

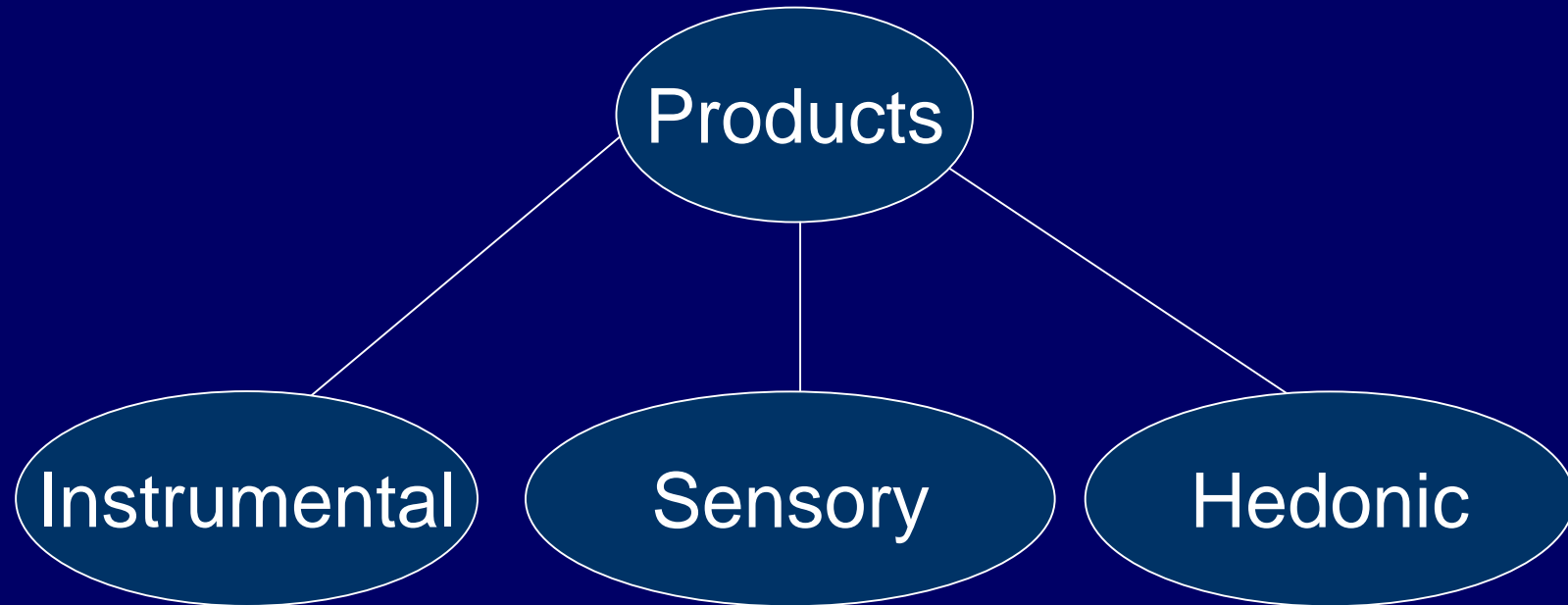
# Hierarchical Multiple Factor Analysis: application to the comparison of sensory profiles

S. Le Dien and J. Pagès

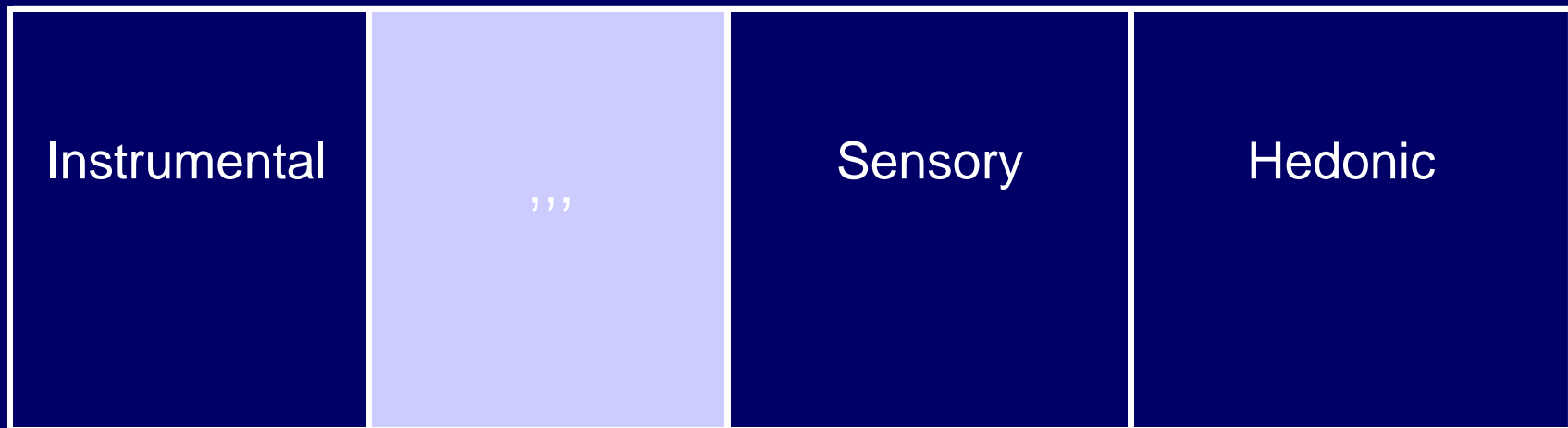
Unité de mathématiques appliquées, ENSAR

