PN-AAK-679
ISN 1964

Postharvest Institute for Perishables

SEMINAR AND WORKSHOP ON
SENSORY EVALUATION OF FOOD
IN NICARAGUA

Report to
Laboratorio de Tecnologia de Alimentos (LABAL)
Ministerio de Industria
and to USAID/Managua

by
Dr. Ruth E. Baldwin
for the
Postharvest Institute for Perishables

GTS Report No. 81-04B
January, 1982

University of Idaho

in cooperation with
United States Agency for
International Development

Project Title: Storage and Processing of Fruits and Vegetables
Project No. AID/DSAN-CA-0265
Washington, D.C., U.S.A. 20523

Seminar and Workshop on Sensory Evaluation of Food in Nicaragua

Report to
Laboratorio de Tecnologia de Alimentos (LABAL)
Ministerio de Industria
and to
USAID/Managua

by
Dr. Ruth Baldwin
Department of Food Science and Nutrition
University of Missouri - Columbia
Columbia, Missouri 65211

January, 1982

LIST OF TABLES

TABLE		PAGE
1.	Classification of tests	4
2.	Summary record	18

LIST OF FIGURES

FIGURE		
1.	Schematic sequence for sensory tests	2
2.	Outline of general procedure for sensory tests	3
3.	Master sheet for a sensory test	6
4.	Sensory panel in progress	7
5.	Decoding data for rank analysis	8
6.	Planning the score sheet	10

TABLE OF CONTENTS

PAGE

I. Executive Summary i

II. Seminar-Workshop on Sensory Evaluation of Food 1

A. Introduction 1
1. Definition 1
2. Types of sensory evaluation 1
3. Classification of sensory test 1

B. General procedures for setting up and controlling discriminative tests 1

C. Analytical sensory tests 5
1. Discriminative tests 5
2. Affective tests 9

D. Selecting and training judges 9

E. Validity of sensory tests 11

III. Evaluation of Seminar-Workshop 13

IV. Recommendations 13

V. List of Participants 15

VI. Appendices
A. Schedule of activities 16
B. Paired difference test 18
C. Duo-trio difference test 22
D. Triangle test 25
E. Ranking test 28
F. Statistical analyses for difference tests 32
Chi-square for paired, duo-trio or triangle tests 33
Z-value 33
Analysis of scores for ranking 34
Testing for homogeneity among replications for triangle or paired tests 39
G. Scoring test 46
H. Statistical analyses for scoring test
Analysis of variance 49
Duncan's New Multiple Range Test 52
T-Test for paired experiments 55

I. EXECUTIVE SUMMARY

The USAID Mission in Managua, Nicaragua was requested by the Laboratorio de Tecnologia de Alimentos (LABAL), in the Nicaraguan Ministry of Industry, to obtain assistance from Dr. Ruth Baldwin of the Department of Food Science and Nutrition, University of Missouri. Dr. Baldwin was requested to spend six days in Managua working with LABAL:

1. To help Nicaraguan staff develop working norms and operating procedures for newly-acquired sensory evaluation equipment at a food laboratory.
2. To train staff in operation of equipment and interpretation of results. This equipment would be used to test new products to be introduced to market, such as bread with soya flour or fortified baby foods.

In response to telegram USAID/Managua 4212, Dr. Baldwin spent the period January 13-18, 1982 in Managua conducting a seminar and workshop on sensory evaluation, types of tests and analyses of data. Prior to the opening of the seminar a meeting was held with Dr. Gonzalo Pertz, Director of LABAL, to plan the entire program.

The seminar was opened with introductory remarks including a definition of sensory evaluation and a classification of sensory evaluation and a classification of tests as formulated by the Sensory Evaluation Divison of the Institute of Food Technologists. The role of sensory evaluation in the food industry was addressed briefly. Gustation and olfaction were discussed and group experience was provided. The remaining three days were devoted to seminar-workshop activities with preliminary discussion of tests followed by conducting the test and analyzing the data. Emphasis was placed on the most commonly used analytical sensory tests including difference tests and descriptive scoring. Four tests were completed during the workshop and others were discussed. Discussions also were directed toward procedures for selecting and training panelists and factors affecting validity of sensory tests. The seminar-workshop provided participants with a sound basis for use of sensory evaluation as a research tool and as a quality control measure.

There were 29 Nicaraguan participants in the program, representing 11 sections in LABAL or the National University. Dr. Pertz (LABAL) translated and assisted with the seminar-workshop. It was a valuable "hands on" experience for the Nicaraguan participants, and they appeared to benefit greatly from it. Therefore, they now have the working knowledge to begin sound sensory evaluation program. Because of the diverse interests of the participants, there should be wide dissemination of the information.

Another visit by a consultant may be desirable in the near future in order to provide additional interaction with those involved in conducting sensory tests. Also, it would be very beneficial for selected LABAL personnel to visit the U.S. for a few weeks association with a university and some food industries, or to pursue an advanced degree.

II. SEMINAR-WORKSHOP ON SENSORY EVALUATION OF FOOD

A list of activities by days is presented as Appendix A. The following is a review of content of the seminar-workshop as it was conducted.

A. Introduction

Sensory evaluation is a relatively new field which has become a valuable and essential part of the food industry. Sensory evaluation is used in product development, matching competitor's products, product improvement, evaluation of processing or ingredient changes, quality assurance, and in assessing consumer acceptance.

1. Definition. The Sensory Evaluation Division (SED) of the Institute of Food Technologists was formed about 10 years ago with approximately 90 members. This Division now has several hundred members. In 1975, the SED formulated the following definition of sensory evaluation:

   "A scientific discipline used to evoke, measure, analyze, and interpret reactions to those characteristics of foods and materials as they are perceived by the senses of sight, smell, taste, touch, and hearing."

2. Types of Sensory Evaluation. Sensory evaluation falls into four categories as follows:
   1) Bench screening: informal evaluations by researchers to screen out products or treatments that have little merit. (The researcher must exercise caution against eliminating products, due to a personal bias, which might ultimately receive consumer acceptance.)
   2) Expert evaluations: Buying and selling coffee, tea, whiskey, etc. is to a large extent determined by expert tasters. (There is a growing interest in use of sensory panels as opposed to experts.)
   3) Analytical testing by laboratory panels conducted under controlled conditions with selected and trained judges.
   4) Affective testing with large numbers of untrained panelists. Affective testing is always preceeded by analytical testing.

In general, sensory evaluation in the food industry follows the scheme presented in Figure 1. The focus group may or may not be included. The function of a focus group is to identify products or treatments which have potential for success. Generally, this is accomplished through informal discussions with consumers.

3. Classification of Sensory Tests. A summary of the various sensory tests and the kind of information obtained is presented in Table 1.

B. General procedures for setting up and controlling discriminative tests.

An outline of the steps involved in setting up and conducting discriminative tests is presented in Figure 2. First, a master, or control, sheet is prepared, and secondly sample containers must be coded. Either, simultaneously or after this, codes can be recorded on the score sheets. A controlled system for portioning samples can be patterned after the master sheet. Finally, the samples are assembled for each judge, and data are collected, analyzed and interpreted.

