

## A chemometrics study of analytical panels in virgin olive oil. An approach for evaluating panels in training.

By R. Aparicio, F. Gutiérrez and J. Rodríguez.

The authors are with The Instituto de la Grasa y sus Derivados. Avda. Padre García Tejero 4.  
Sevilla 41012, Spain.

### RESUMEN

**Un estudio quimiométrico de paneles analíticos sobre el aceite de oliva virgen. Una aproximación a la evaluación de paneles en entrenamiento.**

Un procedimiento matemático que incluye Correlación Canónica, Análisis Factorial y de "clusters", y Regresión Múltiple fué diseñado para estudiar si dos paneles analíticos, con diferentes niveles de entrenamiento, podrían ser considerados idénticos. Para evaluar si un panel en o después de un entrenamiento opera tan bien como uno experto, se analizaron las correlaciones entre los atributos de ambos paneles, las matrices de correlaciones parciales, la varianza explicada por los factores más importantes y la similitud entre los aglomerados ("clusters") entre otras. Además, los tres factores más importantes fueron correlacionados con las valoraciones finales de calidad, para conocer la interrelación entre la hoja de perfil y la tabla de puntuación. Finalmente, un análisis de regresión múltiple permite conocer como es el modelo que los paneles han diseñado para muestras de aceite de oliva virgen evaluadas. También se sugieren las tendencias en la investigación futura.

*PALABRAS-CLAVE: Aceite de oliva virgen – Estudio quimiométrico – Panel analítico.*

### SUMMARY

**A chemometrics study of analytical panels in virgin olive oil. An approach for evaluating panels in training.**

A mathematical procedure including Canonical Correlation, Factor and cluster analysis and Multiple Regression was designed to study whether two different sensory panels, with different levels of training, could be considered as a unity. Among others, correlations between each attribute of both panels, matrices of partial correlations, variance explained by the most important factors and similarity among clusters were computed as the procedure to evaluate if a Panel in training works as well as an expert. In addition, every one of the factors was correlated with the final evaluation of quality for analysing the interrelationship between the profile notes and the grading table of virgin olive oil quality from a mathematical point of view. Finally, a stepwise multiple correlation analyses how the panels modelled the evaluated samples. Promising directions for future research are also suggested.

*KEY-WORDS: Analytical panel – Chemometrics study – Virgin olive oil.*

### 1. INTRODUCTION

In 1974 Stoner et al. published the quantitative descriptive analysis method (QDA) opening a new way in sensorial evaluation. The consequence was an increasing use of the descriptive and quantitative multiscalar test in the sensorial analysis of foods that has become as widely used as Bengtsson's triangular test (Bengtsson, 1943).

However, it was not possible to analyse mathematically the amount of information gathered by QDA without statistical multivariate algorithms. Undoubtedly, sensory and quality problems are generally multivariate problems (Resurreccion, 1988b), (Powers, 1988). Thus, factor analysis procedures such as principal components or principal factor (Clapperton et al, 1979), stepwise discriminant analysis (Ennis et al., 1980), different kinds of cluster programs (Sheppard, 1980), (Resurreccion, 1988a) and multiple regression (Resurreccion, 1988b) have been some of the statistical methods widely used in reaching conclusions.

As far as the authors are aware, those statistical studies have been generally used to find interrelations among attributes, between their intensities and the final evaluation or between objectives measures and a sensory measure. This work has followed a different approach, discussing a methodology that allows everyone to know whether the level of training of a panel is similar to expert, using a panel with a large experience as pattern and another of less as instrument to calibrate. The authors consider that the panel, constituted by human panelists, is the sensory measurement instrument and try evaluating it.

Many authors, Tabashnick and Fidell (1983), Dillon and Goldstein (1984), Harris (1985) and Bizquerra (1989), have suggested, by arborescent structures, the methodology to select the best statistical technique,

according to the research goal and the number of DVs, IVs and covariates. However, the authors' intention is not only to discuss differences between the panels as to try to understand where and why there are differences and how to correct them, if it is possible. Then, the authors' approach follows a stepwise methodology in which the results of the statistical procedures of principal components, regression and cluster analysis have been used as clues to assess the next step.

