Technical News Feature

Progress Report of the AOCS Flavor Nomenclature and Standards Committee[1]

A.E. WALTKING, Best Foods Research and Engineering Center, CPC International Inc., 1120 Commerce Ave., Union, NJ 07083

ABSTRACT

Over the past 5 years, 2 collaborative studies have been organized to evaluate the gas liquid chromatographic (GLC) methods for flavor of edible oils. The most recent of these studies compared the results from 5 GLC procedures to the evaluations of the flavor panels of 8 laboratories. While the GLC procedures proved to be more precise than the panels, it was shown that separate correlation equations must be developed for each type of oil, each degree of hydrogenation or blending, and each manner of storage or abuse of the samples. Thus, either a flavor panel must always be available to establish a reference point for any study to be made, or GLC flavor methods must be recognized as providing a direct, but relative, evaluation for "oil quality."

The Flavor Nomenclature and Standards Committee was founded in 1967 as a sub-committee of the AOCS Standards Committee. While the committee's original purpose was to establish nomenclature, it has since expanded to that of defining methodology for oil flavor evaluation.

The original committee was composed of representatives of 10 organizations, denoted in Table I as charter members.

In the intervening years, the representation has broadened to a total of 28 members from 20 organizations. Since 1976, because of the introduction of gas liquid chromatography (GLC) into the analysis of flavor attributes, both analytical and sensory specialists from the same organizations have been accepted as members of the committee. Thus, 8 of the organizations last year had 2 members each.

Historically, this is the third progress report to be given at an annual meeting by a chairman of the committee. The first report was delivered by E. Hammond in 1970 in Chicago and the second by T. Smouse at the 1973 meeting in New Orleans.

To set the stage for a review of the committee's program since 1973, the unpublished highlights of the earlier activities will be summarized.

The initial program of the committee was a survey of the abilities of flavor panels to define flavor of edible vegetable oils. Three samples of soybean oil were evaluated by a total of 76 participating panelists. The results of the individual labs relative to the sample means are shown in Figure 1.

The laboratories were able to determine the differences among the samples, but there was significant variation among judges within a laboratory and among laboratories. While the above conclusions were derived from the flavor intensity evaluation, when it came to comparisons of the descriptions of flavor, the results were so diverse that they could not readily be evaluated by a statistical method. Bar graphs were drawn to demonstrate the descriptions most frequently given. For example, in Figure 2, sample 1 was more frequently rated watermelony and beany. Having demonstrated that the flavor of aged soybean oils is extremely complex, the committee attempted to zero-in on the characteristic flavors by collaboratively evaluating pure chemicals in a bland carrier. Preliminary screening of oils settled on the use of mineral oil as a carrier, to which 3 levels of each chemical were added.

Very little success was achieved with this approach. In fact, the vast number of flavor and odor descriptions given proved there was little agreement about how to describe the flavor sensation of pure chemicals, except for diacetyl. Since it was also observed that there was better agreement in flavor descriptions within laboratories than among laboratories, it was indicated that there might be a common flavor experience among flavor panel members in a given laboratory. Therefore, it was apparent that the diacetyl preparation gave better agreement because of the common experience of its general presence in butter and margarine.

Attempts were then made using specially prepared soybean oils treated under standard conditions to yield predictable flavor sensations as shown in Table II. Eight laboratories with 86 panelists evaluated this series. Mean values

TABLE I

1979 Members of the Flavor Nomenclature and Standards Committee

Anderson Clayton (2)[a]	Ralston Purina
Best Foods–CPC (2)	Rutgers University[a]
Campbell[a]	Swift
Central Soya (2)	USDA–NRRL (2)[a]
Fritzsche Dodge & Allcott	USDA–SRRL (2)
Glidden-Durkee (2)[a]	Foreign Groups
Hunt-Wesson (2)	
Iowa State University[a]	Lesieur-Cotelle (France)
Kraft Inc. (2)[a]	Inst. des Corps Gras (France)
Lipton	Inst. of Chem. Tech. (Czech.)
Procter & Gamble[a]	Unilever (Netherlands)[a]

[a]Charter member organizations.