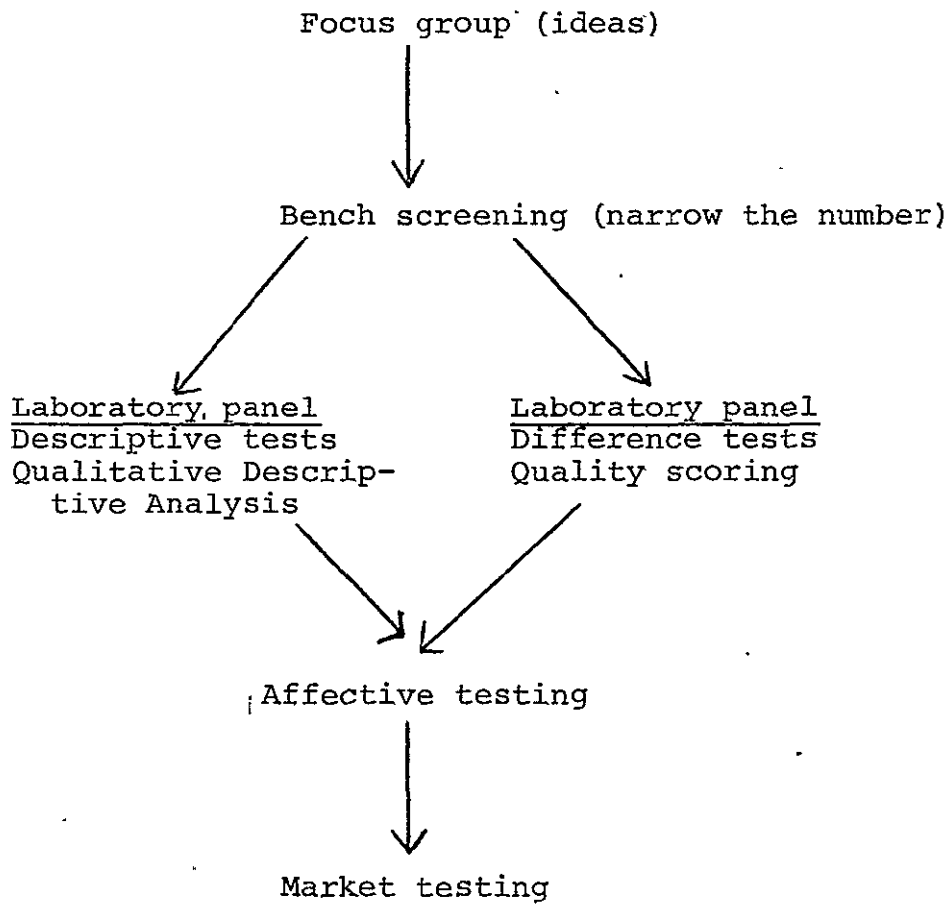


Fig. 1. Schematic sequence for sensory tests.

GENERAL PROCEDURE FOR SENSORY TESTS

1. Prepare master sheet
   a. Number of judges and/or sets
   b. Order of presentation of samples
   c. Random number assignment
2. Code sample containers (can be done as random numbers are assigned)
3. Prepare score sheets
4. Prepare samples
   a. Mark table
   b. Assemble sample containers
   c. Put samples in containers
5. Set up trays and serve
6. Evaluate samples
7. Decode score sheets
8. Determine statistical significance of test

Fig. 2. Outline of general procedure for sensory tests.

Table 1. Classification of tests.

Type of test	Specific test	Type of Panel	Information obtained
Analytical			
Difference	Paired stimuli	Trained	Product A = B Product A ≠ B
	Paired comparison	Trained	Product A > B Product A < B Product A = B
	Duo-Trio	Trained	Product A > B Product A < B
	Triangle	Trained	Product A > B Product A < B
Descriptive	Scoring	Trained	Score for one or several products on the basis of a predetermined scale for specific attributes.
	Profile	Highly Trained	Word description of components of aroma, flavor, and/or texture.
	Quantitative Descriptive Analysis	Highly Trained	Score for components of aroma, flavor and/or texture. schematically presented.
Affective	Paired	Un-trained	Product A > B Product A < B
	Ranking	Un-trained	Two or more products ranked according to preference.
	Food Action Rating Scale	Un-trained	Indication of how often a product would be eaten or would be purchased.

Essentially, the master sheet, Fig. 3, is a record of the judges, the sample codes and the sequence of serving samples. For some tests, this sheet can provide a convenient form for recording and summarizing the data collected during the test.

Fig. 4 is a photograph of a sensory panel in progress during the workshop.

C. Analytical sensory tests.

The seminar-workshop provided time for discussion and practice with several analytical type sensory tests with examples of both discriminative and descriptive tests.

1. Discriminative tests. Typical of the discriminative test are difference tests including the paired test (Appendix B), the duo-trio test (Appendix C), the triangle test (Appendix D), and the ranking test (Appendix E). During the seminar-workshop, the paired test, the triangle test and a ranking test were planned, conducted, and data were analyzed (Appendix F). Data were summarized (Fig. 5) for all participants to see. The duo-trio test was discussed but no laboratory experience was provided with this test. Of the descriptive tests, scoring (Appendix G) was the only one considered in detail.

DATE_____

TEST_____

PRODUCT_____

MASTER SHEET

JUDGE		SAMPLE CODES AND ORDER OF SERVING										TOTAL CORRECT
TOTAL CORRECT												

Fig. 3. Master Sheet for a sensory test.

MISSING PAGE

NO. 7 & 8

during the workshop. Brief mention was made of magnitude estimation, profiling and quantitative descriptive analysis (QDA).

A score sheet was developed for tomato catsup by workshop participants planning in small groups (Fig. 6). These were presented to the entire group for evaluation. The test was conducted, data were analyzed by groups of workshop participants according to analysis of variance (ANOV) and Duncan's New Multiple Range test (Appendix H). A t-test (Appendix H) was suggested for use when only two products or treatments are scored.

The multiple sample tests in which samples are scored in relation to a reference (R) was addressed briefly. The problem arising in selecting an appropriate R was noted.

2. Affective tests. Affective tests were discussed briefly. Emphasis was on the importance of simplicity for these tests and the need for a large number of panelists. The difficulty of conducting this kind of test in Nicaragua was addressed. It is more than likely that it will be necessary for LABAL employees to act as the affective panel at least for the present time.

D. Selecting and training judges.

Criteria for selecting panelists for analytical sensory tests were discussed. Qualifications for judges were enumerated as follows: interest, motivation, cooperativeness, health, available time, freedom from anosmias

MISSING PAGE

NO. 10

and ageusias, age (20-50 years), either male or female, willingness to forgo perfumed cosmetics and to refrain from smoking for a period prior to panels.

General suggestions for training panelists were given. These included familiarization of panelists with the phenomena of gustation, olfaction, kinesthetics, fatigue or adaptation and with general techniques of sensory testing and test setup.

Further training was suggested to permit panelists to become familiar with the characteristics of the product and with the range of variation that might be encountered. The importance of reproducibility and consistency of judges was stressed. It was recommended that judges should be kept informed of their performance and that practice sessions with a variety of tests should be conducted. Use of coded duplicate samples can be used to assess the reproducibility of judges.

The limitations of time, money and sample as they influence the extent of training was mentioned. Also, a simple reward system was suggested for panelists.

E. Validity of sensory tests.

The validity of sensory tests was stressed in summarizing the seminar-workshop. The necessity of stating the objective and selecting an appropriate test was emphasized. Invalidity can enter a sensory test at any point. A brief summary record (Table 2) was recommended as a final step in any sensory evaluation.

Table 2. Summary record.

Element	Information
Objective	Clear, concise statement
Test	Type of test and any remarks relative to it.
Samples	Description
Replications	Number
Analysis of data	Description
Results	Summary table and/or graph
Recommendations	Based on results, what action should be taken?
Project leader	Signature and date

III. EVALUATION OF SEMINAR-WORKSHOP

It is believed that this was a successful "hands on" experience for LABAL personnel and was informative for representatives from the Ministry of Health. Participants appeared to be interested, were willing to become actively involved, were enthusiastic and cooperative.

It was an informal atmosphere with questions posed freely. LABAL personnel now have the working knowledge to begin a sound sensory evaluation program. Their facilities are limited in size, but are functional. There appears to be some difficulty in obtaining supplies which we tend to take for granted in the U.S.A.

Much of the success of the seminar-workshop was, undoubtedly, due to the input of Sr. Gonzalo Pertz of LABAL. Sr. Pertz translated and facilitated in every way possible.

