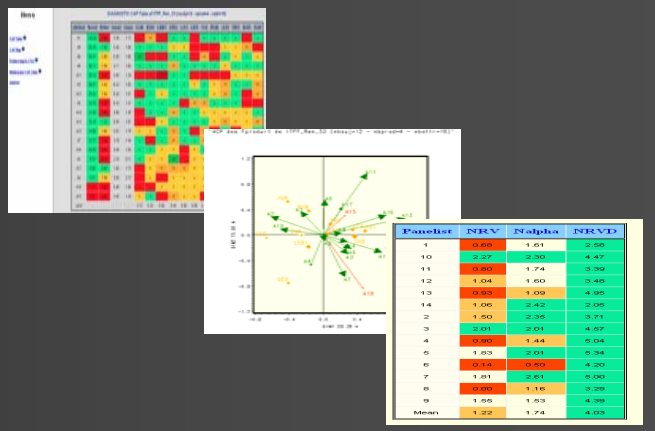


Panelist	NRV	Nalpha	NRVD
1	0.66	1.61	2.58
10	2.27	2.20	4.47
11	0.80	1.74	3.39
12	1.04	1.60	3.46
13	0.93	1.09	4.95
14	1.06	2.42	2.05
2	1.50	2.35	3.71
3	2.01	2.01	4.57
4	0.90	1.44	5.04
5	1.83	2.01	5.34
6	0.14	0.60	4.20
7	1.81	2.61	5.00
8	0.60	1.16	3.29
9	1.55	1.53	4.39
Mean	1.22	1.74	4.03

Multivariate analysis of a sensory database to learn about panel performances

Pascal Schlich

Centre Européen des Sciences du Goût
Dijon, France

Plan

I. Introducing the analysis of quantitative descriptive sensory data

II. Presenting results from ANOVA, MANOVA and CVA

III. Assessing panel homogeneity by RV coefficients and STATIS

IV. Introducing the SensoBase project

Part I.
Introducing the analysis of quantitative descriptive sensory data

1. Quantitative descriptive sensory analysis:

   a multivariate exercise !

2. Analysis of descriptive data:

   - panel performances

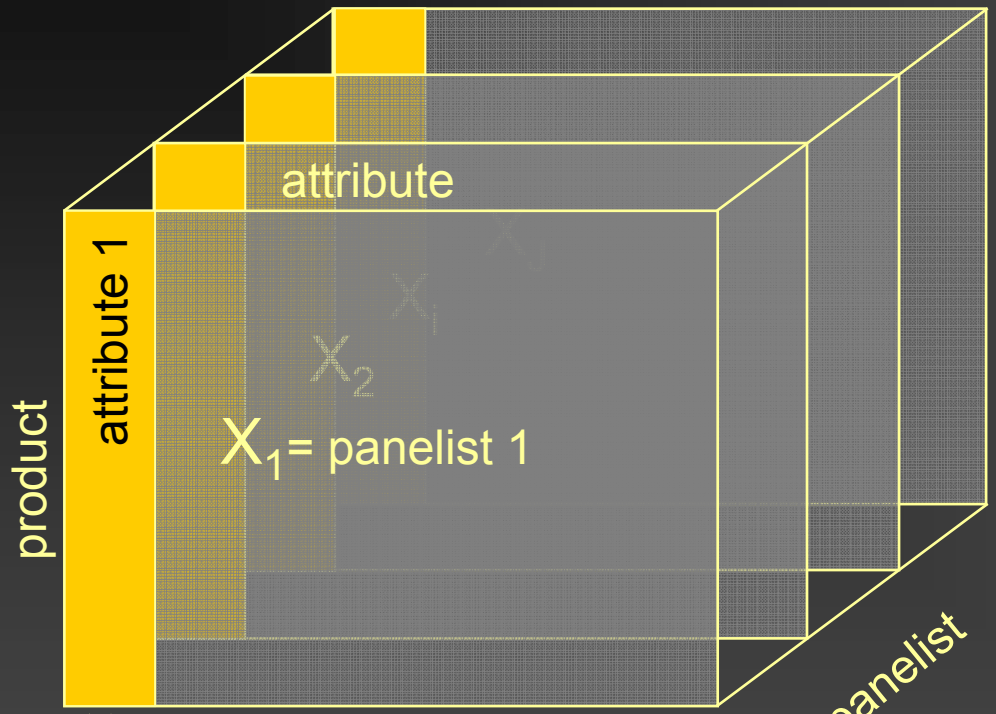
   - product differences

3. From ANOVA to MANOVA:

   - to quantify panel performances

   - to exhibit product differences

Sensory profiling data



Univariate and multivariate panel performances

Panel proficiency

Panelist proficiency

Individual univariate performances

Individual multivariate performances

Comparing individual univariate and multivariate performances

Product differences

3 basic performances

- Reliability - Repeatability
  - to give the same scores to the same products
    - replicates required
- Validity - Accuracy
  - for an individual: to be in agreement with the panel
    - In agreement with a “mean” of the panel
  - for the panel: not to include too many individuals who disagree
    - Concept of “panel homogeneity”
- Discrimination
  - to give scores different “enough” to different products
    - “Enough” means relatively to repeatability at individual level
    - “Enough” means relatively to panel homogeneity at panel level

No individual reliability ⇒ No panel validity ⇒ No product discrimination

4 cases to fill in

	Univariate	Multivariate
Panelist	Monitoring the panel Panel leader everyday job	
Panel	Validating the panel Panel leader objective	

The use of ANOVA, MANOVA and CVA

	Univariate	Multivariate
Panelist	**ANOVA model: Prod** ➢ Repeatability: $\sqrt{MS_E}$ (Root mean square of error) ➢ Validity: correlation with panel mean ➢ Discrimination: $F = MS_{Prod} / MS_E$	**MANOVA model: Prod** ➢ Discrimination: MANOVA test of Prod Canonical Variate Analysis (CVA) ➢ Dimensionality of product configuration Stepwise selection of attributes ➢ Saliency of product differences
Panel	**ANOVA model: Prod + Pan + Prod*Pan** ➢ Repeatability: $\sqrt{MS_E}$ (Root mean square of error) ➢ Validity: $F = MS_{Prod*Pan} / MS_E$ ➢ Discrimination: $F = MS_{Prod} / MS_{Prod*Pan}$	**MANOVA model: Prod + Pan + Prod*Pan** ➢ Homogeneity: MANOVA test of Prod*Pan ➢ Discrimination: MANOVA test of Prod Canonical Variate Analysis (CVA) ➢ Dimensionality of product configuration ➢ Product confidence ellipses on CVA map to contrast homogeneity and discrimination