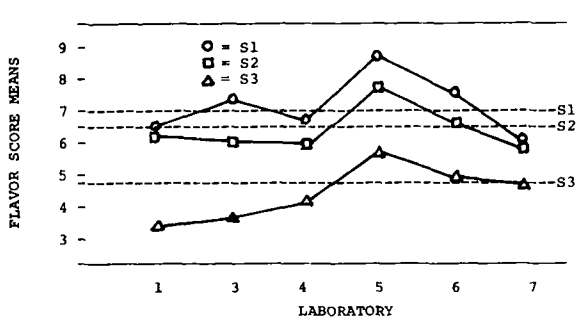

FIG. 1. Variation of laboratory means. Figure reproduced from slide in 1973 report of T. Smouse.

[1]Presented at AOCS national meeting in New York City, May 1980.

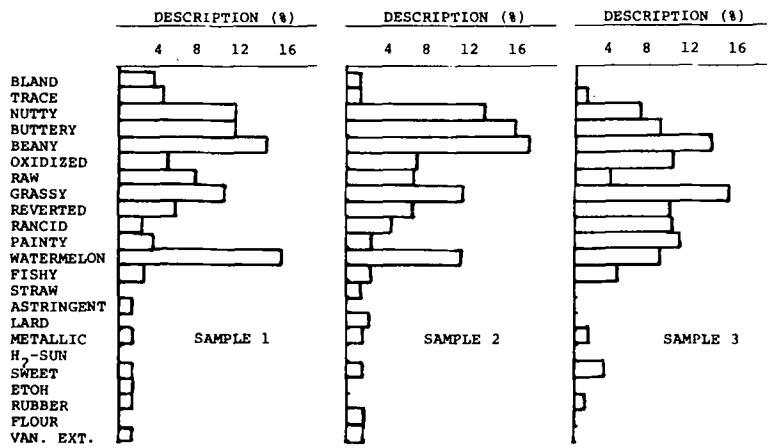

FIG. 2. Flavor descriptions of soybean oil collaborative samples. Reproduced from slide in 1973 report of T. Smouse.

for the laboratories were well within the expected scoring ranges and not significantly different from the grand means. The flavor descriptions obtained were sufficiently consistent to conclude that such oils could be used to train a panel not only for flavor grade, but possibly also for identifying oil type.

About this time, Harold Dupuy was beginning to have success in GLC separations of volatile flavor components and he was relating these to flavor panel perceptions (1-3). The committee decided that, before embarking on finalizing the procedure for flavor characterization, it would be of great value if an objective flavor technique could also be established. Thus, in 1973, the committee report ended with these recommendations: (a) to standardize panels: utilize specially prepared oils with known flavor descriptions and assigned flavor scores; (b) to continue studying the flavor effects of known chemical mixtures evaluated in a bland medium; and (c) to continue studying objective methods to replace subjective organoleptic evaluations. Since that time, efforts have concentrated on recommendation c.

Because of instrumental limitations at the time, direct use of Dupuy's technique was difficult and this led to several modifications. Thus, of the 5 laboratories participating in the GLC portion of the next collaborative study, only one of them used Dupuy's method and no 2 methods were identical (4-6). As had been done before, the collaborative study included soybean oils of varying levels of abuse. The samples were prepared as blends of fresh soybean oil and oil from the same lot which had been extensively heat- and light-abused. The data in Table III show a comparison of the laboratory's mean flavor scores to the grand mean score for all flavor panelists and indicate a ranking based on the difference between the laboratory's mean and the grand mean. Equations were obtained correlating the GLC data generated by each of the 5 laboratories to the flavor panel grand means of the individual samples. All equations showed good correlation with R^2 over .90. The GLC calculated flavor scores were compared to the grand means for the 4 samples and the laboratories were ranked according to the agreement between the flavor scores and the GLC scores in Table IV. Comparing the rank of the flavor panels with the rankings of the GLC data, it was notable that the organizations whose panels ranked higher also ranked higher in the GLC correlations. The overall conclusions from an analysis of variance of the results from this study indicated that there was no significant difference at a 95% confidence level among panelists; the samples were significantly different at the 99% confidence level and the correlation equations did an excellent job of predicting the flavor results using the instrumental data. A comparison of standard deviations of the flavor scores and GLC scores found in Table V shows that the GLC scores were less variable, even though a 2-standard-deviation rejection was used to remove outliers from the flavor scores.

