NOTICE: This standard has either been superseded and replaced by a new version or discontinued.
Contact ASTM International (www.astm.org) for the latest information.

Designation: D 4270 – 95

Standard Guide for Using Existing Practices in Developing and Writing Test Methods[1]

This standard is issued under the fixed designation D 4270; 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 is intended to assist task groups in the preparation of a test method. It can help the task group use existing statistically related ASTM practices more effectively during the development and writing of the test method.

1.2 Some of the essential activities that should occur during the development of a new test method are not in existing or proposed practices under the jurisdiction of ASTM Subcommittee D13.93 on Statistics. This guide includes a brief explanation of how such activities should be carried out.

1.3 This guide is applicable to properties that are evaluated by both parametric and nonparametric methods of estimation. The instructions on properties that are best evaluated by nonparametric methods are less complete than those for properties that are evaluated by parametric methods.

2. Referenced Documents

2.1 *ASTM Standards:*

D 123 Terminology Relating to Textiles[2]

D 2904 Practice for Interlaboratory Testing of a Textile Test Method That Produces Normally Distributed Data[2]

D 2905 Practice for Statements on Number of Specimens for Textiles[2]

D 2906 Practice for Statements on Precision and Bias for Textiles[2]

D 3777 Practice for Writing Specifications for Textiles[3]

D 4271 Practice for Writing Statements on Sampling in Test Methods for Textiles[3]

D 4356 Practice for Establishing Consistent Test Method Tolerances[4]

D 4467 Practice for Interlaboratory Testing of a Test Method That Produces Non-Normally Distributed Data[3]

D 4853 Guide for Reducing Test Variability[3]

D 4854 Guide for Estimating the Magnitude of Variability from Expected Sources in Sampling Plans[3]

D 4855 Practice for Comparing Test Methods[3]

E 456 Terminology Relating to Quality and Statistics[4]

E 691 Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method[4]

2.2 ASTM Adjuncts:[5]

TEX-PAC

NOTE 1—TEX-PAC is a group of PC programs on floppy disk.

3. Terminology

3.1 *Definitions:*

3.1.1 *moving range, MR, n*—the difference without regard to sign between two successive observations.

3.1.2 *nonparametric, adj—in statistics*, a term referring to a technique that is not based on assumptions about the nature of the underlying frequency distribution. (Compare to *parametric*.)

3.1.3 *parametric, adj—in statistics*, a term referring to a technique that assumes the nature of the underlying distribution is known. (Compare to *nonparametric*.)

3.1.4 *practice, n*—a definitive procedure for performing one or more specific operations or functions that does not produce a test result.

3.1.4.1 *Discussion*—A practice is not a down-graded test method. Examples of practices include procedures for conducting interlaboratory testing programs or other statistical procedures; for writing statements on sampling or precision and accuracy; and for selection, preparation, application, inspection, necessary precautions for use or disposal, installation, maintenance, and operation of testing equipment.

3.1.5 *test method, n*—a definitive procedure for the identification, measurement, or evaluation of one or more qualities, characteristics, or properties of a material, product, system, or service that produces a test result.

3.1.6 For the definitions of other terms that appear in this standard, refer to Terminology D 123 and Terminology E 456.

4. Summary of Guide

4.1 The guide is summarized in Fig. 1, which shows the steps in test method development, the sections of this guide which apply, and other references that may be used as aids in a specific step in the development of the test method. As Fig. 1 points out, the references should show to which of the previous steps the development should revert when the results

[1] This guide is under the jurisdiction of ASTM Committee D-13 on Textiles and is the direct responsibility of Subcommittee D13.93 on Statistics.

Current edition approved May 15, 1995. Published July 1995. Originally published as D 4270 – 83. Last previous edition D 4270 – 90.

[2] *Annual Book of ASTM Standards*, Vol 07.01.

[3] *Annual Book of ASTM Standards*, Vol 07.02.

[4] *Annual Book of ASTM Standards*, Vol 14.02.

[5] PC programs on floppy disk are available through ASTM. For a 3½ in. disk request PCN: 12-429040; for a 5¼ in. disk request PCN: 12-429041-18.

Copyright © ASTM, 100 Barr Harbor Drive, West Conshohocken, PA 19428-2959, United States.

When the results of a step are less than completely satisfactory, the reference practice(s) should show to which of the previous steps the development of the test method should revert.

