Designation: E1885 – 04 (Reapproved 2011)

Standard Test Method for Sensory Analysis—Triangle Test[1]

This standard is issued under the fixed designation E1885; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This test method covers a procedure for determining whether a perceptible sensory difference exists between samples of two products.

1.2 This test method applies whether a difference may exist in a single sensory attribute or in several.

1.3 This test method is applicable when the nature of the difference between the samples is unknown. It does not determine the size or the direction of the difference. The attribute(s) responsible for the difference are not identified.

1.4 Compared to the duo-trio test, the triangle test can achieve an equivalent level of statistical significance with fewer assessors. For details on how the triangle test compares to other three-sample tests, see Refs **(1)**, **(2)**, **(3)** and **(4)**.[2]

1.5 This test method is applicable only if the products are homogeneous. If two samples of the same product can often be distinguished, then another method, for example, descriptive analysis, may be more appropriate.

1.6 This test method is applicable only when the products do not cause excessive sensory fatigue, carryover or adaptation.

1.7 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 *ASTM Standards:*[3]

E253 Terminology Relating to Sensory Evaluation of Materials and Products

E456 Terminology Relating to Quality and Statistics

E1871 Guide for Serving Protocol for Sensory Evaluation of Foods and Beverages

2.2 *ISO Standard:*

ISO 4120 Sensory Analysis – Methodology – Triangular Test[4]

3. Terminology

3.1 *Definitions*–For definition of terms relating to sensory analysis, see Terminology E253, and for terms relating to statistics, see Terminology E456.

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 α *(alpha) risk*—probability of concluding that a perceptible difference exists when, in reality, one does not. (Also known as Type I Error or significance level.)

3.2.2 β *(beta) risk*—probability of concluding that no perceptible difference exists when, in reality, one does. (Also known as Type II Error.)

3.2.3 p_c—probability of a correct response.

3.2.4 p_d *(proportion of discriminators)* —proportion of the population represented by the assessors that can distinguish between the two products.

3.2.5 *product*—material to be evaluated.

3.2.6 *sample*—unit of product prepared, presented, and evaluated in the test.

3.2.7 *sensitivity*—general term used to summarize the performance characteristics of the test. The sensitivity of the test is rigorously defined, in statistical terms, by the values selected for α, β, and p_d.

3.3 *triad*—three uniquely coded samples given to an assessor in the triangle test; two samples are alike (that is, of one product) and one is different (that is, of the other product).

4. Summary of Test Method

4.1 Clearly define the test objective in writing.

4.2 Choose the number of assessors based on the level of sensitivity desired for the test. The sensitivity of the test is, in part, a function of two competing risks: the risk of declaring

[1] This test method is under the jurisdiction of ASTM CommitteeE18 on Sensory Evaluation and is the direct responsibility of Subcommittee E18.04 on Fundamentals of Sensory.

Current edition approved Aug. 1, 2011. Published August 2011. Originally approved in 1997. Last previous edition approved in 2004 as E1885 – 04. DOI: 10.1520/E1885-04R11.

[2] The boldface numbers given in parentheses refer to a list of references at the end of the text.

[3] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[4] Available from American National Standards Institute (ANSI), 25 W. 43rd St., 4th Floor, New York, NY 10036, http://www.ansi.org.

the samples different when they are not (that is, α-risk) and the risk of not declaring the samples different when they are (that is, β-risk). Acceptable values of α and β vary depending on the test objective and should be determined before the test (see Appendix X3).

4.3 Assessors receive a triad and are informed that two of the samples are alike and that one is different. The assessors report which they believe to be the different, or "odd," sample, even if the selection is based only on a guess.

4.4 Results are tallied and significance determined by reference to a statistical table.

5. Significance and Use

5.1 This test method is effective for the following test objectives:

5.1.1 To determine whether a perceivable difference results or a perceivable difference does not result, for example, when a change is made in ingredients, processing, packaging, handling or storage; or

5.1.2 To select, train and monitor assessors.

5.2 This test method itself does not change whether the purpose of the triangle test is to determine that two products are perceivably different versus that the products are not perceivably different. Only the selected values of p_d, α, and β change. If the objective of the test is to determine if there is a perceivable difference between two products, then the value selected for α is typically smaller than the value selected for β. If the objective is to determine if the two products are sufficiently similar to be used interchangeably, then the value selected for β is typically smaller than the value selected for α and the value of p_d is selected to define "sufficiently similar."