Part II.
Presenting results from ANOVA, MANOVA and CVA

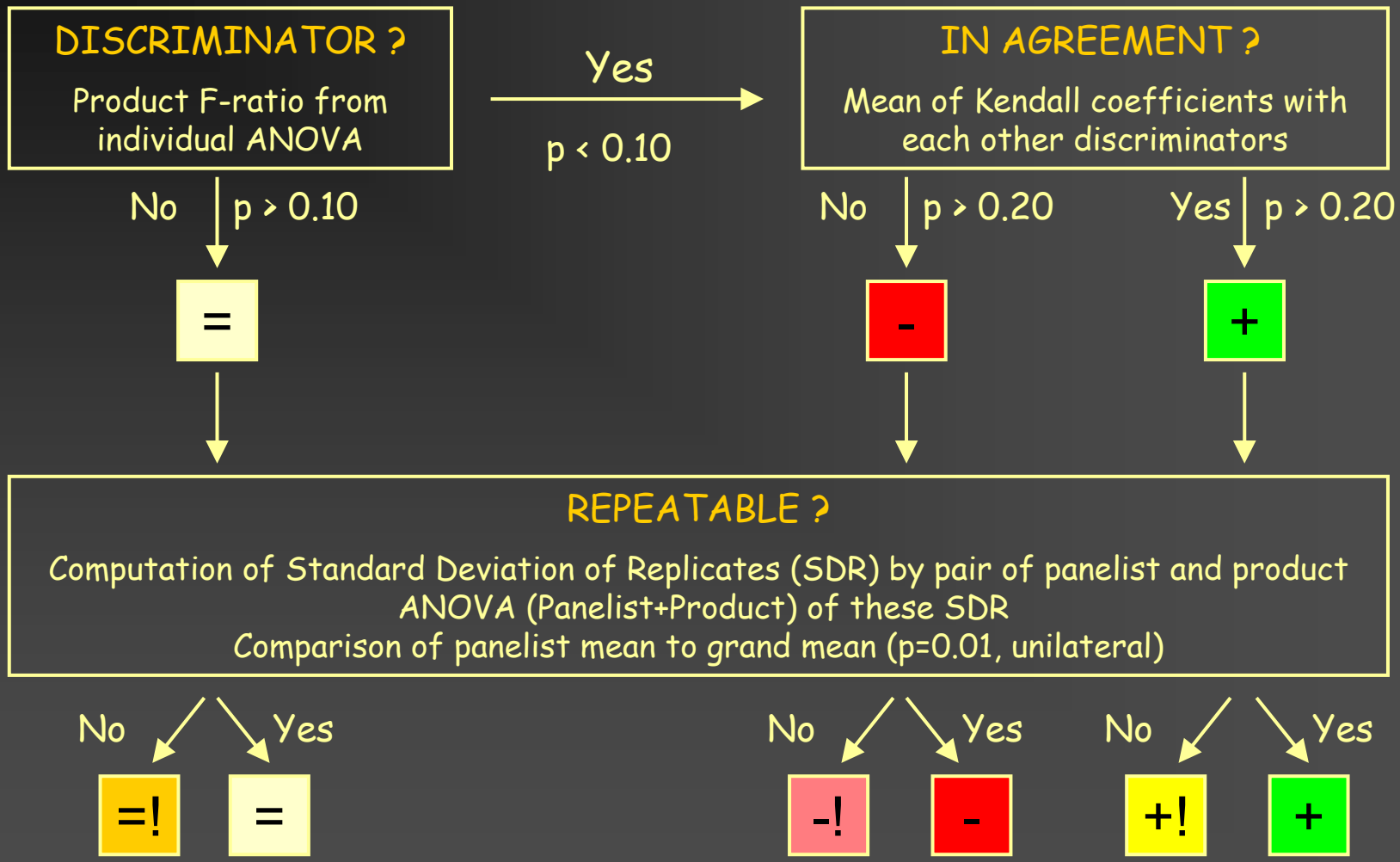
1. The CAP table:
   displaying panelist performances in a single table

2. The CAP map:
   mapping panelist performances on a single plot

3. The FLASH table:
   displaying product differences in a single table

4. The CVA plot:
   mapping product differences on a single map

CAP: Control of Panelist Performances



6 different diagnostics possible for each pair of panelist and attribute

(Schlich, 1997)

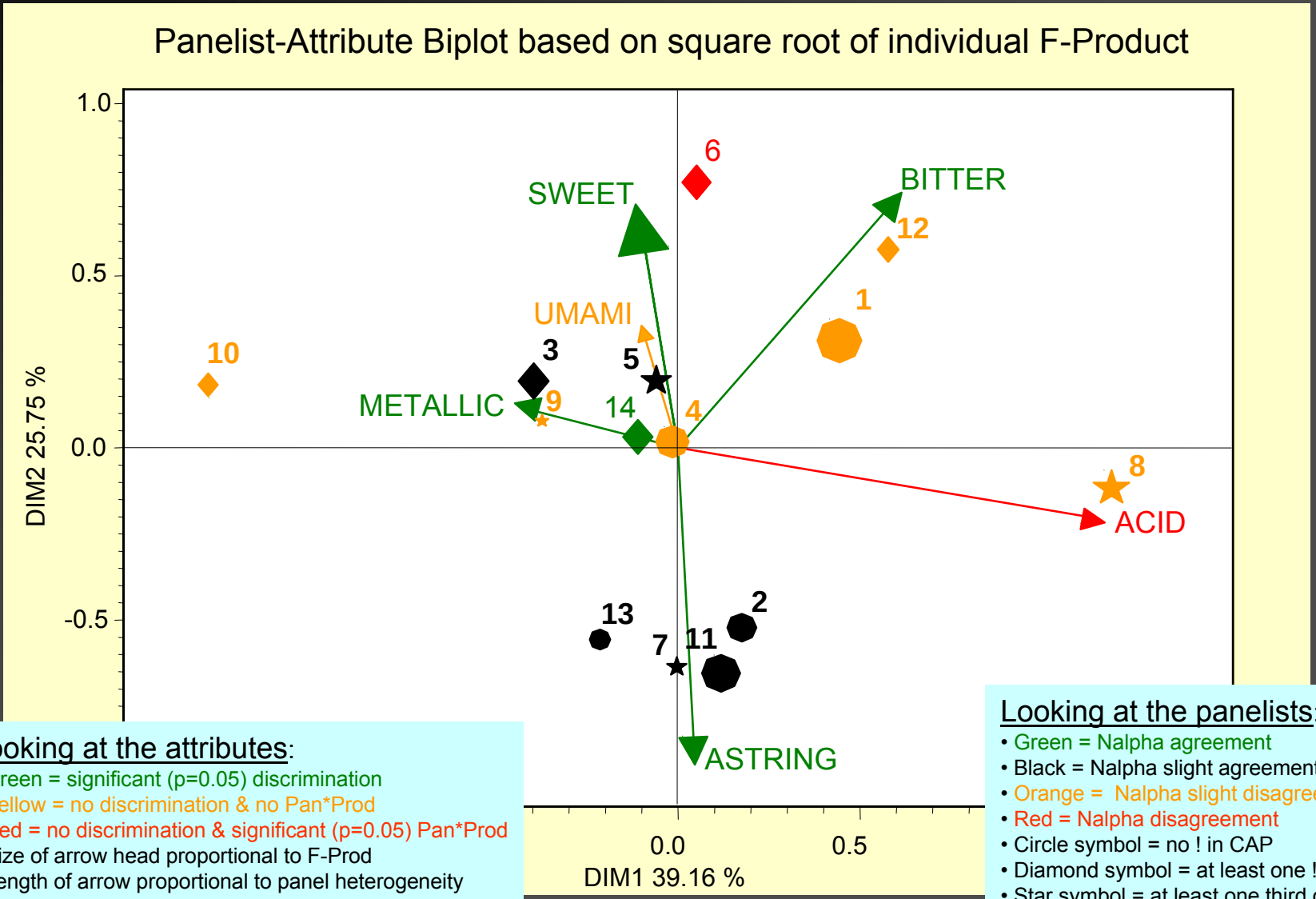
The CAP table

16 steamed potatoes profiled in duplicate - 10 texture attributes - 14 panelists

					PANELIST														
	MEAN	RMSE	FPROD	FINTER	7	5	4	10	11	9	3	8	1	14	12	2	13	6	+ / - / !
RANKF					3.73	5.09	5.36	6.00	6.73	7.27	7.45	7.64	8.45	8.64	8.82	8.91	9.64	11.27	
FIRM	37.50	18.80	17.17	1.51	+	+	+	+	=	+	+	+	+	+	+	+	=	=!	11/0/1
MASHABLE	40.20	25.50	13.92		+	=	=	=	+	+	+	=!	=	+	=	+	=	=	6/0/1
MOIST	46.90	21.10	11.65		+	+	+	+	+	+	+	+	=	=!	=	=	=	+	9/0/1
GRAINY	26.90	20.60	11.58		+	+	+	=	=	+	+	=	=!	+	+	=	=	=!	7/0/2
MEALY	42.30	25.20	11.45		+	=	+	=	=	=	=	=	=!	=	=	+	=	=	3/0/1
COMPACT	48.60	18.60	8.58	1.77	+	+	+	+	-	=	=!	+	=	+	=!	+	+	=!	8/1/3
HOMOGEN	59.00	22.40	7.41		+	+!	+	=	+	+	+	+	=	=!	=	=	+	=	8/0/2
GREASY	15.30	15.70	5.27	1.47	+!	+	+	+	=	=!	=!	=!	=	=	=	=	=	=	4/0/4
PASTY	34.70	22.80	3.68	1.32	+	=!	+	-	+	=!	+	=	+	=!	=!	=	=	=!	5/1/5
STICKY	27.60	22.50	2.95	1.31	=	+	+	-	=	=	=	+	=	=!	=	=	=	=	3/1/1
ASTRING	29.30	21.90	2.58		-	=!	-	=	+	=!	=	=!	-	=	=!	-	+	=	2/4/4
+ / - / !					9/1/1	7/0/3	9/1/0	4/2/0	5/1/0	5/0/3	6/0/2	5/0/3	2/1/2	4/0/4	2/0/3	4/1/0	3/0/0	1/0/4	

- MEAN is the mean scores over products and panelists.
- RMSE is the root mean square of errors in the 2-way mixed model: PROD+PAN+PROD*PAN
- FPROD and FINTER are PRODUCT and PROD*PAN F-ratio in that model, printed when significant (p=0.05)
- RANKF is the mean of individual FPROD ranks over the attributes.
- Attributes and subjects are sorted from the most to the least discriminative power
- + / - / ! = number of + / number of - / number of !

