~~Designation:E1488–08~~ Designation: E 1488 – 08a

An American National Standard

Standard Guide for Statistical Procedures to Use in Developing and Applying Test Methods[1]

This standard is issued under the fixed designation E 1488; 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 guide identifies statistical procedures for use in developing new test methods or revising or evaluating existing test methods, or both.

1.2 This guide also cites statistical procedures especially useful in the application of test methods.

2. Referenced Documents

2.1 *ASTM Standards:* [2]

E 177 Practice for Use of the Terms Precision and Bias in ASTM Test Methods
E 178 Practice for Dealing With Outlying Observations
E 456 Terminology Relating to Quality and Statistics
E 691 Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method
E 1169 Practice for Conducting Ruggedness Tests
E 2282 Guide for Defining the Test Result of a Test Method
E 2489 Practice for Statistical Analysis of One-Sample and Two-Sample Interlaboratory Proficiency Testing Programs
E 2554 Practice for Estimating and Monitoring the Uncertainty of Test Results of a Test Method in a Single Laboratory Using a Control Sample Program
E 2587 Practice for Use of Control Charts in Statistical Process Control

2.2 *ISO Standards:*

ISO 17025 General Requirements for the Competence of Testing and Calibration Laboratories[3]
ISO Guide to the Expression of Uncertainty in Measurement [3]

3. Terminology

3.1 *Definitions:*

3.1.1 *statistical procedures*, *n*—the organized techniques and methods used to collect, analyze, and interpret data.

3.1.1.1 *Discussion*—Statistical procedures include the sampling considerations or the experiment design for the collection of data, or both, and the numerical and graphical approaches to summarize and analyze the collected data.

3.2 For all other formal definitions of statistical terms, see Terminology E 456.

4. Significance and Use

4.1 All ASTM test methods are required to include statements on precision and bias.[4]

4.2 Since ASTM began to require all test methods to have precision and bias statements that are based on interlaboratory test methods, there has been increased concern regarding what statistical experiments and procedures to use during the development of the test methods. Although there exists a wide range of statistical procedures, there is a small group of generally accepted techniques that are very beneficial to follow. This document is designed to provide a brief overview of these procedures and to suggest an appropriate sequence of carrying out these procedures.

[1] This guide is under the jurisdiction of ASTM Committee E11 on Quality and Statistics and is the direct responsibility of Subcommittee E11.20 on Test Method Evaluation and Quality Control.

Current edition approved ~~May 15,~~Oct. 1, 2008. Published ~~July~~October 2008. Originally approved in 1992. Last previous edition approved in ~~2002~~2008 as ~~E1488–02e1.~~E 1488 – 08.

[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, 11 W. 42nd St., 13th Floor, New York, NY 10036.~~

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

[4] See the Form and Style Manual for ASTM Standards that specifies, when possible, precision statements shall be estimated based on the results of an interlaboratory test program.