The final collaborative study in the committee program involved the comparison of 4 levels of oil abuse for 3 types of oil. The manner of presenting the samples, the scoring systems and specified flavor descriptions which are shown in Table VI were all the same as those used in the last study to relate the results with those of the previous study. Thus,

TABLE II

Treated Soybean Oil for Predictive Flavors[a]

No.	Description expected (obtained)	Flavor grade expected (obtained)	Formulation[b]
1	Nutty, buttery (Bland, butter, nutty)	7-8 (7.2)	RBHBWD SBO
2	Raw, hydrogenated (Buttery, hydrogenated)	5-6 (5.5)	90% (1) + 10% RBHB SBO
3	Bitter (Bitter, buttery)	4-5 (4.9)	99.6% (1) + 0.4% Tween 20
4	Painty, rancid (Painty, rancid)	3-4 (1.6)	90% (1) + 10% O_2 SBO

[a]Reproduced from slide in 1973 report by T. Smouse.
[b]R = Refined, B = bleached, H = hydrogenated, W = winterized, D = deodorized.

TABLE III

Comparison of Laboratory's Mean Flavor Score to the Grand Mean Score for All Panelists

	Sample					
	05	51	55	69	$[(x - \bar{x})^2]$ [a]	Ranking
Grand mean	5.50	3.94	7.84	6.79	0	–
Lab 1	7.08	5.75	8.08	7.42	6.08	12
2	5.56	4.33	8.11	6.33	0.44	2
3 A	6.67	5.25	8.54	8.13	5.37	11
B	5.86	5.00	8.09	7.59	1.96	8
4 A	4.75	2.75	6.92	5.33	4.96	10
B	4.92	3.09	6.75	5.17	4.87	9
5 A	5.00	3.88	8.25	7.25	0.63	4
B	5.00	3.81	8.18	7.00	0.43	1
6	5.50	3.83	9.00	7.33	1.65	7
7	5.95	4.33	7.48	7.10	0.58	3
8	5.42	3.17	8.08	6.88	0.67	5
9	4.62	3.29	7.83	6.54	1.26	6

[a]As $\Sigma (x - \bar{x})^2$ approaches 0, the difference between the laboratory's means and the grand means approach 0 or no difference.

TABLE IV

Comparison of GLC Scores to the Grand Mean Score from All Panelists

	Sample						
	05	51	55	69	$[(x - x^{GLC})^2]$ [a]	GLC ranking	Flavor panel ranking
Grand mean	5.50	3.94	7.84	6.79	0	–	–
Lab 10	6.0	4.2	8.1	6.1	0.86	3	–
3	6.6	5.0	8.4	7.2	2.82	5	11&8
4	6.3	5.1	7.7	6.8	2.00	4	10&9
5	5.9	3.7	8.3	7.0	0.47	1	4&1
7	6.2	4.4	7.7	6.6	0.76	2	3

[a]As $\Sigma (x - x^{GLC})^2$ approaches 0, the difference between the GLC scores and the grand means approach 0 or no difference.

we hoped to establish whether the same correlation equations could be used for evaluating the soybean oil, even though the samples were now prepared individually with different levels of light and heat abuse, and not by blending single good and bad samples to different levels. Other comparisons to be made were between the unhydrogenated soybean oils and a set of partially hydrogenated soybean oils, as well as with a set of corn oils. Finally, the samples were prepared with less overall abuse to test the ability of both panels and GLC to discriminate less dramatic flavor differences. Most of the same laboratories participated.

As before, a ranking of all flavor panels relative to the deviation from the mean scores was obtained. Table VII shows the values for the soybean oil set. Underlined values indicate flavor scores which are out of order relative to the abuse applied to the samples.

Even the grand mean value for one of the samples is denoted as being incorrect. This is a result of misranking of this sample by 4 of the 8 panels.

Regression equations using both linear and natural log form were established for predicting flavor scores for each of the 5 different GLC procedures with each of the 3 sets of oil samples, as well as for a combination of all 3 sets. The coefficient of determination for the best equation for a given series is listed in Table VIII. The correlations were done against the grand mean values, excluding the misranked sample in the SBO series. Calculated flavor results were then obtained and compared to the grand mean values.

