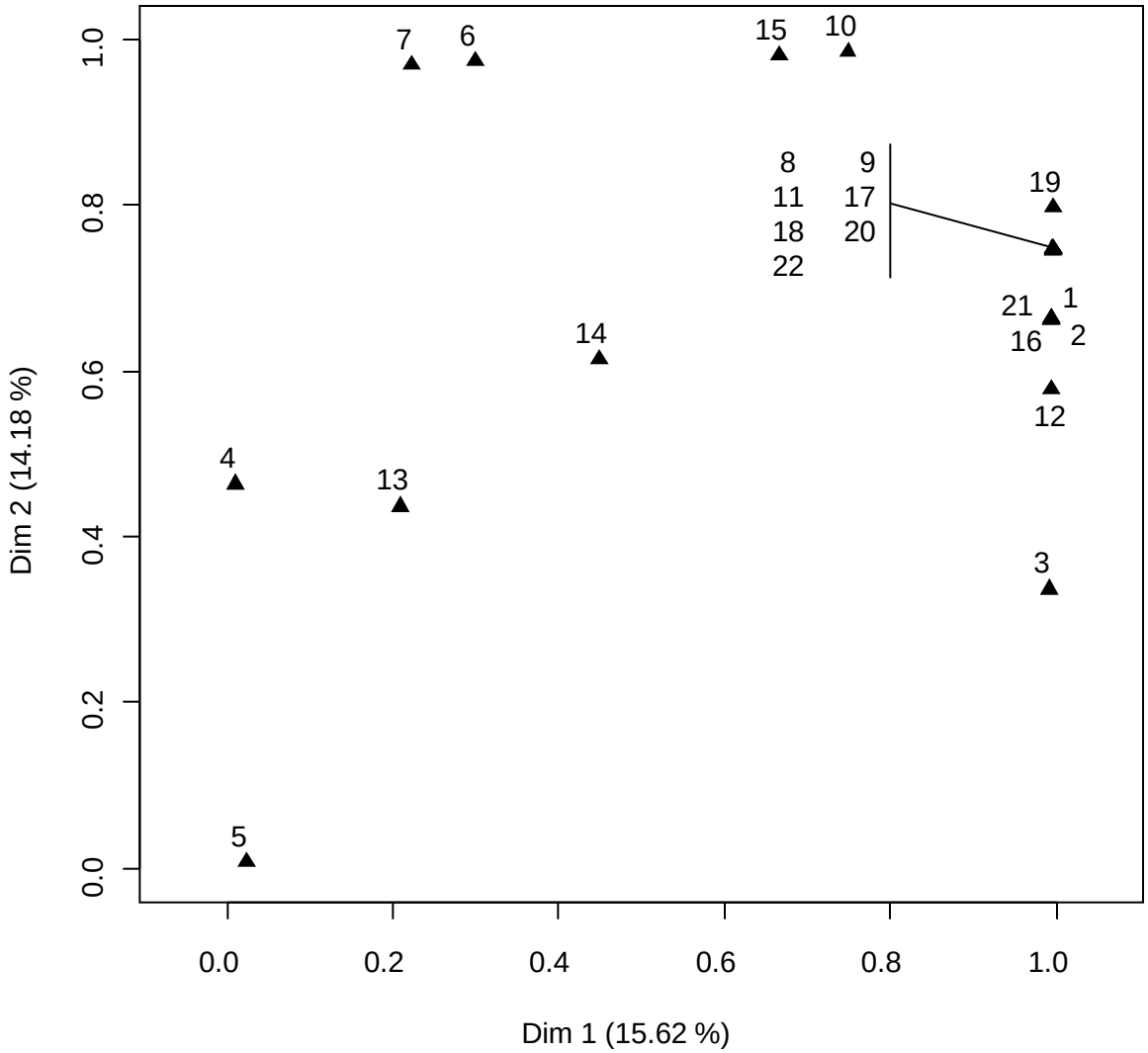
with Q_j the number of level of subject *j*

z_s the axis *s*

V_k the qualitative variable

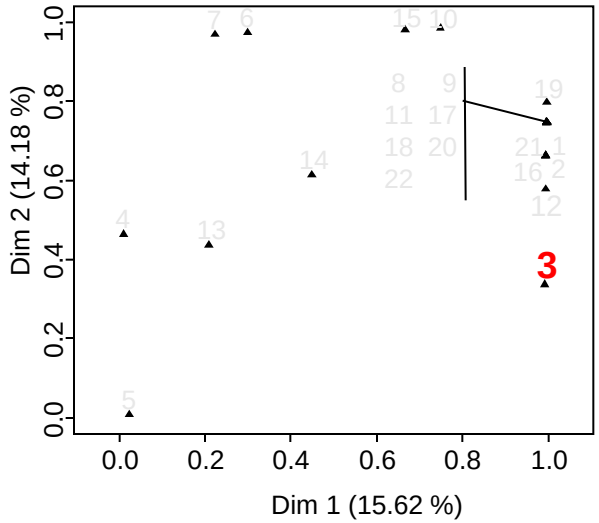
A subject will have a coordinate all the more important than he saw the objects oppositions highlighted by the dimension at an upper level