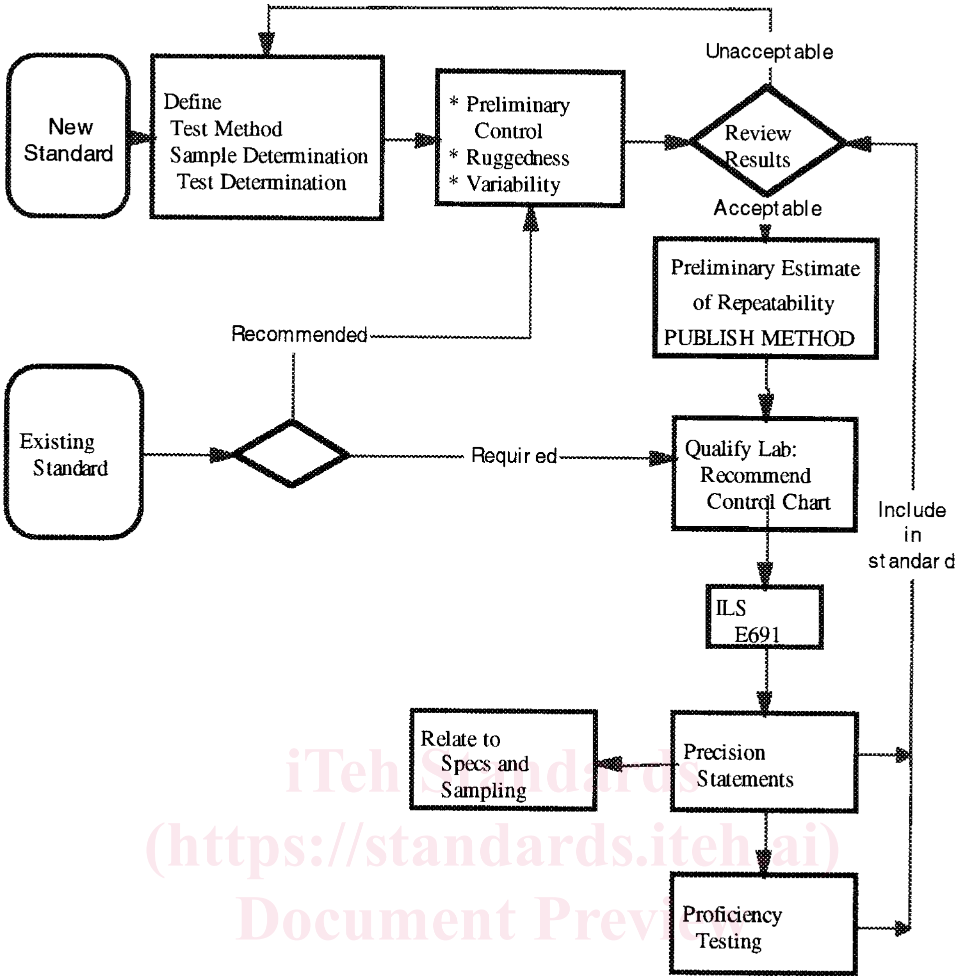

FIG. 1 Sequence of Steps

4.3 Statistical procedures often result in interpretations that are not absolutes. Sometimes the information obtained may be inadequate or incomplete, which may lead to additional questions and the need for further experimentation. Information outside the data is also important in establishing standards and in the interpretation of numerical results.

5. Summary of Guide

5.1 Outlined below is a suggested sequence of four phases useful in the development of a test method. A flowchart is provided in Fig. 1. Such a sequence of analyses may need to be modified in specific situations. The assistance of a qualified statistician is recommended at each review phase.

5.2 *Design Phase*:

5.2.1 This phase includes the formalization of the scope and the significance and use sections. It may include determining the purpose and describing a general approach to the test method but usually does not involve statistical studies.

5.3 *Development Phase*—

5.3.1 Studies may be conducted to evaluate the basic performance of the method. The draft test method is prepared and sampling requirements and the test result (see Guide E 2282) are clearly defined.

5.3.2 A flow chart is extremely valuable to identify the sequence of operations involved in a test method, for example, the sampling steps required to obtain the test specimens, definition of the test determination, how a test result is to be computed, and running the tests on the specimens.

5.4 Validation Phase

5.4.1 The test method is examined for such concerns as its stability, ruggedness, statistical control and the contributions to variability. The completion of this phase should result in preliminary estimates of precision and the identification and suggested ways to estimate potential contributors to uncertainty.

5.4.2 *Evaluation of Short Term Control of Test Method* —A test method must exhibit an ability to provide consistent results at least over short time periods. Preliminary studies or a pilot test should be conducted to evaluate the short term stability of the test method. A small series of repeated tests should be conducted.

5.4.3 *Analysis of Variability*—Statistically designed experiments conducted in one or two laboratories can be used to assess the relative magnitudes of different sources or potential contributors to variability of the test results. Such studies can provide estimates of intermediate measures of precision.

5.4.4 *Ruggedness Test*—A ruggedness test (see Practice E 1169) is a statistically designed experiment that helps identify problems in running the test method, clarifies errors, and points out possible environmental conditions, which may adversely affect the test method or point out need for tightening requirements. The ruggedness test can assist in locating ways of reducing variability in the test method.