6. Apparatus

6.1 Carry out the test under conditions that prevent contact between assessors until the evaluations have been completed for example, booths that comply with STP 913 (**5**).

6.2 Sample preparation and serving sizes should comply with Practice E1871. See Refs (**6**) or (**7**).

7. Assessors

7.1 All assessors must be familiar with the mechanics of the triangle test (the format, the task, and the procedure of evaluation). Experience and familiarity with the product and test method may increase the sensitivity of an assessor and may therefore increase the likelihood of finding a significant difference. Monitoring the performance of assessors over time may be useful for increased sensitivity.

7.2 Choose assessors in accordance with test objectives. For example, to project results to a general consumer population, assessors with unknown sensitivity might be selected. To increase protection of product quality, assessors with demonstrated acuity should be selected.

7.3 The decision to use trained or untrained assessors should be addressed prior to testing. Training may include a preliminary presentation on the nature of the samples and the problem concerned. If the test concerns the detection of particular taint, consider the inclusion of samples during training that demonstrate its presence and absence. Such demonstration will increase the panel's acuity for the taint but may detract from other differences. See STP 758 for details (**8**). Allow adequate time between the exposure to the training samples and the actual triangle test to avoid carryover.

7.4 During the test sessions, avoid giving information about product identity, expected treatment effects, or individual performance until all testing is complete.

7.5 Pooling multiple evaluations by the same assessor is not recommended because results are less representative of the population and the risk of incorrect conclusion is greater.

8. Number of Assessors

8.1 Choose the number of assessors to yield the level of sensitivity called for by the test objectives. The sensitivity of the test is a function of three values: the α-risk, and the β-risk, and the maximum allowable proportion of distinguishers, p_d.[5]

8.2 Prior to conducting the test, select values for α, β and p_d. The following can be considered as general guidelines.

8.2.1 For α-risk: A statistically significant result at:

8.2.1.1 10 to 5 % (0.10 to 0.05) indicates "slight" evidence that a difference was apparent;

8.2.1.2 5 to 1 % (0.05 to 0.01) indicates "moderate" evidence that a difference was apparent;

8.2.1.3 1 to 0.1 % (0.01 to 0.001) indicates "strong" evidence that a difference was apparent; and

8.2.1.4 Below 0.1 % (<0.001) indicates "very strong" evidence that a difference was apparent.

8.2.2 For β–risk: The strength of the evidence that a difference was not apparent is assessed using the same criteria as above (substituting "was not apparent" for "was apparent").

8.2.3 For p_d: the maximum allowable proportion of distinguishers, p_d, falls into three ranges:

8.2.3.1 $p_d < 25$ % represent small values;

8.2.3.2 $25\ \% < p_d < 35\ \%$ represent medium sized values; and

8.2.3.3 $p_d > 35$ % represent large values.

8.3 Having defined the required level of sensitivity for the test using 8.2, use Table A1.1 to determine the number of assessors necessary. Enter Table A1.1 in the section corresponding to the selected value of p_d and the column corresponding to the selected value of β. The minimum required number of assessors is found in the row corresponding to the selected value of α. Alternatively, Table A1.1 can be used to develop a set of values for p_d, α and β that provide acceptable sensitivity while maintaining the number of assessors within practical limits. The approach is presented in detail in Ref (**9**).

