

A Summary of ASTM Methods for Interlaboratory Testing

Developed in 1992 by AATCC Committee RA102.

I. ASTM Standard D 2904, Standard Practice for Interlaboratory Testing of a Textile Test Method that Produces Normally Distributed Data, and ASTM Standard D 2906, Standard Practice for Statements on Precision and Bias for Textiles, are guides for planning interlaboratory tests to evaluate proposed test methods and for using the test results to write precision statements of normally distributed data. It is intended that these methods as well as ASTM Standard D 4467, Interlaboratory Testing of a Textile Test Method that Produces Non-Normally Distributed Data, be used as guides in the development of information for precision and bias statements in AATCC test methods.¹ This monograph is a summary of the vital parts of the ASTM standards and are the minimum conditions that should be used in developing precision and bias statements for AATCC test methods. Components of variance for single operator, within laboratory, and between laboratory effects are determined. Critical differences calculated from the component variances indicate, for averages of “n” specimens from different samples, the minimum differences which are statistically significant.

II. Parameters in the Interlaboratory Test

A. Materials: There should be a minimum of two materials representing the range of interest of the property being tested. Subsamples of materials should be as homogeneous as possible. Where possible, values for each material should be established by alternative methods to determine if there is a variable bias between the proposed method and the referee method at different levels of the property.

B. Laboratories: There should be a minimum of five laboratories.

C. Operators: A minimum of two operators within each laboratory is recommended; however, testing by a single operator is acceptable.

D. Specimens: There should be a minimum of two specimens tested by each op-

erator in each laboratory *from each material*. The number of specimens tested per operator should be determined from the established variance for the test (determined from tests on one material by one operator in one laboratory) and the smaller systematic effect it is desired to be able to detect. Procedures for calculation of the required number of specimens are detailed in Section 5.5 of ASTM Standard D 2905. It is advisable to test a larger number of materials in more laboratories with the number of operators per laboratory and the number of tests per operator at a minimum of two each. The order of tests should be randomized to eliminate any storage or time effects.

E. Instruments: This should not be an effect included in the statistical analysis. When multiple instruments are used in a laboratory, ascertain if differences among instruments exist, and if they do, use known standard samples to obtain correction factors.

III. Procedure

A. A preliminary within-laboratory ruggedness test should be performed prior to interlaboratory testing. It is also advisable to conduct a pilot scale interlaboratory test before full scale testing.

B. Obtain an adequate number of samples and code for distribution to laboratories. The materials should be completely randomized for assignment to laboratories. It may be necessary in some cases to employ partial randomization. For example, in testing yarn from different spinning frames, specimens from each frame may be allocated to each laboratory. Determine subsample homogeneity before distribution to laboratories.

C. Conduct tests at each laboratory according to the procedure of the proposed test method.

IV. Analysis

A. An analysis of variance (ANOVA) is used to determine the significance of effects (operators, laboratories) in the interlaboratory test. This procedure assumes uniformity of variances. If variances are not uniform, data transformations such as those suggested in Section 11 of ASTM Standard D 2904 may be required.

B. ANOVA for single materials.

1. Prepare a separate ANOVA for each material, using a specially designed statistics package from ASTM or

ANOVA procedures available in other statistical software packages (SAS, SPSS). In the latter case, the effects in the model should include laboratory, operator nested within laboratory and specimen nested within operator and laboratory as the sources of variation. The analysis will yield F-values for each effect, and these can be used to determine if significant differences exist between operators and/or between laboratories. Alternatively, the variances can be calculated by hand using the formulas in Annex A2 of ASTM Standard D 2904.

2. Determine the components of variance using either the ASTM or other statistical package. This calculation is part of the ASTM program, but it will require an additional procedure if other standard statistical packages are used (e.g., VARCOMP in SAS). Formulas for calculation of the components of variance are also given in Annex A2 of ASTM Standard D 2904.

3. Calculate the critical differences for each of the effects using the components of variance. The ASTM program provides these critical differences for selected numbers of specimens, or they may be calculated using the equation in Sections 8.2 and 8.4 of ASTM Standard D 2906. The critical differences for each material should be compared to determine if they are sufficiently similar to combine the data for all materials into a single ANOVA table. An engineering decision should be made on the practical importance of the observed variation in the critical differences for materials. The auxiliary tests listed in Section 15 of ASTM Standard D 2904 are helpful in making this determination.

C. ANOVA for all materials.

1. If the critical differences for all materials are sufficiently similar, prepare an ANOVA table which includes all materials. Determine significant effects with F-tests. The ASTM program will perform these analyses directly. For other programs, the effects included in the model should be: material, material*laboratory interaction, operator within laboratory, material*operator within laboratory interaction and specimen within operator and laboratory.

2. Calculate the components of variance and critical differences as before. The component of variance for materials is usually not calculated as the materials were deliberately chosen to exhibit different levels of the property of interest.

¹ASTM Standards D 2904, D 2906 and D 4467 available from ASTM, 100 Barr Harbor Dr., West Conshohocken PA 19428; tel: 610/832-9500; fax: 610/832-9555; web site: www.astm.org.

3. If neither materials*laboratory or material*operator within laboratory is significant, do not include these in the precision statement. However, if either of these effects is significant, different components of variance apply to situations involving one material or more than one material. This means that the test method evaluates materials differently when tested in different laboratories or by operators within laboratories. In these cases *both* single material and multi-material components of variance should be calculated, with the latter including the components of variance for the materials interactions. (See Section A2.14.2.2 of ASTM Standard D 2904 for further explanation.)

V. Report or Precision Statement

A. Describe the interlaboratory test in terms of numbers of materials, laboratories, operators within laboratories and specimens tested by each operator.

B. Report the components of variance

(as variances or as standard deviations) and the critical differences for selected numbers of specimens, for each of the effects determined from the analysis including all materials. If the material interactions are not significant, report critical differences between specimens tested by each operator, between operators and between operators within laboratories.

C. If the materials interactions are significant, report both single and multi-material components of variance and critical differences, for single operator, within laboratory and between laboratory effects.

D. In the case of arbitrary grades, or of other grades that are limited and not continuous, or for which meaningful transformations may not be practicable (such as the AATCC Gray Scale), Recommended Text 8—Special Cases of Ratings in ASTM Standard D 2906 may be used. Most rating scales in AATCC methods are limited and not continuous.

Appendix

I. Definitions of terms (these definitions are taken from ASTM Standards D 123 and E 456)

A. **nested experiment**, n.—an experiment to examine the effect of two or more factors in which the same level (version) of a factor cannot be used with all levels (versions) of other factors.

B. **ruggedness test**, n.—a planned experiment in which environmental factors of test conditions are deliberately varied in order to evaluate the effects of such variation.

C. **standard deviation**, n.—the positive square root of the variance.

D. **specimen**, n.—a specific portion of a material or laboratory sample upon which a test is performed or which is taken for that purpose. (Syn. *test specimen*.)

E. **variance**, n.—a measure of the squared dispersion of observed values or measurements expressed as a function of the sum of the squared deviations from the population mean or sample average.