5.4.5 *Preliminary Estimates of Precision*— From the various studies conducted in accordance with 5.4.2–5.4.4, preliminary estimates of repeatability standard deviations should be developed and published in this test method. Until an interlaboratory study is performed, these estimates generally are considered to be provisional. Information on how a lab should develop uncertainty estimates should also be provided.

5.4.6 *Statistical Control*—A test method must show capability of performing in a consistent way over time. The use of control charts (see Manual 7)[4] to monitor a proposed, or existing, test method over time is one recommended way to examine the controllability or stability of a test method. This statistical control should be demonstrated in one or two laboratories using homogeneous material (test specimen).

5.5 *Evaluation Phase*:

5.5.1 The test method is subjected to interlaboratory studies to provide estimates of within-laboratory repeatability and between-laboratory reproducibility. Additional information is supplied from proficiency studies when conducted.

5.5.2 *Interlaboratory Study (ILS)*—In accordance with ASTM Form and Style Manual, whenever feasible, an interlaboratory study must be conducted. This procedure will provide specific estimates of variation anticipated when using the test method.

5.5.3 Protocol for the ILS, Practice E 691 provides a guide for developing the ILS for the test method. A first step is the writing of an ILS Protocol, which will set out what needs to be done before the test specimens (or test materials) are distributed to the participating laboratories.

5.5.4 *Precision Statements*—Using the estimates of variation obtained in the interlaboratory test, one may prepare precision statements using Practices E 691 and E 177 or equivalent procedures.

6. Development of Test Method — Sampling and Test Result

6.1 Proposed standards that are under development should be treated in a formal manner following as many of the suggested procedures as possible. Standards that are already in existence as approved test methods or in general practice require periodic review that would include selected procedures.

6.2 *Under Development*—The development stage involves test methods that are in the preliminary stages during which equipment may not have been fully tested, practices are not agreed upon, and operators have yet to be adequately trained. Often this stage also applies to standards that have not yet been approved.

6.2.1 It is essential that tests for statistical control, ruggedness, and variability analyses be conducted prior to any interlaboratory test programs.

6.2.2 After all major environmental contributors have been identified, controlled, and incorporated into the test method, and after adequate standardized equipment is available, an interlaboratory test can be conducted. The interlaboratory test program must be completed prior to the first 5-year review. The committee should strive to have interlaboratory results as soon as possible.

6.2.3 After evaluating data from ruggedness tests, variability analysis, or an interlaboratory test program, changes to the test method may be suggested.

6.2.4 If major changes are made to the test method, a repeat of the various steps is usually necessary. Precision and bias statements should reflect the most current version of the test method.

6.3 *Existing Standards*—These standards comprise test methods that are in common use for which standard equipment may exist and for which experienced operators have been trained and are available.

6.3.1 Control charting, ruggedness tests, and variability analyses will be useful, especially if they have not previously been conducted. Such tests may provide better information about variation and necessary tolerances than has previously been available.

6.3.2 If precision estimates have not been established through an actual interlaboratory test program, then such a program should be initiated.

7. Data and Sampling

7.1 *Sample Determination*:

7.1.1 The sampling section of a standard should indicate clearly what constitutes the primary sampling unit, how that sampling unit is further subdivided, and how multiple test values are designated.

7.1.2 In considering the implication of test results as they relate to the material, the test method should be clear as to whether the sampling method or the test is destructive or nondestructive.

7.1.3 The user of the test method should be aware of whether the standard calls for a random sample. In some standards, as for example in sampling from coils or rolls of material, samples may be taken only from certain portions of the material.

7.2 *Test Result Determination*—The procedure for determining a test result must be clear and unambiguous.

7.2.1 An observation leads to an observed value.

7.2.2 Several observed values may lead to a test determination. The observed values need not be the same type of measurements (for example, they may consist of three readings such as length, width, and mass).

7.2.3 Several Test determinations may lead to a test result, as by averaging three test determinations.

7.2.4 A test result is the consequence of a single execution of the entire test method.

7.3 *Type of Data*—The kind of data that results from the application of the test method determines the types of statistical analyses to be performed.