IV. RECOMMENDATIONS

It is likely that another visit by a consultant would be beneficial in the not too distant future. Training panelists can be extremely frustrating. A consultant might be able to assist at this point. Although literature and information for this phase of their development will be sent to LABAL by mail, a need for additional interaction with those involved in conducting sensory tests is probable.

It also seems that if some provision could be made for LABAL personnel to come to the U.S.A. for training

in various areas, it would be extremely beneficial. In some instances a few weeks association with a university would be appropriate along with some prearranged contacts with the food industry. In other instances, it appears that individuals from LABAL might pursue an advanced degree with the stipulation that they return to LABAL so that their program will be enhanced.

MINISTERIO DE INDUSTRIA

Al Contestar Refiérase A

V. LIST OF SEMINAR PARTICIPANTS

G. Pertz	Director, LABAL
A. Hernández	Technical Information, Library, LABAL
Julio Aguilar	Head, Food & Vegetables Section, LABAL
Ma. E. Salgado	Fruits & Vegetables Section, LABAL
R. Herrera	Fruits & Vegetables Section, LABAL
G. Granados	Seafoods Section, LABAL
V. González	Seafoods Section, LABAL
J. Vallecillo	Dairy & Meats Section, LABAL
R. Lazo	Dairy & Meats Section, LABAL
H. Valbuena	Dairy & Meats Section, LABAL
J.J. Collán	Cereals & Oils Section, LABAL
R. Arguello	Cereals & Oils Section, LABAL
J. López	Cereals & Oils Section, LABAL
M. Ruíz	Analytical Chemistry Section, LABAL
S. Salazar	Analytical Chemistry Section, LABAL
D. Reyes	Analytical Chemistry Section, LABAL
M.J. Pérez	Analytical Chemistry Section, LABAL
Y. Flores	Analytical Chemistry Section, LABAL
D. Villareal	Analytical Chemistry Section, LABAL
R. Ma. Prayia	Microbiology Section, LABAL
F. González	Microbiology Section, LABAL
S. Quintana	Cereals & Oils Section, LABAL
A. Rivera	Food Control Lab., Ministry of Health
M. Fajardo	Food Control Lab., Ministry of Health
E. Zuñiga	Food Control Lab., Ministry of Health
F. Juarez	Food Control Lab., Ministry of Health
J. Landrian	Food Inspector, Ministry of Health
L. Toruño	Coll. of Agriculture, National University
J.B. Franco	Coll. of Agriculture, National University

Appendix A: Schedule of Activities by Days

MINISTERIO DE INDUSTRIA

Al Contestar Refiérase A

SCHEDULE OF ACTIVITIES

I. January 12, 1982: Travel

II. January 13, 1982

Morning - Discussion and organization.

Afternoon-Seminar:

- Introduction and definition.
- Gustation and olfaction.
- Classification of sensory tests.

III. January 14, 1982

Seminar and workshop.

- Paired test.
- Duo-trio test.
- Triangle test.

IV. January 15, 1982

Seminar and workshop.

- Scoring.
- Analysis of variance.

V. January 16 and 17, 1982

- Planning for final session, analysis of data, preparation of preliminary report.
- Becoming acquainted with Nicaragua.

.....

VI. January 18, 1982

- Seminar and workshop.
- Discussion of data collected January 15, 1982.
- Completion of statistical analysis.
- Ranking test.
- Discussion of selection and training of panelists.
- Summary.

VII. January 19, 1982: Travel

REB/bdm

Appendix B: Paired Difference Test

PAIRED DIFFERENCE TEST

I. Description

A psychometric or psychophysical method in which stimuli (samples) are presented in pairs for comparison on the basis of some defined criterion such as intensity or degree of a defined quality.

The paired difference test can be classified as paired comparison or paired stimuli. The questions answered by these two tests are as follows:

1. Paired comparison:
   Which sample is stronger, sweeter, etc.?

2. Paired stimuli:
   Do the two samples differ?

If question 1 is asked, the possible permutations for the two samples (A and B) are AB, BA. If question 2 is asked, the possible permutations for the two samples are AA, AB, BA, BB.

Common applications of the paired difference test are in comparing samples processed by new and by old processing procedures, as a quality control test, and for determining thresholds. This test should not be applied if the samples are different upon visual appraisal. This method is an excellent procedure for testing judge performance when applied as a paired stimuli.

II. Analysis of data

Data for a paired difference test may be analyzed as follows:

1. For a quick test to determine if there is a statistically significant difference between treatments or samples, consult a table of significance for paired tests (one-tailed). (See Table D, page 525, Amerine et al, 1965.) This table lists the number of correct judgments required to indicate a significant difference with confidence levels of 95%, 99% or 99.9% (columns headed respectively .05, .01, .001).

2. For a more precise analysis, the chi-square method is appropriate.

3. Calculation of a z-value yields the exact probability of the number of correct identifications being chosen by chance and the confidence level.

I. References

Amerine, M. A., Pangborn, R. M., and Roessler, E. B. 1965. "Principles of Sensory Evaluation of Food." Academic Press, New York City NY. pp. 440-443, 521, 523.

Larmond, E. 1977. "Laboratory Methods for Sensory Evaluation of Food." pp. 23-24, 29-31, 64.

Dawson, E. H., Brogdon, J. L. and McManus, S. 1963. Sensory testing of differences in taste. Food Tech. 17(9):45-48, 51.

Peryam, D. R. 1958. Sensory difference tests. Food Tech. 12(5):231-236.

Schwartz, N. and Pratt, C. H. 1956. Simultaneous vs. successive presentation in a paired comparison situation. Food Research 21(1):103-108.

CODES: Taken from table of random numbers.

ORDER OF SERVING (✓) determined by table of random permutations

DATE 1-13-82

TEST PAIRED

PRODUCT Apple juice

A= control

B= 1% sucrose added

MASTER SHEET

JUDGE		SAMPLE CODES AND ORDER OF SERVING										TOTAL CORRECT
		A		B								
1		395		463 ✓								
2		067		695 ✓								
3		681 ✓		279								
⋮												
TOTAL CORRECT												

SET_____

JUDGE_____

DATE_____

PAIRED COMPARISON

TASTE SAMPLES IN THE ORDER THEY ARE LISTED BELOW, STARTING WITH THE ONE ON YOUR LEFT. RINSE YOUR MOUTH WITH WATER BEFORE TASTING EACH SAMPLE. CIRCLE THE NUMBER CORRESPONDING TO THE CODE FOR THE SAMPLE WITH THE SWEETER TASTE.

_____ _____

Appendix C: Duo-trio Test

DUO-TRIO DIFFERENCE TEST

I. Description

The duo-trio difference test is a modification of the paired difference test. One of a pair of samples (randomly selected) is identified as a reference (R) and presented first. (However, some researchers designate the control (or normal sample) as R rather than randomly selecting the sample designated as R. After the evaluation of R, the two differing samples (A and B) are presented in random order as coded samples. The panelists determine which of the coded samples is identical to the reference. The possible permutations for the coded samples are the same as for paired comparison. Possible permutations including the reference are A-AB, A-BA, B-AB, B-BA.

This test is useful in quality control and as a method for selecting judges with superior abilities in discrimination. This test is not suited to determining differences in intensity as is the paired comparison. Also, it is not applicable if samples' differences are distinguishable by visual appraisal.

II. Analysis of data

The analysis of data is as described previously for paired difference tests.

III. References

Amerine et al. 1965. "Principles of Sensory Evaluation of Food." pp. 276, 333-334, 547.

Larmond, E. 1977. "Laboratory Methods for Sensory Evaluation of Food." pp. 29-31, 64.

Gridgeman, N. T. 1955. Taste comparisons: Two-samples or three? Food Tech. 9(3):1-8.