65 rue de Saint Briec, 35042 Rennes, France

# Introduction

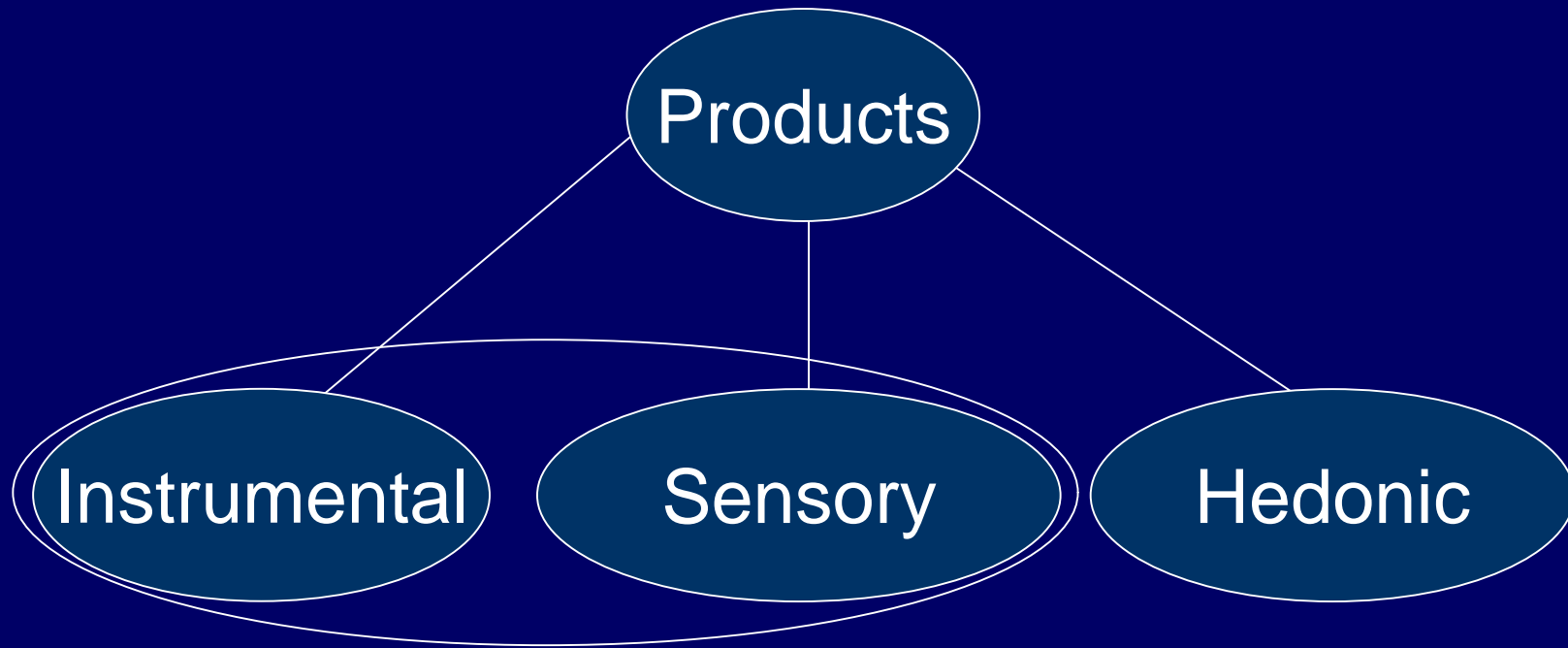


# Multi-block methods

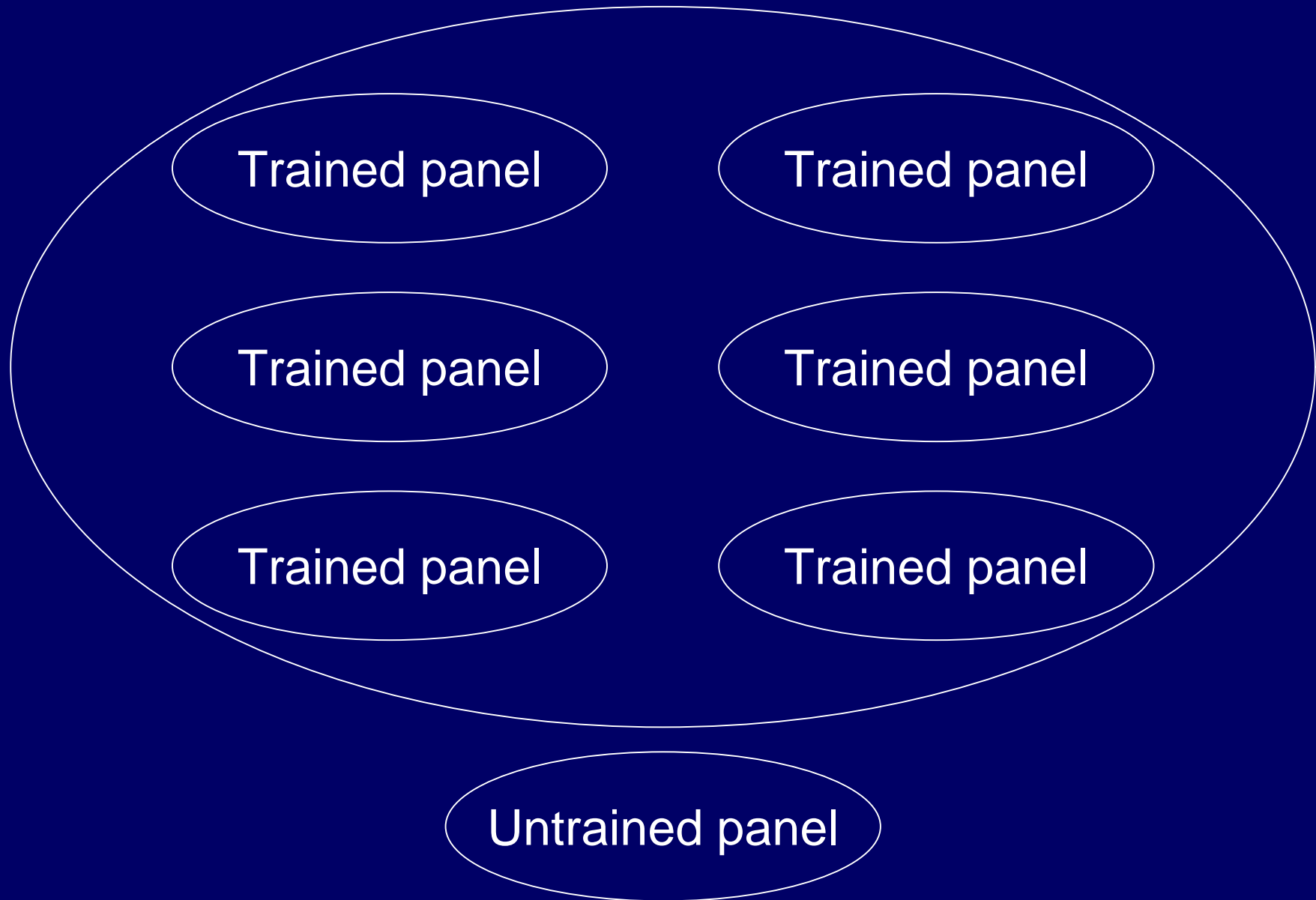


- Generalised Procrustes Analysis (GPA)
- STATIS method
- Multiple Factor Analysis (MFA)

# Hierarchy structure



# Hierarchy structure



# Data

| Commercial brand | % Cocoa | % Fat Content | % Saccharose |
|------------------|---------|---------------|--------------|
| Excellence Noir  | 70      | 42            | 27           |
| Qualité amère    | 60      | 34            | 38           |
| Mi-doux          | 45      | 35            | 40           |
| Amazonie         | 62      | 38            | 36           |
| Pâtissier        | 61      | 40            | 37           |
| Extra Supérieur  | 50      | 34            | 47           |

# Protocol

- 6 French laboratories
- 14 common descriptors
- 11 box intensity scale from 0 to 10
- Training : 10 sessions
- Each product was evaluated by each judge
- Balanced designs for first-order carry-over effects

**7th panel of 29 students**

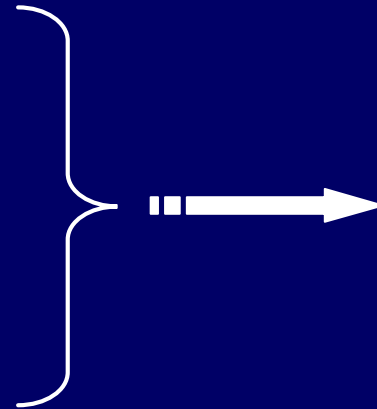
# Data

|            | Panel 1  |     | Panel 6 |  | Untrained Panel |
|------------|----------|-----|---------|--|-----------------|
| Chocolates | 1 ... 14 |     |         |  | 1 ... 14        |
|            |          | ... |         |  |                 |
|            |          |     |         |  |                 |

- 5 chocolates
- $7 \times 14 = 98$  variables
- Difficult test

# “Which value can be granted to profiles provided by Untrained Panels?”

- Consensus
- Discrimination
- Repeatability
- Reproducibility



Internal validation

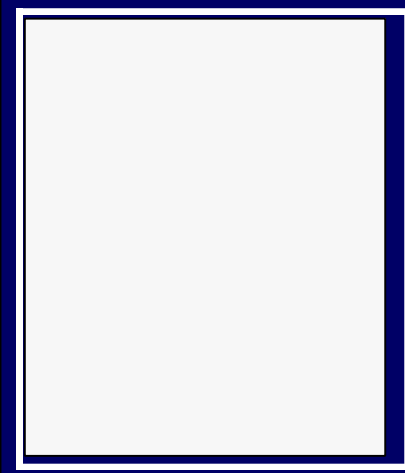
“Trained” **versus** “Untrained”



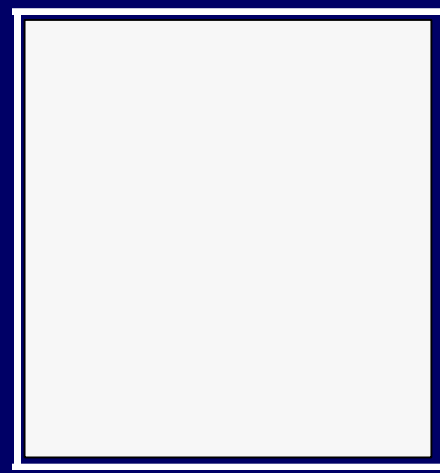
External validation  
(Clapperton, Pigott, 1979)