7.3.1 *Numerical versus Categorical/Attribute Data*—Most of the statistical procedures referred to in this standard deal with numerical data. Control charts are available for all types of data, but all interlaboratory test procedures currently in use depend on numerical data.

7.3.2 *"Normally" Distributed Data*—Most of the statistical procedures referred to in this guide consider that the unknown distribution of the test results can be modeled by a normal distribution.

8. Sources of Variability

8.1 Experimental Realization of a Test Method

8.1.1 A realization of a test method refers to an actual application of the test method to produce a test result as specified by the test method. The realization involves an interpretation of the written document by a specific test operator , who uses a specific unit and version of the specified test apparatus, in the particular environment of his testing laboratory, to evaluate a specified number of test specimens of the material to be tested. Another realization of the test method may involve a change in one or more of the above emphasized experimental factors. The test result obtained by another realization of the test method will usually differ from the test result obtained from the first realization. Even when none of the experimental factors is intentionally changed, small changes usually occur. The outcome of these changes may be seen as variability among the test results.

8.1.2 Each of the above experimental factors and all others, known and unknown, that can change the realization of a test method, are potential sources of variability in test results. Some of the more common factors are discussed in Sections 8.2 to 8.6

8.2 Operator

8.2.1 *Clarity of Test Method*— Every effort must be made in preparing an ASTM standard test method to eliminate the possibility of serious differences in interpretation. One way to check clarity is to observe, without comment, a competent laboratory technician, not previously familiar with the method, apply the draft test method. If the technician has any difficulty, the draft most likely needs revision.

8.2.2 *Completeness of Test Method* — It is necessary that technicians, who are generally familiar with the test method or similar methods, not read anything into the instructions that is not explicitly stated therein. Therefore, to ensure minimum variability due to interpretation, procedural requirements must be complete.

8.2.3 *Differences in Operator Technique* — Even when operators have been trained by the same teacher or supervisor to give practically identical interpretations to the various steps of the test method, different operators (or even the same operator at different times) may still differ in such things as dexterity, reaction time, color sensitivity, interpolation in scale reading, and so forth. Unavoidable operator differences are thus one source of variability between test results. The test method should be designed and described to minimize the effects of these operator sources of variability.

8.3 Apparatus

8.3.1 *Tolerances*— In order to avoid prohibitive costs, only necessary and reasonable manufacturing and maintenance tolerances can be specified. The variations allowed by these reasonable specification tolerances can be one source of variability between test results from different sets of test equipment.

8.3.2 *Calibration*— One of the variables associated with the equipment is its state of calibration, including traceability to national standards. The test method must provide guidance on the frequency of verification and of partial or complete recalibration; that is, for each test determination, each test result, once a day, week, etc, or as required in specified situations. In some test methods the calibration may also depend on the levels. Linearity and constancy of variation may depend on the range of levels..

8.4 Environment

8.4.1 The properties of many materials are sensitive to temperature, humidity, atmospheric pressure, atmospheric contaminants, and other environmental factors. The test method usually specifies the standard environmental conditions for testing. However, since these factors cannot be controlled perfectly within and between laboratories, a test method must be able to cope with a reasonable amount of variability that inevitably occurs even though measurement and adjustment for the environmental variation have been used to obtain control (see 11 [will become section 12]). Thus, the method must be both robust to the differences between laboratories and require a sufficient number of test determinations to minimize the effect of within-laboratory variability.

8.5 Sample (Test Specimens)

8.5.1 A lot (or shipment) of material must be sampled. Since it is unlikely that the material is perfectly uniform, sampling variability is another source of variability among test results. In some applications, useful interpretation of test results may require the measurement of the sampling error.

8.5.2 In interlaboratory evaluation of test methods to determine testing variability, special attention is required in the selection of the material sample) in order to obtain test specimens that are as similar as possible. A small residual amount of material variability is almost always an inseparable component of any estimate of testing variability.

8.6 Time