TABLE V

Standard Deviations among Laboratories

	Sample			
	05	51	55	69
Flavor score variation[a]	0.60	0.71	0.34	0.63
GLC score variation	0.27	0.58	0.33	0.42

[a]Outliers that were found to be 2 SD from the mean were rejected.

TABLE VI

Flavor Grading Scale

Flavor grade	Description of flavor[a]
10	Completely bland
9 (Good)	Trace of flavor but not recognizable
8	Nutty, sweet, bacony, buttery
7 (Fair)	Beany, hydrogenated
6	Raw, oxidized, musty, weedy, burnt, grassy
5 (Poor)	Reverted, rubbery, watermelon, butter
4	Rancid, painty
3 (Very poor)	Fishy, buggy
2	Intensive flavors and objectionable
1 (Repulsive)	

[a]Flavor intensity of descriptor is rated slight for grade level indicated and distinct at next lower grade level.

In almost all cases, the results were better when the 3 individual equations were used rather than the combined equation. This can be seen by comparing the average correlation value of 3 sets to the value for the combined equation.

Table IX is a comparison of the results from the best correlation equation from each laboratory for each sample set to the abuse levels originally imparted to the samples. A tabulation of the number of samples misranked by the GLC in comparison to the flavor panels shows the GLC results, on the average, to be superior. One GLC method never misranked a sample. Two other methods misranked only one pair of the hydrogenated soybean oil samples, which was not found to be significantly different by the flavor panels. The other 2 GLC procedures misranked another of the hydrogenated soybean oil samples that the panels found to be significantly different. However, all GLC procedures properly ranked another 5 samples(3 SBO, 1 corn, 1 hyd SBO) that individual flavor panels misranked.

As indicated earlier, these samples were intentionally prepared to have flavor differences that were small in comparison to the 10-point flavor scale used. When the differences were less than 0.5 of a flavor unit, the flavor panels did not identify them as different, as denoted by the same letter in the significance code, but with about a one-flavor unit difference, they were identified as significantly different.

TABLE VII

Comparison of Laboratory's Mean Flavor Score to the Grand Mean Score for All Panelists

	SBO[a]				SBO rank	Corn rank	Hyd SBO rank	Combined rank
Grand mean	5.41	7.97	6.49	7.82	–	–	–	
Lab 1	5.78	8.22	7.56	8.33	7	1	4	4
3	4.56	7.50	5.94	7.88	5	3	1	2
5	5.29	7.88	–	7.83	1	2	3	1
6	6.71	8.29	7.29	8.14	8	6	6	7
7	4.90	7.45	6.50	7.50	3	5	2	3
8	5.36	8.91	5.41	8.27	2	8	5	5
9	7.10	8.10	7.40	7.20	6	7	7	8
11	4.50	7.88	6.38	7.38	4	4	8	6
Abuse level	3	1	2	0				

[a]Underlined values are out of order relative to treatment abuse.

TABLE VIII

Correlation Values of Equations for Predicting Flavor Score by GLC

	R^2 of oil samples[a]				
Lab no.	SBO	Corn	Hyd SBO	Average of 3 sets	Combined
3	0.84	0.83 (ln)	0.61	0.76	0.76
5	0.96	0.93	0.83	0.91	0.84 (ln)
7A	0.88	0.99 (ln)	0.96	0.94	0.84 (ln)
7B[b]	0.96 (ln)	0.95	0.94	0.95	0.90 (ln)
10[b]	0.90	0.93	0.90	0.91	0.83 (ln)

[a]Equation linear unless otherwise designated.
[b]Analysis of 2,4 decadienal peak area. All others total area.

TABLE IX

Comparison of GLC and Flavor Panel Scores to the Grand Mean Score from All Panelists and the Abuse Level of the Samples

SBO series abuse	3	1	2	0	
Grand mean	5.41	7.97	6.49	7.82	
Flavor panels misranked	0	4	1	1	Total 6
GLC misranked	0	0	0	0	Total 0
Significance code	C	A	B	A	
Corn series abuse	1	0	3	2	
Grand mean	6.20	6.55	5.54	5.93	
Flavor panels misranked	0	0	1	3	Total 4
GLC misranked	0	0	0	1	Total 1
Significance code	A&B	A	C	B&C	
Hyd SBO series abuse	0	2	3	1	
Grand mean	7.83	6.38	5.48	7.46	
Flavor panels misranked	0	1	1	4	Total 6
GLC misranked	0	0	2	2	Total 4
Significance code	A	B	C	A	

Underlined values are misranked relative to abuse. Samples with the same significance code are not statistically different.