## Partition $P_1$

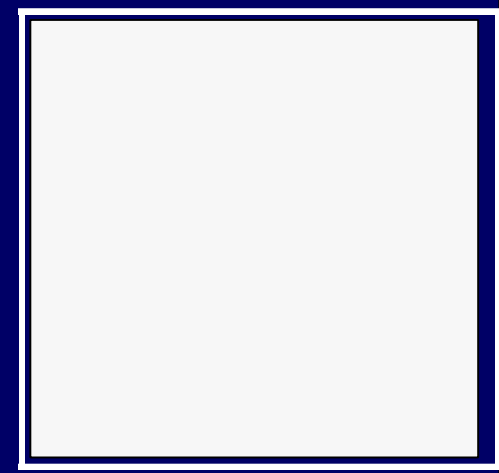
Panel 1



Panel 6

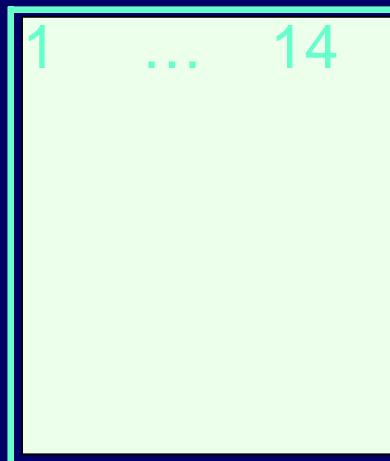


Untrained Panel

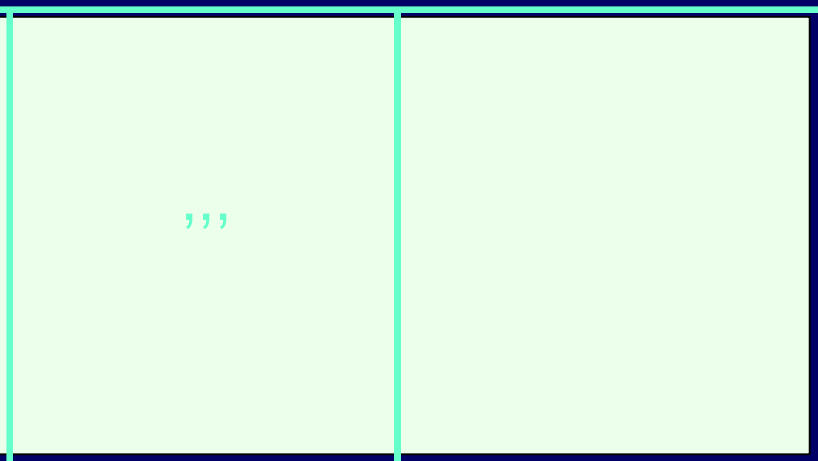


## Partition $P_2$

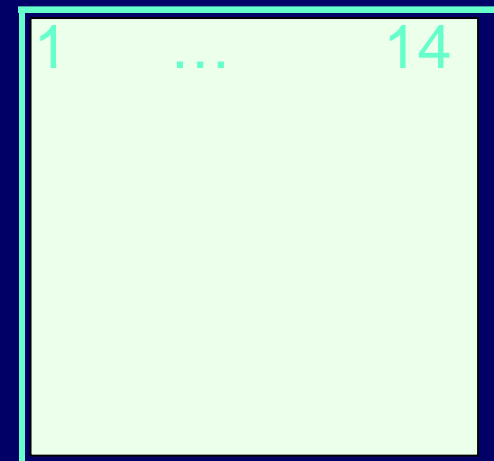
Panel 1



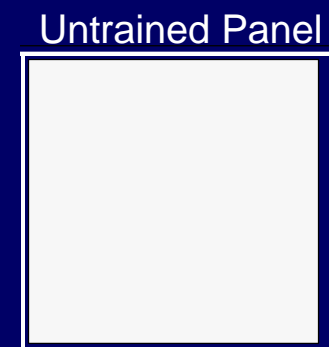
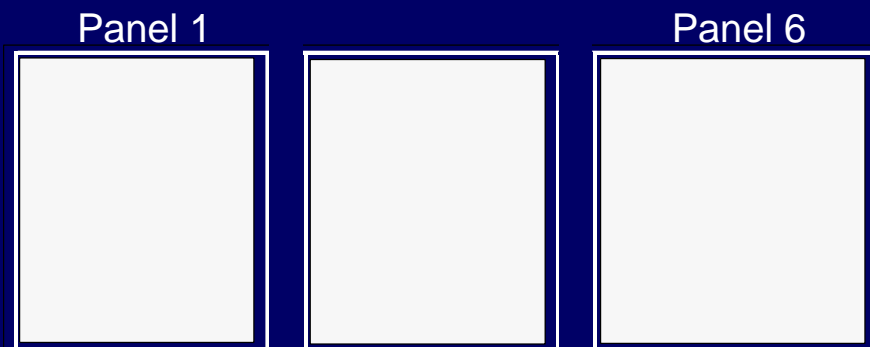
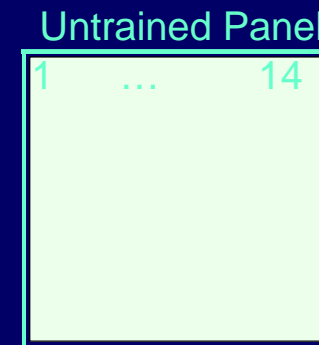
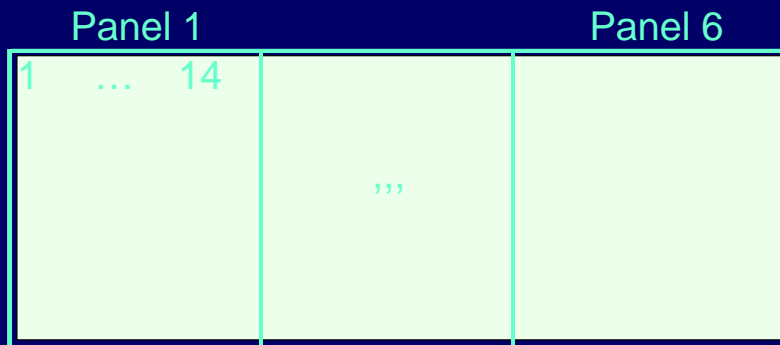
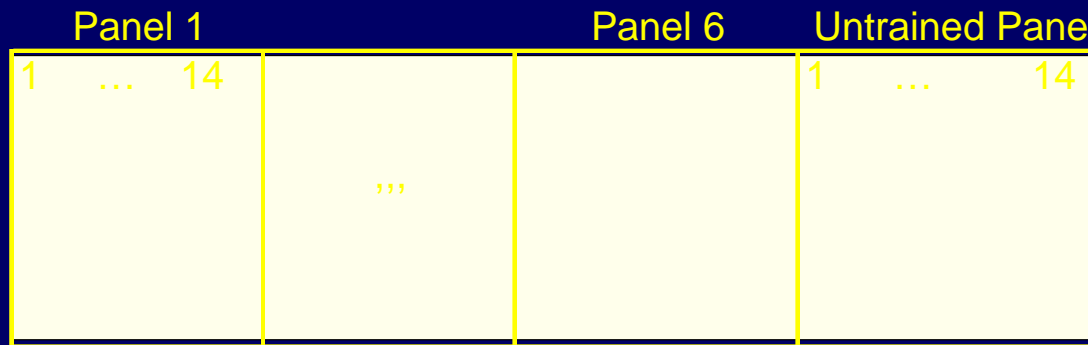
Panel 6