8.4 Often in practice, the number of assessors is determined by material conditions (for example, duration of the

[5] In this test method, the probability of a correct response, p_d is modeled as $p_c = 1^{\alpha/\beta} p_d + (1/3)^{\alpha/\beta} (1-p_d)$, where p_d is the proportion of the entire population of assessors who can distinguish between the two products. It is a strictly statistical "guessing model" of the assessor's behavior. It is not a psychometric model of the assessor's decision process, such as the Thurstone-Ura model that could also be applied in discrimination testing.

experiment, number of available assessors, quantity of product). However, increasing the number of assessors increases the likelihood of detecting small proportions of distinguishers. Thus, one should expect to use larger numbers of assessors when trying to demonstrate that products are similar compared to when one is trying to prove they are different. Often 18 to 36 assessors are used when testing for a difference. For comparable sensitivity when testing for similarity, 42 to 78 assessors are needed.

9. Procedure

9.1 Prepare worksheets and scoresheets (see Appendix X1 – Appendix X3) in advance of the test so as to utilize an equal number of the six possible sequences of two products, A and B. Distribute these at random in groups of six among the panelists. The six sequences are:

ABB	AAB	ABA
BAA	BBA	BAB

9.2 Sometimes the final number of assessors does not end up as a multiple of six. For example, if a test was planned for 36 assessors and only 34 actually participated, there would be five complete series of the six sequences and one incomplete set of four in which two of the six triads were randomly dropped.

9.3 It is critical to the validity of the test that assessors cannot identify the samples from the way in which they are presented. For example, in a test evaluating flavor differences, one should avoid any subtle differences in temperature or appearance caused by factors such as the time sequence of preparation. It may be possible to mask color differences using light filters, subdued illumination, or colored serving containers. Code the serving containers containing the samples in a uniform manner, preferably using three-digit numbers, chosen at random for each test. Prepare samples out of sight and in an identical manner: same apparatus, same serving containers, and same quantities of products (see ASTM Serving Protocols).

9.4 Present each triad simultaneously if possible, following the same spatial arrangement for each assessor (on a line to be sampled always from left to right, in a triangular array, etc.) Within the triad, assessors are typically allowed to make repeated evaluations of each sample as desired. If the conditions of the test require the prevention of repeat evaluations for example, if samples are bulky, leave an aftertaste, or show slight differences in appearance that cannot be masked, present the samples sequentially and do not allow repeated evaluations.

9.5 Each scoresheet should provide for a single triad of samples. If a different set of products is to be evaluated by an assessor in a single session, the completed scoresheet and any remaining product should be returned to the test administrator prior to receiving the subsequent triad. The assessor cannot go back to any of the previous samples or change the verdict on any previous test.

9.6 Do not ask questions about preference, acceptance, or degree of difference after the initial selection of the odd sample. The selection the assessor has just made may bias the reply to any additional questions. Responses to such questions may be obtained through separate tests for preference, acceptance, degree of difference, etc. (see Manual 26) (**10**). A comment section asking why the choice was made may be included for the assessor's remarks.

9.7 The triangle test is a forced-choice procedure; assessors are not allowed the option of reporting "no difference." An assessor who detects no difference between the samples should be instructed to randomly select one of the samples as being the odd one and can indicate that the selection was only a guess in the comments section of the scoresheet.

10. Analysis and Interpretation of Results

10.1 Use Table A1.2 to analyze the data obtained from a triangle test. The actual number of assessors can be greater than the minimum value given in Table A1.1. If the number of correct responses is greater than or equal to the number given in Table A1.2, conclude that a perceptible difference exists between the samples. If the number of correct responses is less than the number given in Table A1.2, conclude that the samples are sufficiently similar. Again, the conclusions are based on the risks accepted when the level of sensitivity (that is, p_d, α, and β) was selected in determining the number of assessors.

10.2 If desired, calculate a confidence interval on the proportion of the population that can distinguish the samples. This method is described in Appendix X4.