Two panels of the Instituto de la Grasa, both trained following the same method, one with more than ten years experience evaluating the quality of virgin olive oil, and the other with only three months, have been used for developing the proposed methodology. The paper initially considers the hypothesis "The two panels are different to one another" and, as it was a null hypothesis, will attempt to demonstrate that there is only one panel. Other conclusions, that have been taken from the statistical procedures, are related to the sensory terms of the kind of food evaluated (virgin olive oil) and the similarity and dissimilarity between the panels. Thus, the intention of the paper is to present research results, and at the same time to explain the applied statistical methodology.

## 2. MATERIALS AND METHODS

### 2.1. Sensory evaluation of oils

A data set of 24 samples of virgin olive oil belonging to more than one crop was collected from different zones all over Spain. The samples were evaluated according to the standard COI/T20 Doc. 3 (C.O.I., 1987) by two panels. The panelists evaluated each sample in duplicate, at different dates and making only one evaluation per session (Gutierrez et al, 1989). One of the panels (henceforth, panel A) was constituted by twelve panelists who had more than ten years experience; the other (panel B) was constituted by fifteen panelists with only three months experience after training.

Figure 1 shows, on the left, the olfactory-gustatory-tactile notes of virgin olive oil. Eighteen attributes with five levels of intensity can be detected and evaluated in virgin olive oil by the panelists, besides other free sensory variables. The attributes which can have a positive influence on the quality of oil are described at the top of the figure, while those with a negative influence are at the bottom. The nine points of the grading table are shown on the right. The whole rank is clustered in five great groups according to the defects and characteristics of the evaluated virgin olive oil.

All panelists were trained to detect whether one or more of those attributes could be perceptible in every sample from a barely level to an extreme one. They have also been taught to assign the best point

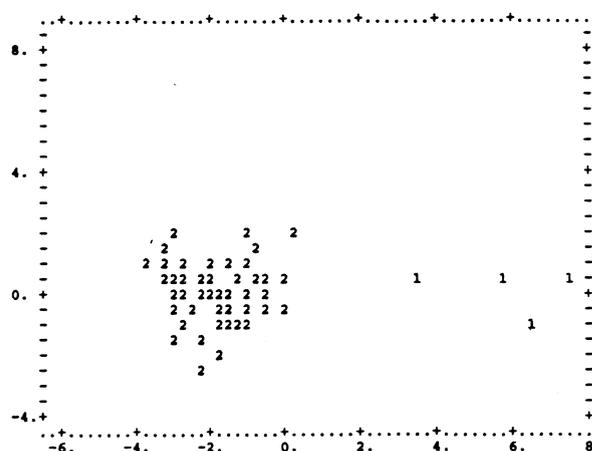


Figure 1  
Standard of profile sheet and grading table to quality evaluation of Virgin Olive Oil (Olive Oil Council Authority).

of the grading table concerning their filled-up profile sheet, following easy expert rules attained by the experience of the trainers, some of which are shown in the defects and characteristics of the grading table by fuzzy descriptions.

The data that the paper analyses are the mean intensity of each one of the attributes for each sample by every panelist. Only four attributes of the whole set have not been detected in the samples. They are sour, vinegary, acid and humidity.

### 2.2. Statistic programs

Multivariate data evaluation was made by BMDP (BMDP, 1981) and SPSS (SPSS, 1986) packages running under VMS on a VAX8550. Six different programs were used, Canonical Correlation Analysis, Bivariate (Scatter) Plots, Stepwise Regression, Factor Analysis, Cluster Analysis of Cases and K-Means Clustering.

Factor Analysis was run with the following options: Principal Components Analysis (PCA), Kaiser's normalization Varimax (VMAX). Three tests, Barlett's, Kaiser-Meyer-Olkin and anti-image covariance (AIC), were previously computed to know whether this procedure could be applied.

The assigned options on K-Means Clustering were: two clusters and their centers and distances were computed from the standardized data by the pooled within-cluster covariance matrix, using Euclidean distances.

Cluster Analysis of Cases was performed to compute the similarity matrix on the basis of the Euclidean distance among the variables, which were the factors obtained by PCA after double cross-validation.