Mitchell, J. W. 1956. The effect of assignment of testing materials to the paired and odd position in the duo-trio taste difference test. Food Tech. 10(4):169-171.

CODES: Taken from table of random numbers
ORDER OF SERVING(✓) determined by table of random permutations.

DATE 2-13-82
TEST Duo-trio
PRODUCT Apple juice
A = Control
B = 1% Sucrose added

MASTER SHEET

JUDGE		SAMPLE CODES AND ORDER OF SERVING										TOTAL CORRECT
		R										
		A	B		A		B					
1			R✓		897✓		740					
2		R✓			478		752✓					
3			R✓		817✓		836					
4			R✓									
⋮												
⋮												
TOTAL CORRECT												

SET_____

JUDGE_____

DATE_____

DUO-TRIO DIFFERENCE TEST

RINSE YOUR MOUTH WITH WATER BEFORE TASTING EACH SAMPLE. BEGIN TASTING WITH THE SAMPLE MARKED R. ONE OF THE CODED SAMPLES IS THE SAME AS THE REFERENCE (R). TASTE THE CODED SAMPLES BEGINNING WITH THE ONE ON YOUR LEFT. CIRCLE THE CODE NUMBER CORRESPONDING TO THE SAMPLE WHICH IS THE SAME AS R.

_____ _____

Appendix D: Triangle Test

TRIANGLE DIFFERENCE TEST

I. Description

The triangle difference test consists of three coded samples which are presented simultaneously. Two of the samples are identical and one is different. The panelist is required to identify the like samples or to indicate the one sample which is different. The possible permutations are AAB, ABA, ABB, BAA, BAB, BBA.

To adapt the master sheet for use with the triangle test, head one column for order of serving and two columns, each, for A and B samples. The order of serving will indicate which A and B columns should be assigned random numbers. The table mark-up should also have two columns, each, for A and B samples.

The triangle test can be applied to determine differences due to processing or ingredient changes, as a quality control test, and as a screening and training test for panelists. It is not suitable if samples are distinguishable by visual appraisal, when there is a carry-over flavor, or if fatigue is a factor.

II. Analysis of data

1. For a quick test to determine if there is a statistically significant difference between treatments or samples, consult a table of significance for triangular tests. (See Table E, page 526, Amerine et al. 1965.

2. For a more precise analysis, the chi-square method is applied.

3. Calculation of the z-value yields the exact probability of the number of correct identifications being chosen by chance and the confidence level.

II. References

Amerine et al. 1965. "Principles of Sensory Evaluation of Food." pp. 276, 335-338, 563.

Byer, A. J. and Abrams, D. 1953. A comparison of the triangular and two-sample taste-test methods. Food Tech. 7(4):185-187.

Larmond, E. 1977. "Laboratory Methods for Sensory Evaluation of Food". pp. 22-23, 63.

AAB = 1
ABA = 2
ABB = 3
BAA = 4
BAB = 5
BBA = 6

CODES: Taken from table of random numbers

Determine order by table of random permutations

MASTER SHEET

DATE 2-13-82

TEST TRIANGLE

PRODUCT Apple juice

A = control

B = 1% sucrose added

JUDGE	ORDER	SAMPLE CODES AND ORDER OF SERVING										TOTAL CORRECT
			A	A		B	B					
1	ABA		614	760		264						
2	BBA		465			739	939					
3	AAB		419	807		925						
4	ABB		396			869	602					
5	BAB											
6	BAA											
7	ABB											
⋮												
TOTAL CORRECT												

SET _____

JUDGE _____

DATE _____

TRIANGLE TEST

RINSE YOUR MOUTH WITH WATER BEFORE TASTING EACH SAMPLE. START TASTING WITH THE SAMPLE ON YOUR LEFT AND PROGRESS TO THE RIGHT. TWO SAMPLES ARE IDENTICAL AND ONE IS DIFFERENT. CIRCLE THE CODES CORRESPONDING TO THE IDENTICAL SAMPLES.

_____ _____ _____

Appendix E: Ranking Test

RANKING TEST

I. Description

The ranking test is a psychometric method that may be used in comparisons of a series of randomly coded samples evaluated simultaneously. The panelist is required to place the samples in order according to a designated characteristic such as intensity. The series of samples may include a coded control or standard. If more than one characteristic for one product is evaluated, a separate ranking test should be conducted for each characteristic.

This test can be applied to determine amount of a new ingredient required to equal the intensity of the original ingredient in terms of a specific quality such as sweetness, thickness, color, etc.

II. Analysis of data

A. For a quick analysis of ranks, proceed as follows:

1. Assign numerical values (1 to n) starting with one for the sample ranked as least intense.

2. Total these numerical values to obtain a rank sum for each treatment or sample.

3. Consult a table of rank totals required for significance. (See Table I-1, pages 536-537, Amerine et al. 1965.) Enter the left column of the table according to the number of judges (replications). Move across the table to the appropriate column for the number of treatments (or samples) in the test.

   a. Determine if rank total fall within the top two figures of the block (no significant ($P<0.05$) difference). If rank totals fall outside the values in the table (i.e. above and below), they are significantly ($P<0.05$) different.

   b. If two rank totals fall within the range of the top two figures in the block, they may be compared by re-ranking as 1 and 2, depending on their magnitude. Then, enter the table for 2 treatments and use the lower two figures, since these are predetermined treatments or samples. Any two rank totals may be compared in this manner.

B. For a more precise analysis of the data, follow the handout titled "Analysis of Scores for Ranking". (See statistical analyses.)

II. References

Amerine et al. 1965. Principles of Sensory Evaluation of Food. pp. 276, 350-354, 536, 558.

Kramer, A. 1960. A rapid method for determining significance of difference from rank sums. Food Technol. 14(11):576.

Kramer, A. 1963. Revised tables for determining significant differences. Food Technol. 17:1596.

(Another reference is listed on "Analysis of Scores for Ranking."

CODES: Taken from table of random numbers
ORDER OF SERVING (1,2,3,4) determined by table of random permutations.

DATE 2-14-83
TEST Ranking
PRODUCT Solutions

MASTER SHEET

JUDGE		SAMPLE CODES AND ORDER OF SERVING											TOTAL CORRECT
		Sucrose											
		0.5%		1%		2%		4%					
1		4 252		3 712		1 497		2 430					
2		1 151		2 973		3 607		4 490					
3													
⋮													
TOTAL CORRECT													

DATE _____

JUDGE _____

RANKING FOR SWEETNESS

DIRECTIONS: RINSE YOUR MOUTH WITH WATER BEFORE TASTING EACH SAMPLE. TASTE THE SAMPLES IN THE ORDER PRESENTED, STARTING WITH THE ONE ON YOUR LEFT AND PROGRESSING TO THE RIGHT. RANK THE SAMPLES IN ORDER OF INCREASING SWEETNESS AND RECORD THE SAMPLE CODES BELOW TO INDICATE YOUR RANKING.

_____ LEAST SWEET

_____

_____

_____ MOST SWEET

STATISTICAL ANALYSIS OF DATA

Chi-square for Paired, Duo-Trio or Triangle Tests

Null hypothesis (Ho) - There is no difference between the two products (or treatments).

Alternate hypothesis (Ha) - There is a difference between the two products (or treatments).

$$X^2 = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2}$$

O_1 = observed number of correct identifications
O_2 = observed number of incorrect identifications
E_1 = expected number of correct identifications
E_2 = expected number of incorrect identifications

Note: For paired difference test and the duo-trio difference test, E_1 and E_2 would be 1/2 of the total number of observations. The chance of a correct identification by chance alone is 0.5. For the triangle difference test, E_1 = 1/3, E_2 = 2/3 of the total number of observations. The chance of a correct identification by chance alone in this test is 0.333.

Since you are determining only if there is a significant difference between two products (or treatments), there is only one degree of freedom in either case.