Untrained Panel

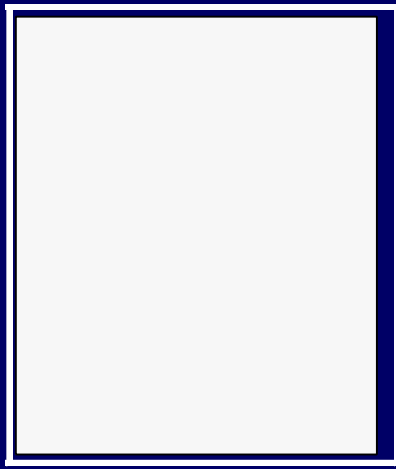


# Taking into account of a hierarchy structure

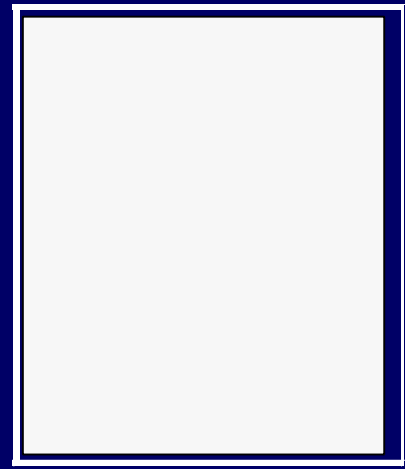
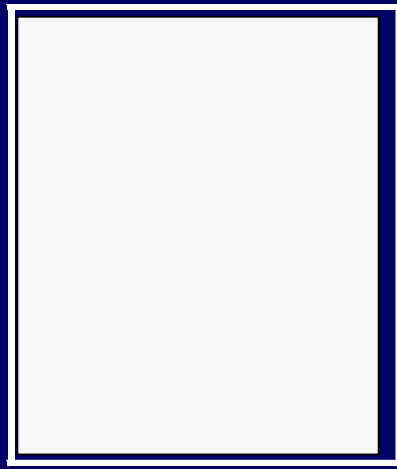


# Multiple Factor Analysis

- J groups of variables  $K_j$
- J matrices of data  $X_j$



$X_1$

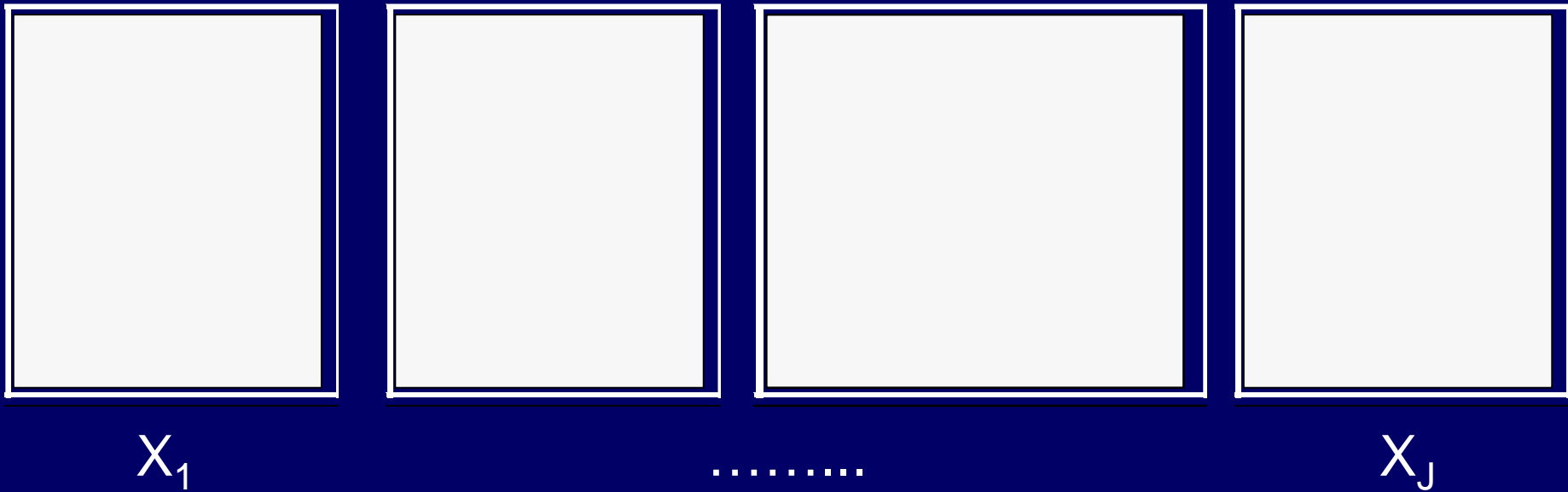


$X_j$

.....

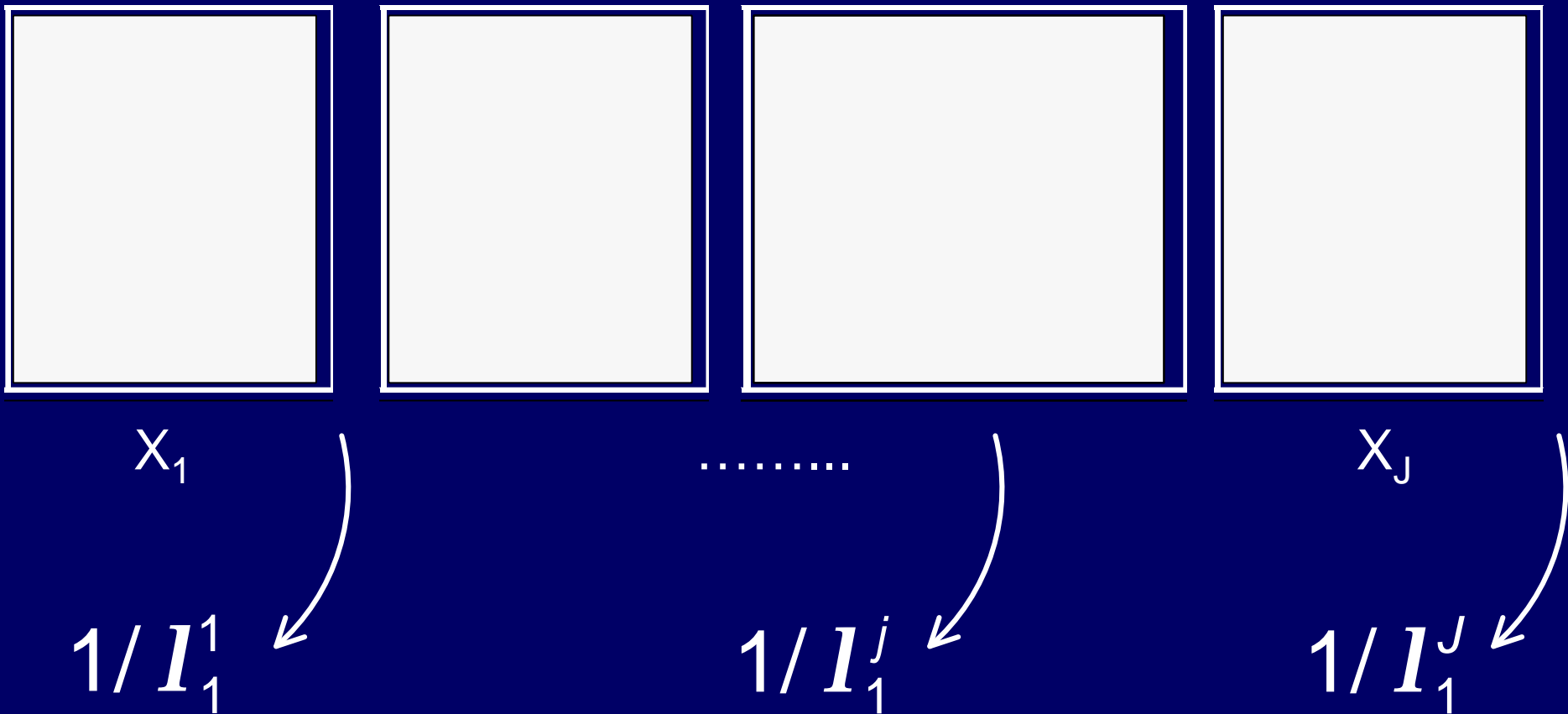
$$F_s^j = \sum_{k \in K_j} a_k^s v_k$$

# Multiple Factor Analysis



- **PCA**: directions of maximum inertia
- **Canonical Analysis**: common factors

# Multiple Factor Analysis Weighting



# Multiple Factor Analysis Weighting

- No group can generate all by itself the first dimension of the consensus
- The more “multidimensional” a group is, the more dimensions it contributes to in the construction of the consensus

# Carroll's generalised canonical analysis

$$z_s / \sum_j R^2(z_s, K_j) \text{ is max}$$

$$\text{Var}(z_s) = 1 \text{ and } \text{Cor}(z_s, z_t) = 0, \forall t < s$$

Problem of identification when the variables of a same group are highly **correlated**

