Designation: E2262 – 03 (Reapproved 2020)

Standard Practice for Estimating Thurstonian Discriminal Distances[1]

This standard is issued under the fixed designation E2262; 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 practice describes procedures to estimate Thurstonian discriminal distances (that is, *d*' values) from data obtained on two samples. Procedures are presented for four forced-choice methods (that is, the triangle, the Duo-Trio, the 3-alternative-forced-choice (or 3-AFC) and the 2-AFC (also called the directional difference test)), the A/Not-A method, the Same-Different method, and for data obtained from ordered category scales. Procedures for estimating the variance of *d*' are also presented. Thus, confidence intervals and statistical tests can be calculated for *d*'.

1.2 The procedures in this practice pertain only to the unidimensional, equal-variance model. Other, more complicated Thurstonian models, involving multiple dimensions and unequal variances exist but are not addressed in this practice. The procedure for forced-choice methods is limited to dichotomous responses. The procedure for the A/Not-A method assumes equal sample sizes for the two samples. The procedure for the Same-Different method assumes equal sample sizes for the matched and unmatched pairs of samples. For all methods, only unreplicated tests are considered. (Tests in which each assessor performs multiple (that is, replicated) evaluations require different analyses.)

1.3 Thurstonian scaling is a method for measuring the perceptual difference between two samples based on a probabilistic model for categorical choice decision making. The magnitude of the perceived difference, δ, can be estimated from the assessors' categorical choices using the methods described in this practice. (See Appendix X3 for a more detailed description of Thurstonian scaling.)

1.4 In theory, the Thurstonian δ does not depend on the method used to measure the difference between two samples. As such, δ provides a common scale of measure for comparing samples measured under a variety of test conditions. For example, Thurstonian scaling can be used to compare products measured under different test conditions, to compare panels (trained, consumer or both) that have evaluated the same samples (using the same or different test methods) and to compare test methods on their ability to discriminate samples that exhibit a fixed sensory difference.

1.5 *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.*

1.6 *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:*[2]

E253 Terminology Relating to Sensory Evaluation of Materials and Products

E456 Terminology Relating to Quality and Statistics

2.2 *ASTM Manual:*[2]

Manual 26 Sensory Testing Methods, 2nd Edition

2.3 *ISO Standard:*[3]

ISO 5495 Sensory Analysis—Methodology—Paired Comparison

3. Terminology

3.1 *Definitions:*

3.1.1 For definitions of terms relating to sensory analysis, see Terminology E253. For terms relating to statistics, see Terminology E456.

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *δ*—the Thurstonian discriminal distance is the distance between the means of the distributions of sensory magnitudes of the two samples in the test (see Appendix X3).

[1] This practice is under the jurisdiction of ASTM Committee E18 on Sensory Evaluation and is the direct responsibility of Subcommittee E18.03 on Sensory Theory and Statistics.

Current edition approved Aug. 1, 2020. Published September 2020. Originally approved in 2003. Last previous edition approved in 2014 as E2262 – 03 (2014). DOI: 10.1520/E2262-03R20.

[2] 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.

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

3.2.2 *d'*—the statistic used to estimate δ based on the data obtained from the test.

3.2.3 *choice proportion* (P_c)—the expected proportion of responses from a forced-choice method. (For example, if there is no perceptible difference between the samples in a triangle test, P_c = 1/3. If there is a perceptible difference, P_c > 1/3.)

3.2.4 *observed choice proportion* (p_c)—the statistic used to estimate choice proportion, P_c, where $p_c = x/n$, where *x* is the observed number of correct responses and *n* is the sample size.

4. Summary of Practice

4.1 Determine the type of data collected on the two samples: data from a forced-choice test, an A/Not-A test, a Same-Different test or an ordered category scale.

4.2 For forced-choice tests, reference the table that corresponds to the test method (that is, triangle test (Tables X1.1 and X1.2), Duo-Trio test (Tables X1.3 and X1.4), 3-AFC test (Tables X1.5 and X1.6), or 2-AFC test (Tables X1.7 and X1.8)). Identify the entry in the table closest to the observed choice proportion (p_c) from the test. Read the estimated value of δ (that is, *d'*) from the corresponding row and column headings of the table. Estimate the variance of *d'* by referencing the appropriate table for the test method. Find the value of *B* that corresponds to the value of *d'* obtained in the first step (see Note 1). The estimated variance of *d'* is $S^2(d') = B/n$, where *n* is the sample size. Use the estimates *d'* and $S^2(d')$ to construct confidence intervals and tests of hypotheses related to the objectives of the research.

NOTE 1—The variance of *d'* is a complicated function of the true value of δ and the decision rule when associated with the test method being used (see Appendix X3). However, regardless of the test method, the variance of *d'* can always be expressed as $S^2(d') = B/n$, where the parameter *B* captures all of the information concerning the test method, and *n* is the sample size. The values of *B* have been tabulated to make the calculation of the variance of *d'* a simple task.