11. Report

11.1 Report the test objective, the results, and the conclusions. The following additional information is recommended:

11.1.1 The purpose of the test and the nature of the treatment studied;

11.1.2 *Full Identification of the Samples*—Origin, method of preparation, quantity, shape, storage prior to testing, serving size, temperature. (Sample information should communicate that all storage handling, and preparation was done in such a way as to yield samples that differ only due to the variable of interest, if at all);

11.1.3 The number of assessors, the number of correct selections, and the result of the statistical evaluation;

11.1.4 *Assessors*—Age, gender, experience in sensory testing, with the product, with the samples in the test;

11.1.5 Any information and any specific instructions given the assessor in connection with the test;

11.1.6 The test environment: use of booths, simultaneous or sequential presentation, light conditions, whether the identity of the samples was disclosed after the test and the manner in which it was done; and

11.1.7 The location and date of the test and the name of the panel leader.

12. Precision and Bias

12.1 Because results of sensory difference tests are functions of individual sensitivities, a general statement regarding the precision of results that is applicable to all populations of assessors cannot be made. However, adherence to the recommendations stated in this standard should increase the reproducibility of results and minimize bias.

13. Keywords

13.1 difference testing; discrimination test; sensory analysis; similarity testing; triangle test

ANNEX

(Mandatory Information)

A1. NUMBER OF ASSESSORS AND CORRECT RESPONSES NEEDED FOR A TRIANGLE TEST

TABLE A1.1 Number of Assessors Needed for a Triangle Test (9)

NOTE 1—Entries are the minimum number of assessors required to execute a triangle test with a prespecified level of sensitivity determined by the values of p_d, α, and β. Enter the table in the section corresponding to the chosen value of p_d and the column corresponding to the chosen value of β. Read the minimum number of assessors from the row corresponding to the chosen value of α.

		β				
α		0.20	0.10	0.05	0.01	0.001
0.20	p_d = 50 %	7	12	16	25	36
0.10		12	15	20	30	43
0.05		16	20	23	35	48
0.01		25	30	35	47	62
0.001		36	43	48	62	81
	p_d = 40 %					
0.20		12	17	25	36	55
0.10		17	25	30	46	67
0.05		23	30	40	57	79
0.01		35	47	56	76	102
0.001		55	68	76	102	130
	p_d = 30 %					
0.20		20	28	39	64	97
0.10		30	43	54	81	119
0.05		40	53	66	98	136
0.01		62	82	97	131	181
0.001		93	120	138	181	233
	p_d = 20 %					
0.20		39	64	86	140	212
0.10		62	89	119	178	260
0.05		87	117	147	213	305
0.01		136	176	211	292	397
0.001		207	257	302	396	513
	p_d = 10 %					
0.20		149	238	325	529	819
0.10		240	348	457	683	1011
0.05		325	447	572	828	1181
0.01		525	680	824	1132	1539
0.001		803	996	1165	1530	1992

TABLE A1.2 Number of Correct Responses Needed for Significance in a Triangle Test (10)

NOTE 1—Entries are the minimum number of correct responses required for significance at the stated α level (that is, column) for the corresponding number of assessors, *n* (that is, row). Reject the assumption of "no difference" if the number of correct responses is greater than or equal to the tabled value.

NOTE 2—For values of *n* not in the table, compute the missing entry as follows: Minimum number of responses (*x*) = nearest whole number greater than $x = (n/3) + z\sqrt{2n/9}$, where z varies with the significance level as follows: 0.84 for α=0.20; 1.28 for α=0.10; 1.64 for α=0.05; 2.33 for α=0.01; 3.10 for α=0.001.

	α						α				
n	0.20	0.10	0.05	0.01	0.001	*n*	0.20	0.10	0.05	0.01	0.001
6	4	5	5	6	. . .	32	14	15	16	18	20
7	4	5	5	6	7	33	14	15	17	18	21
8	5	5	6	7	8	34	15	16	17	19	21
9	5	6	6	7	8	35	15	16	17	19	22
10	6	6	7	8	9	36	15	17	18	20	22
11	6	7	7	8	10	37	16	17	18	20	22
12	6	7	8	9	10	38	16	17	19	21	23
13	7	8	8	9	11	39	16	18	19	21	23
14	7	8	9	10	11	40	17	18	19	21	24
15	8	8	9	10	12	41	17	19	20	22	24
16	8	9	9	11	12	42	18	19	20	22	25
17	8	9	10	11	13	43	18	19	20	23	25
18	9	10	10	12	13	44	18	20	21	23	26
19	9	10	11	12	14	45	19	20	21	24	26
20	9	10	11	13	14	46	19	20	22	24	27
21	10	11	12	13	15	47	19	21	22	24	27
22	10	11	12	14	15	48	20	21	22	25	27
23	11	12	12	14	16	54	22	23	25	27	30
24	11	12	13	15	16	60	24	26	27	30	33
25	11	12	13	15	17	66	26	28	29	32	35
26	12	13	14	15	17	72	28	30	32	34	38
27	12	13	14	16	18	78	30	32	34	37	40
28	12	14	15	16	18	84	33	35	36	39	43
29	13	14	15	17	19	90	35	37	38	42	45
30	13	14	15	17	19	96	37	39	41	44	48
31	14	15	16	18	20	102	39	41	43	46	50