Consult a table for chi-square distribution (See Table C-2, page 524, Amerine et al. 1965.) to determine if the calculated X^2 value exceeds the 99 or 95% confidence levels (A= .01 or A= .05, one degree of freedom). The paired difference and triangle are one-tailed tests; therefore, enter the table at A which is twice that of the probability level you are reporting. For example, for a 99 or 95% confidence level, A= .02 or .10, respectively. If the calculated X^2 value exceeds the value listed in the table, the difference is considered statistically significant at the particular confidence level, and the Ho is rejected and the Ha is accepted.

Appendix F: Statistical Analyses for Difference Tests

1. Chi-square for paired, duo-trio or triangle tests
2. Z-value
3. Analysis of scores for ranking
4. Testing for homogeneity among replications for triangle or paired tests

STATISTICAL ANALYSIS OF DATA

z values for paired or triangle tests:

$$z = \frac{(X - np) - 0.5}{\sqrt{npq}}$$

X = Number of correct responses

n = Total number of judgments

p = Probability of correct decision by chance
Paired test = 1/2 (1-tailed difference test)
Triangle test = 1/3

q = 1 - p

0.5 = continuity factor

Consult Table for z values to determine the probability of this choice being made by chance. (See Table A, page 521, Amerine et al. 1965.)

Enter the left column of the table utilizing the z value (accurate to the first decimal place). Then, locate the column headed by the second decimal place of the z value. The exact probability of the number of correct responses occuring by chance alone will be located at the intersection of the column and row.

Example: z value = 1.53
probability = 0.0630

If the z value is negative, enter the Table as explained above. Subtract the probability from 1 to obtain the probability of obtaining the number of correct responses by chance alone.

ANALYSIS OF SCORES FOR RANKING

See Example I for Steps 1 through 4.

1. Assign a numerical value (1 to n) to each sample, starting with one for the sample ranked as least intense.

2. Add ranks to obtain a total rank score (Σ) for each treatment. List these total rank scores in order of magnitude.

3. Use the following formula to calculate a statistic called chi-square for ranks (χ_r^2):

$$\chi_r^2 = \frac{12}{np(p+1)} \times (\Sigma_1{}^2 + \Sigma_2{}^2 + \ldots) - 3n(p+1)$$

p = number of treatments

n = number of judgments

$\Sigma_1{}^2$ = (sum of treatment 1)^2

12 and 3 are constants

$p - 1$ = degrees of freedom (df)

4. Look up significance of calculated statistic on a table for chi-square (χ^2). (p. 524 Amerine et al. 1965, use A = .05)

   a. Enter table on left according to df.

   b. For sensory tests, $P < 0.05$ or $P < 0.01$ are generally applied to the data. Therefore, if the calculated χ^2 is equal to or exceeds the value in the table, the difference in ranks is statistically significant at the designated level of probability. If the null hypothesis is accepted, no further analysis is needed.

See Examples II through VII for Step 5.

5. If this test shows that treatments differ significantly, individual pairs of treatments may be tested by this method.

   a. Compare the smallest and the largest rank totals.

      1. Recalculate ranks for these pairs as 1 and 2 according to their magnitudes.

      2. Calculate a statistic for χ_r^2 as before. (df = 1)

      3. Determine level of significance. (This must be statistically different if you obtained a significant difference in 4.b. above.)

   b. Continue comparing rank totals in this way until all comparisons have been made.

6. For a fast, less accurate method of comparing total rank scores, see pages 476, 536 and 538 of Amerine et al. (Consider judges as replications.)

Calculations for Analysis of Ranks (χ_r)

Example I.

List rank totals from master sheet in order from smallest to largest as follows:

Treatments	Order of Magnitude	Rank total	n = 5 judges
A	1	8	
B	2	10	
C	3	13	
D	4	19	

$$\chi_r^2 = \frac{12}{5x4(4+1)} \times (8^2+19^2+10^2+13^2) - 3x5(4+1)$$

$$\chi_r^2 = \frac{12}{100} \times 694 - 75$$

$$\chi_r^2 = 83.28 - 75 = 8.3$$

Critical value from χ^2 table, 3 degrees of freedom, $P<0.05 = 7.82$. Therefore, there is a signigicant ($P<0.05$) difference.

Example II.

Compare smallest (treatment A) and largest (treatment D) ranks and recalculate their ranks as 1 and 2 according to their magnitudes as follows:

Ranks		Recalculated Ranks	
Treatment A	Treatment D	Treatment A	Treatment D
2	4	1	2
1	3	1	2
1	4	1	2
1	4	1	2
3	4	1	2
total 8	19	5	10

$$\chi_r^2 = \frac{12}{5x2(2+1)} \times (5^2 + 10^2) - 3x5\ (2+1)$$

$$\chi_r^2 = 5$$

Critical value from χ^2 table, 1 degree of freedom, $P<0.05 = 3.84$. Therefore treatments A and D are significantly ($P<0.05$) different.

Example III.

Compare ranks for treatments A and C and recalculate their ranks according to magnitude as follows:

	Ranks		Recalculated Ranks	
	Treatment A	Treatment C	Treatment A	Treatment C
	2	1	2	1
	1	4	1	2
	1	3	1	2
	1	3	1	2
	3	2	2	1
total	8	13	7	8

$$\chi_r^2 = \frac{12}{5x2(2+1)} \times (7^2 + 8^2) - 3x5\ (2+1)$$

BEST AVAILABLE

$$\chi_r^2 = 0.2$$

Critical value from χ^2 table, 1 degree of freedom, $P<0.05 = 3.84$. Therefore, treatments A and C are not significantly different.

Example IV.

Compare ranks for treatments A and B and recaluclate their ranks according to their magnitudes, as follows:

	Ranks		Recalculated Ranks	
	Treatment A	Treatment B	Treatment A	Treatment B
	2	3	1	2
	1	2	1	2
	1	2	1	2
	1	2	1	2
	3	1	2	1
total	8	10	6	9

$$\chi_r^2 = \frac{12}{5x2(2+1)} \times (6^2 + 9^2) - 3x5\ (2+1)$$

$$\chi_r^2 = \frac{1.2}{3} \times (36 + 81) - 15x3$$

$$\chi_r^2 = .4 \times 117 - 45$$

$$\chi_r^2 = 1.8$$

Critical value at $P<0.05 = 3.84$. Therefore, treatments A and B do not differ significantly.

BEST AVAILABLE

Example V.

Compare ranks of treatments B and C and recalculate their ranks according to their magnitude as follows:

	Ranks		Recalculated Ranks	
	Treatment B	Treatment C	Treatment B	Treatment C
	3	1	2	1
	2	4	1	2
	2	3	1	2
	2	3	1	2
	1	2	1	2
otal	10	13	6	9

Recalculated rank totals are the same as for the comparison of treatments A and B in Example IV. Therefore, no calculations are necessary here. Treatments B and C do not differ significantly.

Example VI.

Compare ranks for treatments B and D and recalculate their ranks according to their magnitudes as follows:

	Ranks		Recalculated Ranks	
	Treatment B	Treatment D	Treatment B	Treatment D
	3	4	1	2
	2	3	1	2
	2	4	1	2
	2	4	1	2
	1	4	1	2
tal	10	19	5	10

Recalculated rank totals are the same as for the comparison of treatments A and D in Example II. Therefore, no calculations are necessary here. Treatments B and D differ significantly.

BEST AVAILABLE

Example VII.

Compare ranks for Treatments C and D and recalculate their ranks according to their magnitudes as follows:

	Ranks		Recalculated Ranks	
	Treatment C	Treatment D	Treatment C	Treatment D
	1	4	1	2
	4	3	2	1
	3	4	1	2
	2	4	1	2
	2	4	1	2
otal	13	19	6	9

Recalculated rank totals are the same as for the comparison of treatments A and B in Example IV. Therefore, no calculations are needed here. Treatments A and B do not differ significantly.