# Multiple Factor Analysis

$$L_g(z, K_j) = \frac{1}{I_j} \sum_{k \in K_j} \text{Cor}^2(z, K_j)$$

$$L_g(z, K_j) = 1 \Leftrightarrow z \text{ is the 1}^{\text{st}} \text{ p.c. } K_j$$

$$z_s / \sum_j L_g(z_s, K_j) \text{ is max}$$

$$\text{Var}(z_s) = 1 \text{ and } \text{Cor}(z_s, z_t) = 0, \forall t < s$$

# Multiple Factor Analysis

Canonical  
Analysis

" $L_g$ "

Weighted  
PCA

" $1/I_1^j$ "

Multiple Factor Analysis

## Practically...

- Consensus
- Superimposed representation of partial individuals
- Graphical display of the groups
- Indicators related to common structures

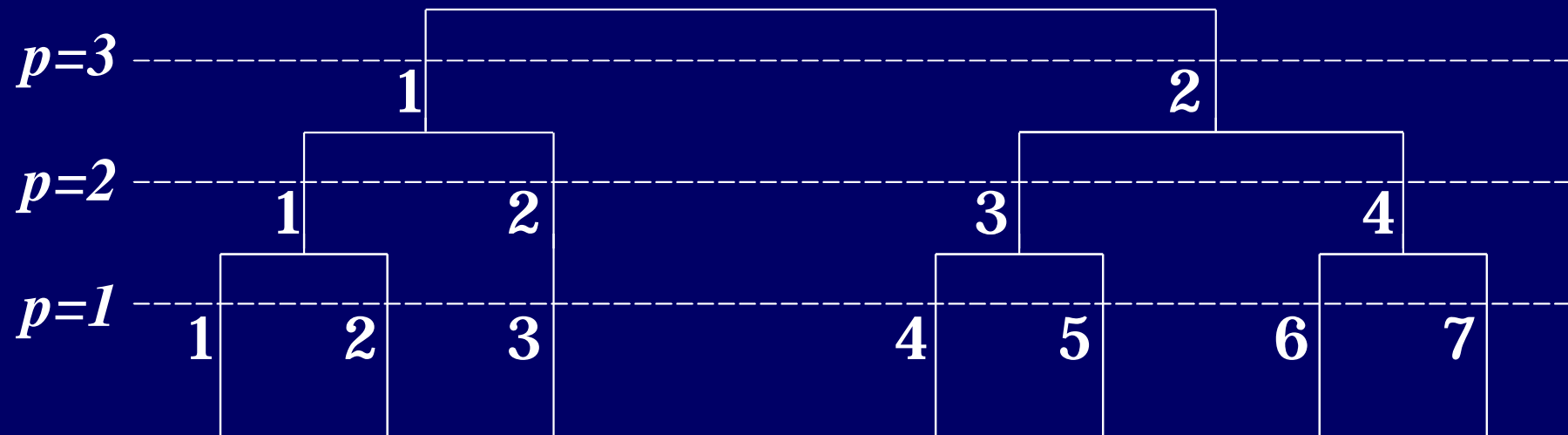
# Hierarchical Multiple Factor Analysis

“HMFA”

# “Multi-nested-block data”

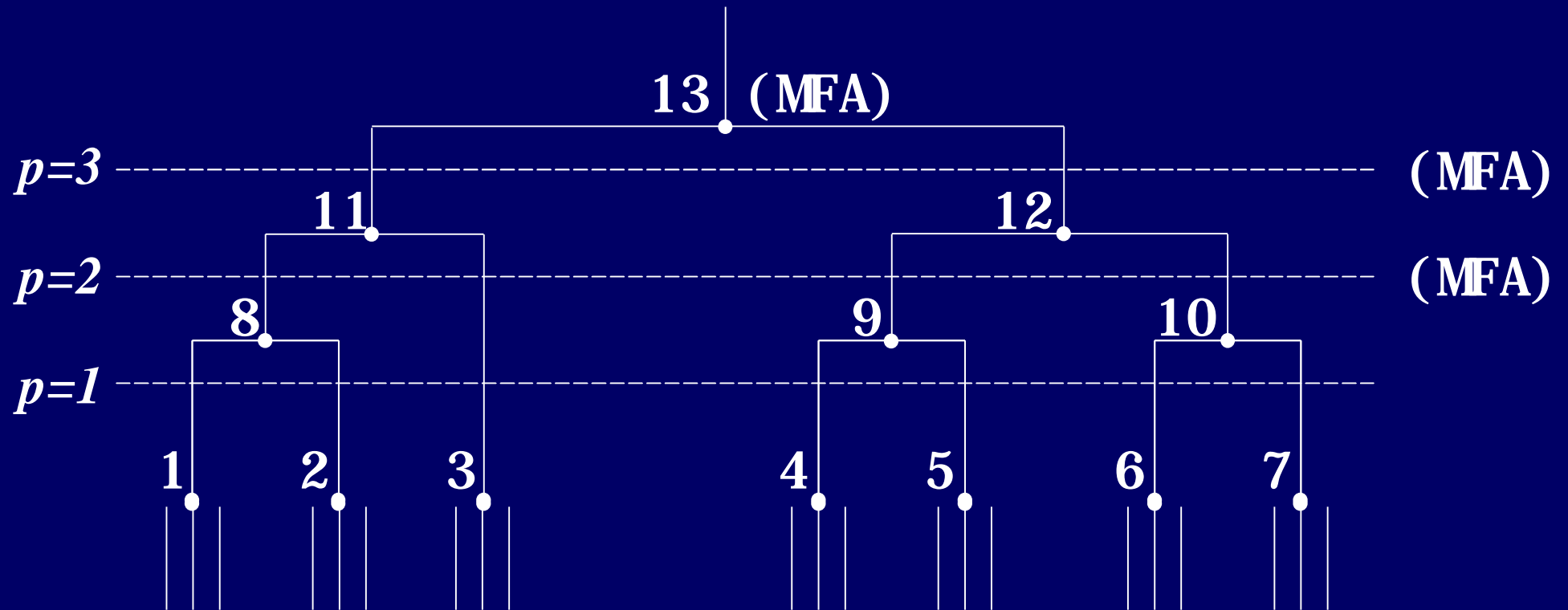
|    |         |         |         |         |         |         |         |
|----|---------|---------|---------|---------|---------|---------|---------|
| P3 | $X_1^3$ |         |         | $X_2^3$ |         |         |         |
| P2 | $X_1^2$ |         | $X_2^2$ | $X_3^2$ |         | $X_4^2$ |         |
| P1 | $X_1^1$ | $X_2^1$ | $X_3^1$ | $X_4^1$ | $X_5^1$ | $X_6^1$ | $X_7^1$ |