APPENDIXES

(Nonmandatory Information)

X1. TRIANGLE TEST TO CONFIRM THAT A DIFFERENCE EXISTS: IMPROVED NON-ALCOHOLIC BEER

X1.1 *Background* —A brewery has developed a process to reduce an unwanted grainy flavor characteristic in it's nonalcoholic beer. The process requires an investment in new equipment. Before proceeding to a preference test involving consumers, the head brewer wants to confirm that the experimental non-alcoholic beer is different from the company's current non-alcoholic beer. The head brewer is willing to take only a small chance of concluding that a difference exists when in reality one does not. However, because he has alternate ways of operating the new process, he is willing to accept a greater risk of missing a difference that does exist.

X1.2 *Test Objective*—To confirm that the experimental beer, "A," can be distinguished from the current nonalcoholic product, "B," in order to justify testing with consumers.

X1.3 *Number of Assessors*—To protect the head brewer from falsely concluding that a difference exists, the sensory analyst proposes α=0.05. Also, in order to keep the amount of tasting within reasonable limits, she suggests setting p_d at 50 % with β=0.20. These values are agreed to by all parties concerned with the test. The analyst consults Table A1.1 in the sections corresponding to p_d=50 % and the column corresponding to β=0.20. Then reading from the row corresponding to α=0.05, she finds that a minimum of 16 assessors are needed for the test. In order to balance the order of presentation of the samples, the analyst decides to use 18 assessors.

X1.4 *Conducting the Test*—27 glasses of "A" and 27 glasses of "B" are coded with unique random numbers. Each of the triads ABB, BAA, AAB, BBA, ABA, and BAB is presented three times so as to cover the 18 assessors in a balanced random order. An example of the scoresheet used is shown in Fig. X1.1.

X1.5 *Analysis and Interpretation of Results*—Ten panelists correctly identify the odd sample. In Table A1.2, in the row corresponding to 18 assessors and the column corresponding to

TRIANGLE TEST

Assessor No. _____ Name _____ Date _____

Instructions

Evaluate samples from left to right. Two are alike. Mark an "X" in the box for the sample which differs from the others. If no difference is apparent, you must guess.

263	941	792
☐	☐	☐

Remarks: _____

FIG. X1.1 Scoresheet for Triangle Difference Test in Appendix X1

α=0.05, the sensory analyst finds that ten correct responses is sufficient to conclude that the two beers are perceptibly different.

X1.6 *Report and Conclusions*—The sensory analyst reports that the experimental beer could, in fact, be distinguished from the current product given the sensitivity levels chosen for the test (p_d=50 %, α=0.05, β=0.20). Evaluation of the experimental beer can proceed to testing with consumers.

X2. TRIANGLE TEST TO CONFIRM THAT TWO SAMPLES ARE SIMILAR CHOICE OF FOIL AND PAPER WRAPPING FOR CANDY BAR

X2.1 *Background* —A confections manufacturer wants to confirm that the substitution of new packaging material for the customary material does not affect the sensory character of its candy bar after three months of shelf storage. The manufacturer realizes that it is impossible to prove that two products are identical, but he wants to be very sure that only a reasonably small proportion of the population will be able to detect a difference if one exists. On the other hand the manufacturer is not greatly concerned if he incorrectly concludes that the products are different when they are not, because the current package is acceptable and the new one is being considered only because it offers greater flexibility for label graphics.

