Designation: E1885 – 18

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, health, and environmental practices and determine the applicability of regulatory limitations prior to use.*

1.8 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

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

E2262 Practice for Estimating Thurstonian Discriminal Distances

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.

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

Current edition approved Aug. 15, 2018. Published August 2018. Originally approved in 1997. Last previous edition approved in 2011 as E1885 – 04 (2011). DOI: 10.1520/E1885-18.

[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.

3.2.8 δ—Thurstonian measure of sensory difference (effect size) relative to perceptual noise (standard deviation) (see Practice E2262).

3.2.9 *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 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.

[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.

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 If one wishes to use Thurstonian δ as a measure of sensory effect size, use Tables A2.1 and A2.2 to convert between p_d and δ. See Ref (**10**) for further discussion on the relationship between Thurstonian δ and p_d.

8.5 Often in practice, the number of assessors is determined by material conditions (for example, duration of the 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) (**11**). 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;