Stepwise Regression was run to estimate the parameters of a multiple linear regression by entering or removing variables according to the RSWAP method, being  $F_{to\_enter}$  4.30 and  $F_{to\_remove}$  4.28. While Canonical Correlation procedure performed the pre-assigned conditions of BMDP6M (BMDP, 1981).

### 3. RESULTS AND DISCUSSION

The panelists were independently taught and multivariate methods were used in training them. The outliers among panelists were previously identified and deleted from subsequent analysis.

To determine whether there were different opinions about the attributes between both panels, a correlation analysis was performed. Table I shows, using every sample, the correlation coefficient between the mean level of intensity of each attribute, described in Figure 1, with its level of significance. A good correlation was found between the following taste and smell attributes of both panels: fruity, green, bitter, pungent, mustiness, muddy, fusty and rancid, which are the most important in evaluating the quality of virgin olive oil. This fact is even more interesting, due to the mean

Table I  
Table of correlation coefficients and their level of significance.

Attributes	Correlation coefficients	Significance levels
Fruity	0.8151	0.001
Apple	0.5646	0.002
Other	0.4514	0.020
Green	0.8342	0.001
Bitter	0.9020	0.001
Pungent	0.9327	0.001
Sweet	0.3927	0.045
Winey	0.6502	0.001
Rough	0.2411	0.238
Metallic	0.5020	0.008
Mustiness	0.7097	0.001
Muddy	0.7952	0.001
Fusty	0.8880	0.001
Rancid	0.9061	0.001

intensities of the perceptible attributes having similar ranks with every sample evaluated by both two panels. The difference between each mean evaluation of the attributes by both panels was less than 1.0 generally. Only 3.4% of the evaluations were greater than that value but always less than 1.5.

#### 3.1. Canonical Correlation

The first selected statistical procedure was canonical correlation. The goal was to analyze the relationships between two sets of variables and to know how the two panels relate to each other. The first pair of canonical variates –linear combination of attributes– maximizes the correlation between a linear combination of panel A and a linear combination of panel B. Successive pairs explain the correlation after the variance due to previous pairs has been removed. Three canonical variates were considered necessary at the 0.01, after applying Bartlett's test. One of the canonical correlations were 0.99 ( $p < 0.001$ ), representing 99% overlapping variance between each pair of canonical variates.

The procedure also detected the attributes with the higher squared multiple correlations in the panel A with the chosen canonical variates of panel B, they were pungent (0.83,  $p = 0.04$ ) and mustiness (0.85,  $p = 0.03$ ). While the attributes with the higher correlation in the panel B with the canonical variates in panel A were bitter (0.93,  $p = 0.001$ ) and also pungent (0.89,  $p = 0.008$ ).

However, Tabachnick and Fidell (1983) suggest to reconsider the use of canonical correlation when a group of variables within a set are identifiable but correlated with one another. In this study, fruity ( $R^2 = 0.94$ ), in panel A, and fusty ( $R^2 = 0.92$ ), in panel B, can be considered nearly a linear combination of others in their panels.

#### 3.2. Principal Components.

At this point, authors thought there was enough background for considering that both panels were, in fact, a unity, from a mathematical point of view. Even though, a null hypothesis was proposed: "There are two different panels", as the hypothesis to be verified (approved or rejected) and every statistical program was designed to verify it.

The selected statistical procedure was Factor Analysis. Factor analysis is the most commonly used technique to reduce a large number of variables to the smallest set of factors with which to explain the variance in the experiment. Thus, principal components analysis was used as the method of factor extraction, and the resulting configuration was rotated applying varimax.

Table II  
Matrix of partial correlations of each pair of attributes of panel A.