CONCLUSION:

Treatment A and Treatment B differ significantly ($P<0.05$) from treatment D. No other significant differences.

Reference:

Wilcoxin, F. 1949. "Some Rapid Approximate Statistical Procedures," American Cyanimid Co., 30 Rockefeller Plaza, New York. 16pp.

TESTING FOR HOMOGENEITY AMONG REPLICATIONS FOR TRIANGLE OR PAIRED TESTS

A test for homogeneity, chi square (χ^2), is applied to data collected from different replications (each having the same number of judges) of a triangle or paired test to determine whether or not these data can be pooled (added). If the χ^2 value for homogeneity is statistically significant, we accept the alternate hypothesis that the replications are heterogeneous. If the replications are heterogeneous, data cannot be pooled or added because there is more variation among replications than there is among treatments.

Formula: $$\chi^2 = \frac{\Sigma(f_o - f_e)^2}{f_e}$$

f_o = frequency (number) of correct responses observed

f_e = frequency (number) of correct responses expected

degrees of freedom (df) = number of replications - 1

<u>Comparison of two treatments</u>

The number of correct responses is determined for each replication. Next, the mean of the correct responses for all replications is calculated. This mean value is the expected value in the χ^2 analysis.

<u>Example</u>: Three replications were conducted with triangle tests to determine if treatments A and B differed significantly. The following data were obtained:

Replications	Number of Correct Responses
1	18
2	17
3	13
Total	48
Mean (48/3)	16

Calculations: $\chi^2 = \frac{(18-16)^2}{16} + \frac{(17-16)^2}{16} + \frac{(13-16)^2}{16}$

$\chi^2 = 0.875$

Table value: $\chi^2 (P<0.05)$, 2 df = 3.84

Since the calculated χ^2 value (0.875) does not equal or exceed the table value (3.84), it is not statistically significant. Therefore, we accept the null hypothesis and conclude that the replications are homogeneous.

Comparison of three or more treatments

If more than two treatments are tested in the experiment, χ^2 values are calculated for each pair of treatments. For example, if there are three treatments, namely A, B and C, and four replications, a χ^2 value is calculated over the four replications for the comparisons "A vs. B", for "A vs. C", and for "B vs. C". These three χ^2 values are added. Also, a χ^2 value is determined for these pooled data. The difference between the value obtained by addition of the three χ^2 values and the χ^2 for the pooled data is designated as the χ^2 for homogeneity.

Example: Four replications were conducted with triangle tests to determine if significant differences existed among treatments A, B and C. Data are presented in Table 1.

Table 1. Number of correct responses from 18 panelists.

Comparison	Replication 1	2	3	4	Total
A vs. B	17	13	14	17	61
A vs. C	7	8	9	7	31
B vs. C	7	9	6	7	29

In a triangle test, it would be expected that one-third of the judgements would be correct for each replication for each set of triangles. With 18 panelists, the expected correct judgements would be 6.

Formula:

χ^2 for homogeneity = (χ^2 for A vs B + χ^2 for A vs C + χ^2 for B vs C) - χ^2 for pooled data.

Calculation for χ^2 A vs B:

$$\chi^2 = \frac{(17-6)^2}{6} + \frac{(13-6)^2}{6} + \frac{(14-6)^2}{6} + \frac{(17-6)^2}{6}$$

$$= 59.167$$

Calculation for χ^2 for A vs C:

$$\chi^2 = \frac{(7-6)^2}{6} + \frac{(8-6)^2}{6} + \frac{(9-6)^2}{6} + \frac{(7-6)^2}{6}$$

$$= 2.500$$

Calculation for χ^2 for B vs C:

$$\chi^2 = \frac{(7-6)^2}{6} + \frac{(9-6)^2}{6} + \frac{(6-6)^2}{6} + \frac{(7-6)^2}{6}$$

$$= 1.833$$

Calculations for χ^2 for pooled data:

$$\chi^2 = \frac{(61 - 24)^2}{24} + \frac{(31 - 24)^2}{24} + \frac{(29 - 24)^2}{24}$$

$$= 60.123$$

Note: The f_e for this χ^2 is 24. The 6 expected responses for the 4 replications also are pooled. df = rep. - 1

Table 2. χ^2 values

Source	df	χ^2
A vs B	3	59.17
A vs C	3	2.50
B vs C	3	1.83
Total	9	63.50
Pooled	3	60.12
Homogeneity	6	3.38

The calculated value for homogeneity is 3.38.

Table value: χ^2 ($P<0.05$), 6 df = 12.6.

Since the calculated χ^2 value (3.38) does not equal or exceed the table value (12.6), it is not statistically significant. Therefore, we accept the null hypothesis and conclude that the replications are homogeneous.

Paired tests.

For paired tests, the same method of calculation would be employed. However, in paired tests the expected frequency of correct responses would be one half of the total responses.

Reference: Snedecor, G.W. and Cochran, W.G., 1967, "Statistical Methods", 6th ed. Iowa State Univ. Press, Ames, Iowa. p. 228.

This test for homogeneity can also be applied to data from ranking. Examples of this application follow:

Example 1:

Rank totals for 8 judges for 4 products

Product	Replication 1	2	3	4	Total
A	8	10	8	11	37
B	20	30	22	24	96
C	22	24	22	16	84
D	30	16	28	29	103
total Rank sum	80	80	80	80	320

$$\frac{320}{4} = 80$$

$$\text{Expect rank} = \frac{80}{4 \text{ product}} = 20$$

For A: $\chi^2 = \frac{(8-20)^2}{20} + \frac{(10-20)^2}{20} + \frac{(8-20)^2}{20} + \frac{(11-20)^2}{20}$

For B: $\chi^2 = \frac{(20-20)^2}{20} + \frac{(30-20)^2}{20} + \frac{(22-20)^2}{20} + \frac{(24-20)^2}{20}$

For C: $\chi^2 = \frac{(22-20)^2}{20} + \frac{(24-20)^2}{20} + \frac{(22-20)^2}{20} + \frac{(16-20)^2}{20}$

For D: $\chi^2 = \frac{(30-20)^2}{20} + \frac{(16-20)^2}{20} + \frac{(28-20)^2}{20} + \frac{(29-20)^2}{20}$

A: $\chi^2 = 7.2 + 5.0 + 7.2 + 4.05 = 23.45$

B: $\chi^2 = 0 + 5 + 0.2 + 0.8 = 6.0$

C: $\chi^2 = 0.2 + 0.8 + 0.2 + 0.8 = 2.0$

D: $\chi^2 = 5.0 + 0.8 + 3.2 + 4.05 = 13.05$

$$\text{Pooled } \chi^2 = \frac{(37-80)^2}{80} + \frac{(96-80)^2}{80} + \frac{(84-80)^2}{80} + \frac{(103-80)^2}{80}$$

$$23.11 + 3.2 + .2 + 6.61 = 33.12$$

χ^2 Values

Source	df	χ^2
A	3	23.45
B	3	6.0
C	3	2.0
D	3	13.05
Total	12	44.50
Pooled	3	33.12
Homo.	9	11.38

Table χ^2 ($P<0.05$, 9df = 16.9)

Since calculated χ^2 = 11.38 does not exceed 16.9, data for replications can be added.

Appendix G: Scoring Test

Example 2:

Product	Replication 1	Replication 2	Total Sums
A	17	17	34
B	13	12	25
C	6	7	13
Sums	36	36	72

χ^2 for individual = $\chi^2 = \frac{(fe-fo)^2}{fe}$ when $fe = \frac{36}{3} = 12$

Rank total = 36 and each of 3 products would expect to receive 1/3, hense $\frac{36}{3} = 12$.

χ^2 for pooled = $\chi^2 = \frac{(fe-fo)^2}{fe}$ when $fe = \frac{72}{3} = 24$

Reason - same as above

For A, $\chi^2 = \frac{(17-12)^2}{12} + \frac{(17-12)^2}{12} = \frac{50}{12} = 4.17$

B, $\chi^2 = \frac{(12-12)^2}{12} + \frac{(12-13)^2}{12} = \frac{1}{12} = 0.08$

C, $\chi^2 = \frac{(7-12)^2}{12} + \frac{(6-12)^2}{12} = \frac{61}{50} = 5.08$

9.33

Pooled $\chi^2 = \frac{(34-24)^2}{24} + \frac{(25-24)^2}{24} + \frac{(13-24)^2}{24} = \frac{222}{24} = 9.25$

Source	df	χ^2
A	1	4.17
B	1	.08
C	1	5.08
Total	3	9.33
Pooled	1	9.25
Homo.	2	0.08

Table value

2 df (α=0.05) = 5.99

Since 0.08 is less than 5.99, there is no significant difference between reps.

SCORING

I. Description

Scoring consists of rating the properties of foods on a scale according to a numerically defined criterion. Scoring is one of the most frequently used sensory tests because it is simple and applicable to all types of products.

Factors which are scored should be listed on the score sheet in the sequence they will be sensed. Scores should reflect a reproducible variation in the product. The scale should be long enough to represent the variation in the product, but no longer. Five to 7 points should be adequate. Scales may be structured or unstructured. Structured scales are those which have each position on the scale identified by an appropriate descriptive term. Unstructured scales are those which are anchored by descriptive terms at the beginning and end only. Scoring can be on an absolute basis or may be applied to measure deviation from the reference, or control, sample.

Scoring is a useful tool in product development, quality control, and storage stability tests. It can be employed for evaluation of a single or several samples. Judge reproducibility can be measured by scoring deviation from the reference when a coded sample identical to the reference is scored.

II. Analysis of data

Generally, panelists' scores for each replication are averaged and these means are analyzed. A t-test is employed if there are only two treatments, or samples, and analysis of variance (ANOV) is used of there are more than two. If a significant F-value is obtained by ANOV, Duncan's New Multiple Range Test can be applied to determine which means differ significantly (see handouts).

II. References

Amerine et al. 1965. Principles of Sensory Evaluation of Food. Academic Press, New York City NY. pp. 277, 354-366, 559.

Larmond, E. 1977. Laboratory Methods for Sensory Evaluation of Food. Canada Department of Agriculture Publication 1637. pp. 41-47.

CODES: Taken from table of random numbers
ORDER OF SERVING (1,2,3) determined by a table of random permutations.

DATE 2-15-82
TEST SCORING
PRODUCT TOMATO CATSUP
PRODUCTS A, B, C

MASTER SHEET

JUDGE		SAMPLE CODES AND ORDER OF SERVING										TOTAL CORRECT
		A		B		C						
1		691 (3)		782 (1)		453 (2)						
2		279 (2)		631 (3)		482 (1)						
TOTAL CORRECT												

JUDGE _____

DATE _____

SAMPLE CODE _____

SCORING

PLEASE RINSE YOUR MOUTH BEFORE TASTING THE SAMPLE. EVALUATE THE SAMPLE FOR _____, _____, AND _____. INDICATE YOUR SCORES BY CIRCLING A DIVIDER ON THE APPROPRIATE SCALES BELOW.

Weak |———|———|———|———| Intense

Weak |———|———|———|———| Intense

Weak |———|———|———|———| Intense

COMMENTS:

Appendix H: Statistical Analyses for Scoring Test

1. Analysis of variance and Duncan's New Multiple Range Test
2. T-test for paired experiments

ANALYSIS OF VARIANCE

The analysis of variance is a procedure for partitioning the variation in a set of data according to the sources of variation. The statistic that is calculated is the F-value. This value indicates whether or not the variance due to the treatment (product) or replication is enough larger than the error variance (uncontrolled variance) to be considered statistically significant. If the F-value is not statistically significant, the null hypothesis is accepted. (Ho = There is no difference among treatments or replications.) If the F-value is statistically significant ($P<0.05$), the alternate hypothesis is accepted. (Ha = There is a difference due to treatment or replications.)

$$F = \frac{\text{Variance due to treatment or replication}}{\text{Variance due to error}}$$

Classify data as follows for each attribute scored. (Compute an ANOV for each attribute.)

Replication*	Treatment					Total
	T_1	T_2	T_3	$T...$	T_n	
Rep_1	X_1					ΣR_1
Rep_2	X_2					ΣR_2
$R...$	.					$\Sigma R...$
.	.					.
.	.					.
					X_{n-1}	
Rep_n					X_n	ΣR_n
Total	ΣT_1	ΣT_2	ΣT_3	$\Sigma T...$	ΣT_n	Grand total
Number of reps/trt.						
$\overline{X}$						

*When computing ANOV for one replication, substitute judges' scores for replication and individual judge scores for X values.

X = Panel mean for replication

$\bar{X}$ = Mean sensory score for treatment

G = Grand total

N = Total number of X values

Correction factor (CF) = $\frac{G^2}{N}$

Total sum of squares (SS) = $\Sigma(X_1^2 + X_2^2 + \ldots X_{n-1}^2 + X_n^2) - CF$

Treatment SS = $\frac{\Sigma T_1^2}{rep/trt} + \frac{\Sigma T_2^2}{rep/trt} + \ldots + \frac{\Sigma T_n^2}{rep/trt} - CF$

Replication SS = $\frac{\Sigma R_1^2}{trt/rep} + \frac{\Sigma R_2^2}{trt/rep} + \ldots + \frac{\Sigma R_n^2}{trt/rep} - CF$

Error SS = Total SS - Treatment SS - Rep SS

Summary of ANOV

Source of variation	DF	SS	Mean square MS	F
Treatment				
Replication				
Error				
Total				

Replication DF = Number of replications -1

Treatment DF = Number of treatments -1

Total DF = [(trt)(rep)] -1 OR N-1

Error DF = (trt -1)(rep -1) OR

Total df - trt df - rep df

$$MS = \frac{SS}{DF} \qquad F \text{ (for treatment)} = \frac{\text{Treatment MS}}{\text{Error MS}}$$

$$F \text{ (for replication)} = \frac{\text{Replication MS}}{\text{Error MS}}$$

Consult a Table of F-Distribution (Table F-1, Amerine et al., 1965, p. 528). To determine the critical value, select the column according to degrees of freedom for treatment (trt - 1) or replication (rep - 1) and the row according to degrees of freedom for error. If the F-value is equal to or exceeds the Table value, it is statistically significant and Ha is accepted.

REFERENCES

Amerine et al. 1965. "Principles of Sensory Evaluation of Food." Academic Press, New York. pp. 451-473.

Larmond, E. 1977. "Laboratory Methods for Sensory Analysis." Research Branch, Canada Dept. of Agr. Publication 1637. pp. 32-36.

Snedecor, G. W. and Cochran, W. G. 1973. "Statistical Methods." 6th ed. The Iowa State University Press. Ames IA.

DUNCAN'S NEW MULTIPLE RANGE TEST

If a significant F-value is obtained by ANOV, Duncan's New Multiple Range test can be computed to locate the significant differences among means for the various treatments.

Calculate the standard error of the mean ($S_{\overline{X}}$).

$$S_{\overline{X}} = \sqrt{\frac{\text{Error mean square}}{r}}$$

r = number of observations per treatment

The least significant range (LSR) for the specified number of means is calculated by multiplying the significant studentized ranges (SSR) by $S_{\overline{X}}$. Calculate an LSR for the total number of means and for all the possible comparisons that can be made. A Table of Studentized Ranges is Xerographically reproduced and attached for your convenience. Enter the table according to error df (from ANOV) on the left and according to the number of treatment means compared under p at the top of the table.