# Hierarchy structure

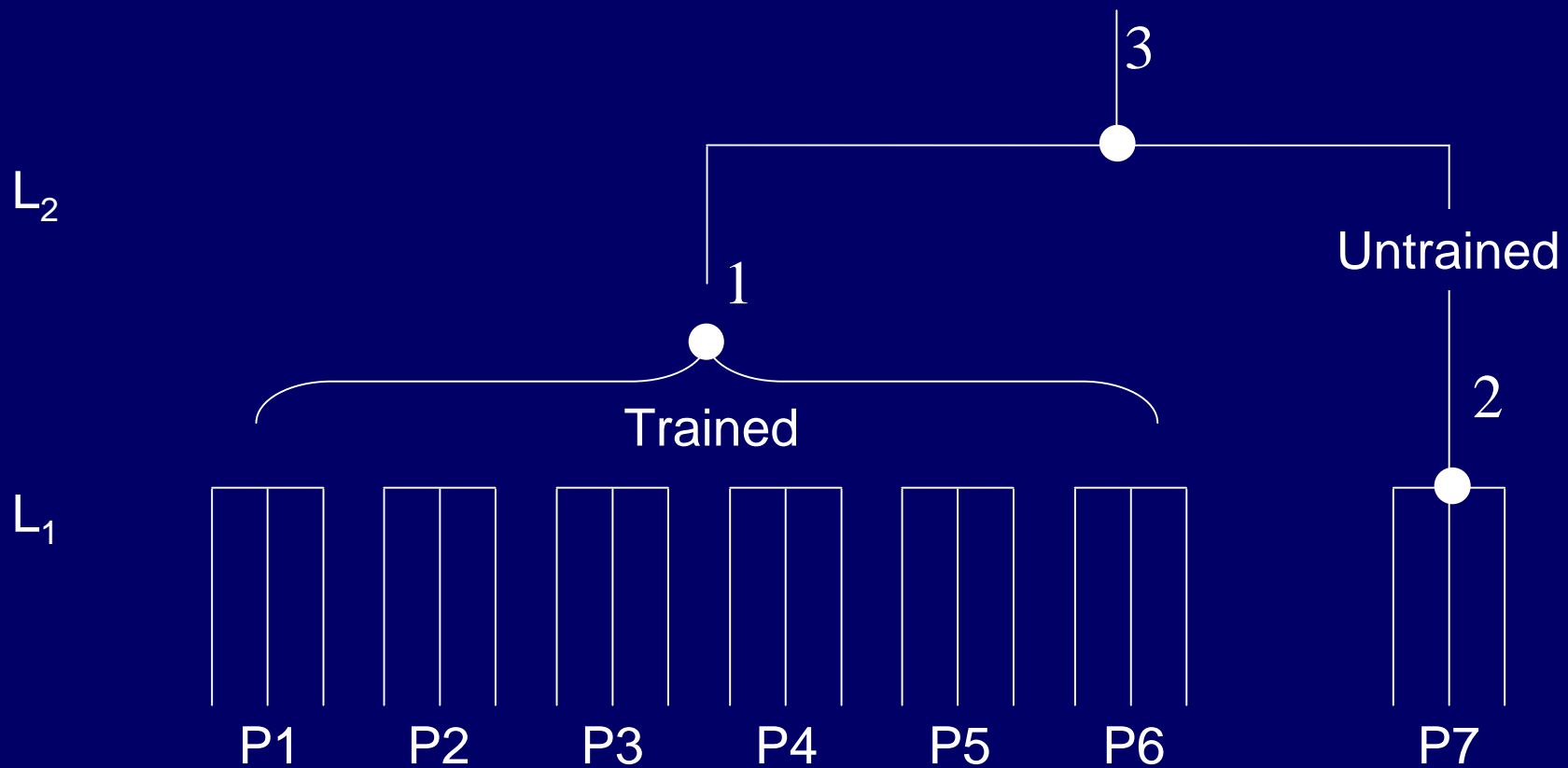


Hierarchical Multiple Factor Analysis

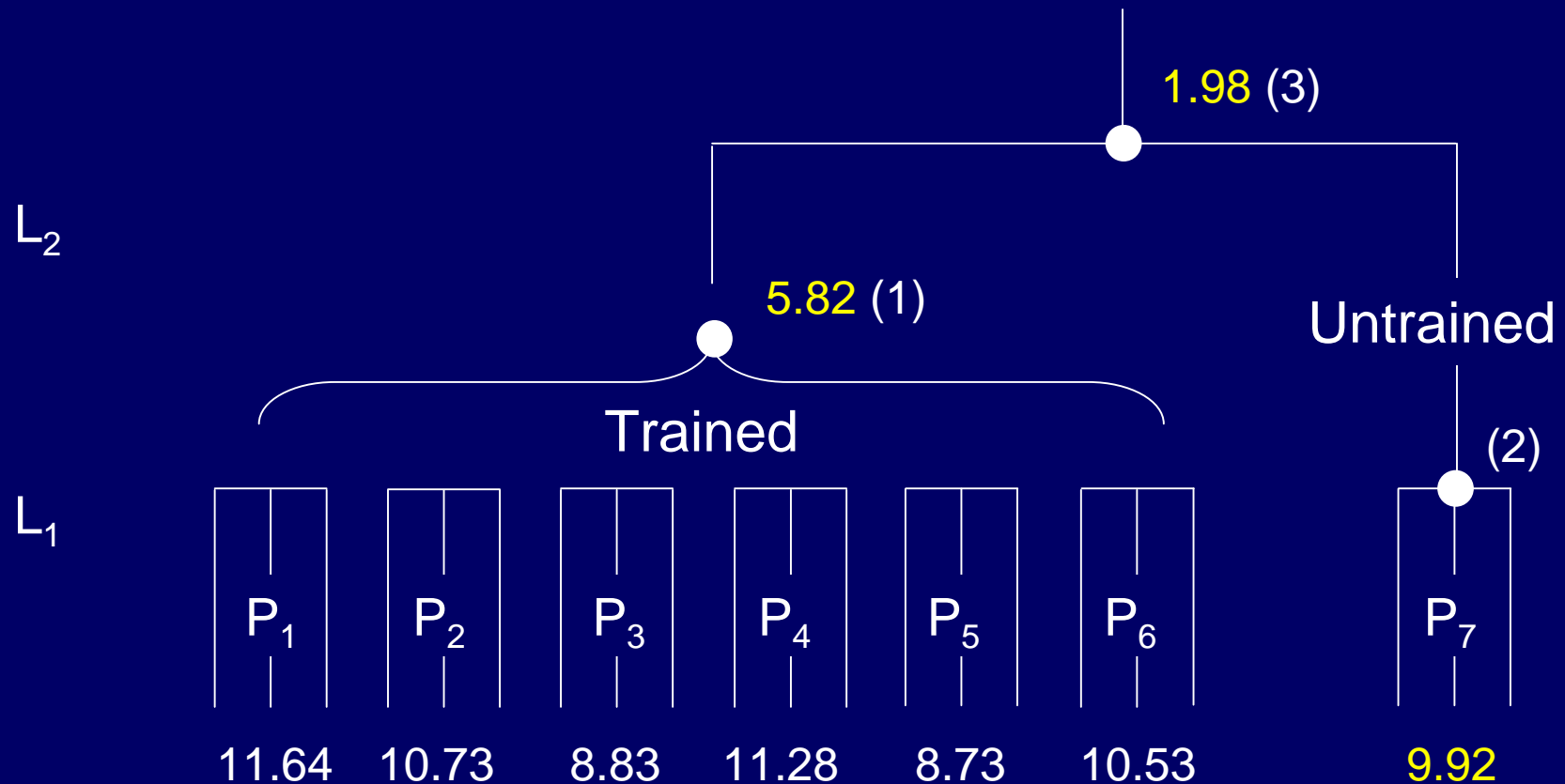
# What you have in mind...