X2.2 *Test Objective*—Determine if product stored for three months in the new packaging material is the same as product stored for three months in the current packaging material.

X2.3 *Number of Assessors*—The sensory analyst works with the manufacturer to decide on the levels of risk that are appropriate for the test. It is decided that the maximum allowable proportion of discriminators should be p_d=30 %. The manufacturer is only willing to take a β=0.05 chance of failing to detect that level of discriminators. On the other hand, because of the acceptability of the current package, the manufacturer and the sensory analyst agree that it is reasonable to accept an α=0.20 risk of concluding that there is a difference when there is not. The analyst enters Table A1.1 in the section corresponding to p_d=30 % and the column corresponding to β=0.05. Then, in the row corresponding to α=0.20 she finds that 39 assessors are needed for the test.

X2.4 *Conducting the Test*—The sensory analyst uses the worksheet shown in Fig. X2.1 and the scoresheet shown in Fig. X2.2 to run the test. She cycles through the six possible triads: AAB, ABA, BAA, BBA, BAB, and ABB six times with the first 36 assessors. She then randomly selects three triads to serve assessors 37, 38, and 39.

X2.5 *Analysis and Interpretation of Results*—14 of the 39 assessors correctly identify the odd sample in the test. Referring to Table A1.2, the analyst finds that 16 correct responses from the 39 respondents are required for significance at the α=0.20 level. The analyst concludes that the new packaging meets the manufacturer's criterion of 95 % certainty (that is, β=0.05) that no more than p_d=30 % of the population are able

Date: *Oct. 4, 1993* Test code: *587-FF03*

Triangle Test Sample Order and Serving Protocol

Post this sheet in the area where trays are prepared. Code scoresheets and serving containers ahead of time.

*Product Type: Candy Bars

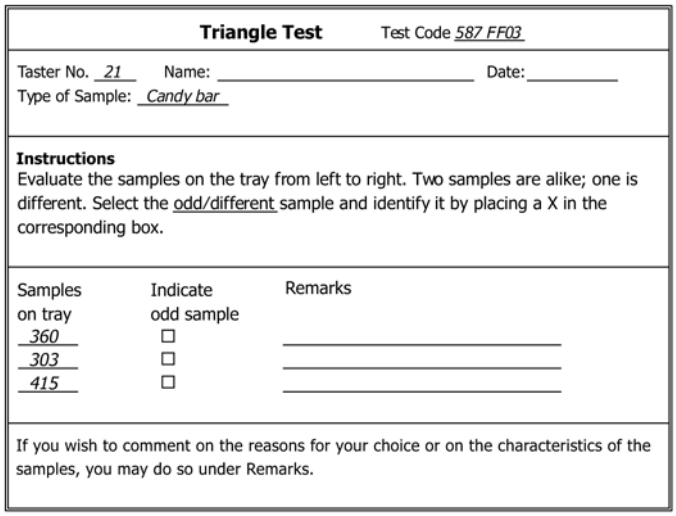
Sample identification:
Sample 1 = Package 4736 (current) Sample 2 = Package 3987 (new)

Code serving containers as follows:

PANELIST	SAMPLE-CODE	PANELIST	SAMPLE-CODE
1	1-100 1-795 2-140	19	1-745 1-247 2-724
2	1-189 2-168 1-733	20	1-344 2-370 1-355
3	2-718 1-437 1-488	21	2-360 1-303 1-415
4	2-535 2-231 1-243	22	2-134 2-401 1-305
5	2-839 1-402 2-619	23	2-185 1-651 2-307
6	1-145 2-296 2-992	24	1-508 2-271 2-465
7	1-792 1-280 2-319	25	1-216 1-941 2-321
8	1-167 2-936 1-180	26	1-494 2-783 1-414
9	2-689 1-743 1-956	27	2-151 1-786 1-943
10	2-442 2-720 1-213	28	2-423 2-477 1-164
11	2-253 1-444 2-505	29	2-570 1-772 2-887
12	1-204 2-159 2-556	30	1-398 2-946 2-764
13	1-142 1-325 2-632	31	1-747 1-286 2-913
14	1-472 2-762 1-330	32	1-580 2-558 1-114
15	2-965 1-641 1-300	33	2-345 1-562 1-955
16	2-582 2-659 1-486	34	2-385 2-660 1-856
17	2-429 1-884 2-499	35	2-754 1-210 2-864
18	1-879 2-891 2-404	36	1-574 2-393 2-753
		37	1-793 1-308 2-742
		38	1-147 2-395 1-434
		39	2-396 2-629 1-957