Complete the following table:

$$LSR = SSR \times S_{\overline{X}}$$

Number of treatment means	2	3	4	etc.
SSR				
LSR ($P<0.05$)				

Record the treatment means ($\overline{X}$ from ANOV) in order of magnitude from smallest to largest. See example below for treatments I, II, and III with means 2.0, 1.9, and 2.8, respectively.

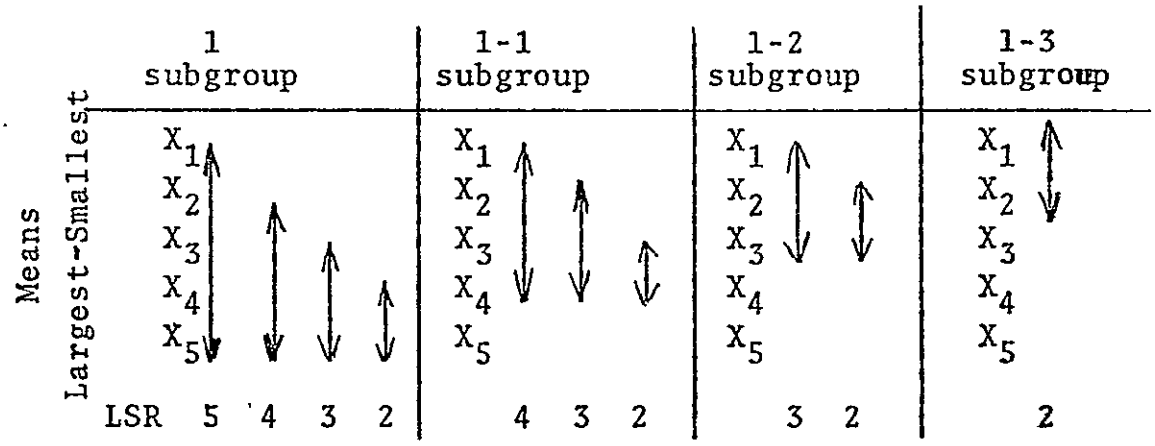
Treatment identification	Treatment II	Treatment I	Treatment III	etc.
Mean ($\overline{X}$)	1.9	2.0	2.8	

Compare means as follows:

1. Determine the difference between the largest and smallest mean. If this difference is equal to, or exceeds, the LSR for the total number of means, the difference is statistically significant ($P<0.05$).

2. Determine the difference between the largest and next smallest mean. If this difference is equal to, or exceeds, the second-smallest LSR, the difference is statistically significant ($P<0.05$).

3. Continue comparing means in this manner, until all possible comparisons have been made. A suggested pattern for making these comparisons is shown below.



Once a nonsignificant difference has been found within any subgroup, no further comparisons need be made.

4. Indicate significant differences by one of the following methods:

   a. Use letters following each mean. Means followed by the same letter do not differ significantly (see example on the following page).

   b. Draw a continuous line under means which do not differ significantly (see example on the following page).

   c. Present results graphically (see example on the following page).

REFERENCE

Amerine et al. 1965. "Principles of Sensory Evaluation of Food". Academic Press. New York, NY. pp. 454-457.

Suggested Methods of Presenting Results of Duncan's Multiple Range Test

BEST AVAILABLE

le 2—Mean panel scores[a] for sensory attributes of meat loaves ed in a loaf pan or on a rack

tments	Sensory attributes: Beef aroma	Beef flavor	Bean flavor	Juiciness
trol				
n	3.2	3.8	1.4	3.9
:k	3.2	3.6	1.1	3.6
-substituted				
n	2.4	2.5	2.9	3.3
:k	2.4	2.4	3.0	3.0
kpea-substituted				
n	1.5	1.5	4.4	2.7
:k	1.3	1.4	4.5	2.7

= 48. Range of scores: 1, low intensity; 5, high intensity. Means nin an attribute (column) differ significantly (P < 0.05) if the ters differ Enclosed by the same line do not differ significantly.

Table 3—Mean values[a] for cooking losses of meat loaves

Treatments	Cooking losses: Evaporation %	Drip %	Total %
Control			
pan	5.00c	9.07a	14.07a
rack	9.82a	7.21b	17.03b
TSP-substituted			
pan	6.12c	1.16cd	7.27d
rack	7.45b	1.88c	9.34c
Chickpea-substituted			
pan	6.04c	0.40d	6.44d
rack	5.87c	0.82d	6.69d

[a] N = 6. Means within an attribute (column) differ significantly (P < 0.05) if the letters differ.

2—Mean panel scores[a] for sensory attributes of meat loaves d in a loaf pan or on a rack

ments	Sensory attributes: Beef aroma	Beef flavor	Bean flavor	Juiciness
rol				
n	3.2a	3.8a	1.4a	3.9a
k	3.2a	3.6a	1.1a	3.6a
substituted				
n	2.4b	2.5b	2.9b	3.3b
k	2.4b	2.4b	3.0b	3.0b
kpea-substituted				
n	1.5c	1.5c	4.4c	2.7c
k	1.3c	1.4c	4.5c	2.7c

48 Range of scores: 1, low intensity; 5, high intensity. Means nin an attribute (column) differ significantly (P < 0.05) if the ers differ.

Table 3—Mean values[a] for cooking losses of meat loaves

Treatments	Cooking losses: Drip %	Total %
Control		
pan	9.07	14.07
rack	7.21	17.03
TSP-substituted		
pan	1.88	9.34
rack	1.16	7.27
Chickpea-substituted		
pan	0.40	6.44
rack	0.82	6.69

[a] N=6. Means enclosed by the same line do not differ significantly.

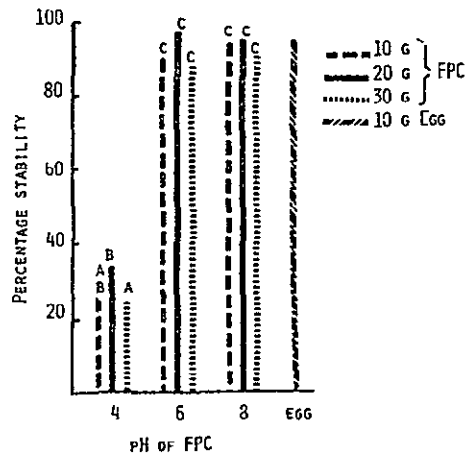


Fig. 1—Means for stability of emulsions containing different amounts of 10% solubilized FPC, adjusted to three pH levels, and mayonnaise containing liquid whole egg. Where letters above bars differ, the means differ significantly (P < 0.05) from each other (N=5).

The t-Test for Paired Experiments

Analysis for difference between paired samples. One analysis should be computed for each attribute which is scored.

Judge	Sample A	Sample B	Difference (d)

d is calculated for each pair of samples by subtracting the score for sample B from that for sample A. Some d values may be negative.

$\Sigma d =$ (Add algebraically)

N = number of differences or pairs

D.F. = number of pairs - 1

$$\bar{d} = \frac{\Sigma d}{N}$$

$$S_d = \sqrt{\frac{(d_1{}^2 + d_2{}^2 + d_3{}^2 \ . \ . \ . \ ) - \frac{(\Sigma d)^2}{N}}{N - 1}}$$

$$S_{\bar{d}} = \frac{S_d}{\sqrt{N}}$$

$$t = \frac{\bar{d}}{S_{\bar{d}}}$$ (t may be either positive or negative)

Consult a Table of significant differences for t (Steel, R. G. D. and Torrie, J. H. 1960. "Principles and Procedures of Statistics". McGraw Hill Book Co., New York, NY. p. 433. Enter the left column according to the D.F. and determine the critical value for a 1-tailed test ($P<0.05$). If the calculated value for t exceeds the critical value, there is a difference at this level of significance.