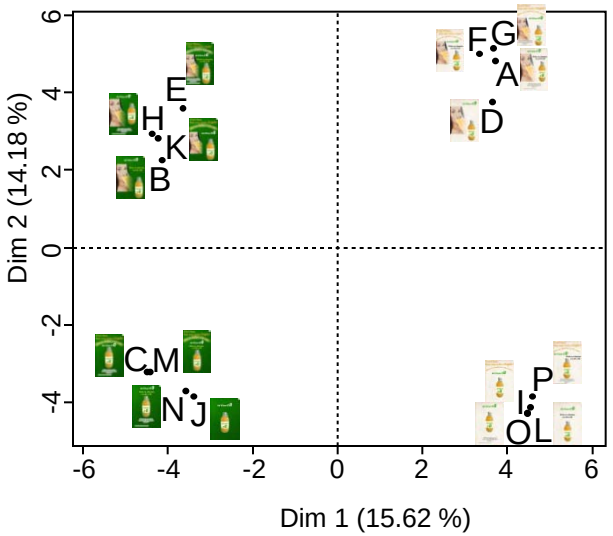
Subjects representation



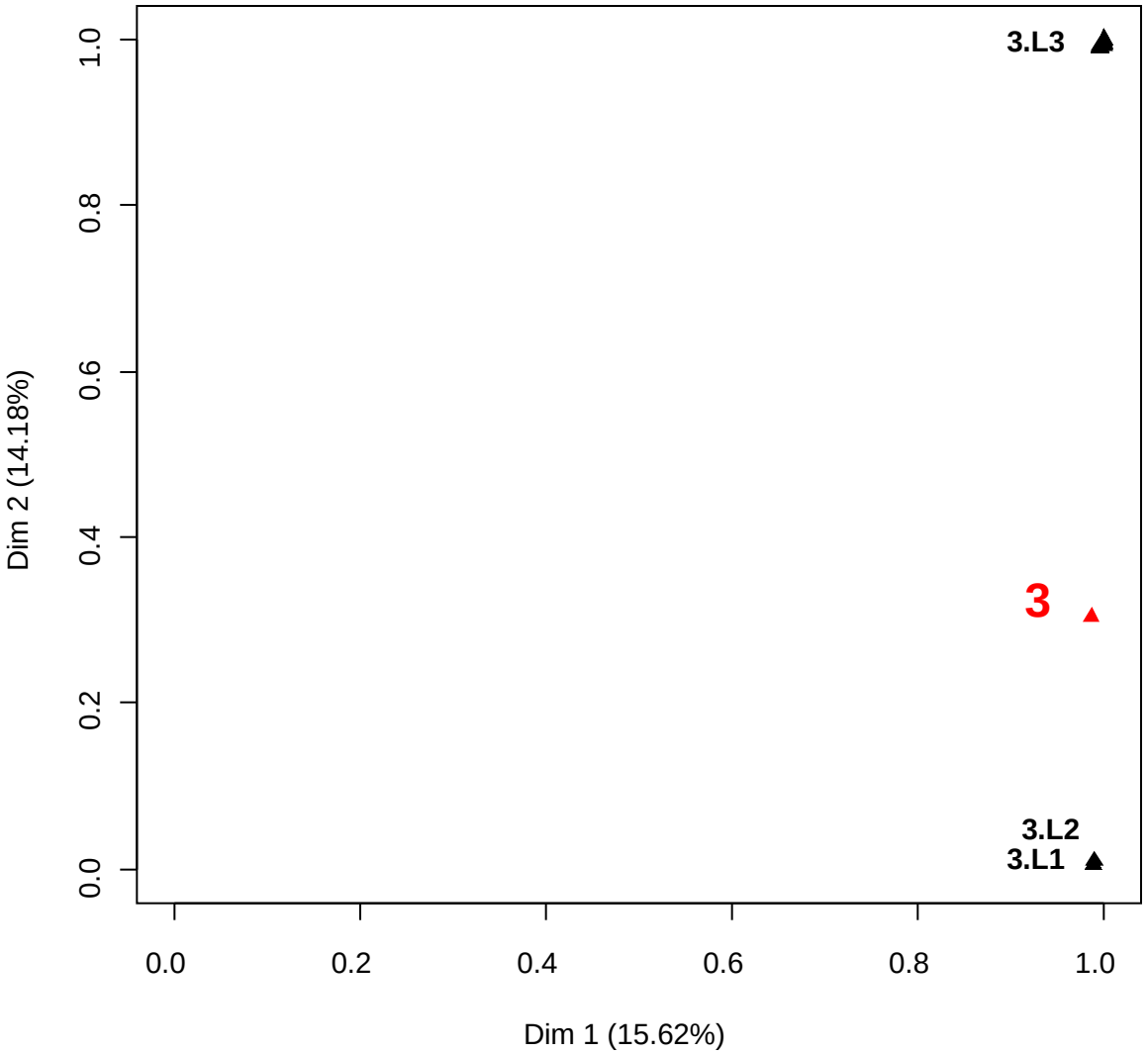
Subject 3


L1
L2
L3

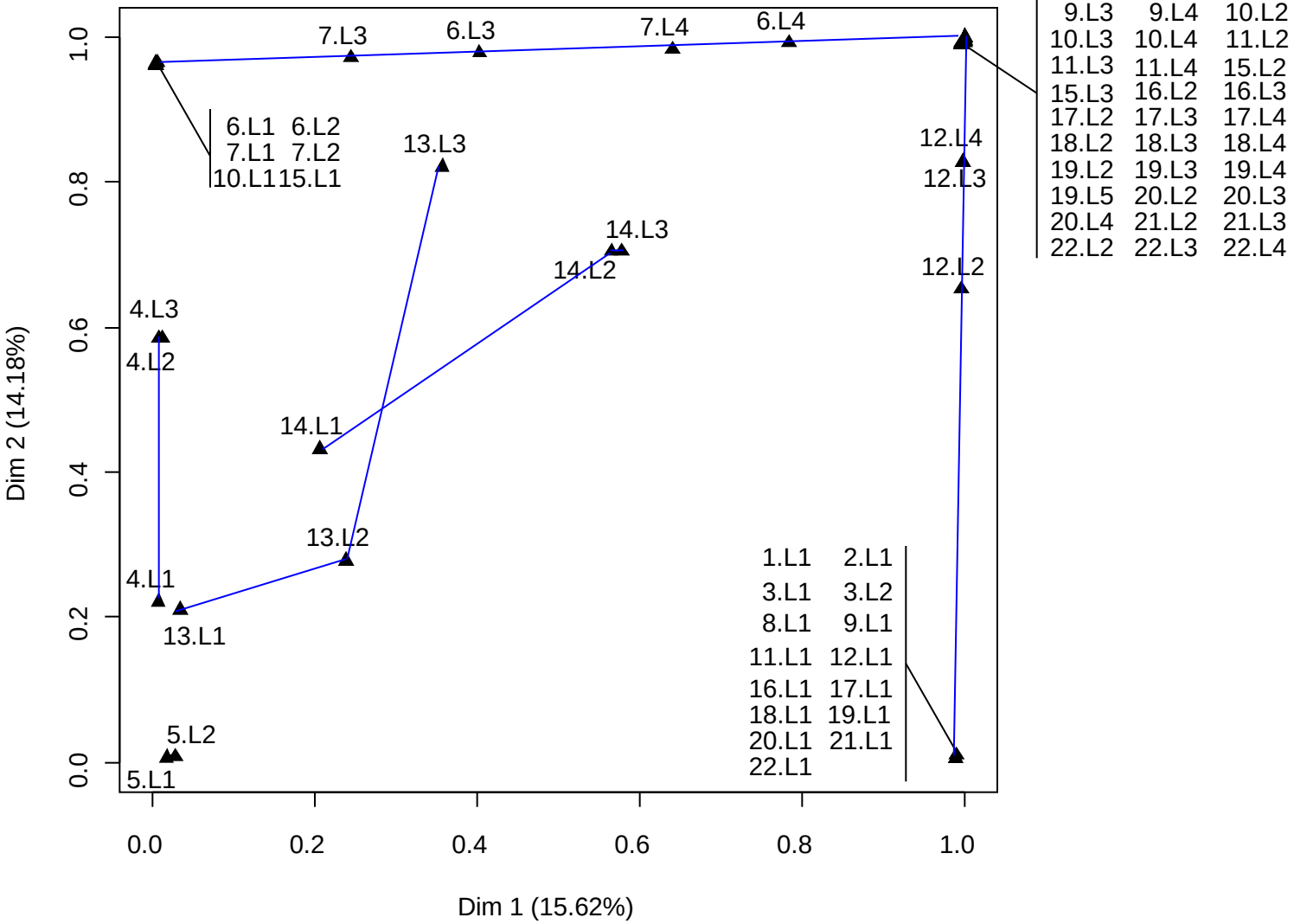

Dim 2 (14.18 %)
Dim 1 (15.62 %)
3


Dim 2 (14.18 %)
Dim 1 (15.62 %)
E
H
K
B
F
G
A
D
C
M
N
J
P
I
O
L

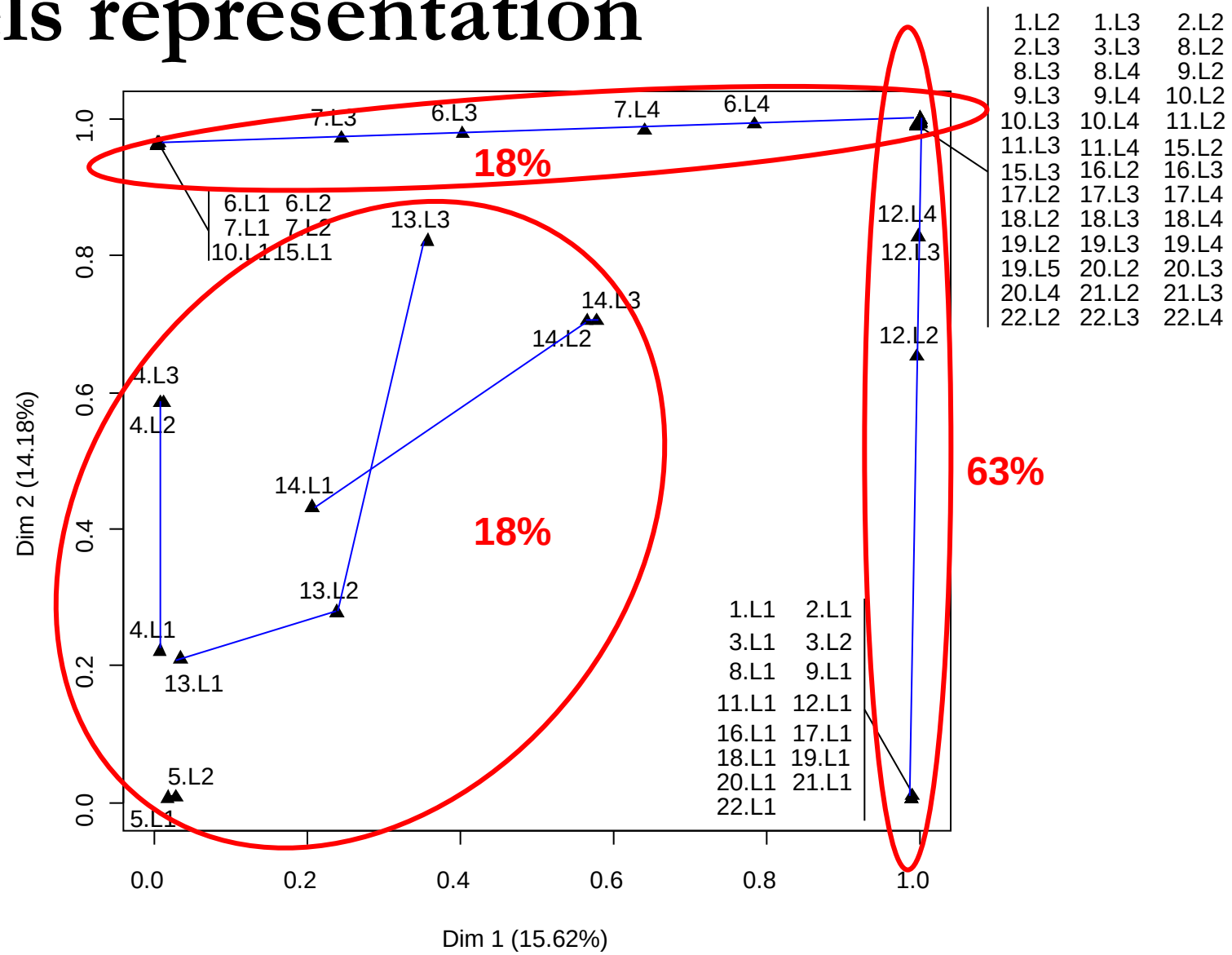
Levels representation of subject 3



Levels representation



Levels representation



Conclusion

- Allows the simultaneous taking into account of hierarchies of partitions in a same analysis
- Methodology providing rich and interpretable results
  - Allows also a words representation
- Another application:
  - *Poster 33: Analysing trees issued from a hierarchical sorting task using HMFA*

Conclusion

- Observing consumers performing napping, sorting and hierarchical sorting lead us naturally to consider what we have called a **“free holistic approach”** where the assessors are free to use the holistic approach they feel more comfortable with
- To be analyzed with HMFA…

http://sensominer.free.fr

Journal of sensory studies **SensoMineR a package for sensory data analysis**

http://factominer.free.fr

Journal of statistical software **FactoMineR: an R package for multivariate analysis**

Levels representation

- Superimposed to subjects representation
- Coordinate of level V_k on axis s:

$$\eta^2(z_s, V_k)$$

- 2 consequences:
  - For one subject, the coordinates of its levels are ordered along each axis
  - A subject is at the barycentre of his levels