SENSORIALIS
ÉVALUATION SENSORIELLE

The right number of consumers to be enrolled in a liking test strongly depends on the level of sensory complexity among products

Nadra MAMMASSE (*), Pascal SCHLICH

Food and Behaviour Research Center, UMR6265 CNRS, UMR1324 INRA, Université de Bourgogne, Agrosup Dijon, Dijon, France

10th Sensometrics - July 25-28 Rotterdam

Background (1/4)

The number of panelists to be enrolled in a hedonic test ranges in literature from 50 to 100 at least if no liking segmentation is sought

- **C**hambers & Baker Wolf (1996) : 100 people is usually considered adequate for most of problems handled in small consumer tests

- **S**tone & Sidel (2004) : 25-50 subjects per product in laboratory testing

Background (2/4)

❑ **H**ough et al. (2005) : presented the basic concepts to estimate the number *N* of consumers for sensory acceptability on 128 previous consumer studies using

- **σ :** variability data reported as the root mean square of error in the anova model (no details about the model). **σ =** 0.23 with a standard deviation of 0.037
- d : the difference in means that is sought in the experiment
- α : type I error , β : type II error

$$H_0 \text{ "}\mu_1 = \mu_2 = \cdots = \mu_p\text{" Versus } H_1 \text{"}\exists\,(i,j)\ ,\ \mu_i \neq \mu_j\text{"}$$

Result : Table of values of N were computed using these parameters

Background (3/4)

Two way anova model

$$X_{i,j} = \mu + \alpha_i + \beta_j + \varepsilon_{i,j}\ ,\ \varepsilon_{i,j} \sim N(0,\sigma_\epsilon^2)$$

- α_i is the effect of the *ith* product
- β_j is the effect of the *jth* panelist
- $\varepsilon_{i,j}$ is the random error

The root of mean square error **RMSE** measures the heterogeneity of consumer liking

Background (4/4)

Recently, the new French standard AFNOR XPV 09-500 has recommended (if σ and d are unknown) 100 panelists in hedonic tests whereas only 60 were recommended in the former standard

Objective

The study aimed to examine whether this range of panel size seems adequate in 7 actual experiments representing different products

Material & Methods (1/5)

Panelists were required to give a liking score on products using a 9-point hedonic scale

Product	Number of Products tested	Number of sensory dimension	Product description	Sensory laboratory	Panel Size
Cake	5	1	Fat variation	lab1	150
				lab2	150
Stewed apples	5	1	Sugar variation	lab3	150
				lab4	150
Crisps	6	2	Fat variation (different oils) + Salt variation	lab 5	150
Smoked herring	6	3	Salt variation + texture variation + other	lab 6	150
Sausage	6	2	3 with natural casings + 3 with artificial casings + taste variation	lab 7	150

Material & Methods (2/5)

Complete Panel *N*

Reduced Panel *n=N-k*

Estimate ***k*** the **maximum number of subjects** that can be removed from the original panel without **loss of information** regarding the product comparison

For each **k**, **k=0,…,130** by **10**, 1 000 incomplete panels are simulated by resampling **n=N-k** assessors among **N** with replacement

Material & Methods (3/5)

To measure the loss of information due to panel size reduction, 4 criteria were defined :

❑ **Correlation Approach**

Correlation coefficient between the vector of product mean scores of the complete panel and the vector of product mean scores of the reduced one

❑ ***Rv coefficient between product configurations***

Measure of similarity between the product configuration obtained from the complete panel and the product configuration obtained from the reduced one

Material & Methods (4/5)

❑ Anova Approach

Based on the two-way Anova model ***Product + Panelist***

Compute the *Fisher intraclass correlation coefficient* ***ICC***

$$\text{ICC} = \frac{MS_{Prod}}{MS_{Prod} + MS_E}$$

❑ *Concordance Rate in product pair comparison*

Based on the number of concordant pairs of products in the complete panel and in the reduced one

N		N - k	
		concordance = 1	concordance = 0
A ≠ B	A > B	A > B	A ≤ B
	A < B	A < B	A ≥ B
A = B		A = B	A ≠ B

Material & Methods (5/5)

Decision rule for the estimation of k

For each k, 1000 values per criterion are computed

- No more than 10% of correlation coefficients *R* having a value less than 0.8
- No more than 10% of *Rv* coefficients having a value less than 0.8
- No more than 10% of fisher interclass correlation coefficients *ICC* having a 10% loss of discrimination ($ICC_N - ICC_{N-k} = 10\ \%$)
- No more than 10% of samples having q/p pairs discordant (3/10 for 5 products and 4/15 for 6 products)

All computations with SAS® software

Scores were transformed into values within [0;1]