The CAP map to summarize panelist performances
Potato taste



Looking at the attributes:
- Green = significant (p=0.05) discrimination
- Yellow = no discrimination & no Pan*Prod
- Red = no discrimination & significant (p=0.05) Pan*Prod
- Size of arrow head proportional to F-Prod
- Length of arrow proportional to panel heterogeneity towards discrimination

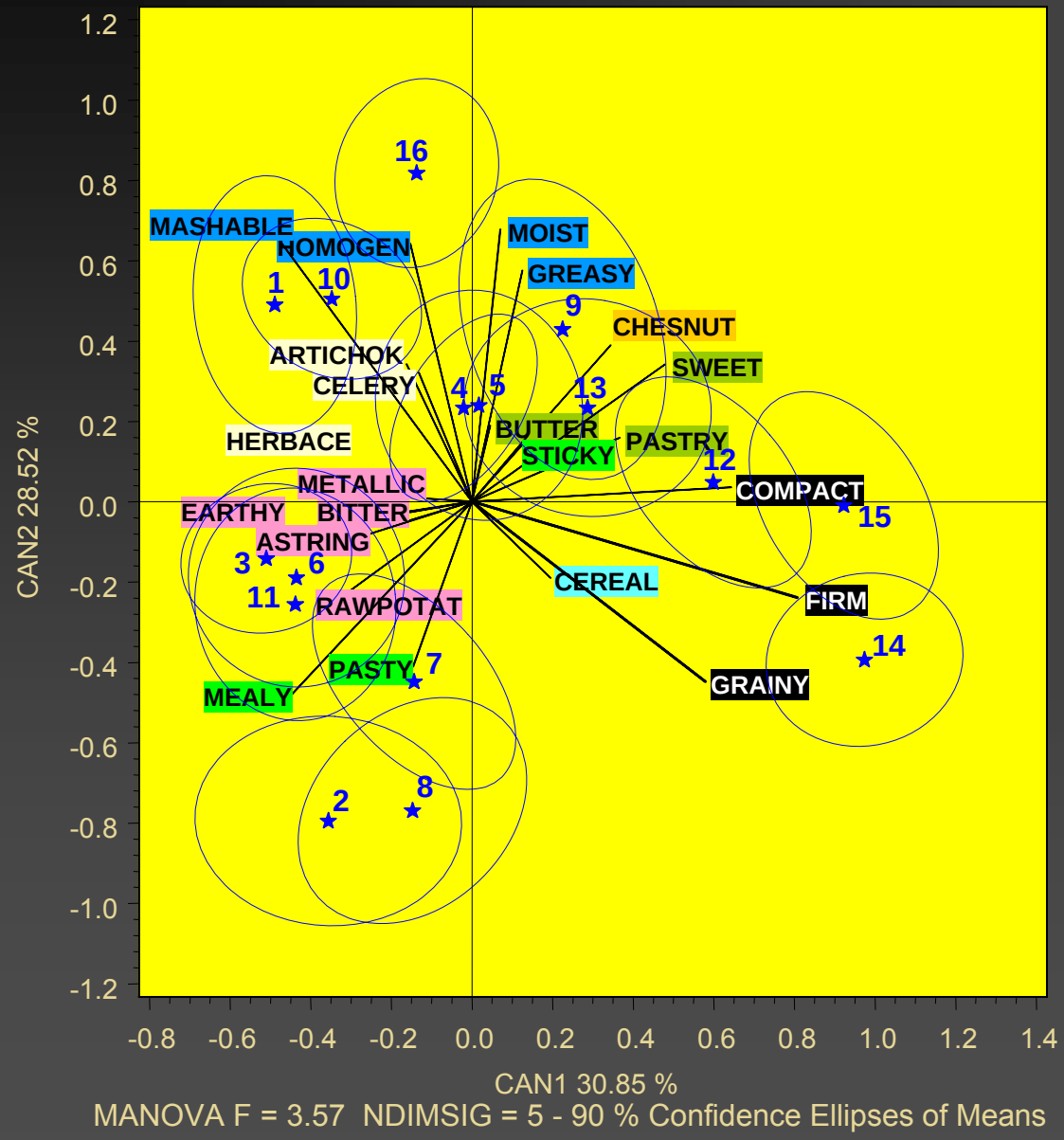
Looking at the panelists:
- Green = Nalpha agreement
- Black = Nalpha slight agreement
- Orange = Nalpha slight disagreement
- Red = Nalpha disagreement
- Circle symbol = no ! in CAP
- Diamond symbol = at least one ! in CAP
- Star symbol = at least one third of ! in CAP
- Symbol size proportional to F-Prod in MANOVA of the 6 attributes

The FLASH Table:
a practical display of product differences

16 steamed potatoes - 14 panelists - 23 texture, flavor and taste attributes

ATTRIBUTE	F	PROB	GMEAN	16	3	10	1	6	2	8	5	11	4	13	9	7	12	15	14
FIRM	17.20	0.0000	37	19-	21-	21-	22-	22-	24-							49+	63+	69+	77+
COMPACT	8.58	0.0000	49		29-		40-	27-	39-	36-			59+				60+	73+	73+
GRAINY	11.60	0.0000	27	6-	14-	12-	15-	18-		52+							39+	51+	46+
PASTRY	4.43	0.0000	19				10-			10-						4-	28+	44+	
SWEET	9.49	0.0000	42	54+	30-	56+	27-	32-		29-	33-			54+	56+	22-	51+	69+	
BUTTER	1.74	0.0470	17								25+							25+	9-
CHESNUT	6.02	0.0000	41						29-	21-		30-		53+	58+			63+	58+
CEREAL	2.33	0.0040	26				17-												44+
ARTICHOK	2.63	0.0010	40			52+			23-										28-
CELERY	1.86	0.0290	17			29+									24+				
HERBACE	3.23	0.0000	31						20-			49+				42+		20-	13-
EARTHY	5.78	0.0000	30				40+	42+		41+		43+		21-		47+	19-	14-	
ASTRING	2.58	0.0010	29													44+	18-		
BITTER	2.88	0.0000	27						18-	36+				17-		46+			
RAWPOTAT	2.81	0.0010	23				32+			38+								10-	
METALLIC	1.78	0.0400	15				23+	26+											
PASTY	3.68	0.0000	35				22-	47+	56+								23-	24-	
MEALY	11.40	0.0000	42	29-	60+			54+	78+	55+	31-	54+		32-			26-	27-	31-
STICKY	2.95	0.0000	28	37+	16-							17-			47+				
MASHABLE	13.90	0.0000	40	69+		58+	66+		29-	25-	51+					26-	27-	16-	11-
HOMOGEN	7.41	0.0000	59	78+		72+	70+		42-	31-	69+		70+		69+			45-	
GREASY	5.27	0.0000	15	26+	7-			8-	2-	7-	26+	7-	25+		25+	8-			
MOIST	11.70	0.0000	47	68+		63+	69+	38-	17-	29-	58+	35-				33-	56+		34-