1. Label plates with the unique 3 digit random numbers and arrange in serving order by panelist.
2. To serve, place samples and a coded scoresheet on a serving tray.
3. Decode whether reply was correct or incorrect by referring back to the worksheet.

FIG. X2.1 Worksheet for Appendix X2



Triangle Test Test Code 587 FF03

Taster No. 21 Name: _____ Date: _____
Type of Sample: Candy bar

Instructions
Evaluate the samples on the tray from left to right. Two samples are alike; one is different. Select the odd/different sample and identify it by placing a X in the corresponding box.

Samples on tray	Indicate odd sample	Remarks
360	☐	_____
303	☐	_____
415	☐	_____

If you wish to comment on the reasons for your choice or on the characteristics of the samples, you may do so under Remarks.

FIG. X2.2 Scoresheet for Appendix X2

to detect a difference. The new packaging can be substituted for the current.

X3. TRIANGLE TEST THAT BALANCES THE RISKS OF DIFFERENCE AND SIMILARITY PROCESS COST REDUCTION VERSUS QUALITY CONTROL

X3.1 *Background* —A sensory analyst is planning a study to help determine if a new cost-saving process can be implemented without perceptibly degrading the quality of the finished product. The manager of quality assurance wants to be at least 95 % certain (that is, β=0.05) that no more than 20 % of consumers would be able to detect a difference if the change is made. At the same time, the director of operations is anxious to reap the benefits of the lower cost process, so he does not want to take more than a 5 % chance (that is, α=0.05) of concluding that there is a difference between the samples when there is not.

X3.2 *Test Objective*—To determine if a perceptible difference exists between the current product and product produced using the new, cost-saving process.

X3.3 *Number of Assessors*—In order to determine the minimum number of assessors needed for the test, the sensory analyst enters Table A1.1 in the section corresponding to p_d=20 % and the column corresponding to β=0.05. In the row corresponding to α=0.05, the sensory analyst finds that a minimum of n=147 assessors are necessary. The sensory analyst knows, however, that only 50 assessors are readily available to participate in the test. The sensory analyst informs the project team that obtaining the required number of assessors will cost additional time and money, neither of which is available. Therefore, the sensory analyst, the manager of quality assurance, and the director of operations renegotiate the test sensitivity parameters to provide the maximum possible risk protection with the number of available assessors. Consulting Table A1.1 again, it is determined that p_d=30 %, α=0.10, β=0.10, and n≥43 provides acceptable sensitivity given the number of assessors available for the test. The analyst prepares the test for 48 assessors in order to balance the orders of presentation of the triads.

X3.4 *Analysis and Interpretation of Results*—Although 48 triads were prepared, only 45 assessors actually participated in the test (that still meets the minimum 43 required to achieve the level of sensitivity chosen for the test). A total of 18 of the 45 assessors correctly selected the odd sample in the test. Consulting Table A1.2, the analyst finds in the row corresponding to n=45 assessors and in the column corresponding to α=0.10 that 20 correct responses are required to reject the null hypothesis assumption of "no difference." Since the observed number of correct responses (18) is less than the number required to determine that a difference exists (20), it is decided that the products are sufficiently similar.

X3.5 *Conclusion* —The sensory analyst concludes with at least 90 % confidence (β=0.10) that no more than 30 % of the population of assessors (p_d=30 %) can differentiate between products produced using the current and the proposed processes. Product made using the new, cost-saving process is deemed sufficiently similar to that of the current process, so the recommendation is made to implement the new process.