4.3 For the A/Not-A method, tally the observed choice proportions of "A" responses for the A sample and the "A" responses for the Not-A sample. Read the value of *d'* from Table X1.9 in the column that corresponds to the observed choice proportion of the "A" responses for the Not-A sample (p_{na}) and the row that corresponds to the observed choice proportion of the "A" responses for the A sample (p_a). The same method is used to estimate the variance of *d'*, $S^2(d')$, using Table X1.10.

4.4 For the Same-Different method, tally the proportion of "same" responses for the matched pairs of samples (that is, A/A or B/B) and the proportion of "same" responses for the unmatched pairs of samples (that is, A/B or B/A). Read the value of *d'* from Table X1.11 in the column that corresponds to the observed proportion of "same" responses for the unmatched pairs ($p_s/_u$) and the row that corresponds to the observed proportion of the "same" responses for the matched pairs ($p_s/_m$). The same method is used to estimate the variance of *d'*, $S^2(d')$, using Table X1.12.

4.5 For ordered category scales, a rapid, table-look-up approach is used. For each sample, the category scale data are collapsed into two categories. One sample is selected to be the "A" sample and the other sample is selected to be the "Not-A" sample. Choice proportions are tallied for each sample and the values of *d'* and its variance, $S^2(d')$, are obtained from Tables X1.9 and X1.10, respectively, by the same techniques used in the A/Not-A method.

5. Significance and Use

5.1 Under the assumptions of the model, the Thurstonian model approach to measuring the perceived difference between two samples (whether overall or for a specific attribute) is independent of the sensory method used to collect the data. Converting results obtained from different test methods to *d'* values permits the assessment of relative differences among samples without requiring that the samples be compared to each other directly or that the same test methods be used for all pairs of samples.

5.2 Thurstonian scaling has been applied to:

5.2.1 Creating a historical database to track differences between production and reference samples over periods in which different test methods were used to measure the difference,

5.2.2 Comparing the relative sensitivities of different user groups and consumer segments,

5.2.3 Comparing trained panels that use different measuring techniques,

5.2.4 Comparing the relative sensitivities of consumers versus trained panels,

5.2.5 Comparing different methods of consumer testing (for example, CLT versus HUT, preference versus hedonic scales, etc.), and

5.2.6 Comparing different discrimination test methods.

6. Procedure

6.1 *Forced-choice Methods*—The relationship between δ and the expected choice proportion, P_c, is different for different forced-choice methods because the decision rule used by the assessors varies from one method to another (see Appendix X3). As a result, different tables are required to estimate δ depending on the method used. Tables for the four most commonly used methods are presented. The estimated value of δ (that is, *d'*) is obtained as follows:

6.1.1 Compute the observed choice proportion as $p_c = x/n$, where *x* is the observed number of correct responses and *n* is the sample size.

6.1.2 Obtain *d'* by entering the table in Appendix X1 that corresponds to the test method used: triangle test (Table X1.1), Duo-Trio (Table X1.3), 3-AFC (Table X1.5), or 2-AFC (Table X1.7). Find the entry in the table that is closest to the observed value of p_c. The value of *d'*, accurate to one decimal place, is the row-label of the table corresponding to the selected entry. The second decimal place of *d'* is the column-label of the table corresponding to the selected entry.

6.1.3 Obtain the estimated variance of *d'* as follows. Enter the appropriate table in Appendix X1: triangle test (Table X1.2), Duo-Trio (Table X1.4), 3-AFC (Table X1.6), or 2-AFC (Table X1.8). Find the value of *B* in the row and column that correspond to the value of *d'* obtained in 6.1.2. Compute the estimated variance of *d'* as $S^2(d') = B/n$, where *n* is the sample size. Use the estimates d' and $S^2(d')$ to construct confidence intervals and tests of hypotheses related to the objectives of the research.

6.2 *A/Not-A Method*—Compute the choice proportions of the two samples, $p_a = x_a/n$ and $p_{na} = x_{na}/n$, where x_a is the number of times the "A" sample is chosen as being "A," x_{na} is the number of times the "Not-A" sample is chosen as being "A" and n is the sample size.

NOTE 2—This practice only considers the case where the number of "A" samples equals the number of "Not-A" samples, $n = n_a = n_{na}$.

6.2.1 Read the value of d' from Table X1.9 in Appendix X1 in the column that corresponds to the observed choice proportion of the "Not-A" sample (p_{na}) and the row that corresponds to the observed choice proportion of the "A" sample (p_a).

6.2.2 To obtain an estimate of the variance of d', read the value of B from Table X1.10 in Appendix X1 using the same technique as in 6.2.1. The variance estimate is $S^2(d') = B/n$, where n is the sample size.