	fruity	apple	other	green	bitter	pungent	sweet	winey	rough	metal.	musti.	muddy	fusty	rancid
fruity	1.000													
apple	0.834	1.000												
other	0.130	0.250	1.000											
green	0.783	0.642	0.024	1.000										
bitter	0.767	0.678	0.139	0.816	1.000									
pungent	0.848	0.657	0.058	0.637	0.764	1.000								
sweet	-0.417	-0.507	-0.114	-0.504	-0.632	-0.428	1.000							
winey	-0.380	-0.298	0.129	-0.399	-0.315	-0.340	0.345	1.000						
rough	-0.628	-0.557	-0.176	-0.489	-0.426	-0.485	0.103	-0.231	1.000					
metallic	-0.524	-0.499	-0.435	-0.378	-0.263	-0.365	-0.152	0.207	0.434	1.000				
mustiness	-0.750	-0.613	-0.296	-0.636	-0.677	-0.777	0.384	0.382	0.451	0.469	1.000			
muddy	-0.832	-0.718	-0.232	-0.643	-0.706	-0.810	0.450	0.372	0.518	0.303	0.830	1.000		
fusty	-0.748	-0.533	-0.205	-0.599	-0.593	-0.678	0.280	0.476	0.322	0.360	0.782	0.804	1.000	
rancid	-0.462	-0.440	-0.306	-0.363	-0.123	-0.237	0.096	0.014	0.562	0.622	0.231	0.174	0.127	1.000

Table III  
Matrix of partial correlations of each pair of attributes of panel B.

	fruity	apple	other	green	bitter	pungent	sweet	winey	rough	metal.	musti.	muddy	fusty	rancid
fruity	1.000													
apple	0.688	1.000												
other	0.431	0.501	1.000											
green	0.788	0.628	0.301	1.000										
bitter	0.662	0.529	0.516	0.768	1.000									
pungent	0.840	0.564	0.419	0.734	0.806	1.000								
sweet	-0.426	-0.172	-0.109	-0.438	-0.592	-0.587	1.000							
winey	0.031	0.032	-0.194	-0.082	-0.043	-0.086	-0.233	1.000						
rough	-0.362	-0.341	-0.234	-0.156	-0.304	-0.421	0.206	0.097	1.000					
metallic	-0.496	-0.171	-0.114	-0.287	-0.385	-0.537	0.369	0.057	0.377	1.000				
mustiness	-0.713	-0.382	-0.372	-0.588	-0.654	-0.779	0.447	0.004	0.507	0.566	1.000			
muddy	-0.708	-0.436	-0.260	-0.596	-0.650	-0.709	0.476	-0.151	0.474	0.641	0.729	1.000		
fusty	-0.670	-0.411	-0.339	-0.488	-0.569	-0.710	0.233	0.264	0.723	0.653	0.667	0.776	1.000	
rancid	-0.573	-0.601	-0.177	-0.486	-0.184	-0.312	0.172	-0.253	-0.077	0.147	0.149	0.084	0.009	1.000

However, three different tests were previously computed to know whether the factor analysis could be applied. Thus, the results of Kaiser-Meyer-Olkin's test (0.76 panel A, 0.69 Panel B) showed the matrices were adequate to implement factorial analysis. Besides, the results of Barlett's test of sphericity rejected the null hypothesis (260.1 panel A, 243.7 panel B) and only there were 26 (14.3%) off-diagonal elements of Anti-Image-Correlation Matrix greater than 0.09 in panel B and even lesser in panel A, 14 (7.7%).

The measure of sampling adequacy displayed that the attributes "other ripe fruit", panel A, and "winey", panel B, had a low value. However, they have not been removed.

The first information given by this procedure was the covariance matrix. Table II and III show the matrices obtained with the independent analysis of the panels, in agreement with the proposed null hypothesis. As far as the authors can see, the level of the values are very similar in general.

However, some drawbacks have been found. The correlations between the pair of attributes rough-fusty, rough-rancid, metallic-fusty, metallic-rancid and metallic-mustiness are rather different. This discrepancy could be due to the characteristics of two of those attributes: rough (tactile note) and metallic (olfatory-

Table IV  
Proportion of variance explained by each factor.

	PROPORTION OF VARIANCE EXPLAINED			
	IN DATA SPACE		IN FACTOR SPACE	
	Panel A	Panel B	Panel A	Panel B
Factor 1	0.5204	0.4911	0.7015	0.6801
Factor 2	0.1414	0.1309	0.1905	0.1813
Factor 3	0.0801	0.1002	0.1080	0.1386
Total	0.7419	0.7222	1.0000	1.0000

gustatory-tactile note). This second attribute has a clear relation with the storage, fustiness and the kind of technological process used, which is quite unlike the first attribute. Thus, those correlations given by panel A are closer to the real meaning of both attributes.

At first glance, the discrepancy indicates that panel B has not attained an adequate knowledge about those attributes, due to the fact that they are not easy to understand and the evaluation of a great number of samples with those attributes is required to gain a good experience. More studies could be made of other relationships such as, for instance, sweet-vinegary, even though the level of confidence on the conclusions would probably be lower.

Table V  
Sorted rotated factor loadings of both panels.

	Panel A			Panel B		
	FACTOR	FACTOR	FACTOR	FACTOR	FACTOR	FACTOR
	1	2	3	1	2	3
fruity	0.691	-0.472	-0.459	-0.605	0.709	0.000
apple	0.644	-0.465	-0.323	0.000	0.828	0.000
other	0.000	-0.596	0.000	0.000	0.542	-0.464
green	0.711	-0.263	-0.387	-0.454	0.734	0.000
bitter	0.802	0.000	-0.400	-0.643	0.539	0.000
pungent	0.683	-0.270	-0.482	-0.754	0.533	0.000
sweet	-0.697	0.000	0.000	0.529	0.000	-0.536
winey	0.000	0.000	0.852	0.000	0.000	0.811
rough	-0.601	0.653	0.000	0.691	0.000	0.269
metallic	0.000	0.863	0.326	0.757	0.000	0.000
mustiness	-0.491	0.388	0.626	0.805	-0.305	0.000
muddy	-0.645	0.291	0.558	0.853	0.000	0.000
fusty	-0.384	0.271	0.736	0.878	0.000	0.297
rancid	0.000	0.787	0.000	0.000	-0.805	-0.280

Many analysts use the value of 1.0 for the explained variance to decide how many factors are adequate to model the data. However, we applied double cross-validation to define the number of significant factors; by this means, it was found that the first three factors, explaining more than 70.0% of the total variance, were significant enough in sensory analysis. Table IV shows that the successive addition of only three factors gives that account of the variance explained in data space, and 1.0 in factor space. Moreover, the first factor of each panel explains 52.04% (panel A) and 52.89% (panel B) of the total variance. In this way, the variance explained by panel B is only a little greater than that by panel A, showing a certain degree of mathematical similarity between the panels, against the proposed hypothesis.

The loadings for each of the attributes on the three factors are given in Table V. These loadings were obtained after applying a varimax rotation that diminishes the mathematical rumour by maximizing the variance of the loadings across attributes within factors and also tending to reapportion the variance among factors so that they become relatively equal in importance.

In Table V, loadings less than 0.25 have been replaced by zero. Factor 1 of panel B appears to increase as similarity to the attributes described as "fusty", "muddy" and "mustiness". These, which belong to a non-positive group of attributes, have a closed relation. Besides, the sign of the attributes were chosen in a mode opposite to logic and in consequence the positive notes have negative coefficients, contrary to the negative ones. In the authors' opinion, the panelists were worried looking for negative notes in the samples rather than positive ones, which can be a psychological problem of beginners. Other partial conclusions could be taken from factor 2 or 3, such as the signs of factor 2 follow the logic now, or factor 3 is composed by negative attributes excepting "other ripe fruit" (here "other").

In the other panel, Table V, the positive and negative attributes elected by all three factors are more balanced, due to its large experience. In contrast to panel B, factor 1 appears to be strictly related to the positive attributes with closed profile, so a virgin olive oil with the attributes fruity and green logically has the attributes bitter and pungent, as the loadings of panel A show on this factor.

Despite these differences, the general performance of both panels seems similar, because they are as a figure and its mirror-image. Panel A evaluates the presence of positive attributes and the absence of negative ones, while panel B evaluates the presence of negative attributes and the absence of the positives. Therefore, a cluster analysis could help us to accept or reject the null hypothesis, due to its being

a general descriptor for procedures that groups cases according to some measure of similarity. This kind of analysis can provide important clues to the most probable groups of the samples, independently of the panel that evaluated them.

Before that, it was important to know if there were multivariate outliers in panel A and B and delete them from subsequent analysis.

The ratio  $CHISQ/DF$ , BMDP (1981), is relevant for identifying outliers among cases. This ratio contains Mahalanobis distances, evaluated as Chi-Squares, of each case from the centroid of the cases for the original data. Because Chi-Square has been divided by degree of freedom, one can follow the same process for  $p=0.01$  with 14df for each panel. Therefore, values in excess of 2.08 indicate outliers. Applying this criterion, no samples were found as outliers in both panels.

Following a similar procedure, the ratio between the Mahalanobis distance of each case from the centroid of the factor scores and degree of freedom allows to identify the outliers with respect to the solution. The new cutoff value was obtained by chosen number of factors as degrees of freedom. With the use of this new criterion, two cases, sample number 4 in panel A and number 7 in panel B, were found to be near to outliers in the solution space.

### 3.3. Cluster Analysis

The non-hierarchical cluster procedure "K-Means Clustering" was the first attempt to determine if there were two panels, which meant that each panel evaluated the same sample in different mode, or only one panel, which meant that the attributes detected by both panels were similar in general.

Thus, to verify the similarity among the samples, we established two as the fixed number of homogeneous groups of cases to be built, without considering whether the samples were evaluated by either panel A or B. The result is shown in Figure 2. There are the two groups as it was programmed on the statistical procedures; one of them is constituted by four samples and the other by the rest. But the samples of this small group belong to both panels and they are coupled two-by-two, Table VI.

Hence, this result shows that initially there were not two panels because the cluster analysis was not able to cluster panel A independently of panel B. The order of building two clusters was executed, but, in both clusters, each sample of one of the panels is accompanied with its couple in the other panel, despite the sample number 7 was only near to outlier in panel B.

Despite this good result, verifying that panel B was as expertise as panel A, a hierarchical clustering

VIRGIN OLIVE OIL

PROFILE SHEET

GRADING TABLE

OLFACTORY-GUSTATORY-TACTILE NOTES

	(*)	1	2	3	4	5
Olive fruity (ripe and green) ...						
Apple						
Other ripe fruit						
Green (leaves, grass)						
Bitter						
Pungent						
Sweet						
Other allowable attribute(s) (Specify .....						
Sour/winey/vinegary/acid 1/						
Rough						
Metallic						
Mustiness/humidity 1/						
Muddy sediment						
Fusty ("atrojado")						
Rancid						
Other unallowable attribute(s) (Specify .....						

DEFECTS	CHARACTERISTICS	OVERALL MARKS POINTS
None	Olive fruity	9
	Olive fruity and fruitness of other fresh fruit	8
		7
Slight and barely perceptible	Weak fruitness of any type	6
Perceptible	Rather imperfect fruitness, anomalous odours and tastes	5
Considerable, on the border of acceptability	Clearly imperfect, unpleasant odours and tastes	4
Great and/or serious, clearly perceptible	Totally inadmissible odours and tastes for consumption	3
		2
		1

1/ Delete where not applicable

(\*) Perception

- 1 Barely perceptible                      4 Great
- 2 Slight                                      5 Extreme
- 3 Average

REMARKS .....

.....

NAME OF TASTER .....

LEGEND OF SAMPLE .....

DATE .....

Figura 2

A scatter plot of the orthogonal projection of cases into the plane defined by the centers of the two cluster built by a K-Means Clustering Procedure.

Table VI

Samples clustered in group named as 1.

Panel	Case	Weight	Distance
A	7	1.000	3.8364
A	12	1.000	3.4945
B	7	1.000	3.5211
B	12	1.000	1.6216
Average Distance			3.1184

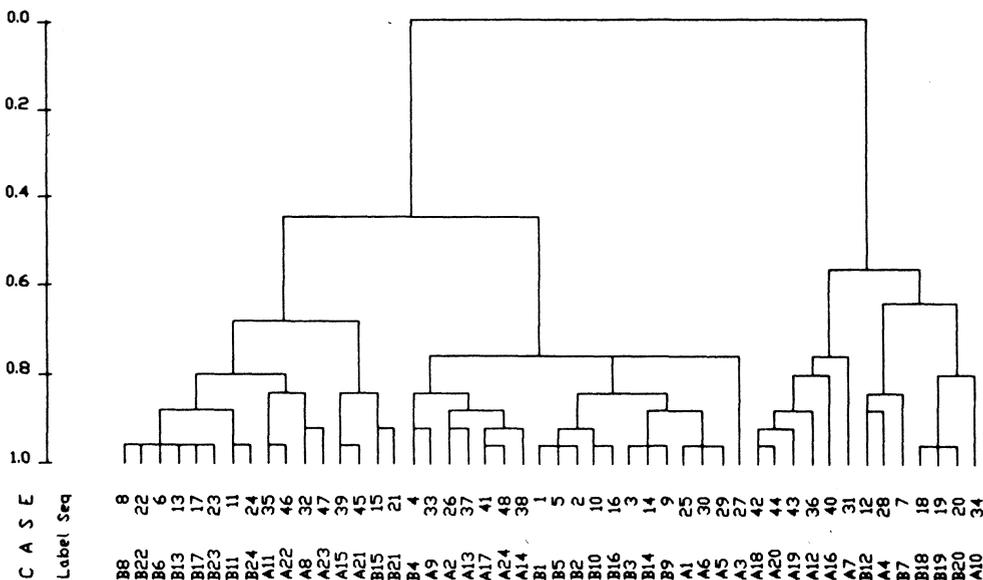


Figure 3

Dendrogram of both panels using complete linkage by SSPS+Statistical library.

find groups constituted only by samples evaluated by the same panel, but at a similarity value between 0.8 and 1.0. In the third group, there are only two samples of panel A without their couple from panel B.

**3.4. Correlation and Regression Analysis.**

Until now, the paper has analysed the panels by the attributes and their intensities detected in the samples. However, this analysis could be considered incomplete without a study of the final evaluation of every sample given by each panel.

A correlation program was run and its result, a correlation coefficient of 0.82 with a level of significance lesser than 0.001, can be considered good enough in general (Powers, 1984). However, the authors considered that some information could be pointed out, matching those values with the information kept in the factors.

Three correlation coefficients were also computed comparing the three factors of each panel with its final evaluation. The results, Table VII, show that the greatest correlation is with the first factor, which explains greater variance than the others, Table IV.

Table VII

Correlation coefficients between the final evaluation and the factors for each panel.

	Panel A		Panel B	
	Correlation coefficients	Significance level	Correlation coefficients	Significance level
factor1	0.9722	0.000	0.876	0.000
factor2	0.3215	0.110	-0.322	0.877
factor3	-0.1964	0.340	0.636	31E-5

The best correlation with the first factor shows that the panelists were good trained to assign the best level of the grading table according to the attributes previously identified and only following fuzzy rules. Thus, it is most important to remark the fuzzy logic and the statistics get similar results in this case. So, the authors are working on new research in which

they hope to explain by fuzzy linguistic terms how the panelists assign the points of the grading table.

Meanwhile, a multiple linear regression was used to improve the correlation coefficients displayed in table VII. It was performed on the basis of the RSWAP method to entering and removing the attributes, the final evaluation as dependent variable and the attri-

Table VIII

Regression models fitted to the data of attributes in each panel.

Panel	Regression model	Adjusted R-squared	Std. error of estimate
A	$y=4.51+.95*fruity-.31*fusty-2.04*metallic+.39*green-0.62*muddy$	0.97	0.25
B	$y=5.94+.86*fusty+2.29*rancid-1.27*rough-.72*apple$	0.92	0.31

butes as independent variables. Table VIII shows the regression model fitted to the data of each panel, besides the multiple R-squared and the standard error of the estimate, which can be used to compute the confidence intervals.

The regression models reflect what the panelists pondered upon the samples. The opinions were imbalances in panel B, the panelists were worried reaching from the negative notes of the samples, while they were more balanced in panel A. The regression model of panel A could be a good formula, after more evaluations of virgin olive oil samples, to get the best score of the grading table considering only the results of the profile sheet. In any case, this new alternative has better mathematical background than other proposals designed without any scientific base.

At this point, considering only attributes, paper has pointed out enough information for knowing whether a panel after a period of training can be considered at the same level of one expertise. The results show that panels can be considered as if they were one or, in other words, the panel with lesser experience evaluates the sample as well as the other with nearly ten years. However, the panelists of panel B has not attained and optimal knowledge about some attributes, rough and metallic and its evaluations are not good balanced.