Step in Development	Section(s)	Other References
Determining Need	6	none
Sources of Methods	7.1	none
Applicability of Potential Methods	7.2	none
Initial Test Method Tolerances	7.3	Practice D 4356
Reducing Variability	7.4	Practice D 4853
Statistical Control	7.5	none
Revised Test Method Tolerances	7.6	Practice D 4356
Initial Draft(s)	8.1	Blue Book D13.91 White Paper
Selecting Procedure to Use	8.2	Practice D 4855
Interlaboratory Testing	9	Practice D 2904 Practice D 4467 Practice E 691
Terminology	10.2	Terminology D 123 Terminology E 456
Uses and Significance	10.3	Practice D 2906
Sampling	10.4	Practice D 2905 Practice D 3777 Practice D 4271 Guide D 4854
Precision and Bias	10.5	Practice D 2906 Blue Book
Review of Draft	11.1	none
Ballots	12.1	ASTM Regulations

FIG. 1 Suggested Steps in Test Method Development

of a specific step are less than satisfactory.

5. Significance and Use

5.1 There are enough existing practices related to the development of test methods or the preparation of one or more sections of new test methods so that even the experienced author may not use them as effectively as possible. This guide shows the person(s) preparing a new test method where and when the existing practices can best be used.

5.2 Using this standard as a guide, a task group should be able to prepare a draft of a new test in less time and with less effort than if this guide were not used. In addition, a well-prepared draft of a new test method is less likely to receive negative votes.

6. Determining Need for a Method

6.1 There is no simple answer to the question: "Is a new test method needed?" Generally, a new test method is needed if (*1*) no ASTM test method exists to identify, measure, or evaluate one or more significant qualities, characteristics, or properties of a material, product, system, or service and (*2*) need for such a method exists on an industry-wide basis. The responsible subcommittee needs to be convinced that a true need exists and be agreed upon the exact nature of the quality, characteristic, or property of interest to be evaluated. It is essential that a test method measure a quality, characteristic, or property that will predict the usefulness of a material, product, system, or service when it is put to its intended use.

7. Evaluating Potential Methods

7.1 *Sources of Potential Methods*—Potential procedures for a proposed new test method usually come from two sources: either one or more companies or laboratories have a method that is being used by them, or someone has decided that a specific procedure is a reasonable one to use even though it has not been used in industry. Existing procedures may be modified by the task group in the light of a consensus of viewpoints.

7.2 *Applicability of Potential Methods*—The first job of the task group is to consider each of the potential methods and determine if the method really measures the property of interest or whether it measures some related property instead. A specific procedure can usually be investigated within a single laboratory. In the light of these findings, a judgment should be made whether each of the potential methods is worthy of further work. In addition, if the method is supposed to predict performance in a later stage in the life of a material or product, the task group needs to verify that the method will really do so. At this point, the task group should eliminate any potential test procedure that does not meet these requirements. If all of the potential procedures are eliminated, the task group should make every reasonable attempt to discover or invent a procedure that does meet these requirements. Failure to do so will mean that the selected test method will not be as good as the users have a right to expect.

7.3 *Initial Test Method Tolerances*—Practice D 4356 discusses the fact that the tolerances specified for each of the measurements of some characteristic of a test specimen determine the exactness with which a test result can realistically be reported. For example, a procedure for determining mass per unit area of a fabric specimen should require that the length and width of the specimen be measured to within some tolerance and the mass to within another tolerance. If these tolerances are broad, it will not be realistic to require reporting the mass per unit area quite exactly. Before starting work on the one or more potential procedures that are still under consideration, the task group should select consistent test method tolerances.

7.4 *Reducing Variability*—It is quite likely that the results for each of the potential test methods include variability that might be eliminated or reduced by changes in the procedure. Guide D 4853 discusses how to locate the sources of such unnecessary variability and how to eliminate them by (*1*) using ruggedness tests, (*2*) evaluating and using components of variance, (*3*) averaging results from more than one specimen, (*4*) compositing samples prior to testing, and (*5*) using physical standards to reduce time related changes in the method or equipment. A brief description of ruggedness testing and a set of references are given in Practice D 2904. The task group should contact a person with statistical experience for help with the methods of eliminating unnecessary variability listed above.

7.5 *Statistical Control*—Useful testing procedures should be in a state of statistical control; that is, the long-term variability should be no greater than is predicted from the short-term variability. Guidance is given in 7.4 for detecting and eliminating assignable causes of increased long-term variability. The existence of a state of statistical control can be demonstrated by control charts prepared as directed in STP 15D.[6]

[6] *Manual on Presentation of Data and Control Chart Analysis, ASTM STP 15D,* ASTM 1976.

7.5.1 Checking the statistical control of a test procedure requires having a number of specimens that are as homogeneous as possible.

7.5.1.1 For test procedures that produce a determination that is a continuous variate or a count of nonconformities, two operators might each test one new randomly assigned specimen daily for twenty or more days.

7.5.1.2 For test procedures that produce a count of nonconforming specimens out of a specified number of specimens, two operators might each test one new block of a constant number of randomly assigned specimens daily for twenty or more days.

7.5.1.3 For test procedures that produce data like rankings or grades which have no known underlying frequency distribution, qualified statistical help is required.

7.5.2 Control charts prepared as directed in STP 15D can be used to demonstrate that (*1*) each operator is in control, (*2*) the data from the two operators have essentially the same control limits, (*3*) there is no bias between the operators, and (*4*) there are no significant trends with time. In the case of test procedures that produce a determination that is a continuous variate, control limits for the averages of the two operators can be also calculated using both Eq 1 and Eq 2:

$$\overline{\overline{X}} \pm 1.880\overline{R} \quad (1)$$

$$\overline{\overline{X}} \pm 2.66\overline{MR} \quad (2)$$

where:

$\overline{\overline{X}}$ = grand average of all observations,

$\overline{R}$ = average range or the average of the differences without regard to sign between the two observations on a single day, and

$\overline{MR}$ = average moving range or the average of the difference without regard to sign between successive daily averages.

7.5.3 If the data from each of the operators are in a state of statistical control, if the data for each operator yield essentially the same control limits, if there is no bias between the operators, and if there is no evidence of significant trends, the procedure for the proposed test method is in a state of control. If any one of these conditions is not true, the cause for the lack of control should be determined and eliminated if possible.

7.6 *Revised Test Method Tolerances*—If any significant improvements are made in the test procedure(s) as a result of 7.4 or 7.5, review and possibly revise the test method tolerances (see Practice D 4356).

8. Selecting the Procedure to be Used

8.1 For each procedure still under consideration, write a draft of the sections on apparatus, conditioning (if applicable), procedure, and reporting. Use the *Form and Style for ASTM Standards* (Blue Book) and the White Paper on preparing standards for Committee D-13 as guides in preparing these sections.[7] These sections should be in the proper form and complete enough for the user to find answers to all procedural questions instead of having to assume an answer that might differ from that of another user.

8.2 If there are still two procedures under consideration, make a choice based on the procedures described in Practice D 4855 and upon relative costs and complexity of the test methods.[8] If there are three or more procedures under consideration, get statistical help in making the choice. If the choice of a procedure is not obvious, defer the final decision on the choice until pilot interlaboratory tests have been run as directed in 9.3.

9. Interlaboratory Testing

9.1 An interlaboratory testing program should not be started until the task group is satisfied that all possible sources of unnecessary variability have been eliminated from each procedure still being considered. An interlaboratory testing program is essentially worthless as a means of detecting and eliminating unnecessary sources of variability. Embarking on such a program before thorough use of such practices as those specified in Section 7 often results in data that are so bad that the interlaboratory test data must be discarded, the procedure(s) improved as directed in Section 7, and the interlaboratory test repeated. This approach is a costly and time consuming error in procedure. Often some of the laboratories included in the first interlaboratory test program may not be able to take part in a second such test program.

9.2 Before any work is done on interlaboratory testing, the task group should be satisfied with the draft sections on apparatus, conditioning, procedure, and reporting for any procedures still being considered. The task group may prefer to submit these draft sections to an informal ballot of the entire subcommittee before considering interlaboratory testing.

9.3 Refer to Practice D 2904, Practice D 4467, or Practice E 691 for the planning and analysis of the interlaboratory test program. If there has been no final choice of alternative procedures, run pilot scale interlaboratory tests on each of the procedures still under consideration. If the procedures in 7.4 have been properly carried out and there is a clear cut choice of a procedure, the primary virtue of a pilot scale interlaboratory test is that it checks out the adequacy of the draft sections on apparatus, conditioning, procedure, and reporting.

9.3.1 Both Practice D 2904 and Practice E 691 are oriented toward test procedures that produce observations best evaluated by parametric methods of estimation. If the data from the interlaboratory test program are expected to be best evaluated by nonparametric methods of estimation, it is imperative for this fact to be considered during the planning of the program. Under these conditions, it is advisable to consult Practice D 4467 and to obtain qualified statistical help in the planning and evaluation of an interlaboratory test program and in the preparation of statements on precision and bias. Recommended Texts **3, 4, 5,** and **8** in Practice D 2906 apply to such data and should be studied before beginning an interlaboratory test program.

9.4 Use Practice D 4356 to determine (*1*) the measurement tolerances specified in the section on procedure, and (*2*) the test result tolerance specified in the section on reporting.

[7] Available from ASTM Headquarters.

[8] Mandel, J. and Stiehler, R.D., *Journal of Reasearch of the National Bureau of Standards*, Vol 53, 1954, p. 155 and same authors, Analytical Chemistry, Vol 29, 1957, p. 17A.

9.5 After the interlaboratory test results are all available, analyze the data utilizing the adjunct TEX-PAC. The output from the program will indicate whether or not the results from all materials can be grouped or if the results can be grouped for certain materials or need to be reported individually.

9.5.1 If the decision is that the materials do not differ enough to require separate reporting by materials, combine the data for all of the materials into an analysis of variance table like that in Table A2.4 of Practice D 2904. Calculate the components of variance using the equations in Table A2.4 of Practice D 2904, tabulate the results as shown in Annex A2 of Practice D 2904, and form estimates of the single-operator component, the within-laboratory component, and the between laboratory component as directed in Annex A2 of Practice D 2904. It may be advisable to get statistical assistance in performing these tasks.

10. Preparation of Complete Draft

10.1 Expand to a complete draft the partial draft specified in 8.1. Prepare the statistically related sections as directed in the following sections.

10.2 Use Terminology D 123, Terminology D 4392, and Terminology E 456 to develop the section on Terminology.

10.3 Use Section 4 of Practice D 2906 to prepare a statement on *Uses and Significance* about the use of the test method for acceptance testing of commercial shipments.

10.4 Use Practice D 3777 to plan the requirements for sampling. Use Practice D 4271 to write a section on sampling. If appropriate, use Practice D 2905 to write instructions on how a user is to calculate the number of specimens per subunit of the laboratory sample.

10.5 Use Practice D 2906 to prepare statements on precision and bias.

11. Review of Draft

11.1 Before sending the draft out to ballot, send a copy of the draft to Subcommittee D13.91 on Editorial Policy for an editorial review. Contact the D-13 Staff Manager at ASTM Headquarters for additional information.

12. Ballots

12.1 Upon agreement by the subcommittee at a regular meeting, send the draft out for a subcommittee ballot. The easiest way to initiate a subcommittee ballot is for the secretary of the subcommittee to send a copy of the draft to the staff person from the Standards Development Division of ASTM who is currently responsible for Committee D-13. The draft should be accompanied by a cover letter requesting a subcommittee ballot and stating whether the responses are to be sent to headquarters for later reporting to the subcommittee or to a member of the subcommittee. Alternatively, the secretary of the subcommittee can prepare and distribute the subcommittee ballot.

12.2 Upon completion of a satisfactory subcommittee ballot, the draft is ready for a Committee D-13 ballot. Requests for a Committee D-13 ballot should be sent by the subcommittee secretary to the responsible staff manager. The appropriate forms are available on request from the D-13 staff manager.

13. Keywords

13.1 developing test methods; statisitcs; writing test methods

The American Society for Testing and Materials takes no position respecting the validity of any patent rights asserted in connection with any item mentioned in this standard. Users of this standard are expressly advised that determination of the validity of any such patent rights, and the risk of infringement of such rights, are entirely their own responsibility.

This standard is subject to revision at any time by the responsible technical committee and must be reviewed every five years and if not revised, either reapproved or withdrawn. Your comments are invited either for revision of this standard or for additional standards and should be addressed to ASTM Headquarters. Your comments will receive careful consideration at a meeting of the responsible technical committee, which you may attend. If you feel that your comments have not received a fair hearing you should make your views known to the ASTM Committee on Standards, 100 Barr Harbor Drive, West Conshohocken, PA 19428.

This standard is copyrighted by ASTM, 100 Barr Harbor Drive, West Conshohocken, PA 19428-2959, United States. Individual reprints (single or multiple copies) of this standard may be obtained by contacting ASTM at the above address or at 610-832-9585 (phone), 610-832-9555 (fax), or service@astm.org (e-mail); or through the ASTM website (http://www.astm.org).