CVA of 16 steamed potatoes



MANOVA F = 3.57 NDIMSIG = 5 - 90 % Confidence Ellipses of Means

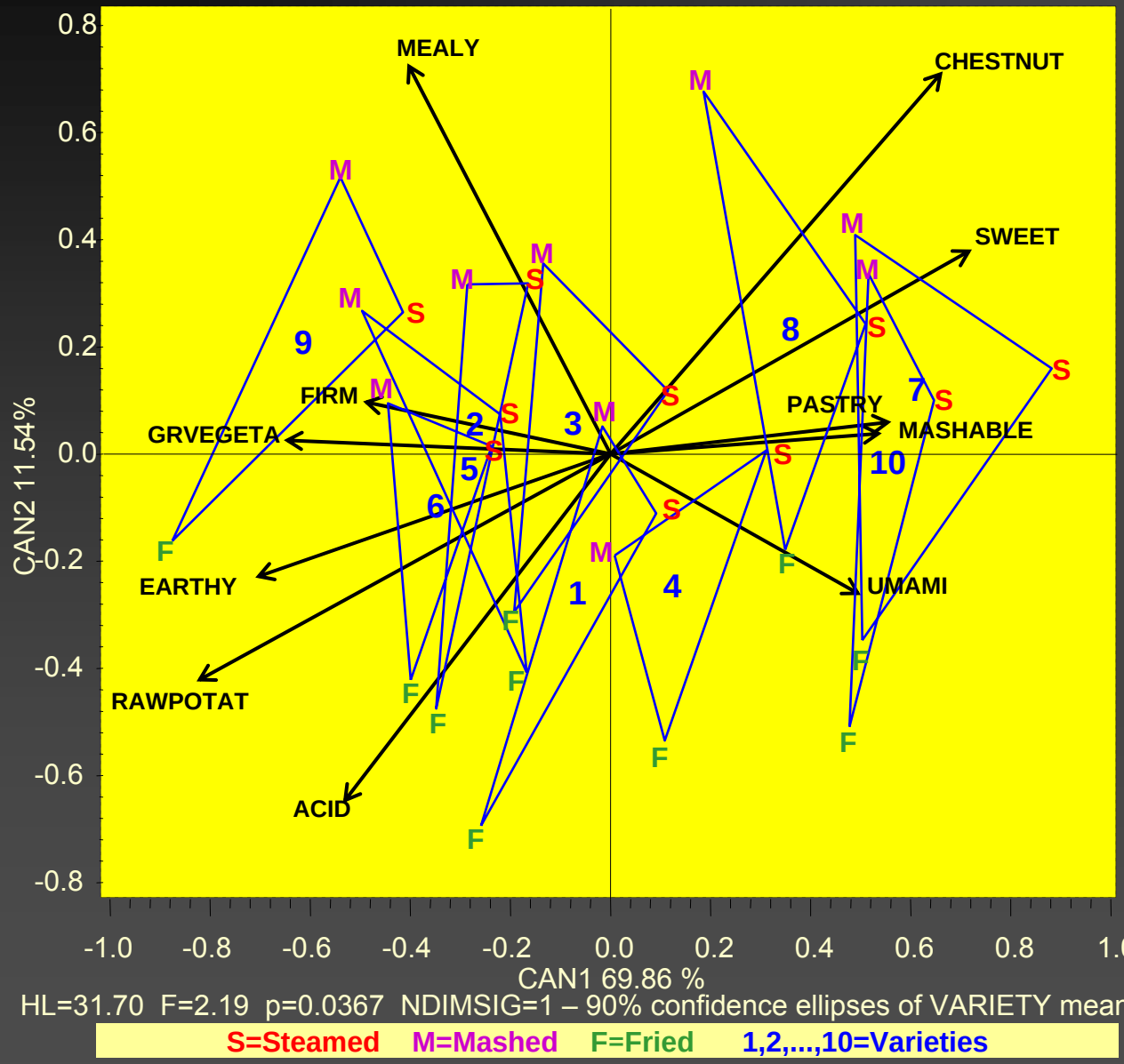
The colors on the attibute names refer to 8 attribute clusters obtained by VARCLUS in SAS

Attributes are located on the canonical map thanks to their correlations with canonical components

The confidence ellipses suggest that product 14 and 15 are not really different, whereas product 12 is less firmer than 14 but not than 15

CVA of Potato Varieties

11 significant (p=0.10) attributes out of 22



HL=31.70 F=2.19 p=0.0367 NDIMSIG=1 – 90% confidence ellipses of VARIETY means

S=Steamed M=Mashed F=Fried 1,2,...,10=Varieties

What are the sensory differences among varieties robust over the cooking modes ?

These differences generate a single dimension only…

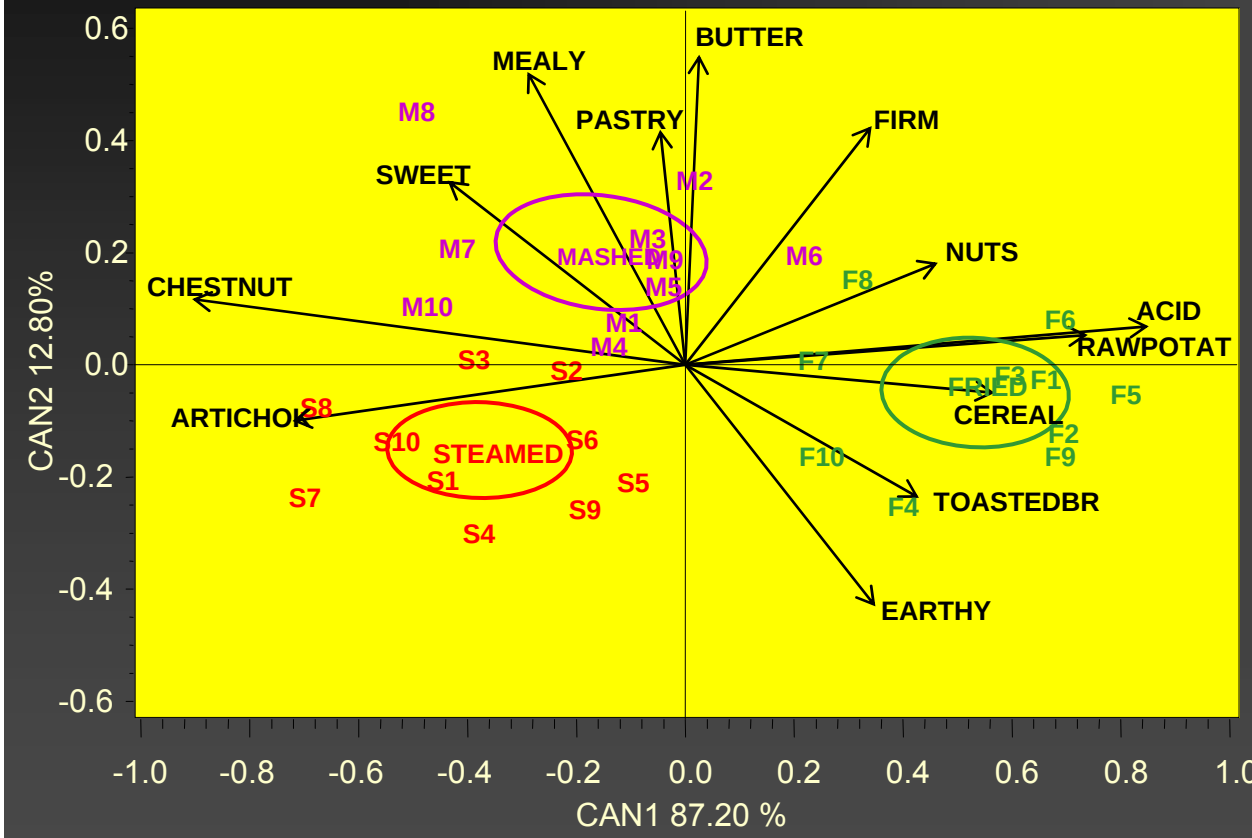
Which goes from firmness and veg on the left to mashable and sweet on the right

And which distinguishes varieties 9 on the left to varieties 7,10 and 8 on the right

Axis 2 is a cooking mode effect and thus was not significant for discrimination among varieties

CVA of Cooking Modes

13 significant (p=0.10) attributes out of 22



HL=51.06 F=10.97 p=0.0015 NDIMSIG=2 – 90% confidence ellipses of COOKING MODE means

S=Steamed M=Mashed F=Fried 1,2,...,10=Varieties

What are the sensory differences among the cooking modes robust over varieties ?

These differences generate a two-dimensional space…

Which distinguishes the fried potatoes as being more acid and less sweet than the others.

On the flavor aspect:
fried ⇒ raw potato and cereal,
mashed ⇒ butter and pastry,
steamed ⇒ artichok and chesnut.

Discrimination and dimensionality of individual potato configurations

	From 14 flavor attributes			From 6 taste attributes			From 10 texture attributes		
Subject	NSEL	MANOVA F	NCAN	NSEL	MANOVA F	NCAN	NSEL	MANOVA F	NCAN
1	5	5.77	3	2	3.04	1	2	2.77	1
2	0	.	0	1	2.60	1	2	2.59	2
3	3	6.48	2	2	2.70	2	3	2.48	2
4	0	.	0	2	2.38	1	2	3.82	2
5	3	2.63	2	4	1.92	1	3	6.24	3
6	6	3.96	3	2	3.60	2	2	2.18	1
7	6	2.13	4	1	2.88	1	3	3.50	3
8	3	3.84	2	1	6.10	1	3	2.68	2
9	1	2.60	1	0	.	0	3	2.86	2
10	2	3.29	2	1	2.54	1	4	7.87	4
11	5	2.72	4	1	2.75	1	2	3.16	2
12	4	3.24	2	1	2.82	1	2	6.41	2
13	1	3.98	1	1	1.97	1	2	3.18	2
14	1	10.67	1	2	2.57	1	4	4.96	3
Panel	**14**	**1.98**	**3**	**6**	**2.68**	**2**	**10**	**5.68**	**3**

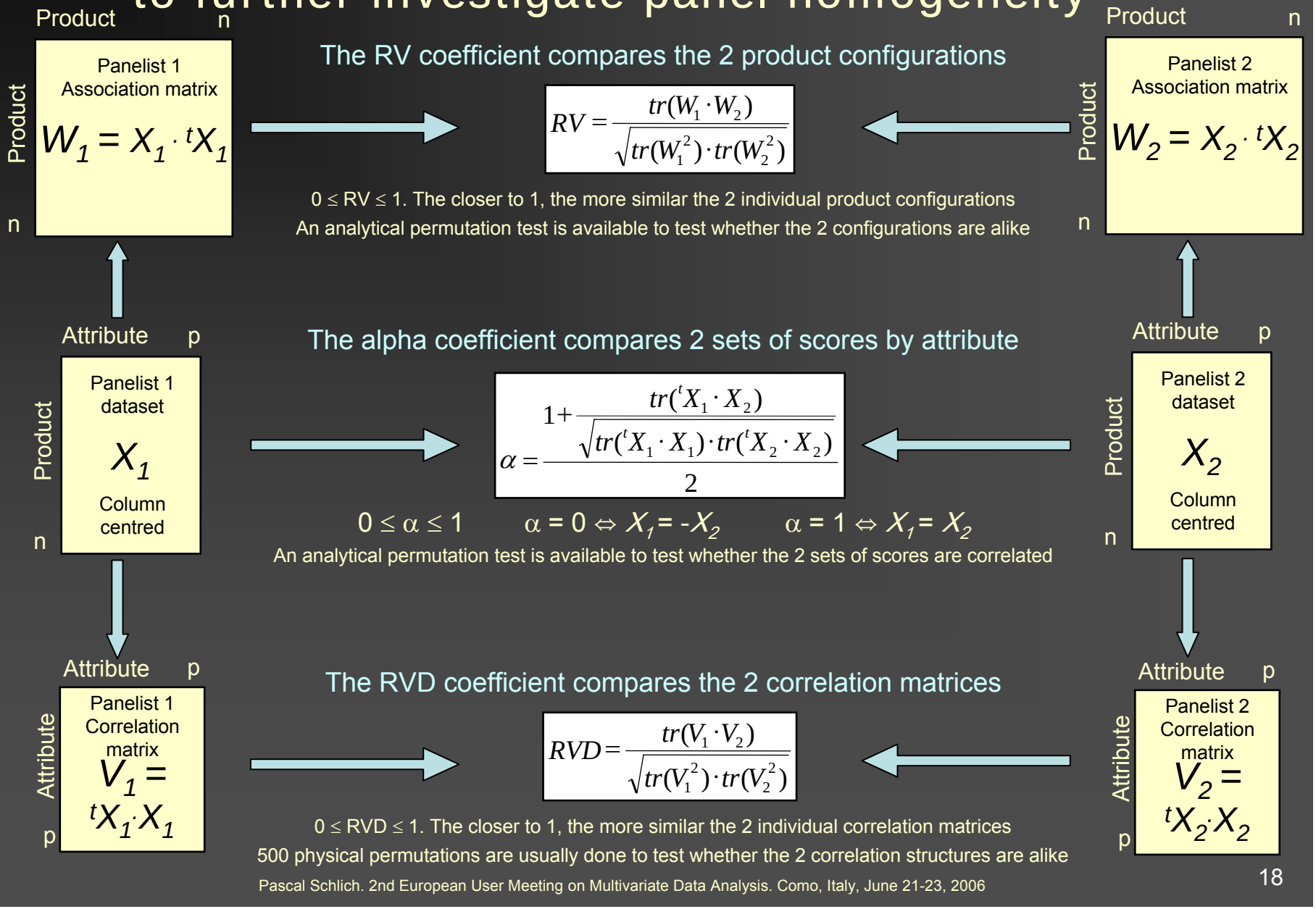
NSEL: number of selected attributes by Proc STEPDISC (SLE=0.15 SLS=0.15) - No selection done at panel level
NCAN: number of significant (p=0.10) canonical axes
MANOVA F: **green if p<0.05**, no color if 0.05<p<0.15 and **in red when non significant**

Part III.
Assessing panel homogeneity by RV coefficient and STATIS

1. Three RV coefficients for comparing panelists based on:
   - product configurations
   - attribute correlations
   - sensory scores

2. Three STATIS methods

The Conf-Scor-Corr RV framework to further investigate panel homogeneity



Conf-Scor-Corr agreement on potato flavor

14 panelists - 16 products - 14 Flavor attributes

Panelist	RV	α	RVD	NRV	Nα	NRVD
1	0.29	0.56	0.53	0.40	1.25	2.46
2	0.38	0.59	0.54	1.32	2.04	2.51
3	0.37	0.59	0.53	1.91	1.74	3.53
4	0.35	0.56	0.51	1.07	1.20	3.12
5	0.37	0.58	0.60	1.35	1.62	3.39
6	0.28	0.51	0.52	-0.02	0.18	2.98
7	0.37	0.61	0.52	1.20	2.46	3.12
8	0.29	0.53	0.41	0.21	0.80	2.45
9	0.33	0.57	0.54	1.17	1.39	2.95
10	0.40	0.61	0.54	2.15	2.27	3.14
11	0.27	0.57	0.52	0.61	1.39	2.56
12	0.32	0.55	0.45	0.75	1.29	2.67
13	0.27	0.54	0.50	0.40	0.72	3.22
14	0.36	0.57	0.44	0.72	1.66	2.14
Mean	**0.33**	**0.57**	**0.51**	**0.94**	**1.43**	**2.88**

- RV, alpha and RVD are indices of agreement in [0,1] in terms of product configuration, product scores and attribute correlations
- NRV, Nalpha and NRVD are standardized deviation of these indices to their distribution under permutation
- NRV and Nalpha are obtained by formulas whereas 500 permutations were drawn for NRVD

Green = Agreement – White = Slight Agreement
Orange = Slight Disagreement - Red = Disagreement