TABLE X

Standard Deviations among Laboratories

	Soybean series			
Abuse level	3	1	2	0
Flavor score variation	0.96	0.47	1.20	0.42
GLC score variation	0.18	0.20	0.20	0.12
	Corn series			
Abuse level	1	0	3	2
Flavor score variation	0.26	0.33	0.58	0.64
GLC score variation	0.10	0.09	0.08	0.15
	Hyd SBO series			
Abuse level	0	2	3	1
Flavor score variation	0.41[a]	1.33	0.71	0.33[a]
GLC score variation	0.21	0.25	0.37	0.37

[a]Outliers that were found to be 2 SD from the mean were rejected.

TABLE XI

Correlations Established for One GLC Method Relative to Results of the Flavor Panel of the Same Organization

	Soybean series				R^2
Abuse level	3	1	2	0	–
Lab flavor panel	4.90	7.45	6.50	7.50	–
GLC value	4.86	7.34	6.57	7.46	0.99
Difference	-0.04	-0.11	+0.07	-0.04	–
Grand mean	5.41	7.97	6.49	7.82	–
GLC value	5.46	7.63	6.50	8.10	0.88
Difference	+0.05	-0.34	+0.01	+0.28	–

Overall, the GLC methods were more accurate; that is, collectively, they misranked the oils fewer times than the group of flavor panels.

The standard deviations of the GLC and flavor panel results in Table X again prove that the GLC methods are more precise than the flavor panels when compared both individually and as a group.

The conclusions from these studies were that: (a) GLC flavor analysis can be more accurate and precise than individual flavor panels or even a group of flavor panels; (b) correlation equations can be obtained which will provide calculated flavor panel scores for various oil samples on an individual group of samples; (c) however, correlations will result in different equations if the samples: contain different oil types; were of the same oil but a different degree of hydrogenation; were of the same oil but stored or abused under different conditions. Therefore, no simple equation or group of simple equations can be established by one GLC system to be used with another GLC system.

Equations for establishing the relationship of general samples appear to be possible only on an individual, case-by-case basis, and not by formal collaboration. In other words, GLC correlation equations established by a given organization will not be valid relative to oils or GLC systems which vary at all with those from which the equations were derived.

As an example of this last statement, Table XI presents the data from one of the higher ranking flavor panels in the last collaborative compared with the GLC results from the same laboratory. Entirely different correlation equations now provide an almost perfect R^2 and a closer match on the same samples than that derived from the grand mean values. This means that both panels and GLC were sensitive enough to detect differences in the handling of samples among laboratories. Therefore, because of its accuracy and precision, GLC still appears to be a feasible tool to reduce the volume of samples required for flavor panel evaluations when the correlations can be established for a closely controlled system. Attempting to standardize the methodology to one GLC procedure would provide no advantage because all of the GLC methods tested could provide more accurate and precise values than the flavor panels. Also, the equations are not dependent on the method alone, but on the oil type, degree of hydrogenation and the temperature and light exposure of the oil. Thus it is the Committee's intention to propose recommended practices for inclusion into the AOCS methods book which will cover general procedures for flavor panel training and operation, as well as for establishing correlations with GLC systems measuring flavor volatiles.

REFERENCES

1. Dupuy, H.P., S.P. Fore and L.A. Goldblatt, JAOCS 46:876 (1971).
2. Dupuy, H.P., S.P. Fore and L.A. Goldblatt, Ibid. 50:340 (1973).
3. Dupuy, H.P., E.T. Rayner and J.I. Wadsworth, Ibid. 53:628 (1976).
4. Dupuy, H.P., E.T. Rayner, J.I. Wadsworth and M.G. Legendre, Ibid. 54:445 (1977).
5. Jackson, H.W., and D.J. Giacherio, Ibid. 54:458 (1977).
6. Waltking, A.E., and H. Zmachinski, Ibid. 54:454 (1977).

[Received October 30, 1981]