# Back to our panellists



# Taking into account of a hierarchy structure



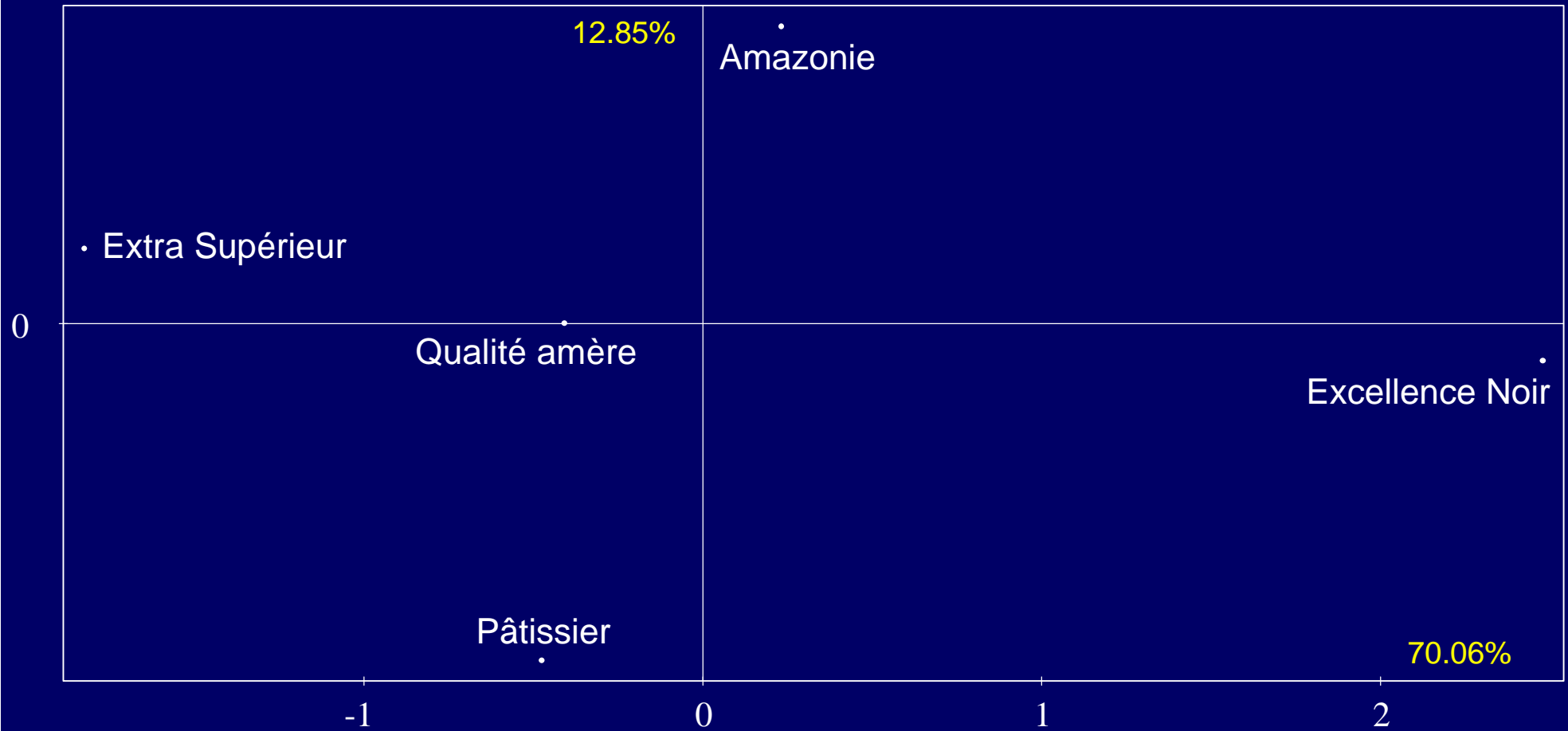
# RV Coefficient

|             | Panel 1 | Panel 2 | Panel 3 | Panel 4 | Panel 5 | Panel 6 | Panel 7 | Panel (1-6) |
|-------------|---------|---------|---------|---------|---------|---------|---------|-------------|
| Panel 1     | 1,000   |         |         |         |         |         |         |             |
| Panel 2     | 0,920   | 1,000   |         |         |         |         |         |             |
| Panel 3     | 0,890   | 0,857   | 1,000   |         |         |         |         |             |
| Panel 4     | 0,909   | 0,918   | 0,953   | 1,000   |         |         |         |             |
| Panel 5     | 0,838   | 0,887   | 0,928   | 0,943   | 1,000   |         |         |             |
| Panel 6     | 0,921   | 0,952   | 0,932   | 0,979   | 0,951   | 1,000   |         |             |
| Panel 7     | 0,925   | 0,937   | 0,884   | 0,929   | 0,859   | 0,954   | 1,000   |             |
| Panel (1-6) | 0,945   | 0,954   | 0,961   | 0,984   | 0,959   | 0,990   | 0,947   | 1,000       |

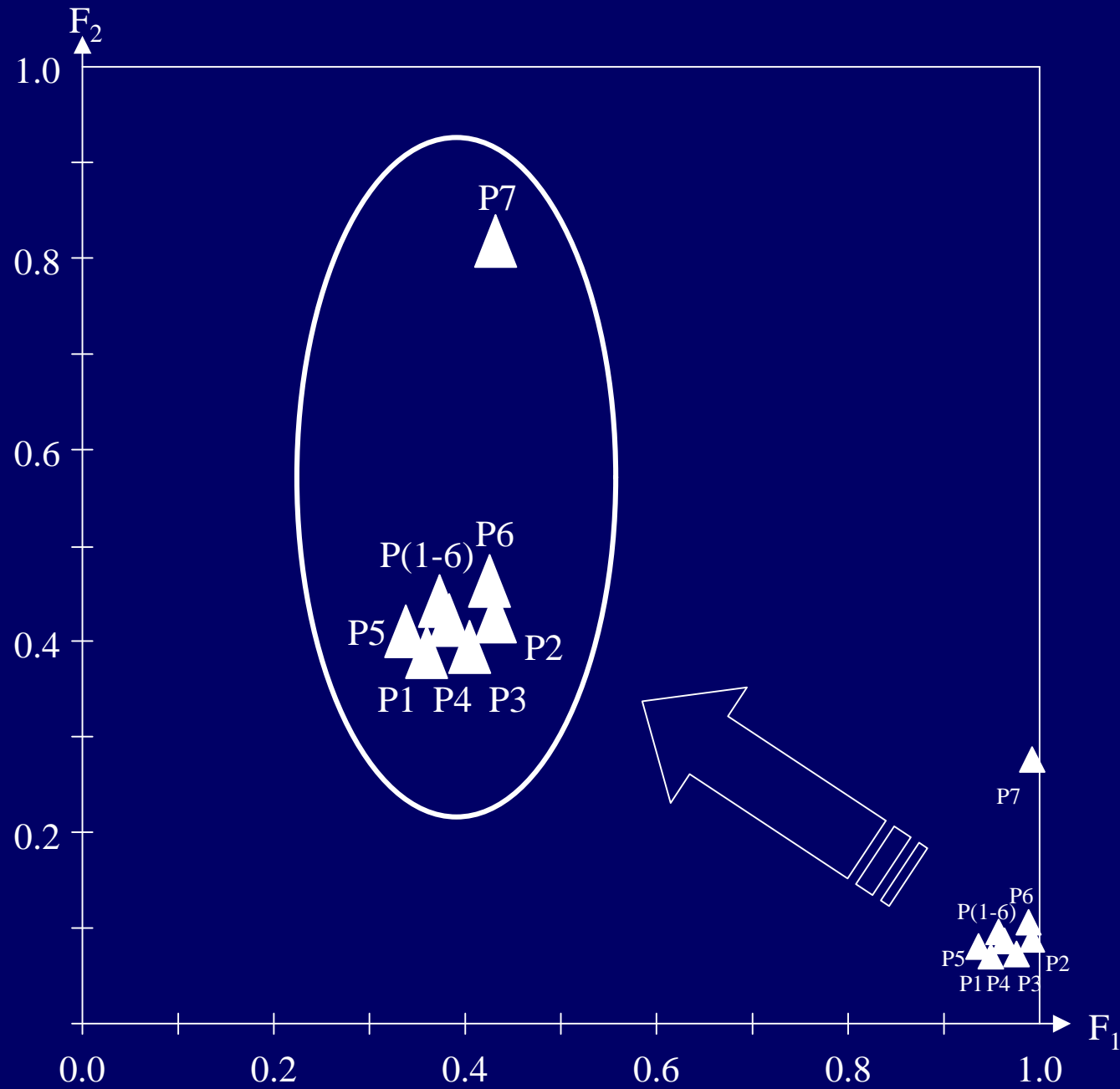
# Canonical variables

|           | F <sub>1</sub> | F <sub>2</sub> | F <sub>3</sub> | F <sub>4</sub> |
|-----------|----------------|----------------|----------------|----------------|
| Panel 1   | 0,977365       | 0,453954       | 0,643247       | 0,540047       |
| Panel 2   | 0,980822       | 0,489264       | 0,802266       | 0,966610       |
| Panel 3   | 0,984523       | 0,568679       | 0,880273       | 0,836682       |
| Panel 4   | 0,988669       | 0,652865       | 0,866238       | 0,463138       |
| Panel 5   | 0,974778       | 0,777716       | 0,977268       | 0,775938       |
| Panel 6   | 0,994726       | 0,918507       | 0,947662       | 0,569389       |
| Panel 1-6 | 0,996470       | 0,986689       | 0,984770       | 0,855531       |
| Panel 7   | 0,996473       | 0,998623       | 0,942714       | 0,435152       |