Agreement on potato flavor, taste and texture

	14 Flavor attributes			6 Taste attributes			10 Texture attributes		
Panelist	NRV	N α	NRVD	NRV	N α	NRVD	NRV	N α	NRVD
1	0.40	1.25	3.24	0.56	1.10	1.00	3.71	3.07	2.68
2	1.32	2.04	3.03	1.07	1.58	1.21	4.50	3.91	4.13
3	1.91	1.74	3.85	1.01	1.55	0.42	3.65	3.49	4.04
4	1.07	1.20	3.11	0.01	1.09	1.62	3.53	3.56	2.66
5	1.35	1.62	4.23	1.16	1.63	1.54	4.27	3.63	3.69
6	-0.02	0.18	2.94	-0.19	0.66	1.30	2.69	2.88	3.68
7	1.20	2.46	2.94	1.75	1.55	0.59	4.20	3.68	3.22
8	0.21	0.80	1.82	0.98	1.19	0.87	3.39	3.44	3.31
9	1.17	1.39	3.57	0.66	1.09	0.40	2.44	2.92	3.35
10	2.15	2.27	3.50	1.50	1.37	0.35	2.83	3.29	3.20
11	0.61	1.39	3.57	0.73	1.63	0.50	2.64	2.75	2.79
12	0.75	1.29	2.48	0.73	1.23	1.42	3.31	3.08	2.95
13	0.40	0.72	3.46	1.53	1.61	1.60	2.08	2.67	3.47
14	0.72	1.66	1.42	1.21	1.86	1.68	3.98	3.56	2.60
Mean	**0.94**	**1.43**	**3.08**	**0.91**	**1.37**	**1.04**	**3.37**	**3.28**	**3.27**

Green = Agreement – White = Slight Agreement
Orange = Slight Disagreement - Red = Disagreement

Defining the STATIS compromise

Panelist κ

Panelist k

RV coefficients among the k panelists

RV

Eigen vector → β =

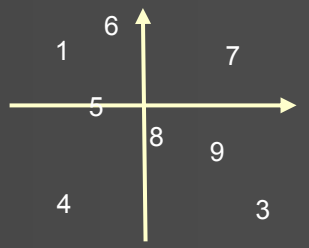
First eigenvector

$$W = \sum_{i=1}^{k} \beta_i \cdot W_i$$

where $W_i = X_i \cdot {}^t X_i$

Eigen vector →

Product Compromise Map



Double centering

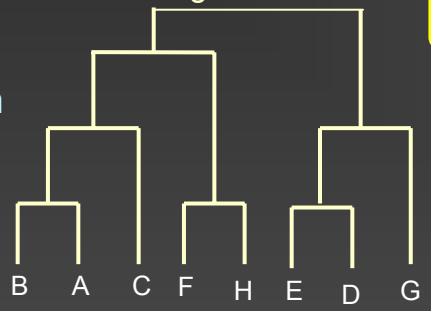
Panelist κ

Panelist k

$rv_{ij} - rv_{i.} - rv_{.j} + rv_{..}$

Classification →

Panelist segmentation

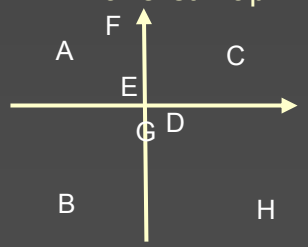


The larger β_i, the stronger the agreement of panelist i with the panel

The STATIS compromise is an average of individual configurations weighted by their agreement with the panel

Eigen vector

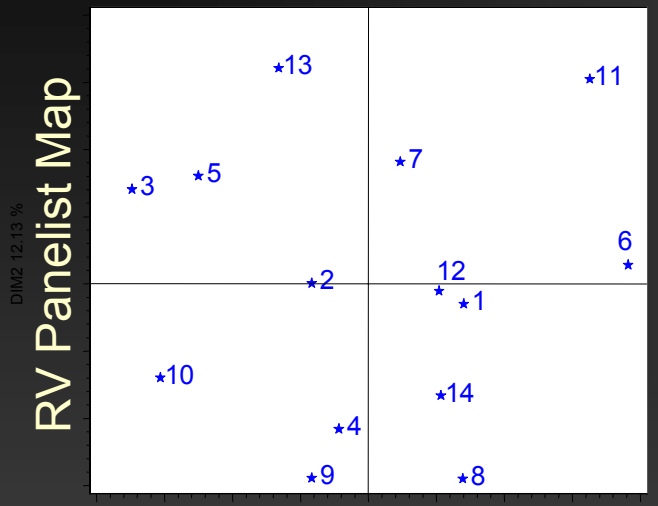
RV Panelist Map

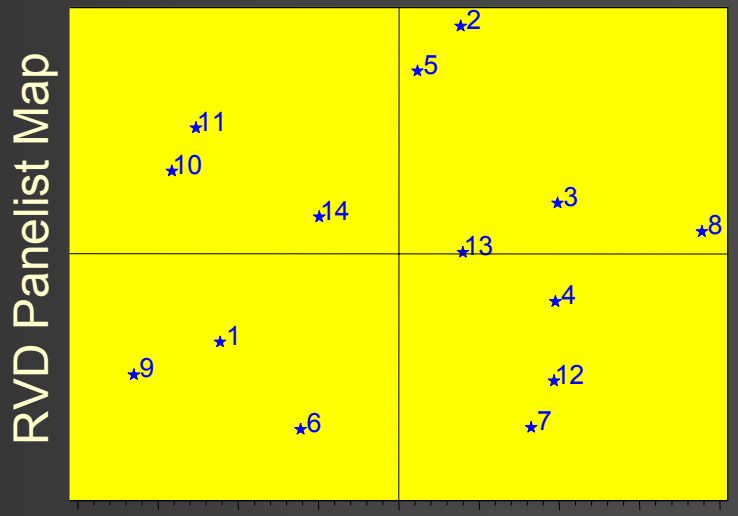


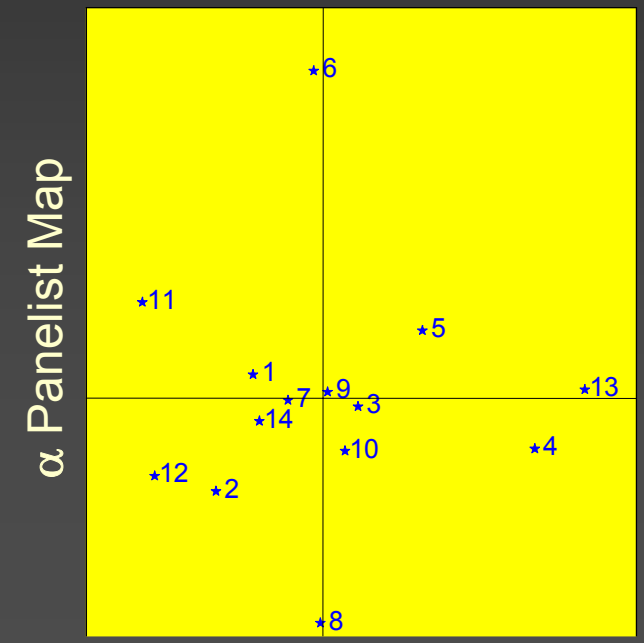
Same analysis can be done on:

- α **coefficients** to do a new STATIS taking into account panelist agreement on common attributes
- **RVD coefficients** to do STATIS dual which compares panelists on attribute correlation and produces an attribute compromise map