X4. CONFIDENCE INTERVALS FOR TRIANGLE TESTS

X4.1 *Background* —If desired, analysts can calculate a confidence interval on the proportion of the population that can distinguish the samples. The calculations are as follows, where c = the number of correct responses and n = the total number of assessors:

$$p_c \text{ (proportion correct)} = c/n$$
$$p_d \text{ (proportion distinguishers)} = 1.5\ p_c - 0.5$$
$$s_d \text{ (standard deviation of } p_d) = 1.5\sqrt{p_c(1-p_c)/n}$$
$$\text{upper confidence limit} = p_d + z_\beta s_d$$
$$\text{lower confidence limit} = p_d - z_\alpha s_d$$

z_α and z_β are critical values of the standard normal distribution. For a 90 % confidence interval, z=1.28; for a 95 % confidence interval, z=1.64, and for a 99 % confidence interval, z_a=2.33.

X4.2 *Analysis and Interpretation of Results*—Consider the data from Appendix X3, above, where c=18 and n=45. It follows that:

$$p_c \text{ (proportion correct)} = 18/45 = 0.40$$
$$p_d \text{ (proportion distinguishers)} = 1.5(0.40) - 0.5 = 0.10$$
$$s_d \text{ (standard deviation of } p_d) = 1.5\sqrt{(0.40)(1-0.40)/45} = 0.11$$
$$\text{upper confidence limit} = 0.10 + 1.28(0.11) = 0.24$$
$$\text{lower confidence limit} = 0.10 - 1.28(0.11) = -0.04$$

The analyst can be 90 % certain that the actual proportion of the population that can distinguish the samples lies somewhere between -4 and 24 %. Since the lower end-point of the interval is negative, p_d=0 % is in the interval and is therefore a possible value, thus supporting the conclusion of no significant difference that was drawn in Appendix X3. The upper limit of the confidence interval (that is, $p_d \leq 24$ %) also supports the conclusion reached in Appendix X3 because it is less than the 30 % cut-off used in the example.

REFERENCES

(1) Ennis, D.M., "The Power of Sensory Discrimination Methods", *Journal of Sensory Stud.*, Vol 8, 1993, pp. 353-370.

(2) Frijters, J.E.R., "Three-Stimulus Procedure in Olfactory Psychophysics: An Experimental Comparison of Thurstone-Ura and Three-Alternative Forced-Choice Models of Signal Detection Theory," *Perception and Psychophysics* , Vol 28, No 5, 1980, pp. 390-397.

(3) MacRae, S. "The Interplay of Theory and Practice Sensory Testing," *Chem and Ind*., Jan. 5, 1987, pp. 7-12.

(4) O'Mahony, M.A.P.D.E., and Odbert, N., "A Comparison of Sensory Difference Testing Procedures: Sequential Sensitivity Analysis and Aspects of Taste Adaptation," *Journal of Food Science*, Vol 50, 1985, pp. 1055-1058.

(5) *Guidelines for Physical Requirements for Sensory Evaluation Laboratories*, *ASTM STP 913*, ASTM.

(6) Herz, R.S. and Cupchik, G.C., "An Experimental Characterization of Odor-evoked Memories in Humans", *Chemical Senses*, Vol 17, No 5, 1992, pp. 519-528.

(7) Todrank, J., Wysocki, C.J. and Beauchamp, G.K., "The Effects of Adaptation on the Perception of Similar and Dissimilar Odors," *Chemical Senses*, Vol 16, No 5, 1991, pp. 476-482.

(8) *Guidelines for the Selection and Training of Sensory Panel Members*, *ASTM STP 758*, ASTM.

(9) Schlich, P., "Risk Tables for Discrimination Tests," *Food Quality and Preference*, Vol 4 1993, pp. 141-151.

(10) Meilgaard, M., Civille, G.V., Carr, B.T., *Sensory Evaluation Techniques, 2nd Edition,* CRC Press, Inc., Boca Raton, FL, 1991.