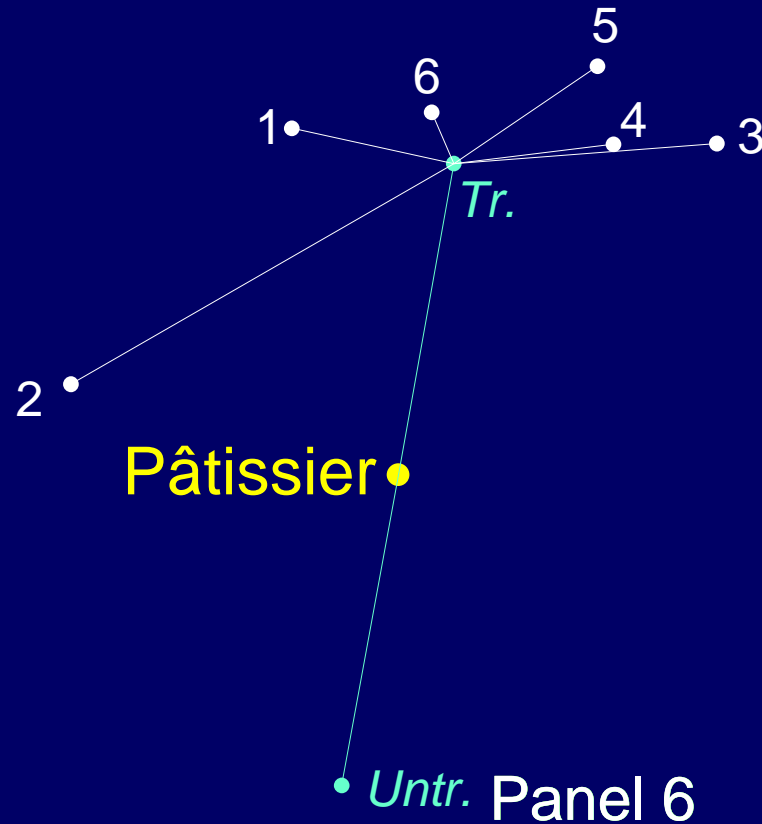
# Graphical representation



# Graphical representation



# Graphical representation

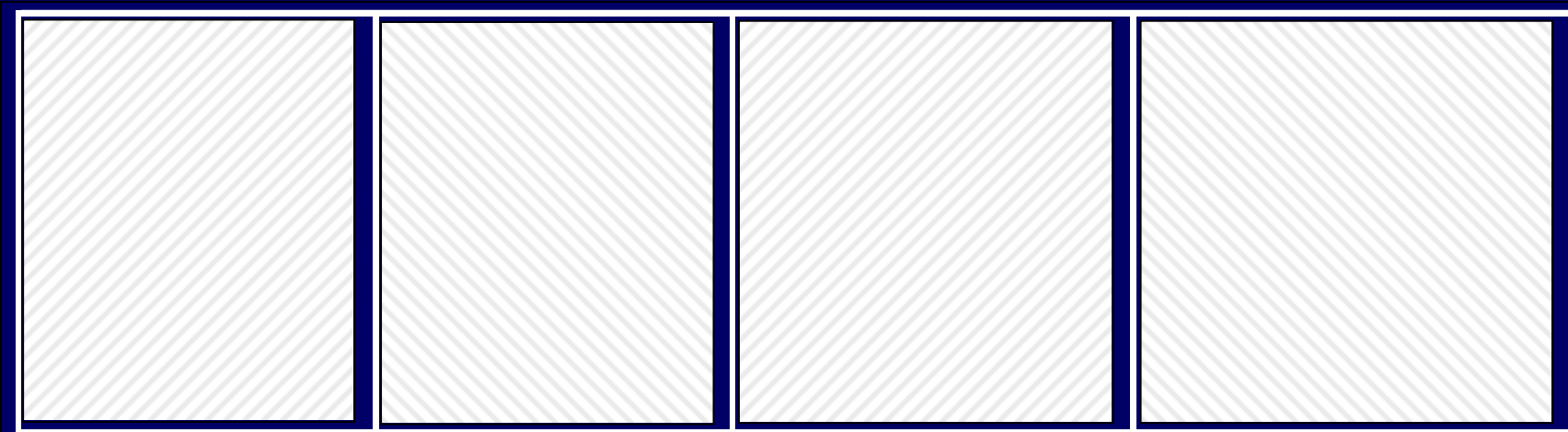


Panel 1

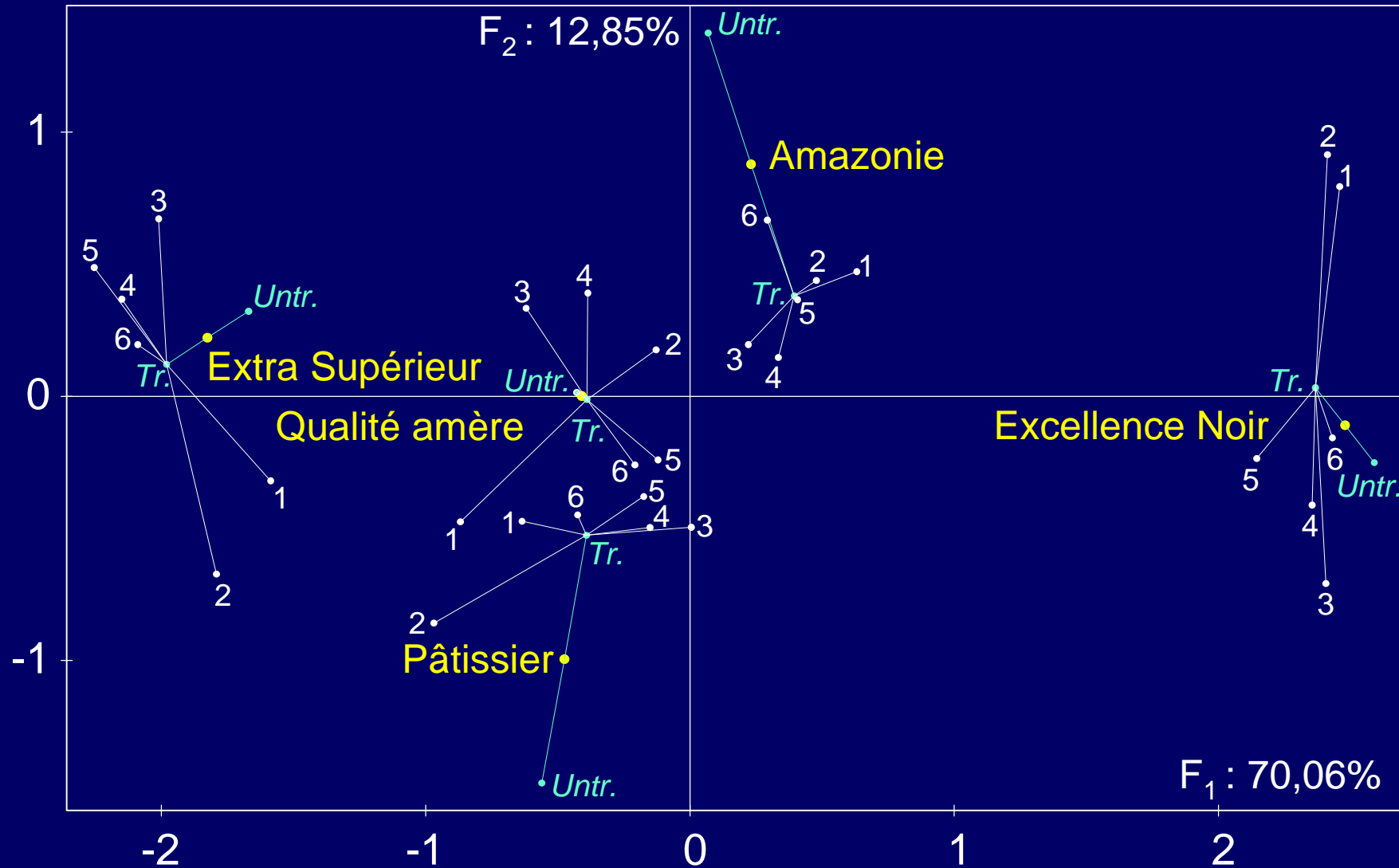
*Untr.* Panel 6

Untrained Panel

Chocolates



# Graphical representation



# Conclusion and perspectives

- **Classic**: satisfactory reproducibility of the trained panels
- **Polemical**: proximity of the sensory profiles established by the untrained panel and the consensus
- **Methodical**: hierarchies are everywhere
- **Perspectives**: extension to other multi-block methods