Flavor panelist maps based on RV, α and RVD

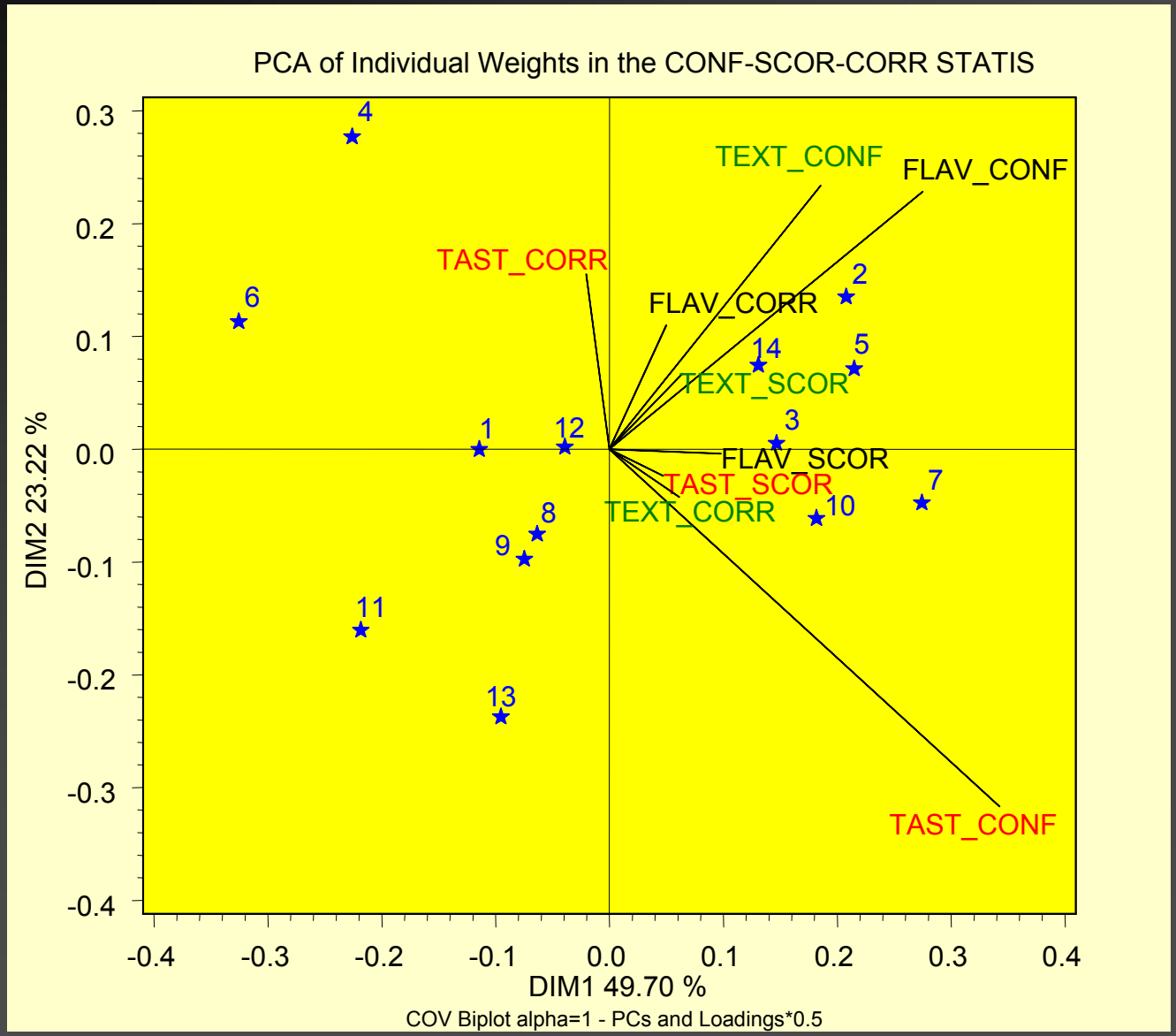






- Maps are clearly different
- Panelists 6, 8 and 13, who were the 3 reds on Nα, are apart on α-map
- RVD map useless since everybody agree
- RV map can be used to segment people with similar product configuration

Comparing individual weights from the 3 STATIS CONF-SCOR-CORR on the 3 sensory aspects



- CONF-SCORE-CORR different weights
- CONF weights more heterogeneous
- 2 groups of panelists along the first bisector
- Panelists 4 and 6 are special

Individual weights in Taste STATIS			
Panelist	CONF	SCOR	CORR
4	0.66	0.98	1.05
6	0.70	0.93	1.06
1	0.94	0.97	0.99
11	0.94	1.02	0.94
9	0.97	0.96	0.89
12	1.00	0.99	1.09
8	1.02	0.98	1.00
2	1.05	1.02	1.07
3	1.06	1.02	0.93
14	1.07	1.06	1.06
13	1.11	1.04	1.01
10	1.13	0.98	0.87
5	1.13	1.03	1.08
7	1.23	1.01	0.95
Panelists sorted by CONF weight			

Part IV.
Introducing the SensoBase project

1. What is SensoBase all about?
2. Current size of the SensoBase
3. First finding from the SensoBase
4. Current research using the SensoBase:

What is the SensoBase project ?

- A database of raw data from quantitative descriptive analysis
- From any sensory lab, on any type of product and attribute…
- Why?
  - To document the practices of sensory analysis
  - To benchmark panelist performances
  - To allow for extensive comparison of statistical analysis techniques
- How?
  - Collecting datasets online (products remaining anonymous)
  - Running the CAP analyses locally
  - Returning tables and plots to the data providers for free
  - Winning data provider loyalty

https://liris.cesg.cnrs.fr

LABORATOIRE D'INTERFACE
RECHERCHE INDUSTRIE
SENSOMETRIE

Home

User Identification :

Login :

Password :

Enter

What is SENSOBASE ?

Methods of exchange

Extract of statistical results and their meaning (French)

Register

Send us your suggestions

- Pangborn communications (2005)
- Drinktec P.Schlich (2005)
- Atelier DTS, RSIG 2005, P.Schlich

Welcome to the SENSOBASE website

At Centre Européen des Sciences du Goût,

with the LIRIS SENSOBASE team who invite you

to "exchange your sensory profiling data
for statistical analyses"

NEW !! The sensobase excel file (V2003 only) is available in english.

Send your sensory profiling data...

Panelist	NRV	Nalpha	NRVD

Today, SensoBase is composed of:

- 421 datasets from 42 sensory labs in 11 different countries
- 5,112 panelists or 1,317 different people
- 3,489 products or 2,758 different products
- 11,554 attributes or 8,706 different attributes
- A total of 2,811 767 scores

The “median dataset” has
12 panelists, 6 products, 2 reps and 26 attributes

Repeatability by type of product and type of attribute

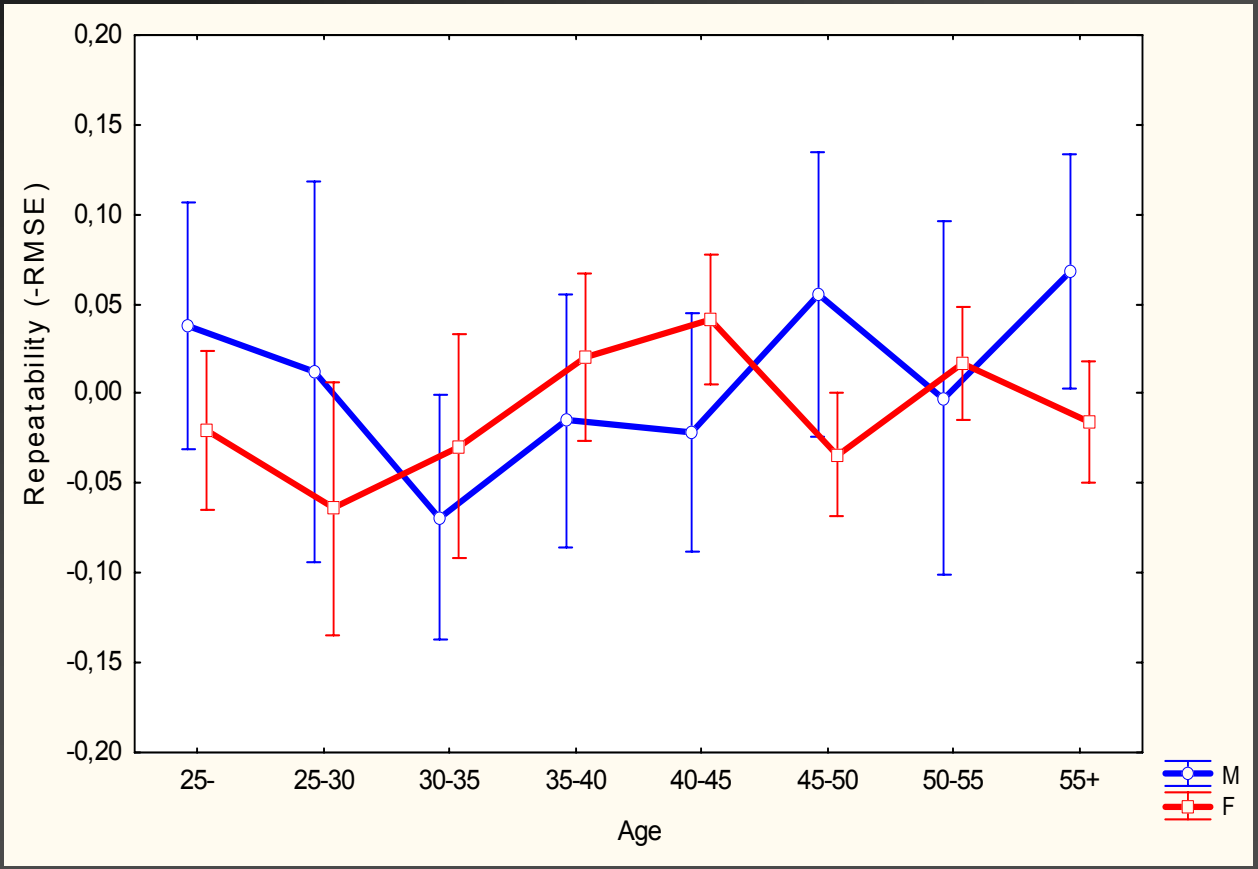
Average root mean square of error (RMSE) from individual one way ANOVAs