#### 4. FUTURE TRENDS.

Even considering there are other statistical methods, (Sharaf et al, 1986), (Lebart et al., 1984), (Powers, 1984), the information taken from those statistical procedures could be considered as a methodology for knowing when this kind of panel has ended its period of training. Special attention might be paid to the information kept in the correlation matrix, where it is possible to find explanations about either mistakes or successes of the panel in training and its panelists.

However, some research on fuzzy linguistic terms and final evaluation are now under development. This is a new way by which it could be possible to remove the current mode of getting the panel's final evaluation, replacing it by an expert system working on fuzzy logic. Following the new approach that tries to join Artificial Intelligence and Sensory Evaluation (Galvin et al., 1990).

#### ACKNOWLEDGEMENTS

The authors would like to acknowledge their indebtedness to all those panelists who have contri-

buted to the development of the sensory evaluations. This work has been supported by CICYT-SPAIN grant, project ALI-88-0187-C02-02.

#### REFERENCES

- Bengtsson, K. (1943).- "Prowmking som analysmetod Statiskik behandling av resultatet".- Sven Bryggaregoren Manadsbl **58** (3), 102-149.
- Bisquerra, R. (1989).- "Introduccion conceptual al Analisis Multivariable".- PPU, Barcelona (Spain).
- BMDP Statistical Software.- 1981 Edition.- Dept. of Biomathematics, University of California, Los Angeles, CA, USA.
- Clapperton, J. F., and Piggott, J. R. (1979).- "Differentiation of ale and lager flavors by principal components analysis of flavor characterization data".- J. Inst. Brew. **85**, 271-274.
- Dillon, W. R. and Goldstein, M. (1984).- "Multivariate Analysis. Methods and Applications".- John Wiley, New York.
- Ennis, D. M., Boelens, H., Haring, H. and Bowman, P. (1980).- "Multivariate analysis in sensory evaluation".- Food Technol. **34** (11) 83-90.
- Galvin, J. R., and Waldrop, H.L. (1990).- "The future of Sensory Evaluation in the Food Industry".- Food Technol. **44** (1) 95-100.
- Gutierrez, F., Albi, M. A., Parma, R., Rios, J. J., and Olias, J. M. (1989).- "Bitter Taste of Virgin Olive Oil: Correlation of Sensory Evaluation and Instrumental HPLC Analysis".- J. Food Sci. **54**, 68-70.
- Harris, R. J. (1985).- "A primer of Multivariate Statistics".- Academic Press, Orlando.
- Lebart, L., Morineau, A. and Warwick, K. M. (1984).- "Multivariate Descriptive Statistical Analysis: Correspondence Analysis and Related Techniques for Large Matrices".- Wiley, New York.
- Olive Oil Council Authority (1987).- "Sensory Evaluation of olive Oil".- COI/T. 20/Doc. num. 3. June, 18.
- Powers, J. (1984).- "Using General Statistical Programs to Evaluate Sensory Data".- Food Technol. **38** (6) 74-82.
- Powers, J. (1988).- "Uses of Multivariate Methods in Screening and Training Sensory Panelists".- Food Technol **42** (11), 123-127.
- Resurrección, A.V.A. (1988a).- "Comparison of Flavor Quality of Peanut Based Pastes and Peanut Butter by Sensory Methods".- J. Food Sci. **53** (6), 1827-1830.
- Resurrección, A.V.A. (1988b).- "Applications of Multivariate Methods in Food Quality Evaluation".- Food Technol. **42** (11), 128-136.
- Sharaf, I. L., Illman, D., and Kowalski, B. R. (1986).- "Chemometrics".- Wiley, New York.
- Sheppard, R. N. (1980).- "Multidimensional scaling, tree-fitting and clustering".- Science **210**, 390-396.
- SPSS User's Guide (1986).- 2nd. Edition.- MacGraw-Hill Book Co.
- Stone, H., Sidel, I., Oliver, S., Woolsey, A., and Singleton, R. C. (1974).- "Sensory Evaluation by Quantitative Descriptive Analysis".- Food Technol. **28** (11), 24-34.
- Tabachnick, B. G. and Fidell, L. S. (1983).- "Using Multivariate Statistics".- Harper and Row, Publishers, New York.

(Recibido: Agosto 1990)