Results & Discussion

I. Two-way Anova results on the complete panel for each data set

II. Panel size recommendation based on the four approaches

I. Two-way Anova Results (1/4)

Data set	MS_{Prod}	RMSE	F_{Prod}	MISD
Cake **lab1**	0.607	0.171	20.43	0.164
Cake **lab2**	0.857	0.173	28.47	0.171
Stewed apples **lab3**	1.005	0.177	32.00	0.182
Stewed apples **lab4**	0.544	0.161	20.85	0.159
Crisps	0.541	0.188	15.21	0185
Smoked Herring	0.331	0.198	8.38	0.190
Sausage	4.129	0.190	113.94	0.213

✓MS_{Prod} Product mean square
✓RMSE : Root mean square of error in the two-way Anova model
✓F_{prod} Fisher value of product effect
✓Mean of individual standard deviations

I. Two-way Anova Results (2/4)

Product mean scores

Cake lab1				Cake lab2			
Product	Designation		Mean	Product	Designation		Mean
3644		A	0.702	3644	A		0.745
3645	B	A	0.670	3643	A		0.707
3643	B	C	0.660	3645	B		0.668
3642		C	0.623	3642	B		0.647
3641		D	0.537	3641	C		0.546

Both of the two panels preferred the same product except a small inversion in rank for the second and the third product.

I. Two-way Anova Results (3/4)

Product mean scores

Stewed Apples lab3				Stewed Apples lab4			
Product	Designation		Mean	Product	Designation		Mean
3731		A	0.784	3731	A		0.768
3730		B	0.717	3732	B		0.713
3732	C	B	0.699	3730	B		0.706
3729	C		0.669	3729	C		0.639
3733		D	0.561	3733	C		0.619
MSProd = 1.005				MSProd = 0.554			

Both of the two panels preferred the same product except a small inversion in rank for the second and the third product.

I. Two-way Anova Results (4/4)

Product mean scores

Smoked Herring

Product	Designation	Mean
4011	A	0.687
4012	A	0.664
4008	B	0.612
4007	B	0.607
4010	B	0.605
4009	C	0.555

MSProd = 0.331

Sausage

Product	Designation	Mean
4029	A	0.711
4031	A	0.704
4028	A	0.681
4032	B	0.427
4033	C B	0.395
4030	C	0.373

MSProd = 4.129

Results & Discussion

II. Panel size recommendations

II. Panel size recommendation (1/3)

Estimation of ***k*** the **maximum number of subjects** that can be removed with reasonable loss of information

*For each **k**, 1000 values per criterion are computed*

II. Panel size recommendation (2/3)

k	Simulation	R
0	1	0.97
0	2	0.92
.	.	.
.	.	.
.	.	.
0	999	0.98
0	1000	0.93
10	1	0.88
10	2	0.90
.	.	.
.	.	.
.	.	.
10	1000	0.87
.	.	.
.	.	.
.	.	.
130	999	0.23
130	1000	0.36

Sort in

Increasing order

k= 10

n=N-k = 140

k	Rank	R
0	1	0.87
0	2	0.88
.	.	
.	.	.
.	.	.
0	100	0.90
0	101	0.92
.	.	.
.	.	.
.	.	.
0	999	0.97
0	1000	0.98
10	1	0.70
10	2	0.71
.	.	.
.	.	.
.	.	.
10	100	0.89
10	101	0.90
.	.	.
.	.	.
10	1000	0.96
20	1	0.68
.	.	.
.	.	
20	100	0.78
.	.	.
.	.	.
.	.	.
130	1000	0.05

II. Panel size recommendation (3/3)

Criterion / Data set	R	RV	ICC	Concordance rate	Maximum Panel size
Cake lab1	30	60	70	80	80
Cake lab2	20	40	50	70	70
Stewed Apples lab3	20	40	50	60	60
Stewed Apples lab4	30	60	60	70	70
Crisps	40	80	80	90	90
Smoked Herring	70	140	100	150	150
Sausage	20	20	20	20	20

Discussion

Data set	Product Description	Number of sensory dimensions	Maximum Panel size	MS_{prod}
Cake lab1	Fat variation	1	80	0.607
Cake lab2			70	0.857
Stewed Apples lab3	Sugar variation	1	60	1.005
Stewed Apples lab4			70	0.544
Crisps	Fat variation (different oils) + salt variation	2	90	0.541
Smoked herring	Salt variation + texture variation +other	3	150	0.331
Sausage	Type of casing was a dominant characteristic	2	20	4.129

Conclusion

- The level of heterogeneity of consumer liking was rather similar over the trials and thus did not impact the panel size recommendations

- The number of panelists to be enrolled in a hedonic test depends **mainly** on the level of complexity of the product space

- It is not possible to define a global number of consumers valid for every studies, as was suggested in the literature and by the standards...

Perspectives

- The sensory analyst should have an idea of the size of differences between product to be compared
- Develop a methodology to target the number of sensory dimensions before carrying out a trial by for instance a sensory profiling
- Underline the link “Sensory complexity” / “Consumers’ preferences”
- What would be the number of panelists to be enrolled if a segmentation is sought ?

THANK YOU FOR YOUR ATTENTION