	Taste	Odour	Visual	Aroma	Texture	MEAN	N	STD
Sauce	0.95	0.98	0.83	0.92	1.04	**0.95 A**	1630	0.43
Meat/Fish	0.92	1.49	1.41	0.94	1.32	**1.19 B**	1919	0.55
Beverage	1.18	1.24	1.10	1.29	1.01	**1.20 BC**	1996	0.65
ReadyMM	1.06	1.05	1.28	1.18	1.47	**1.21 BC**	1274	0.55
Dairy Pdt	1.02	1.18	1.14	1.44	1.39	**1.24 C**	2229	0.68
Fruit/Veg	1.30	1.31	1.21	1.33	1.45	**1.33 D**	1705	0.57
Bakery	1.23	1.34	1.30	1.30	1.75	**1.39 E**	838	0.52
MEAN	**1.09 A**	**1.19 B**	**1.21 B**	**1.22 B**	**1.33 C**	**1.20**	.	.
N	2889	1930	1795	2670	2307	.	11591	.
STD	0.55	0.61	0.55	0.60	0.63	.	.	0.59

Based on a sample of 296 datasets from the Sensobase

No clear age and gender effect on repeatability, although a slight interaction was observed

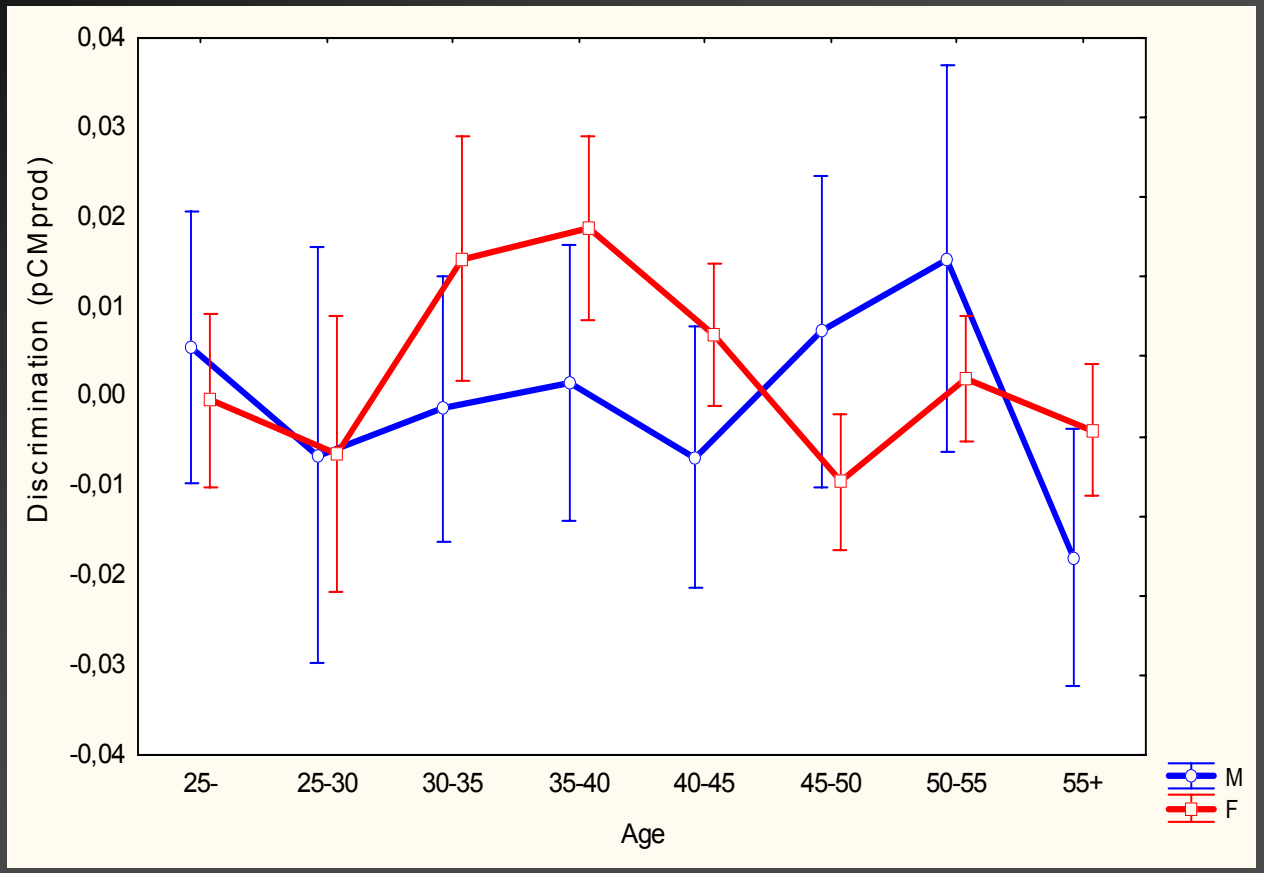
Mean of -RMSE centered by attribute within dataset



Gender: F=1.27, p=0.2602; Age : F=1.09, p=0.3660; Interaction: F=2.17, p=0.0340

Based on a sample of 296 datasets from the Sensobase

An effect of age on discrimination !
And this effect is different between gender



Gender: F=0.80, p=0.3721; Age: F=2.37 p=0.0205; Interaction: F=2.06, p=0.0444

Based on a sample of 296 datasets from the Sensobase

Also, level of education and experience with descriptive analysis were found correlated with discrimination

more at the next Sensometrics Meeting
(Norway, august 2006)

Assessing panel heterogeneity in performances

to be presented at the next Sensometrics Meeting (Norway, August 2006)

$$Y_{asr} = \alpha_a + \beta_a \upsilon_s + \varepsilon_{asr}$$

$$\varepsilon_{asr} \sim N(0, \sigma_a^2)$$

- Level is α_a : assessor main effect. Often very strong, not interesting
- Scaling is β_a : individual scale spread
- $\text{Disagreement}_a = \sqrt{\sigma_a^2 - \sigma_{a0}^2}$

Likelihood difference between usual and assessor models

- Variability is σ_a: individual repeatability
- $\text{Sensitivity}_a = \beta_a^2 / \sigma_a^2$

Brockhoff, P. M., & Skovgaard, I. M. (1994). Modelling individual differences between assessors in sensory evaluations. *Food Quality and Preference*, 5, 215–224.

Brockhoff, P. M. (1998a). Assessor modelling. *Food Quality and Preference*, 9, 87–89.

Based on the Brockhoff's assessor model and 30 datasets, heterogeneity was significant in:

- 65% of the attributes for repeatability
- 85% of the attributes for scaling
- 85% of the attributes for sensitivity

and 38% of the attributes were affected by a "pure disagreement" (adjusted to scaling)

However, correcting for scaling did not seem to result in a significant gain of product discrimination

Conclusion

1. The panel leader does not have time to surf into statistical softwares. He needs outputs easy to read and report-ready

   ⇨ The CAP system is one possible answer to this need

2. Any concept or progress in the univariate analysis of sensory data should be transposed to multivariate analysis

   ⇨ MANOVA-CVA transposes those from ANOVA

3. Taking attributes into account only at individual level makes sense with "untrained panelists". But, one must validate panel homogeneity with "trained panelists"

   ⇨ RV-STATIS is a useful framework for doing so

4. Sensory analysis must validate levels of performance achieved

   ⇨ The SensoBase was launched for that main purpose

5. Although the number of datasets already collected is huge, the SensoBase still needs a **larger number of data providers** to gain an higher representative value …