6.3 *Same-Different Method*—Compute the choice proportions for the matched (m) and unmatched (u) pairs of samples, $p_{s/m} = x_{s/m}/n$ and $p_{s/u} = x_{s/u}/n$, where $x_{s/m}$ is the number of "same" responses for the matched pairs (A/A or B/B) evaluated, $x_{s/u}$ is the number of "same" responses for the unmatched pair and n is the number of matched or unmatched pairs evaluated.

NOTE 3—This practice only considers the case where the number of matched pairs equals the number of unmatched pairs, $n = n_m = n_u$.

6.3.1 Read the value of d' from Table X1.11 in Appendix X1 in the column that corresponds to the observed proportion of "same" responses for unmatched pair ($p_{s/u}$) and the row that corresponds to the observed proportion of "same" responses for the matched pair ($p_{s/m}$).

6.3.2 To obtain an estimate of the variance of d', read the value of B from Table X1.12 in Appendix X1 using the same technique as in 6.3.1. The variance estimate is $S^2(d') = B/n$, where n is the sample size.

6.4 *Ordered Category Scales*—A rapid, table-look-up method is described. The method collapses the category-scale data into two categories, regardless of the number of categories on the physical scale used to collect the data. It is recognized that information detail is lost by collapsing the data into two categories. However, the estimates of d' and its variance, $S^2(d')$, obtained from the technique are accurate. The computational ease offsets the small loss of accuracy incurred.

6.4.1 Tally the frequency distributions of category scale ratings for the two samples. Select the sample with the lower median rating to be the Not-A sample. Select the sample with the higher median rating to be the A sample.

6.4.2 Collapse the frequency data for each sample into two categories as follows. Identify the category in which the median of the Not-A sample occurs. Pool the number of responses from that category and all lower categories for each sample separately and record the totals in the two-by-two table under "Low" (that is, the y_{na} and y_a tallies, below). Pool the number of responses for the remaining, higher categories for each sample separately and record the totals in the two-by-two table under "High" (that is, the x_{na} and x_a tallies, below).

Sample	Low	High
Not-A	y_{na}	x_{na}
A	y_a	x_a

6.4.3 Compute the choice proportions of the two samples, $p_a = x_a/n$ and $p_{na} = x_{na}/n$, where x_a and x_{na} are obtained from the table above and n is the sample size, common to both samples.

6.4.4 Apply the same technique used in the A/Not-A method (see 6.2). Read the value of d' from Table X1.9 in Appendix X1 in the column that corresponds to the observed choice proportion of the Not-A sample (p_{na}) and the row that corresponds to the observed choice proportion of the A sample (p_a).

6.4.5 To obtain an estimate of the variance of d', read the value of B from Table X1.10 in Appendix X1 using the same technique as in 6.4.4. The variance estimate is $S^2(d') = B/n$, where n is the sample size.

6.5 *Statistical Tests and Confidence Intervals*—Often the objective of a sensory discrimination test is to determine if the samples in the test are perceptibly different. In other instances it is of interest to obtain an estimate of the size of the perceptible difference (and to measure the precision of the estimated difference). Because testing for a difference and estimating the size of a difference address different goals, it is not surprising that different statistical methods apply to each. For the purpose of testing if a perceptible difference exists, the binomial and chi-square tests traditionally associated with the test methods discussed in this practice are appropriate. For the purposes of estimating the size of the difference and assessing the precision of that estimate, confidence intervals are appropriate. Because δ is the difference between the means of two normal distributions and d' is an estimate of δ, it can be assumed that d' is approximately normally distributed. Based on this assumption, statistical confidence intervals concerning δ can be constructed using traditional techniques.

6.5.1 A 100(1 - α) % two-sided confidence interval on δ is calculated as: $d' \pm Z_{\alpha/2}S(d')$, where d' is the estimated value of δ, $Z_{\alpha/2}$ is the upper-α/2 percentage point of the standard normal distribution (for example, for a 90 % confidence interval $Z_{\alpha/2}$ = 1.65; for a 95 % confidence interval $Z_{\alpha/2}$ = 1.96; etc.), and $S(d')$ is the standard deviation of d', that is, the square root of, $S^2(d') = B/n$. Similarly, 100(1 - α) % one-sided confidence intervals on δ are calculated as: $d' + Z_\alpha S(d')$ for a one-sided upper confidence interval and $d' - Z_\alpha S(d')$ for a one-sided lower confidence interval, where Zα is the upper-α percentage point of the standard normal distribution (for example, for a 90 % confidence interval Z_α = 1.28; for a 95 % confidence interval Z_α = 1.65; etc.) and d' and $S(d')$ are as defined above.

6.5.2 To test if δ is greater than zero, that is, that the two samples in the test are perceptibly different, use the binomial or chi-square test that is traditionally associated with the discrimination method used.

6.5.3 To test if it is reasonable to believe two δ's have the same value, that is, to test the hypotheses H_0: $\delta_1 = \delta_2$ versus H_a: $\delta_1 \neq \delta_2$ form the ratio: