Designation: C802 – 14

Standard Practice for Conducting an Interlaboratory Test Program to Determine the Precision of Test Methods for Construction Materials[1]

This standard is issued under the fixed designation C802; 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 techniques for planning, conducting, and analyzing the results of an interlaboratory study (ILS) with the objective of developing the precision statement of a test method. It is designed to be used in conjunction with Practice C670. The methods used in this standard are consistent with those in Practice E691.

1.2 This practice is not intended for use in programs whose purpose is to develop a test method or to assess the relative variability of two or more test methods. Refer to Practice C1067 for procedures to evaluate the ruggedness of a test method.

1.3 The system of units for this practice has not been specified. Dimensional quantities in the practice are presented only in examples of calculations.

1.4 *This standard does not purport to address all safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 *ASTM Standards:*[2]

C109/C109M Test Method for Compressive Strength of Hydraulic Cement Mortars (Using 2-in. or [50-mm] Cube Specimens)

C136 Test Method for Sieve Analysis of Fine and Coarse Aggregates

C311/C311M Test Methods for Sampling and Testing Fly Ash or Natural Pozzolans for Use in Portland-Cement Concrete

C670 Practice for Preparing Precision and Bias Statements for Test Methods for Construction Materials

C1067 Practice for Conducting a Ruggedness Evaluation or Screening Program for Test Methods for Construction Materials

E105 Practice for Probability Sampling of Materials

E177 Practice for Use of the Terms Precision and Bias in ASTM Test Methods

E178 Practice for Dealing With Outlying Observations

E456 Terminology Relating to Quality and Statistics

E691 Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method

3. Terminology

3.1 *Definitions:*

3.1.1 For definitions of general statistical terms, refer to Terminology E456.

3.1.2 For definitions of terms associated with precision of test methods for construction materials, refer to Practice C670.

4. Significance and Use

4.1 This practice provides requirements for planning and conducting an interlaboratory study to obtain data to develop single-operator and multilaboratory precision statements for a test method. It includes methods to evaluate data consistency before carrying out the calculations to develop the precision statement. The procedures are compatible with those in Practice E691.

4.2 The ILS data obtained from this practice are intended for developing the precision values for writing single-operator and multilaboratory precision statements in accordance with Practice C670.

4.3 Appendix X1 provides an example to illustrate the calculations to obtain the precision values of the test method from the ILS data. This may be used to check a user-developed electronic spreadsheet for carrying out the calculations.

4.4 Appendix X2 discusses the additional calculations required for an interlaboratory study of a test method that includes making test specimens as part of the procedure. In this case, batch-to-batch variability needs to be considered.

[1] This practice is under the jurisdiction of ASTM Committee C09 on Concrete and Concrete Aggregates. This practice was developed jointly by ASTM Committee C01, C09, D04, and D18, and is endorsed by all four committees.

Current edition approved Dec. 15, 2014. Published June 2015. Originally approved in 1974. Last previous edition approved in 2009 as C802 – 09A. DOI: 10.1520/C0802-14.

[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.

***A Summary of Changes section appears at the end of this standard**

4.5 Appendix X3 discusses the use of analysis of variance as an alternative approach to obtain the precision values from the ILS data.

5. General Requirements

5.1 Certain criteria need to be met before undertaking an interlaboratory study to determine the precision of a test method. If some conditions are not met or are met incompletely, the program will become more complicated to administer and require more work and expense, or may result in impaired information. The requirements outlined in this section are intended to ensure that the test method is free of technical difficulties to the greatest extent possible before an expensive and time-consuming interlaboratory study is undertaken.

5.2 The first requirement is the existence of a valid and well-written test method that has been developed in one laboratory and has been subjected to ruggedness evaluation of the testing procedure and conditions as described in Practice C1067. As a result of the screening procedure and some experience with the test method in the sponsoring laboratory and one or two others, a written version of the test method has been developed (but not necessarily published as a standard) that describes the test procedure in terms that can be followed by a competent operator in any properly equipped laboratory. Critical conditions that affect the test results need to be identified and the proper and realistic degree of control of those conditions have to be specified in the description of the test procedure.

5.2.1 The tolerances established for various conditions in a test method provide reasonable ranges for these conditions and recognize that precise values with small tolerances may not be achievable in practice. Variations in test results due to variations in such conditions contribute to the total variation, which determines the precision of the test method. If the resulting variation is so great that uncertainties in average values obtained by the test method are unacceptably high, the test method itself is at fault and it will need to be improved or replaced by a better one. An expensive and time-consuming interlaboratory study is not recommended for such a test method.

5.2.2 Apparatus required for performing the test must be defined clearly and must be available or able to be produced. If alternative apparatus is permitted, criteria need to be provided on the performance requirements of the apparatus, such as by specifying acceptable limits of measurements on standard reference materials.

5.3 Personnel in laboratories participating in the ILS should have adequate experience with routine laboratory procedures so that they are competent to run the test. The importance of this requirement will vary with the complexity of the method and the degree to which it departs from familiar procedures.

5.4 It is helpful to have preliminary knowledge about how changes in materials and conditions affect the test results. There should be a reasonable degree of certainty that the single-operator variances are the same in different laboratories, and that troublesome interactions do not exist. These conditions are investigated in the initial analysis of the data of an interlaboratory study, and are discussed further in 10.4.

5.5 Facilities and procedures for procurement, preparation, and distribution of samples or test specimens must be available.

5.6 Selection of samples or test specimens must be done by a randomization process, and one person who is familiar with randomization procedures needs to be responsible for seeing that an appropriate randomization technique is used. Refer to Practice E105.

5.7 The precision of the test method should be evaluated on different materials with a range of the characteristic being measured that encompasses the typical use of the method in practice. (See 7.1 and 7.2.)

5.8 Adequate numbers of participating laboratories, operators, and materials must be available. Requirements in these areas are specified in Sections 6 and 7.

5.9 The entire interlaboratory test program should be developed from the beginning with the help and advice of persons familiar with statistical procedures and with the materials involved. The ASTM International Interlaboratory Study Program can support subcommittees in the development of precision statements by assisting in the design of an interlaboratory study, distribution of specimens or samples, data analysis, and preparation of a draft research report. Additional information about the ASTM ILS program can be obtained from the ASTM Website.

5.9.1 It may not always be possible to obtain people who are familiar with the materials involved and who have a sufficient knowledge of the proper statistical techniques and their proper use. In this case, the subcommittee should obtain the services of a statistician who has experience in practical work with data from materials testing, and provide that person with an opportunity for learning something about the particular materials and test method involved. Planners of an interlaboratory study need to avoid the pitfall of assuming that the use of statistical analysis software programs necessarily results in special expertise in manipulating the data or interpreting the results.

5.10 It is important to bear in mind that estimates of the precision of a test method are always based on a particular set of data obtained at a particular time and precision values need to be kept up-to-date. As materials, apparatus, and conditions change, and operators change or gain more experience, the characteristic precision of the results obtained may change, especially if the test method is new. In some cases, it may be desirable to conduct more tests at a later date (though not necessarily a repetition of the complete interlaboratory study) in order to provide a check on estimates previously obtained and either verify them or introduce revisions. When a subcommittee revises a test method, it should consider whether the proposed changes might affect the precision of the method. If there is a possibility that precision will be affected, limited interlaboratory testing is recommended to evaluate whether the existing precision statement is still applicable or if a new ILS needs to be organized to better reflect the precision of the revised method.

6. Laboratories

6.1 Obtaining participating laboratories for an interlaboratory study is often one of the most difficult problems connected with the process. The number of laboratories available is seldom as extensive as one would like, and if the test method is new, complicated, or expensive and time-consuming to run, the problem is further complicated.

6.2 For the purposes of programs using this practice, it is recommended that at least ten competent laboratories be included (**1, 2**).[3] In cases where it is impossible to obtain ten laboratories, the effect of an increased number may be obtained by repeating the program with the same group of laboratories six months later. If this procedure is followed, it is necessary to be sure that the same materials are used, and that their characteristics have not changed in the interim. This approach, however, may not provide a proper measure of the between-laboratory component of variance, unless different operators or equipment, or both, are used for the repeat testing. In any case, six is the absolute minimum number of laboratories for evaluating the precision of a test method. This means that at least seven to eight laboratories should be in the ILS study in case problems are encountered with the data provided by a participating laboratory.

6.3 In general, it is recommended that any laboratory that is considered qualified to run the test in routine testing situations should be permitted and encouraged to participate. "Qualified" implies proper laboratory facilities and testing equipment, competent operators familiar with routine laboratory techniques, a history of reliable testing work, and sufficient time and interest to do a good job. It does not mean, however, that only a select group of laboratories that are considered to be those best qualified for the interlaboratory study should be picked. Precision estimates for inclusion in a test method must be obtained under conditions and through the efforts of laboratories and personnel that are representative of the situations in which the test method will be used in practice (**2**). If a laboratory satisfies all the other requirements, but its personnel has had insufficient experience with the method, the operators in that laboratory should be given an opportunity to familiarize themselves with the method and to practice its application before the interlaboratory study starts.

7. Materials

7.1 *Number*—The number of materials to be included in an interlaboratory study will depend on the following:

7.1.1 The range of the values of the property that may be measured in practice and how the precision varies over that range;

7.1.2 The types of materials to which the test method is to be applied;

7.1.3 The difficulty and expense involved in obtaining, processing, and distributing samples or specimens;

7.1.4 The difficulty of, length of time required for, and expense of performing the tests; and

7.1.5 The uncertainty of prior information on any of these points. For example, if it is already known that the precision is relatively constant or proportional to the average level over the range of values of interest, a smaller number of materials will be needed than if it is known that the precision changes erratically at different levels. A preliminary pilot or screening program may help to settle some of these questions, and may often result in the saving of considerable time and expense in the full interlaboratory study (**1**).

7.2 In general, at least three materials or three different average values of the measured test characteristic is considered acceptable. The materials need to be selected to obtain as broad a range of the test characteristic as is practicable. If the test method is used to determine properties that are used for acceptance testing in a specification, it is particularly important that materials be included in the ILS whose properties are near the specification limits.

7.3 *Specimen Distribution*—The ILS is based on the assumption that all tests are performed on specimens that are as similar as is possible. Generally, two approaches are used for making and distributing the specimens or materials for the ILS.

7.3.1 For a test method that does not involve production of the test specimens as part of the method, specimens are produced at one location from a homogenous sample and then distributed to the participating laboratories. The specimens need to be assigned to the participating laboratories on a random basis. If the characteristic to be measured changes with age, specific instructions on test age need to be provided.

7.3.2 For a test method that involves fabrication of test specimens as part of the method, the raw materials for making the test specimens are shipped to the participating laboratories. In this case, samples of the constituent materials are taken from homogenous blends of the materials. The samples are selected on a random basis for shipment to the participating laboratories. Facilities are needed that have the proper equipment for blending the materials.

7.3.3 In some cases, it is not possible to distribute materials to participating laboratory because of the nature of the material or effects of transportation or age. This may require operators from participating laboratories to convene at one location to test the materials. This procedure is used commonly in developing precision statements for fresh concrete test methods.

8. Estimates of Precision

8.1 In accordance with Practice C670, the procedure described in this practice is designed to provide data to develop two basic estimates of the precision of a test method: (a) single-operator precision, and (b) multilaboratory precision (**1**). (See Note 1.)

8.2 *Single-operator precision* provides estimates of the inherent variability of the test method and the maximum difference that may be expected between replicate measurements made on the same material in the same laboratory by the same operator using the same apparatus within a time span of a few days. The words "may be expected" mean that there is 5 % likelihood that the difference will exceed the stated maximum difference, even if testing conforms to the test method. In

[3] The boldface numbers in parentheses refer to the list of references at the end of this practice.

Practice C670, the maximum acceptable difference is referred to as the "difference limit (d2s)" or "difference limit (d2s%)" if the coefficient of variation is the appropriate measure of precision.

8.3 *Multilaboratory precision* provides estimates of the variability among laboratories and the maximum difference that may be expected between measurements made on the same material in two different laboratories.

8.4 If estimates of precision due to other factors are required for the test method, the ILS needs to be planned to provide data to develop the appropriate statistics for the systems of causes being considered and the appropriate combination of modifiers given in Practice E177 should be used to describe those statistics. The advice of a statistical consultant is recommended for these cases.

NOTE 1—Appendix X2, for example, explains how to analyze ILS data for developing the *single-operator, multi-batch precision* of a test method that involves making the test specimens as part of the procedure. Another example is developing the *single-operator, multi-day precision*, which would involve the additional variability due to testing on different days.

9. Collection of Data

9.1 In order to minimize the problems concerned with analysis of data, a definite form and instructions for obtaining and recording the data have to be developed and distributed to all participating laboratories.

9.2 *Directions to Laboratories*—The directions to the laboratories should deal mainly with reporting of data. No special instructions for performing the tests that differ from those given in the test method should be included. The laboratories must be instructed to conduct tests and report results exactly as specified in the test method, with the one exception as noted in 9.2.2. Often data are disseminated in digital form, but laboratories need to maintain hard-copies of their data to provide documentation in the event of digital data corruption.

9.2.1 *Averaging Test Determinations*—Laboratories should particularly be cautioned against practices such as running a number of tests and selecting the "best" ones or reporting the average of several determinations, except as such averaging is specified in the test method. For example, Test Method C109/C109M specifies three or more test specimens, and requires that the strength of all acceptable test specimens made from the same batch and tested at the same period shall be averaged and reported. In this case, the directions for the interlaboratory study must specify the number of individual determinations to be obtained and reported. If a test result is defined, either in the test method or in the instructions to laboratories participating in an interlaboratory test program, as the average of a particular number of determinations, the individual determinations shall always be reported, in addition to the averages.

9.2.2 *Rounding of Data:*

9.2.2.1 Generally, laboratories need to report all figures obtained in making the measurements, rather than rounding the results before recording them. In some cases, this may result in recording of more digits than is customary or even more than the test method calls for in the section on Reporting. This is necessary because the variation from which information about the precision of the test method comes is contained in the least significant digits, which are often discarded in reporting the results of routine testing (**3**). For example, Test Method C136 calls for reporting the percentage retained on a sieve to the nearest whole number. This is adequate for the usual reporting purposes, but for purposes of determining precision, at least one decimal place is needed. It is better to require the reporting of too many decimal places than too few, because a decision about rounding all data can be made when the analysis is done. If too few places are reported, however, valuable information may be irretrievably lost, and the result might well be the impairment of the entire program.

9.2.2.2 If a test determination is the result of a calculation based on two or more measured quantities, the basic measurements should be used in the calculations without any rounding. The planners of the interlaboratory program will then have to determine how many places need to be reported in order to retain the essential information for determining variability. Sometimes it is advisable to ask the laboratories to report the basic quantities measured instead of, or in addition to, the final calculated result. This enables the final result to be checked, or changes in decisions about the test results to be made, when the data are analyzed. An example would be a strength test for which the measured specimen dimensions along with the ultimate load should be reported so that the reported strength can be verified.

9.3 *The Data Sheet*—This practice is based on the following assumptions: p laboratories each have made n replicate determinations on each of q materials (**4**). Table 1 and Table 2 are examples of data sheets for an individual laboratory and for a summary of data for the entire ILS program with: p = ten laboratories, n = four replicate determinations, and q = five materials. These data sheets suggest the format to be used if an individual determination constitutes a test result. If individual determinations are averaged or otherwise subjected to calculation to produce a test result, the format of the individual

TABLE 1 Data Sheet for an Interlaboratory Test Program for an ASTM Test Method

Laboratory: XX					
Replicate	Material				
	A	B	C	D	E
a	_____	_____	_____	_____	_____
b	_____	_____	_____	_____	_____
c	_____	_____	_____	_____	_____
d	_____	_____	_____	_____	_____

TABLE 2 Summary Data Sheet for an Interlaboratory Test Program for an ASTM Test Method

Laboratory	Replicate	Material				
		A	B	C	D	E
1	a					
	b					
	c					
	d					
2	a					
	b					
	c					
	d					
3	a					
	b					
	c					
	d					
4	a					
	b					
	c					
	d					
5	a					
	b					
	c					
	d					
6	a					
	b					
	c					
	d					
7	a					
	b					
	c					
	d					
8	a					
	b					
	c					
	d					
9	a					
	b					
	c					
	d					
10	a					
	b					
	c					
	d					

laboratory data sheet may be altered or a secondary sheet provided to permit recording the fundamental measurements and the test results.

9.4 *Number of Replicates*—Even if the test method calls for a single determination as a test result, replicate determinations are required in the ILS in order to obtain information for calculating the single-operator precision.

9.4.1 The number of replicate determinations to be made on each material in each laboratory depends largely on the number of laboratories participating, on the homogeneity of the material, and on the expense, difficulty, and time involved in increasing the number of determinations. In order to obtain the necessary information to write a meaningful precision statement, it is often necessary to use more replicates in the interlaboratory study than is required for routine use of the test method. An increase in the number of replicates improves the estimates of single-operator precision but has no effect on between-laboratory precision (**2**). It is recommended that, for 10 to 15 participating laboratories, at least three replicates are required. If it is not possible to obtain 10 participating laboratories, the number of replicates, n, should be at least $(30/p) + 1$. If 30 is not a multiple of p, $30/p$ is rounded to the next higher integer. For more than 15 laboratories, the number of replicates may be reduced to two. This will give an adequate estimate of single-operator precision, but information about multilaboratory precision is not as good as desired with fewer than 10 laboratories.

9.4.2 The variation among replicate determinations is supposed to be representative of the irreducible error variance characteristic of the test method. In some cases, it is possible to take supposedly replicate measurements in such a manner that there is little or no opportunity for chance variation; and the

measurements are in effect simply repetitions of the same measurement. For example, in making a chemical analysis by atomic absorption or some other kind of automatic measuring device, laboratories have been known to take three readings of the meter on the same sample in quick succession. The three readings so taken were almost identical, but were still reported as replicate readings. In cases such as this, three separate readings with different portions of the sample or with separate specimens should be obtained, with the same operator and apparatus to provide meaningful replicate measurements.

9.5 *Outliers*—Section 10.4 describes a procedure for identifying test determinations that have unexpected variations from those obtained by other participants in the ILS. In general, the practice of discarding individual test determinations, which appear to differ by suspiciously large amounts from the others, is to be avoided unless there is clear evidence that there was some physical reason to consider the determination faulty. Discarding results with unexpected deviations but without a proper basis or an assignable cause for the deviations may result in unrealistic precision values that may not be relevant to the test method. On the other hand, retaining invalid results with unexpectedly large variations may result in precision values that tolerate less careful testing. Laboratories must be instructed to report all results in their proper place and include notes describing the conditions surrounding those results that are suspected of being faulty. Sometimes if a test really went wrong, a laboratory should discard the results and repeat the test. Tests are not to be repeated, however, just because the results don't look good. Further guidance on dealing with outliers is given in Practice E178 and in Refs (**2, 5**).

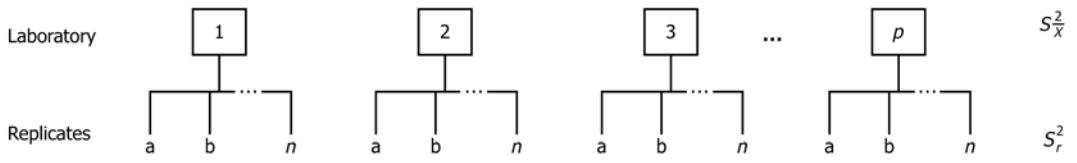
9.6 *Missing Data*—The method of data analysis used in this practices assumes a balanced set of data, which means that each laboratory provides the required number of determinations for all materials. Sometimes individual determinations are missing from the summary because they were discarded, failed to be supplied by a laboratory, or for other reasons. In general, if the number of missing data items from all laboratories constitutes no more than about 3 % of the total number of items, the analysis may be conducted as though the missing items were present. For example, if one out of four required replicate determinations on a given material from laboratory, *i*, out of 10 laboratories is missing, the three reported determinations should be added and divided by three to obtain the average, $\bar{X}_i$. The single-operator variance, s_{ri}^2, should be calculated using three for the number of determinations. From then on, both values should be used as though they were based on four determinations. If the number of missing results exceeds 3 % of the total, some of the tests should be repeated in order to obtain proper measurements for the missing values. Missing values handled in this way must be individual values distributed throughout the mass of data, and are not to be concentrated as a group in one laboratory-material cell (see Note 2). If the latter occurs, the laboratory should provide another group of measurements on the material in question.

NOTE 2—A cell refers to the group of replicate test determinations for a particular combination of laboratory and material. Appendix X3 describes an analysis-of-variance procedure that can be used to analyze unbalanced sets of data. The advice of a statistical consultant should be obtained if such procedures are used.

10. Analysis of Data

10.1 An ILS is a type of experiment design known as a nested design or a hierarchal design (**6**). The general purpose of a nested design is to identify and quantify the sources of variation in a process. In the case of the ILS, the objective is to quantify the single-operator and between-laboratory components of variance. Fig. 1 is a diagram of a single-stage nested design that is representative of the basic ILS described in this practice. The laboratories participating in the ILS are the first stage representing the factor "laboratory." For each laboratory, there are *n* replicate determinations. The replicate determinations are considered to be "nested" within the factor "laboratory."

10.2 The procedure described herein is simplified, and statistical terms are avoided to the greatest extent possible in order to make the practice usable by persons with little statistical background. This exposes the practice to the danger that, although the technique recommended is widely applicable to many situations using many kinds of data, it may be used mechanically in situations for which it is not applicable. For this reason, it is recommended to seek the advice of a person who is familiar with the statistical procedures before undertaking analysis of ILS data by this or any other published procedure. An example of the procedure is given in Appendix X1. For further description of the method, see Ref (**1**). The



$$\text{Between-laboratory component of variance} = S_L^2 = S_{\bar{X}}^2 - \frac{S_r^2}{n}$$

FIG. 1 Diagram of a Single-Stage Nested Experiment Design

TABLE 3 Single-Operator and Between-Laboratory Analysis for Material A[A]

Laboratory	Data (Replicates)				Average, $\bar{X}_{iA}$	Single-Operator Variance, s^2_{riA}
	a	b	c	d		
1	—	—	—	—	$\bar{X}_{1A}$	s^2_{r1A}
2	—	—	—	—	$\bar{X}_{2A}$	s^2_{r2A}
3	—	—	—	—	$\bar{X}_{3A}$	s^2_{r3A}
4	—	—	—	—	$\bar{X}_{4A}$	s^2_{r4A}
5	—	—	—	—	$\bar{X}_{5A}$	s^2_{r5A}
6	—	—	—	—	$\bar{X}_{6A}$	s^2_{r6A}
7	—	—	—	—	$\bar{X}_{7A}$	s^2_{r7A}
8	—	—	—	—	$\bar{X}_{8A}$	s^2_{r8A}
9	—	—	—	—	$\bar{X}_{9A}$	s^2_{r9A}
10	—	—	—	—	$\bar{X}_{10A}$	s^2_{r10A}

[A] p = 10 laboratories
n = 4 replicate test determinations on each material in each laboratory

Overall average $\bar{X}_A = \frac{\Sigma \bar{X}_{iA}}{p}$

Pooled single-operator variance = $s^2_{rA} = \frac{\Sigma s^2_{riA}}{p}$

Variance of laboratory averages = $s^2_{\bar{X}_A} = \frac{\Sigma(\bar{X}_{iA} - \bar{X}_A)^2}{p-1}$

Between-laboratory component of variance = $s^2_{LA} = s^2_{\bar{X}_A} - \frac{s^2_{rA}}{n}$

procedure does not require sophisticated software and can be implemented using an electronic spreadsheet.[4]

10.3 *Single-Operator and Between-Laboratory Components of Variance for Each Material*—Before starting the analysis, plot the data. This can be done by making a scatter plot of the test determinations for each laboratory. A separate plot can be made for each material, or all data can be shown in one plot. These plots will reveal any potential data inconsistencies that will be investigated further in accordance with 10.4. The first step in the analysis is to obtain estimates of single-operator and between-laboratory components of variance for each material. This may be done by setting up the data as shown in Table 3 and using the equations presented in this section. Table 3 is set up as an example using Material A for tests in ten laboratories (p = 10) with four replicate determinations per laboratory (n = 4) to correspond with the example summary data sheet in Table 2. Each row of data represents a particular laboratory-material combination and often called a *cell*. Tables similar to Table 3 would be used for each material in the study. In the equations that follow, the subscript g is used to designate a single test determination for a particular material in one laboratory and goes from 1 to n, where n is the number of replicates for each laboratory. The lower case letters (a, b, c, d) in Table 3 represent the replicate test determinations. The subscript i is used to designate a particular laboratory in the analysis and goes from 1 to p, where p is the total number of laboratories. The subscript j is used to designate the different materials, and goes from 1 to q, where q is the total number of materials. As shown in Table 1 and Table 2, the different materials are identified with capital letters A, B, C, and so forth.

10.3.1 *Single-operator analysis*—The averages, $\bar{X}_{iA}$, and variances, S^2_{riA}, in the last two columns of Table 3 are the single-operator averages and variances for the given material (in this example, it is Material A). These quantities are calculated from the n replicate test determinations within each of the p laboratories as follows:

x_{gij} = single test determination g by laboratory i for material j

$$\bar{X}_{ij} = \frac{\Sigma x_{gij}}{n} \quad (1)$$ = average of n replicate test determinations for laboratory i on material j

$$s^2_{rij} = \frac{\Sigma(x_{gij} - \bar{X}_{ij})^2}{n-1} \quad (2)$$ = single-operator variance of replicate determinations for laboratory i on material j

10.3.2 *Between-laboratory analysis*—From the single-operator averages, Eq 1, and variances, Eq 2, for each laboratory, the following quantities are calculated for the given material: (1) the pooled single-operator variance, (2) the

[4] A statistical software package called Dataplot® is available from the National Institute of Standards and Technology that will perform the plotting and calculations described in this practice. The program can be downloaded from homepage: http://www.nist.gov/itl/sed/dataplot.cfm. Instructions on how to analyze ILS data can be found at this site: http://www.itl.nist.gov/div898/software/dataplot/refman1/auxillar/e691.htm.

overall average, (3) the variance of laboratory averages, and (4) the between-laboratory component of variance. These values are entered at the bottom of Table 3 and are calculated as follows (Note 3):

$$s_{rj}^2 = \frac{\Sigma s_{rij}^2}{p} \quad (3) \quad = \quad \text{pooled single-operator variance for material } j \text{ (Note 4)}$$

$$\bar{X}_j = \frac{\Sigma \bar{X}_{ij}}{p} \quad (4) \quad = \quad \text{overall average for all laboratories for material } j$$

$$s_{\bar{X}_j}^2 = \frac{\Sigma\left(\bar{X}_{ij} - \bar{X}_j\right)^2}{p-1} \quad (5) \quad = \quad \text{variance of laboratory averages for material } j$$

$$s_{Lj}^2 = s_{\bar{X}_j}^2 - \frac{s_{rj}^2}{n} \quad (6) \quad = \quad \text{between-laboratory component of variance for material } j\text{. If the calculated values is negative, the between-laboratory component of variance is taken as zero.}$$

NOTE 3—Appendix X1 includes an example showing how these calculations are made for each material.

NOTE 4—The method of pooling variances used here applies only if the individual variances are based on the same number of replicate tests. In general, a pooled estimate of a variance is not obtained by averaging individual variances if the number of replicate determinations is not the same for all laboratories. Refer to a textbook on basic principles of statistics for the method to pool variances if the number of replicates is not constant.

10.4 *Data Consistency*—Before continuing with the remaining analysis of the ILS data to determine the precision of the test method, it is necessary to check each laboratory's data for consistency in terms of the average and the dispersion of the results. If data from one laboratory are not consistent with data from the other laboratories, it may be necessary to eliminate that laboratory's data before completing the analysis. Inconsistent data may inflate the calculated precision values and thereby encourage laboratories to tolerate less careful testing. The approach for checking data consistency used in this practice is the same as in Practice E691. Two statistics are used to evaluate data consistency: (1) the *h*-value and (2) the *k*-value.

10.4.1 *Check Laboratory Averages*—The *h*-value is used to check whether the average value for a laboratory is consistent with the overall average of the other laboratories for a given material. The *h*-value is calculated for each laboratory and material as follows:

$$h_{ij} = \frac{\bar{X}_{ij} - \bar{X}_j}{s_{\bar{X}j}} \quad (7)$$

where:

$\bar{X}_{ij}$ = average of results for laboratory *i* and material *j* (Eq 1),
$\bar{X}_j$ = overall average of results for material *j* (Eq 4), and
$s_{\bar{X}j}$ = standard deviation of laboratory averages for material *j*, which is the square root of Eq 5.

10.4.2 *Check Laboratory Dispersion*—The *k*-value for each laboratory is used to check the consistency of the single-operator variability for a given material. The *k*-value is calculated for each laboratory and material as follows:

$$k_{ij} = \frac{s_{rij}}{s_{rj}} \quad (8)$$

where:

s_{rij} = single-operator standard deviation of replicate determinations for laboratory *i* and material *j*, which is the square root of Eq 2, and
s_{rj} = pooled single-operator standard deviation for material *j*, which is the square root of Eq 3.

10.4.3 *Critical h- and k-values*—The calculated *h*- and *k*-values are compared with the critical values shown in Table 4, which is extracted from a larger table in Practice E691. The second column gives the critical *h*-value, which depends only on the number of laboratories. The subsequent columns give the critical *k*-values, which depend on the number of laboratories and the number of replicate test determinations. The *h*-values can be positive or negative, while the *k*-values are

TABLE 4 Critical Values of *h* and *k* at the 0.5 % Significance Level[A]

p No. of Labs	Critical value of *h* (±)	Critical values of *k* Number of replicates, *n*				
		2	3	4	5	6
3	1.15	1.72	1.67	1.61	1.56	1.52
4	1.49	1.95	1.82	1.73	1.66	1.60
5	1.74	2.11	1.92	1.79	1.71	1.65
6	1.92	2.22	1.98	1.84	1.75	1.68
7	2.05	2.30	2.03	1.87	1.77	1.70
8	2.15	2.36	2.06	1.90	1.79	1.72
9	2.23	2.41	2.09	1.92	1.81	1.73
10	2.29	2.45	2.11	1.93	1.82	1.74
11	2.34	2.49	2.13	1.94	1.83	1.75
12	2.38	2.51	2.14	1.96	1.84	1.76
13	2.41	2.54	2.15	1.96	1.84	1.76
14	2.44	2.56	2.16	1.97	1.85	1.77
15	2.47	2.57	2.17	1.98	1.86	1.77
16	2.49	2.59	2.18	1.98	1.86	1.77
17	2.51	2.60	2.19	1.99	1.86	1.78
18	2.53	2.61	2.20	1.99	1.87	1.78
19	2.54	2.62	2.20	2.00	1.87	1.78
20	2.56	2.63	2.21	2.00	1.87	1.79

[A] The above critical values for the *h* and *k* consistency statistics were calculated from Student's *t* and the *F*-ratio as described in Practice E691.

always positive. The critical values in Table 4 are the 0.5 % significance level. According to Practice E691, this significance level was chosen on the basis of judgment and experience so that not too many nor too few laboratories are flagged for further investigation. Appendix X1 of Practice E691 provides the basis for the critical values of *h* and *k*. Refer to Practice E691 for applicable values of *h* and *k* if more than 20 laboratories or more than 6 replicate determinations are involved in the ILS.

10.4.4 *Summary of h- and k-values*—The *h*- and *k*-values calculated for each laboratory and each material are assembled in a table. Values that exceed the critical values and values that approach the critical values should be highlighted. The *h*- and *k*-values should also be plotted as bar graphs grouped in two ways: (1) by laboratories and (2) by materials. The critical values should be drawn on the plots. The plots of *h* and *k* and the marked tables give a picture of the overall character of the variability of the test method as well as singling out particular laboratories that should be investigated.

10.4.5 *Plots by Laboratories*—Examples of plots of *h* and *k* by laboratories are shown in Appendix X1 based on the illustrative data. For each laboratory, the materials are grouped in increasing order of the overall average property value. The following guidelines can be used to evaluate differences between laboratories.

10.4.5.1 *h-Plot*—The *h*-plot indicates how the laboratory average property values for each material compare with the overall average for that material. There are several general patterns in these plots. In one, all laboratories have both positive and negative *h*-values among the materials. In the second, individual laboratories tend to have either positive or negative *h*-values for all materials, and the number of laboratories with negative values is approximately the same as the number of laboratories with positive values. Neither of these patterns is unusual or requires investigation, although they may tell something about the nature of the test method variability. In the third pattern, one laboratory has all positive (or negative) *h*-values compared with the other laboratories, which have substantially all negative (or positive) *h*-values. Such a pattern calls for an investigation of that laboratory. Another pattern to look for occurs within one laboratory, in which the *h*-values for materials with low property levels are of one sign and for materials with high property levels the *h*-values are of the opposite sign. If the *h*-values are extreme, investigation is warranted. As described in Practice E691, the investigation should consider examination of potential clerical errors in recording or transcribing data and it should consider examination of the laboratory reports for deviations from the test method or ILS protocol.

10.4.5.2 *k-Plot*—The *k*-plot compares the single-operator variability among the laboratories. As stated, *k*-values are always positive. The primary pattern to look for is whether a laboratory has large (or small) *k*-values for all or most of the materials. Elimination of that laboratory from the analysis may result in a set of data with similar *k*-values for the remaining laboratories. High *k*-values represent high single-operator variability. A check for outliers may be used to examine the data for the particular laboratory-material combination. The case of a small *k*-value is not usually as troublesome as that of a large *k*-value. If one laboratory, however, performs its tests in such a way that the normal causes of variation are not permitted to occur, there may be an unrealistically low single-operator standard deviation. Small *k*-values may indicate an insensitive measurement scale or other measurement problems. If all the *k*-values are erratic, the test method is in trouble. Efforts to develop precision statements from the data should be suspended and further study of the test method should be undertaken to determine the causes for such erratic behavior. The advice of a statistical consultant should be obtained if there is doubt about eliminating a laboratory with a high or low *k*-value.

10.4.6 *Plots by Material*—If a plot by laboratory shows several *h*- or *k*-values near the critical values, look at the corresponding plot by material to see how that laboratory differs from the rest for a given material. Often an *h*-value that seems strong in the plot by laboratory, because of its relation to the values for the other materials, will turn out to be reasonably consistent with the other laboratories for the same material. On the other hand, the *h*- or *k*-value for the one laboratory may be revealed as strongly different from the values for the other laboratories in the plot by material. If so, this behavior should be investigated.

10.4.7 *Interactions*—A common problem with test results obtained from an interlaboratory study is the presence of interactions among laboratories and materials. This means that the pattern of the results obtained on the material by one laboratory differs from the pattern obtained by the other laboratories. In extreme cases, different laboratories may even fail to rate materials in the same order based on the measured average properties. The accepted statistical technique for finding significant interactions is an analysis of variance of the total ILS data (all materials included). A reasonably reliable method for checking to see if troublesome interactions may exist, however, is to make a plot of the averages obtained on the materials by each laboratory. This plot should show similar patterns of change from material to material for all laboratories. If one laboratory shows a noticeably different pattern from the others it should be considered for elimination. If the patterns vary for more than one or two of the laboratories, the test method needs to be investigated, and the causes of the interactions discovered and eliminated. The advice of a statistical consultant should be obtained.

10.5 *Single-Operator and Multilaboratory Variances*—After the analyses for data consistency and interactions have been completed in accordance with 10.4 and a valid data set has been established, the final values of the various variance components are assembled as shown in Table 5. The averages in Column 2 are the overall averages of the measured characteristic for each material, which are arranged in order of increasing value. The component of variance in Columns 3 is the pooled single-operator variance (Eq 3), and the component in Column 4 is between-laboratory component of variance for each material (Eq 6). The single-operator variance in Column 5 is the same as the pooled variance in Column 3. The

TABLE 5 Averages, Components of Variance, and Variances for All Materials

Material (1)	Average[A] (2)	Components of Variance		Variance[B]	
		Single-Operator (3)	Between-Laboratory (4)	Single-Operator, s_r^2 (5)	Multilaboratory, s_R^2 (6)
A	$\bar{X}_A$	s_{rA}^2	s_{LA}^2	s_{rA}^2	$s_{rA}^2+s_{LA}^2$
B	$\bar{X}_B$	s_{rB}^2	s_{LB}^2	s_{rB}^2	$s_{rB}^2+s_{LB}^2$
C	$\bar{X}_C$	s_{rC}^2	s_{LC}^2	s_{rC}^2	$s_{rC}^2+s_{LC}^2$
D	$\bar{X}_D$	s_{rD}^2	s_{LD}^2	s_{rD}^2	$s_{rD}^2+s_{LD}^2$
E	$\bar{X}_E$	s_{rE}^2	s_{LE}^2	s_{rE}^2	$s_{rE}^2+s_{LE}^2$

[A] Listed in increasing order of magnitude.
[B] For single test determination.

multilaboratory variance, s_R^2, shown in Column 6 is obtained by summing the single-operator and the between-laboratory components of variance as follows:

$$s_{Rj}^2 = s_{rj}^2 + s_{Lj}^2 \quad (9)$$

= multilaboratory variance for material j

NOTE 5—The subscripts r and R are used for single-operator and multilaboratory conditions for consistency with Practices E177 and E691, which refer to these as *repeatability (r)* and *reproducibility (R)* conditions.

10.5.1 *Interpretation*—The single-operator and multilaboratory variances in Columns 5 and 6 of Table 5 apply to single test determinations, even though the data from which they are derived involve replicate test determinations. Replicate test determinations are necessary in the ILS to establish single-operator precision. The precision statements based on the variances defined by Eq 3 and Eq 9 will apply to comparisons between two single test determinations within a laboratory and to comparisons of single test determinations obtained in two laboratories.

10.5.2 *Test result is average of multiple determinations*—If the test method defines a test result as the average of m replicate determinations, the single-operator variance for comparison of two test results by the same operator is obtained by dividing the pooled single-operator variance, s_r^2, by m. For comparison of test results between two laboratories, the multilaboratory variance is obtained as follows:

$$s_{Rj}^2 = \frac{s_{rj}^2}{m} + s_{Lj}^2 \quad (10)$$

= multilaboratory variance for test result defined as the average of m replicate test determinations for material j

In this case, however, the pooled single-operator variance is applicable for establishing the maximum expected range among the replicate test determinations obtained by the operator in one laboratory. Refer to Practice C670 for additional guidance on determining the maximum expected range.

10.6 *Estimates of Precision*—The reason for listing the materials in increasing order of the average value of the measured property in Table 5 is to observe whether the precision varies with the level of the property measured, and thus to make a decision about the proper form of the precision statement. For this purpose, the quantities listed in Table 6 are calculated and displayed as shown, still in increasing order of magnitude of the average. Column 2 in Table 6 is the same as Column 2 in Table 5. Columns 3 and 4 contain the square roots of the variances in Columns 5 and 6 of Table 5. Columns 5 and 6 of Table 6 contain the corresponding coefficients of variation, expressed in percentage, that is, the respective standard deviation divided by the corresponding average in Column 1 and multiplied by 100.

10.6.1 *Determination of Form of Precision Statement*—The appropriate form of a precision statement depends on the relationship between the average level of the property measured for the different materials and the single-operator and multilaboratory standard deviations. There are three main forms of the relationship that cover most of the cases that are pertinent to ASTM test methods: (a) cases in which the standard deviation is relatively constant over the range of measured values; (b) cases in which the standard deviation has an approximately linear relationship with the average value and, therefore, the coefficient of variation is relatively constant; and (c) cases where the materials fall into two or more distinct groups within which condition (a) or (b) holds approximately, and for each of which a characteristic precision can be determined. In most cases, the determination of which of these alternatives applies, or whether some more complicated situation exists can be determined for practical purposes by plotting

TABLE 6 Averages, Standard Deviations, and Coefficients of Variation for All Materials

Material (1)	Average[A] (2)	Standard Deviation[B]		Coefficient of Variation[B]	
		Single-Operator (3)	Multilaboratory (4)	Single-Operator (5)	Multilaboratory (6)
A	$\bar{X}_A$	s_{rA}	s_{RA}	CV_{rA}	CV_{RA}
B	$\bar{X}_B$	s_{rB}	s_{RB}	CV_{rB}	CV_{RB}
C	$\bar{X}_C$	s_{rC}	s_{RC}	CV_{rC}	CV_{RC}
D	$\bar{X}_D$	s_{rD}	s_{RD}	CV_{rD}	CV_{RD}
E	$\bar{X}_E$	s_{rE}	s_{RE}	CV_{rE}	CV_{RE}

[A] Listed in increasing order of magnitude.
[B] For single test determination.

the standard deviations and coefficients of variation against the measured average value of the property. Two plots, one for the two standard deviations (single-operator and multilaboratory) and one for the two coefficients of variation, are usually adequate (see Note 6). If more sophisticated techniques are desired, they may be found in other references (**1, 4**). The appropriate measures of precision described in 10.6.2 – 10.6.5 become the indexes of precision used in Practice C670 to develop the precision statements.

NOTE 6—Usually, the same case should be applicable to both single-operator and multilaboratory precision. Sometimes, however, one of the two types of measures of precision depends on the property level and the other is not. In situations like this, it may be possible to select a suitable compromise in order to have the two precision statements in the same form. The advice of a statistical consultant should be obtained.

10.6.2 *Constant Standard Deviation*—In this case, the pooled single-operator standard deviation and the pooled multilaboratory standard deviation over all materials become the basic statistics for writing the precision statements in accordance with Practice C670. The pooled standard deviations are obtained by adding Columns 5 and 6 of Table 5 for the estimates of single-operator and multilaboratory variances, respectively, dividing each of the two totals by the number of materials, *q*, and taking the square roots of the quotients.

10.6.3 *Constant Coefficient of Variation*—In this case, the average single-operator coefficient of variation and the average multilaboratory coefficient of variation over all materials become the basic statistics for writing the precision statements in accordance with Practice C670. Because it is not possible to pool coefficients of variation in the same manner as variances and standard deviations, the simple arithmetic averages of Columns 5 and 6 in Table 6 are used.

10.6.4 *Separate Groups with Constant Standard Deviation or Coefficient of Variation* (see Note 7)—In this case, the single-operator and multilaboratory standard deviations or coefficients of variation are calculated separately for each material or group of materials in the same manner as described in 10.6.2 or 10.6.3. For each group, the range of average values over which the index of precision applies is supplied with the estimates. Refer to X1.3 of Practice C670 for an example of a precision statement for this case.

NOTE 7—Situations of the type described in 10.6.5 and 10.6.5.1 are often indications that something is wrong with the experimental situation or the test method. If the standard deviation and coefficient of variation are so erratic that it is difficult to write an applicable precision statement without giving separate indexes of precision for each material tested in the interlaboratory program, this is very possibly an indication that the test method itself may be subject to erratic variations and may need to be restudied and revised. Also interactions or non-normal distributions may exist in the data (**2**). In cases of erratic precision, a precision statement in the test method may really be more misleading than helpful to persons trying to use or interpret the results of the test method. It may actually provide invalid information about what should be expected if the test method is used.

10.6.5 *Irregular or Nonlinear Relationship Between Standard Deviation, Coefficient of Variation, and Average Level* (see Note 7)—One way of dealing with situations that do not apply to 10.6.2 – 10.6.4 is to use the largest estimate of the standard deviation or coefficient of variation (whichever comes closest to being constant) and to use the abbreviation "max" after the indexes of precision (see 6.2.6 of Practice C670). This practice is discouraged because the resulting indexes of precision are certain to be more lenient than they should be. The maximum limit applies strictly to the level at which the maximum standard deviation or coefficient of variation occurred. Tests done at other levels, for which lower precision limits apply, will be judged on the basis of a wider tolerance than they should be. Also, individual estimates of variance can vary widely from each other yet still be estimates of the same underlying variance, and it may often be that the pooled or averaged estimates are still the most appropriate ones to use, even if upon superficial examination, the individual variances appear to scatter rather wildly. There are statistical methods, such as Bartlett's test, that can be used to establish whether the variance estimates are statistically similar (**7**). It is again emphasized that the advice of a statistical consultant is needed here.

10.6.5.1 Cases where the standard deviation or coefficient of variation is a nonlinear function of the average level are dealt with in Ref. (**1**). Very often the amount of data, especially the number of laboratories and materials, is insufficient to establish the form of such a relationship beyond question, and estimates of precision based on one of the cases already described will serve. In addition, the difficulty of writing a precision statement based on a nonlinear relation, that can be easily understood and applied by the user of a test method indicates that such statements should be avoided if possible.

11. Keywords

11.1 coefficient of variation; components of variance; data consistency; interlaboratory study; multilaboratory precision; single-operator precision; standard deviation

APPENDIXES

(Nonmandatory Information)

X1. EXAMPLE OF DATA ANALYSIS FOR AN INTERLABORATORY STUDY

X1.1 *Data Source*—The following example is based on data from an interlaboratory study conducted by ASTM Subcommittee C09.24 via the CCRL Pozzolan Proficiency Sample Program. Testing was in accordance with Test Methods C311/C311M.

X1.2 *Characteristics of the Study and Data*—The ILS program was not preceded by a preliminary testing program because the test method has been used for over 50 years. Hence, the procedure is well known and most laboratory technicians involved in testing cement and concrete are familiar with performing the test.

X1.2.1 In the main test program, 13 laboratories tested four fly ashes for chemical and physical properties. Data used in this example are the fineness test results. In accordance with Test Methods C311/C311M, fineness is determined by wet-washing a 1.000-gram specimen of the fly ash over a 45-μm sieve. The mass of dry material retained on the sieve is measured to the nearest 0.001 g. Fineness is expressed as percentage of the original mass of the specimen and reported to the nearest 0.1 %. Thus the data used constitute an ILS program with p = 13 laboratories; q = four materials; and n = three replicate test determinations per material.

X1.3 *Analysis of Data*:

X1.3.1 *Step 1: Assembling the Data*—Table X1.1 shows an example of a data sheet from one laboratory. In the planning of an interlaboratory study, complete directions for recording data need to be stated clearly, so that the laboratories submit the required information in a consistent format. In this example, the operators were instructed to report the individual fineness values (% retained) to the nearest 0.01%, which is one additional digit greater than is required by the test method. The four fly ash materials are called A, B, C, D, and each test determination is indicated by the letters a, b, and c.

X1.3.1.1 Table X1.2 is the summary data sheet for all laboratories and all materials.

X1.3.2 *Step 2: Plotting the Data*—After all the ILS data have been assembled into one table, prepare a scatter plot of all the data. Such a plot can reveal gross errors in data entry and may point to potential outliers that need to be subject to closer examination as described in 9.5 and 10.4. Fig. X1.1 is scatterplot of the data grouped by laboratory. Visual examination does not reveal any erratic results, but there are a few combinations of laboratories and materials that seem to deviate from general trends. These will be investigated further in the formal checks for data consistency.

X1.3.3 *Step 3: Single-operator and Between Laboratory Analysis for Each Material*—The procedure outlined in Section 10 is used for preliminary analysis of the single-operator and the between-laboratory components of variance. Data from all

TABLE X1.1 Laboratory 1 Data Sheet for Interlaboratory Test for Fly Ash Fineness

Laboratory: ABC Testing Agency, Washington, DC				
Replicate	Material			
	A	B	C	D
a	13.39	18.30	26.14	38.26
b	13.82	16.92	24.65	38.35
c	13.36	17.90	24.74	38.53

TABLE X1.2 Summary Data Sheet for Interlaboratory Test Program for Test for Fly Ash Fineness

Laboratory	Replicate	Material			
		A	B	C	D
1	a	13.39	18.30	26.14	38.26
	b	13.82	16.92	24.65	38.35
	c	13.36	17.90	24.74	38.53
2	a	13.31	16.95	24.19	37.31
	b	13.41	16.62	23.85	37.61
	c	13.29	16.38	24.18	36.89
3	a	12.44	17.13	23.34	36.63
	b	13.14	17.00	23.94	36.63
	c	13.74	17.36	23.34	37.12
4	a	11.89	16.85	23.64	36.35
	b	12.79	17.93	23.50	35.96
	c	12.48	18.34	24.25	37.07
5	a	12.54	17.38	25.94	37.57
	b	12.04	16.75	25.54	37.76
	c	12.34	16.87	25.64	37.83
6	a	13.02	17.78	24.10	37.94
	b	13.88	18.63	24.54	37.61
	c	13.90	19.76	23.49	38.12
7	a	13.20	16.33	24.12	36.06
	b	13.41	16.24	24.16	36.96
	c	13.51	16.17	23.86	36.59
8	a	12.30	16.88	23.30	37.36
	b	12.40	16.69	23.40	37.14
	c	12.50	17.18	23.40	36.64
9	a	12.98	17.29	24.31	37.43
	b	13.19	16.95	24.25	37.36
	c	12.27	16.80	24.26	37.27
10	a	13.83	16.80	27.43	37.26
	b	14.08	17.24	26.82	37.58
	c	13.80	17.05	26.71	37.10
11	a	13.58	17.45	24.32	37.39
	b	13.60	16.69	24.53	38.19
	c	13.68	16.98	24.58	38.55
12	a	13.37	17.63	24.56	37.29
	b	13.14	17.22	24.27	37.21
	c	12.83	17.29	24.15	37.87
13	a	12.21	18.16	23.84	37.98
	b	11.89	17.56	23.67	36.99
	c	11.96	17.58	23.84	37.29

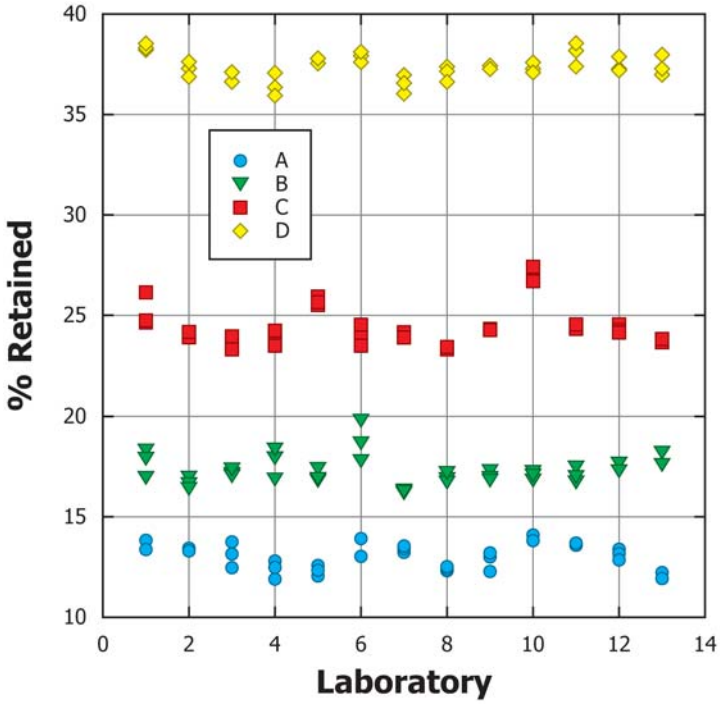


FIG. X1.1 Scatterplot of All Data

laboratories are used in the preliminary analysis. The preliminary analysis is needed to check for data consistency. The replicate data in each material column in Table X1.2 are transposed to create a separate table for each material as shown in Tables X1.3-X1.6. Each row shows the test determinations for each laboratory. The values under the headings a, b, and c are the replicate test determinations for each laboratory for the given material. The last two columns are the average, $\bar{X}_i$, and variance, s_{ri}^2, of the replicate test determinations by each laboratory. At the bottom of each of Tables X1.3-X1.6 are the various parameters calculated as described in 10.3.

TABLE X1.3 Single-Operator and Between-Laboratory Analysis for Material A[A]

Laboratory	Data			Average $\bar{X}_i$	Single-Operator Variance s_{ri}^2
	a	b	c		
1	13.39	13.82	13.36	13.52	0.0662
2	13.31	13.41	13.29	13.34	0.0041
3	12.44	13.14	13.74	13.11	0.4233
4	11.89	12.79	12.48	12.39	0.2090
5	12.54	12.04	12.34	12.31	0.0633
6	13.02	13.88	13.90	13.60	0.2524
7	13.20	13.41	13.51	13.37	0.0250
8	12.30	12.40	12.50	12.40	0.0100
9	12.98	13.19	12.27	12.81	0.2324
10	13.83	14.08	13.80	13.90	0.0236
11	13.58	13.60	13.68	13.62	0.0028
12	13.37	13.14	12.83	13.11	0.0734
13	12.21	11.89	11.96	12.02	0.0283

[A] p = 13 laboratories
n = 3 replicate test determinations on each material in each laboratory
Overall average = $\bar{X}$ = 13.04
Pooled single-operator variance = s_{rA}^2 = 0.109
Variance of laboratory averages = s_{XA}^2 = 0.359
Between-laboratory component of variance = s_{LA}^2 = 0.322

TABLE X1.4 Single-Operator and Between-Laboratory Analysis for Material B[A]

Laboratory	Data			Average $\bar{X}_i$	Single-Operator Variance s_{ri}^2
	a	b	c		
1	18.30	16.92	17.90	17.71	0.5041
2	16.95	16.62	16.38	16.65	0.0819
3	17.13	17.00	17.36	17.16	0.0332
4	16.85	17.93	18.34	17.71	0.5924
5	17.38	16.75	16.87	17.00	0.1119
6	17.78	18.63	19.76	18.72	0.9866
7	16.33	16.24	16.17	16.25	0.0064
8	16.88	16.69	17.18	16.92	0.0610
9	17.29	16.95	16.80	17.01	0.0630
10	16.80	17.24	17.05	17.03	0.0487
11	17.45	16.69	16.98	17.04	0.1471
12	17.63	17.22	17.29	17.38	0.0481
13	18.16	17.56	17.58	17.77	0.1161

[A] p = 13 laboratories
n = 3 replicate test determinations on each material in each laboratory
Overall average = $\bar{X}$ = 17.26
Pooled single-operator variance = s_{rB}^2 = 0.215
Variance of laboratory averages = s_{XB}^2 = 0.381
Between-laboratory component of variance = s_{LB}^2 = 0.309

X1.3.4 *Step 4: Investigation of Data Consistency*—The procedure described in 10.4 is used with the results of the preliminary analysis to check for consistency of the ILS data. Tables X1.7 and X1.8 show the h-values and k-values based on the parameter values calculated in the preliminary analysis, and using Eq 7 and Eq 8. For p =13 and n = 3, the critical values of h and k from Table 4 are ± 2.41 and 2.25, respectively.

TABLE X1.5 Single-Operator and Between-Laboratory Analysis for Material C[A]

Laboratory	Data			Average $\bar{X}_i$	Single-Operator Variance s_{ri}^2
	a	b	c		
1	26.14	24.65	24.74	25.18	0.6980
2	24.19	23.85	24.18	24.07	0.0374
3	23.34	23.94	23.34	23.54	0.1200
4	23.64	23.50	24.25	23.80	0.1590
5	25.94	25.54	25.64	25.71	0.0433
6	24.10	24.54	23.49	24.04	0.2780
7	24.12	24.16	23.86	24.05	0.0265
8	23.30	23.40	23.40	23.37	0.0033
9	24.31	24.25	24.26	24.27	0.0010
10	27.43	26.82	26.71	26.99	0.1504
11	24.32	24.53	24.58	24.48	0.0190
12	24.56	24.27	24.15	24.33	0.0444
13	23.84	23.67	23.84	23.78	0.0096

[A] p = 13 laboratories
n = 3 replicate test determinations on each material in each laboratory
Overall average = $\bar{X}$ = 24.23
Pooled single-operator variance = s_{rC}^2 = 0.122
Variance of laboratory averages = s_{XC}^2 = 0.994
Between-laboratory component of variance = s_{LC}^2 = 0.953

TABLE X1.6 Single-Operator and Between-Laboratory Analysis for Material D[A]

Laboratory	Data			Average $\bar{X}_i$	Single-Operator Variance s_{ri}^2
	a	b	c		
1	38.26	38.35	38.53	38.38	0.0189
2	37.31	37.61	36.89	37.27	0.1308
3	36.63	36.63	37.12	36.79	0.0800
4	36.35	35.96	37.07	36.46	0.3171
5	37.57	37.76	37.83	37.72	0.0181
6	37.94	37.61	38.12	37.89	0.0669
7	36.06	36.96	36.59	36.54	0.2046
8	37.36	37.14	36.64	37.05	0.1361
9	37.43	37.36	37.27	37.35	0.0064
10	37.26	37.58	37.10	37.31	0.0597
11	37.39	38.19	38.55	38.04	0.3525
12	37.29	37.21	37.87	37.46	0.1297
13	37.98	36.99	37.29	37.42	0.2577

[A] p = 13 laboratories
n = 3 replicate test determinations on each material in each laboratory
Overall average = $\bar{X}$ = 37.36
Pooled single-operator variance = s_{rD}^2 = 0.137
Variance of laboratory averages = s_{XD}^2 = 0.321
Between-laboratory component of variance = s_{LD}^2 = 0.275

Calculated *h*- and *k*-values that exceed the limits are underlined in Tables X1.7 and X1.8. Figs. X1.2-X1.5 show *h*- and *k*-plots grouped by laboratory and materials. The critical values of *h* and *k* are shown on the plots. As a review, the *h*-value is an indicator of between-laboratory consistency, and *k* is an indicator of single-operator consistency. A study of Figs. X1.2-X1.5 shows the following:

X1.3.4.1 Laboratory 10 appears to have obtained a higher average value for fly ash C than the other laboratories.

X1.3.4.2 The single-operator variability for Laboratory 1 and fly ash C appears to be greater than for the other laboratories.

TABLE X1.7 *h*-Values[A] for Between-Laboratory Data Consistency

Laboratory	Material			
	A	B	C	D
1	0.81	0.73	0.75	1.80
2	0.50	-0.98	-0.36	-0.16
3	0.11	-0.15	-0.89	-1.00
4	-1.09	0.73	-0.64	-1.59
5	-1.22	-0.42	1.28	0.64
6	0.94	2.38	-0.39	0.94
7	0.56	-1.64	-0.39	-1.45
8	-1.07	-0.55	-1.07	-0.55
9	-0.38	-0.40	-0.16	-0.01
10	1.44	-0.37	2.56	-0.08
11	0.97	-0.35	0.05	1.21
12	0.12	0.20	-0.10	0.17
13	-1.70	0.83	-0.65	0.11

[A] Critical *h*-value for p = 13 equals ± 2.41

TABLE X1.8 *k*-Values[A] for Single-Operator Data Consistency

Laboratory	Material			
	A	B	C	D
1	0.78	1.53	2.39	0.37
2	0.19	0.62	0.55	0.98
3	1.97	0.39	0.99	0.76
4	1.39	1.66	1.14	1.52
5	0.76	0.72	0.60	0.36
6	1.52	2.14	1.51	0.70
7	0.48	0.17	0.47	1.22
8	0.30	0.53	0.17	1.00
9	1.46	0.54	0.09	0.22
10	0.47	0.48	1.11	0.66
11	0.16	0.83	0.39	1.61
12	0.82	0.47	0.60	0.97
13	0.51	0.73	0.28	1.37

[A] Critical *k*-value for p = 13 equals 2.15

X1.3.5 *Step 5: Investigation for Interactions*—Fig. X1.6 is a dot plot of the average % retained for each laboratory and for each material designation arranged in increasing level of % retained. All laboratories showed similar patterns. It can, therefore, be concluded that no serious interactions were present. Based on this and the fact that no single laboratory reported results that were consistently different from the other laboratories, there was no strong basis for excluding data from any of the laboratories. Thus the results of the preliminary analyses indicated by the parameter values at the bottom of Tables X1.3-X1.6 are the final results for developing the precision values.

X1.3.6 *Step 6: Development of the Estimates of Precision*—Table X1.9 shows the components of variance and the single-operator and multilaboratory variances obtained for each material in accordance with 10.5. The single-operator and multilaboratory variances given in Table X1.9 are for individual test determinations as described in 10.5.1. Table X1.10 shows the corresponding single-operator and multilaboratory standard deviations and the corresponding coefficients of variation.

X1.3.7 *Step 7: Determination of the Form of the Precision Statements*—Fig. X1.7 and Fig. X1.8 show, respectively, plots of standard deviation versus average % retained and coefficient of variation versus average % retained. Each plot shows single-laboratory (SO) and multilaboratory (ML) results. Fig.

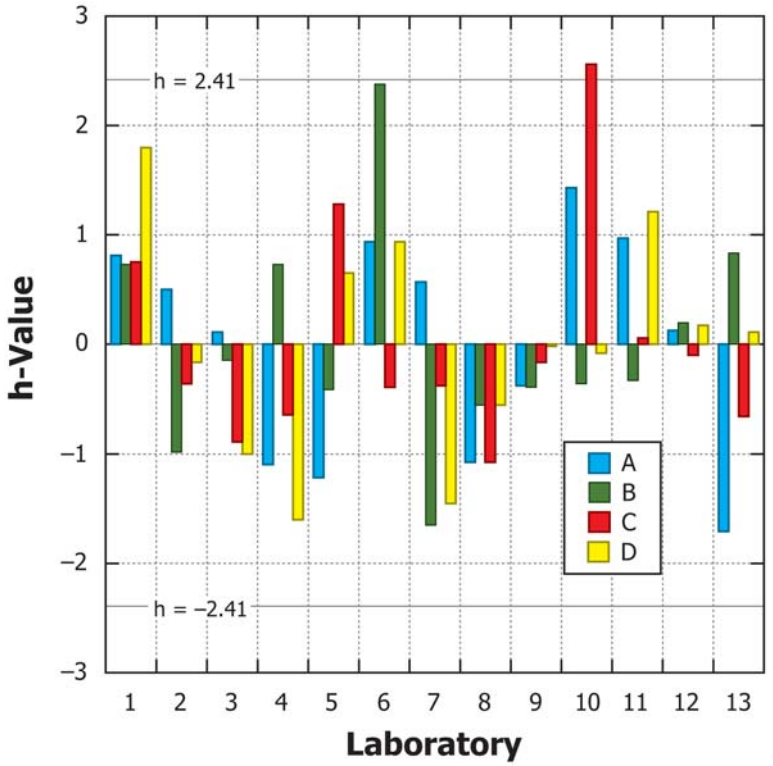


FIG. X1.2 *h*-Values Grouped by Laboratory

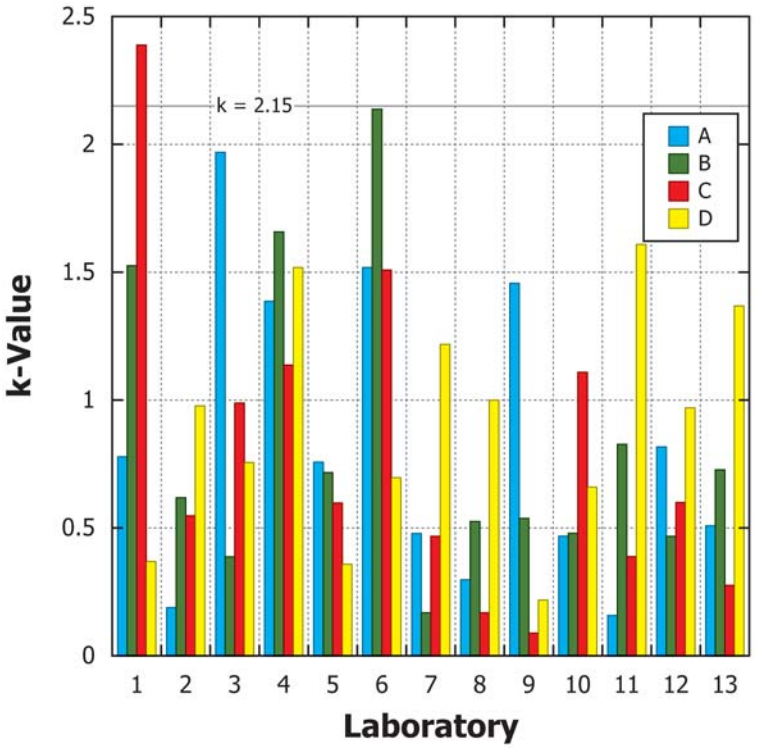


FIG. X1.3 *k*-Values Grouped by Laboratory

X1.7 indicates that the standard deviation tends to be relatively independent of the average % retained, for both single-operator and multilaboratory conditions. On the other hand, Fig. X1.8 indicates that both the single-operator and multilaboratory coefficient of variation tend to decrease with increasing level of % retained. Thus it is concluded that a constant standard

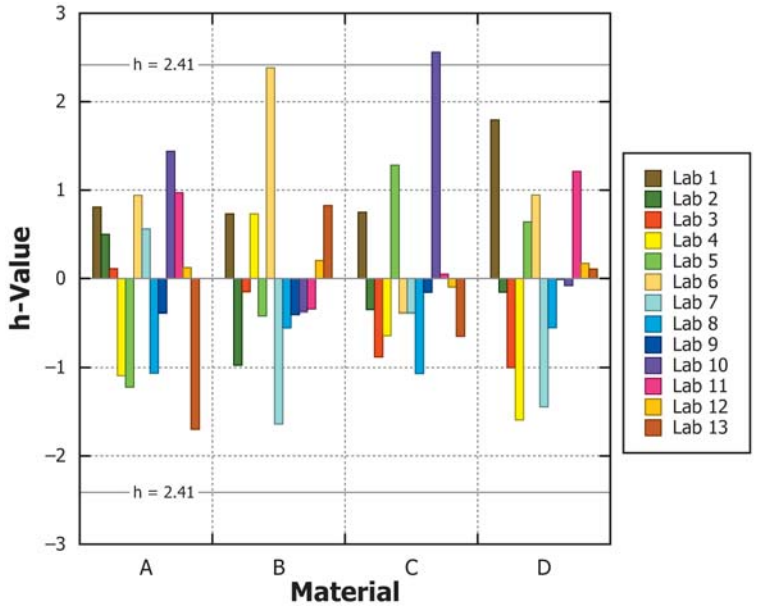


FIG. X1.4 *h*-Values Grouped by Material

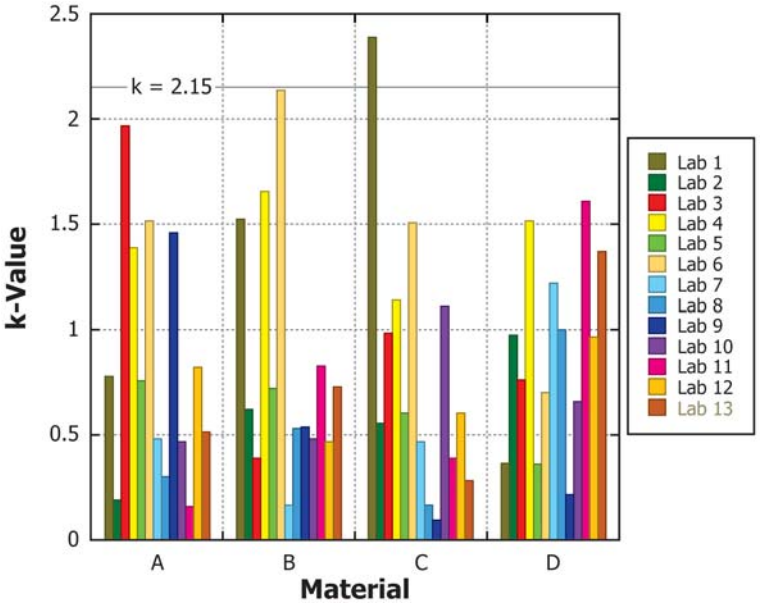


FIG. X1.5 *k*-Values Grouped by Material

deviation is the appropriate basis for developing the precision statements. Accordingly, the pooled single-operator and multilaboratory variances were obtained as the average values for the last two columns in Table X1.9. These average values are 0.146 and 0.611. Taking the square roots of these values results in a single-operator standard deviation of 0.38 % retained and a multilaboratory standard deviation of 0.78 % retained. The maximum expected differences between test results are obtained by multiplying these standard deviations by the appropriate factors as described in Practice C670.

X1.3.8 *Step 8: Writing the Precision Statements*—In accordance with Practice C670, the precision statements based on the ILS data are as follows:

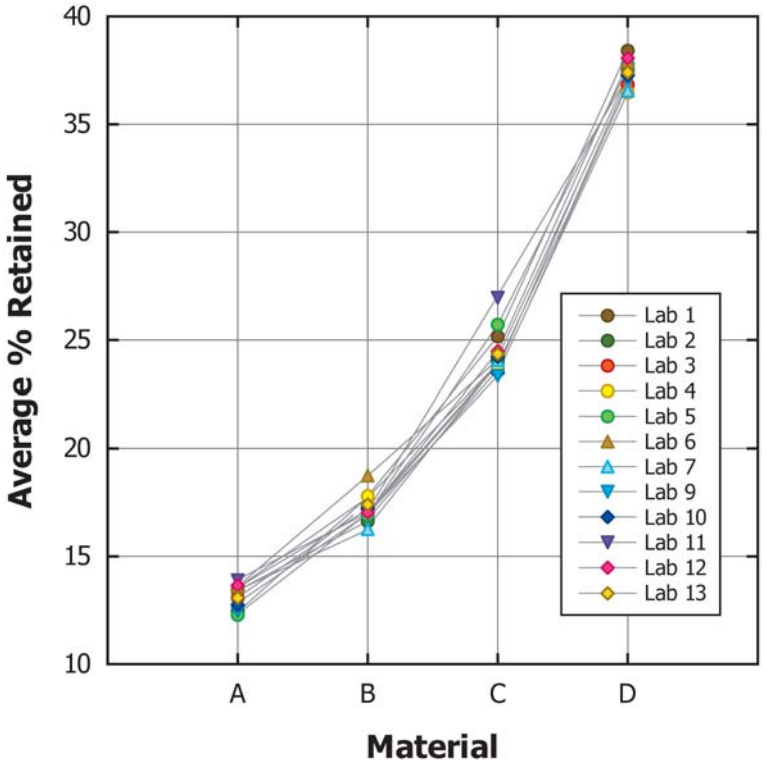


FIG. X1.6 Average % Retained versus Material

TABLE X1.9 Averages, Components of Variance, and Variances for All Materials

Material	Average[A]	Components of Variance		Variances[B]	
		Single-Operator	Between-Laboratory	Single-Operator	Multi-Laboratory
A	13.04	0.109	0.322	0.109	0.431
B	17.26	0.215	0.309	0.215	0.524
C	24.43	0.122	0.953	0.122	1.075
D	37.36	0.137	0.275	0.137	0.412

[A] Listed in increasing order of magnitude.
[B] For single test determination.

TABLE X1.10 Averages, Standard Deviations, and Coefficients of Variation for All Materials

Material	Average[A]	Standard Deviations[B]		Coefficients of Variation[B]	
		Single-Operator	Multi-Laboratory	Single-Operator	Multi-Laboratory
A	13.04	0.330	0.657	2.53	5.03
B	17.26	0.464	0.724	2.69	4.19
C	24.43	0.349	1.037	1.43	4.24
D	37.36	0.370	0.642	0.99	1.72

[A] Listed in increasing order of magnitude.
[B] For single test determination.

X1.3.8.1 *Single-Operator Precision*—The single-operator standard deviation of a single test determination of % retained has been found to be 0.38 %. Therefore, results of two properly conducted tests by the same operator on the same material are not expected to differ by more than 1.1 %.[5]

X1.3.8.2 *Multilaboratory Precision*—The multilaboratory standard deviation of a single test determination of % retained has been found to be 0.78 %. Therefore, results of two properly conducted tests by two different laboratories on specimens of the same material are not expected to differ by more than 2.2 %.[5]

NOTE X1.1—These precision statements are based on an interlaboratory study that involved 13 laboratories, four materials with average % retained of approximately 13 %, 17 %, 24 %, and 37 %, and three replicate tests per operator. Supporting data have been filed at ASTM Headquarters and may be obtained by requesting Research Report RR:C09- XXXX.

[5] These numbers represent the difference limits (d2s) as described in Practice C670.

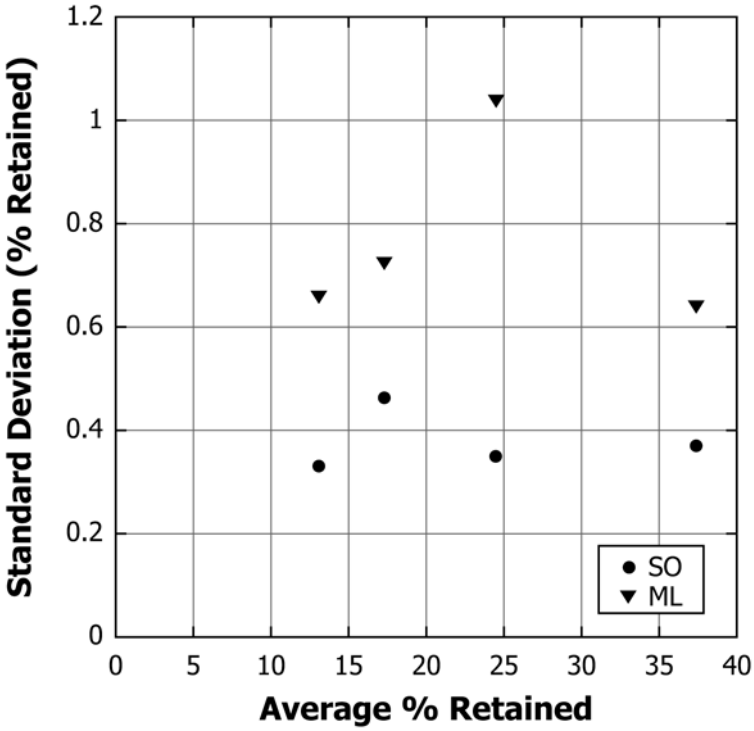


FIG. X1.7 Standard Deviation versus Average

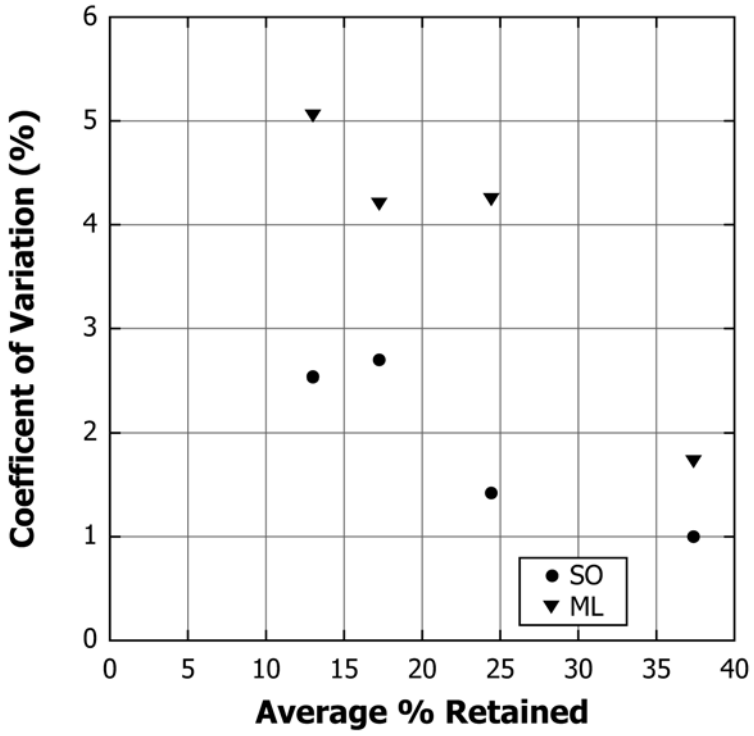


FIG. X1.8 Coefficient of Variation versus Average

X2. ILS FOR TEST METHOD THAT REQUIRES MAKING TEST SPECIMENS

X2.1 *General*—The analysis in Section 10 is based on an ILS that involves testing identical specimens prepared at one location and distributed randomly to the participating laboratories. The analysis is applicable to test methods for testing specimens prepared in accordance with a separate standard practice. As discussed in 10.1, the basic ILS is described as a single-stage nested experiment design. The components of variance are the single-operator variance and the between-laboratory variance. For test methods that require making test specimens as part of the procedure, there is an additional component of variance due to batch-to-batch variability that needs to be investigated. For such test methods, the ILS is treated as a two-stage nested experiment as illustrated in Fig. X2.1. Each laboratory is required to repeat the test with specimens made from two or more separate batches using the same materials. The repeated testing is necessary to analyze the between-batch component of variance associated with producing test specimens as part of the test method. In this two-stage nested design, the second factor "batch" is nested within the primary factor "laboratory." Refer to a textbook on experiment design, such as (**6**), for further information on nested designs. This Appendix describes the analysis of ILS data for a test method that requires making test specimens as part of the procedure. This analysis procedure would be applied to the results for each of the materials in the ILS program.

X2.2 *Notation*—The notation used in this Appendix and in Appendix X3 is defined below:

n_r = number of replicate specimens tested per batch in the ILS,

n_b = number of batches used in the ILS,

p = number of laboratories in the ILS,

a = index for an individual test determination for each batch, 1 to n_r,

b = index for the batch, 1 to n_b,

i = index for the laboratory, 1 to p,

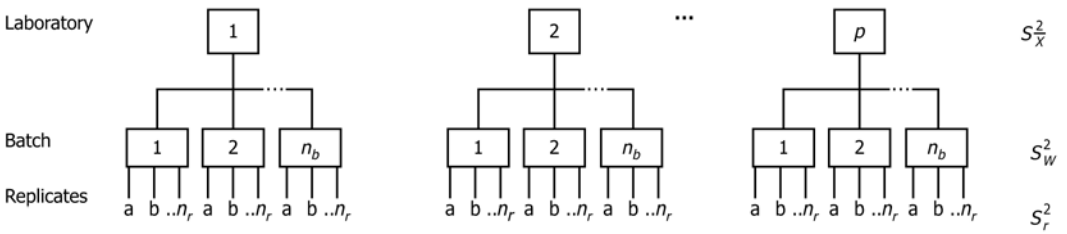
x_{abi} = individual test determination,

$\bar{X}_{bi}$ = average of individual test determinations for batch b by laboratory i,

s_{bi}^2 = variance of individual test determinations for batch b by laboratory i,

$\bar{X}_i$ = average of all individual test determinations by laboratory i,

s_i^2 = variance of the averages from each batch for laboratory i,



$$\text{Between-batch component of variance} = S_b^2 = S_W^2 - \frac{S_r^2}{n_r}$$

$$\text{Between-laboratory component of variance} = S_L^2 = S_{\bar{X}}^2 - \frac{S_W^2}{n_b}$$

FIG. X2.1 Diagram of a Two-Stage Nested Experiment Design

$\bar{\bar{X}}$ = overall average from all laboratories,

$s_{\bar{X}}^2$ = variance of laboratory averages from ILS,

s_r^2 = pooled within-batch (single-operator) variance,

s_b^2 = between-batch component of variance,

s_L^2 = between laboratory component of variance,

s_W^2 = pooled between-batch variance,

s_{WL}^2 = single-operator, multi-batch variance,

s_R^2 = multilaboratory variance,

m_b = number of batches required in the test method,

m_r = number or replicate determinations per batch required in the test method,

MS_e = the mean square of the error,

MS_b = the mean square of between-batch effects, and

MS_L = the mean square of between-laboratory effects.

X2.3 *Single-operator variance*—This is the variance of individual test determinations for specimens made from one batch. It is calculated from the ILS data by pooling the within-batch variances obtained from all the batches and all the laboratories as follows:

$$s_r^2 = \frac{\Sigma s_{bi}^2}{p n_b} \quad \text{(X2.1)}$$

X2.4 *Between-batch component of variance*—The between-batch variance is the variation of the batch averages within a laboratory and it is calculated from the ILS data by pooling the variances of the batch averages from all the laboratories as follows:

$$s_W^2 = \frac{\Sigma s_i^2}{p} \quad \text{(X2.2)}$$

This between-batch variance includes a component due to the single-operator variance and a component, s_b^2, due to batch-to-batch variability. The between-batch component of variance is calculated from the ILS data as follows:

$$s_b^2 = s_W^2 = \frac{s_r^2}{n_r} \quad \text{(X2.3)}$$

X2.5 *Between-laboratory component of variance*—The variance of the laboratory averages, $s_{\bar{X}}^2$, includes a between-batch component and between-laboratory component. The between-laboratory component, s_L^2, is calculated from the ILS data as follows:

$$s_L^2 = s_{\bar{X}}^2 - \frac{s_W^2}{n_r} \quad \text{(X2.4)}$$

X2.6 *Example*—An example is presented to illustrate the determination of the above parameters. The interlaboratory study involved 10 laboratories (p =10). Each laboratory was provided sufficient ingredients to make three batches of each material (n_b = 3). Three replicate test determinations were obtained on specimens made from each batch (n_r = 3). Table X2.1 shows the individual determinations obtained for one material. The calculations are based on the assumption that each laboratory reports the correct number of test determinations per batch and uses the correct number of batches per material. The batch average and the within-batch variance for each batch are shown in Columns 6 and 7, respectively. Column 8 shows the average for each laboratory, and Column 9 shows the variance of the batch averages within each laboratory. The values of the other parameters that have been discussed are shown at the bottom of the table.

X2.7 *Estimates of precision*—The parameters calculated from the ILS data are used to determine the precision indices required to develop the precision statement for the test method. In most test methods, specimens are made from one batch of material. For generality, it will be assumed that the test method requires m_b batches and m_r replicate test determinations per batch. These may or may not be the same values that were used in the ILS.

X2.7.1 *Single-operator precision*—The square root of the pooled within-batch variance, s_r^2, provides an estimate of the single operator standard deviation. This index can be used to establish the maximum expected range of individual determinations among specimens from the same batch as described in Practice C670.

X2.7.2 *Single-operator, multi-batch precision*—The single-operator, multi-batch variance, s_{WL}^2, is calculated from the between-batch component of variance, the single-operator variance, and the number of replicate test determinations per batch required by the test method:

$$s_{WL}^2 = s_b^2 + \frac{s_r^2}{m_r} \quad \text{(X2.5)}$$

The square root of s_{WL}^2 provides the single-operator, multi-batch standard deviation needed to calculate the maximum expected difference between the results (average of m_r replicates) from two batches within a laboratory, or the maximum expected range of the averages from more than two batches. See Practice C670 for sample wording.

X2.7.3 *Multilaboratory precision*—The multilaboratory variance, s_R^2, is calculated from the between-laboratory component of variance, the single-operator, multi-batch variance, and the number of batches required by the test method:

$$s_R^2 = s_L^2 + \frac{s_{WL}^2}{m_b} \quad \text{(X2.6)}$$

This can be used to establish the multilaboratory standard deviation needed to calculate the maximum expected difference between test results obtained by two laboratories. A test result is the average obtained from m_b batches with m_r replicate test determinations for each batch.

TABLE X2.1 Sample Data from an ILS Involving Multiple Batches

Laboratory [1]	Batch [2]	Test Determinations a [3]	b [4]	c [5]	Batch Average $\bar{X}_{bi}$ [6]	Within-batch Variance s_{bi}^2 [7]	Laboratory Average $\bar{X}_i$ [8]	Variance of Batch Averages s_i^2 [9]
1	1	2974	2917	2928	2939.7	914.3	2974.8	4935.7
	2	3076	3040	3051	3055.7	340.3		
	3	2811	2948	3028	2929.0	12043.0		
2	1	3087	3100	3112	3099.7	156.3	3077.0	3229.8
	2	3036	3093	3228	3119.0	9723.0		
	3	2978	3044	3015	3012.3	1094.3		
3	1	2965	2801	2819	2861.7	8089.3	2805.6	2534.7
	2	2751	2765	2856	2790.7	3250.3		
	3	2754	2748	2791	2764.3	542.3		
4	1	2727	2897	2894	2839.3	9466.3	2841.7	410.8
	2	2867	2739	2862	2822.7	5256.3		
	3	2911	2800	2878	2863.0	3249.0		
5	1	2851	2944	2857	2884.0	2709.0	2922.4	46193.9
	2	3104	3142	3216	3154.0	3244.0		
	3	2677	2681	2830	2729.3	7604.3		
6	1	2796	2681	2798	2758.3	4486.3	2774.8	12597.9
	2	2766	2635	2614	2671.7	6784.3		
	3	2997	2918	2768	2894.3	13530.3		
7	1	2779	3039	2957	2925.0	17668.0	3089.8	41923.8
	2	2950	2974	3152	3025.3	12177.3		
	3	3261	3378	3318	3319.0	3423.0		
8	1	3199	3154	3181	3178.0	513.0	3255.3	7772.4
	2	3366	3420	3268	3351.3	5937.3		
	3	3289	3152	3269	3236.7	5476.3		
9	1	3176	3247	3125	3182.7	3754.3	3071.7	23362.1
	2	3098	3156	3151	3135.0	1033.0		
	3	2926	2854	2912	2897.3	1457.3		
10	1	3076	3170	3207	3151.0	4561.0	3128.3	23287.1
	2	2939	2981	2977	2965.7	537.3		
	3	3264	3259	3282	3268.3	146.3		
Sum						149168	29941	166248

Overall average $\bar{\bar{X}} = 29\ 941/10 = 2994$

Pooled single-operator variance, $s_r^2 = 149\ 168/(10 \times 3) = 4972$

Pooled between-batch variance, $s_W^2 = 166\ 248/10 = 16\ 625$

Variance of laboratory averages, $s_{\bar{X}}^2 = 24\ 522$

Between-batch component of variance, $s_b^2 = s_W^2 - \frac{s_r^2}{n_r} = 16\ 625 - 4972/3 = 14\ 968$

Between-laboratory component of variance, $s_L^2 = s_{\bar{X}}^2 - \frac{s_W^2}{n_b} = 24\ 522 - 16\ 625/3 = 18\ 980$

X2.7.4 *Form of precision statement*—The standard deviations obtained for each material would be investigated as discussed in 10.6 to establish whether standard deviation or coefficient of variation is the appropriate measure of precision.

X2.8 *Alternative analysis*—Appendix X3 discusses the use of analysis of variance for determining the components of variance from a two-stage nested design.

X3. USING ANALYSIS OF VARIANCE FOR ILS DATA

X3.1 *General*—The single-operator and the between-laboratory components of variance can be determined from the ILS data using the method known as analysis of variance (ANOVA). This data analysis method is available in statistical analysis computer programs and in spreadsheet programs. This Appendix explains how the replicate test determinations obtained by the different laboratories for each material can be analyzed using a one-way ANOVA for ILS data obtained in a single-stage nested design. For a two-stage nested design, as would be used to investigate the between-batch component of variance, a more general ANOVA technique is applicable and should be undertaken only by experienced analysts.

X3.2 *Data entry*—The data need to be assembled in accordance with the requirements of the computer program being used. In general, it is necessary to group the replicate test determination by laboratory. Table X3.1 is sample test data for Material C taken from Table X1.5. The data are in an acceptable format for a one-way ANOVA using a spreadsheet program. In this example, there are 13 laboratories (p =13) and three replicate results by each laboratory (n = 3).

X3.2.1 As is always the case, the first step is to plot the test data. Fig. X3.1 is a scatter plot of the data grouped by laboratory number. The plot will reveal if there are errors in entering the data into the spreadsheet and will show any

TABLE X3.1 Sample Data from Table X1.5 to Illustrate Use of ANOVA

Laboratory	Data		
	a	b	c
1	26.14	24.65	24.74
2	24.19	23.85	24.18
3	23.34	23.94	23.34
4	23.64	23.50	24.25
5	25.94	25.54	25.64
6	24.10	24.54	23.49
7	24.12	24.16	23.86
8	23.30	23.40	23.40
9	24.31	24.25	24.26
10	27.43	26.82	26.71
11	24.32	24.53	24.58
12	24.56	24.27	24.15
13	23.84	23.67	23.84

extreme test determinations that should be evaluated as potential outliers. The data need to be investigated for data consistency in accordance with 10.4 before completing the ANOVA. This was done in Appendix X1 and is not repeated here.

X3.3 *Results of ANOVA*—In simple terms, ANOVA is used to determine whether different "treatments" have a statistically significant effect on the mean value. In an ANOVA, the variation of the individual test determinations from the overall average is partitioned into different sources of the variation. One source is the "treatments" and the other source is random error. In the case of data from an interlaboratory study, the "treatment" source corresponds to the "between-laboratory" component of variation and the random error corresponds to the single-operator variation.

X3.3.1 Table X3.2 shows the results of a one-way ANOVA using the data in Table X3.1. The term "one-way" is used because the effect of only one factor, in our case "laboratory," is being examined. The column labeled "*SS*" represents the "sum of squares" and column labeled "*MS*" represents the "mean squares." The mean squares are obtained by dividing the sum of squares by the corresponding degrees of freedom, *df*. For this example, the between-laboratory mean square, MS_L, equals 2.982 (rounded value), and the error mean square, MS_e, equals 0.122. The pooled single-operator variance for this material (C) equals the error mean square, that is,

$$s_{rC}^2 = MS_e = 0.122 \tag{X3.1}$$

X3.3.2 The between-laboratory mean square MS_L is related to the between-laboratory component of variance as follows (**4**):

$$MS_L = ns_L^2 + MS_e \tag{X3.2}$$

X3.3.3 Thus the between-laboratory component of variance for Material C is as follows:

$$s_{LC}^2 = \frac{MS_L - MS_e}{n} = \frac{2.982 - 0.122}{3} = 0.953 \tag{X3.3}$$

X3.3.4 The values of these two components of variance are in agreement with values shown at the bottom of Table X1.5. With these two components of variance determined, Eq 9 or Eq 10 can be used to calculate the multilaboratory variance. Note that if MS_L is less than MS_e, the between-laboratory component of variance is set equal to zero.

X3.4 *Missing data*—The use of ANOVA to analyze interlaboratory data permits the calculation of the components of variance if there are missing data or if test determinations are removed because they are determined to be outliers. As an example, Table X3.3 shows the same sample data in Table X3.1, except that there are three missing values.

X3.4.1 The resulting ANOVA table is shown as Table X3.4. The pooled single-operator variance MS_e = 0.045. The between-laboratory component of variance is, however, given by Eq X3.4 and Eq X3.5 (**8**). In Eq X3.5, n_i is the number of replicate test determinations for laboratory *i*.

$$s_{LC}^2 = \frac{MS_L - MS_e}{K} \tag{X3.4}$$

$$K = \frac{\Sigma n_i - \frac{\Sigma n_i^2}{\Sigma n_i}}{p - 1} \tag{X3.5}$$

X3.4.2 With the three missing data, the value of *K* is 2.764. Therefore, the value of the between-laboratory component of variance is given by Eq X3.6:

$$s_{LC}^2 = \frac{MS_L - MS_e}{K} = \frac{2.061 - 0.045}{2.764} = 0.729 \tag{X3.6}$$

X3.4.3 As the number of missing data increases, the above procedure will not provide a correct estimate of the between-laboratory component of variance. In such cases, the method in Ref. (**9**) can be used.

X3.5 *ANOVA for Two-Stage Nested Design*—If the ILS is designed as a two-stage nested experiment to investigate the component of variance due to another factor besides "laboratory," such as the between-batch component of variance, ANOVA can be used to obtain the variance components. In this case, however, care is required to ensure that the analysis is conducted for a "nested" design as opposed to a "crossed" design. A "crossed" design means that the level of each factor appears in combination with the levels of all the other factors. An example of a crossed design is a factorial experiment design, such as the fractional factorial design used in a ruggedness evaluation of a test method as described in Practice C1067.

X3.5.1 For the case of an ILS that requires the *p* laboratories to make n_b batches from the distributed materials, and to perform n_r replicate tests on specimens made from each batch, the degrees of freedom and the variance components for the mean squares are shown in Table X3.5 (**6**).

X3.5.1.1 Thus the components of variance can be calculated from the means squares shown in the table of the ANOVA results using the following relationships:

$$s_L^2 = \frac{MS_L - MS_B}{n_b n_r} \tag{X3.10}$$

$$s_b^2 = \frac{MS_b - MS_e}{n_r} \tag{X3.11}$$

$$s_r^2 = MS_e \tag{X3.12}$$

X3.5.1.2 If the numerators in Eq X3.10 or Eq X3.11 are negative values, the corresponding component of variance is assigned a value of 0.

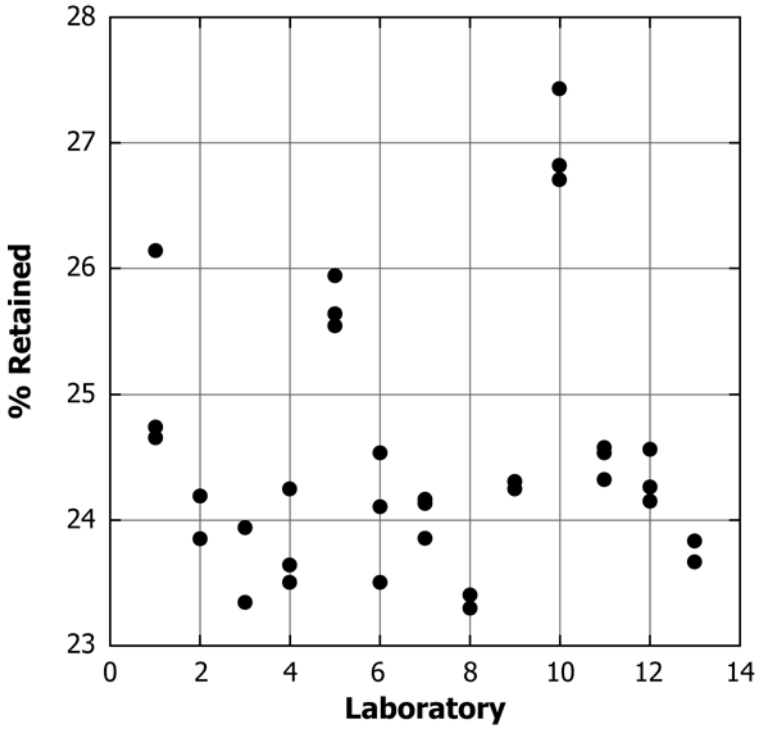


FIG. X3.1 Scatter Plot of Test Determinations for Each Laboratory

TABLE X3.2 ANOVA Table for Data in Table X3.1

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Laboratories	35.78119	12	2.981766	24.37462	4.13E-11	2.147926
Error	3.1806	26	0.122331			
Total	38.96179	38				

TABLE X3.3 Sample Data to Illustrate Use of ANOVA with Missing Data

Laboratory	Data		
	a	b	c
1	–	24.65	24.74
2	24.19	23.85	24.18
3	23.34	23.94	23.34
4	23.64	23.50	24.25
5	25.94	25.54	25.64
6	24.10	24.54	–
7	24.12	24.16	23.86
8	23.30	23.40	23.40
9	24.31	24.25	24.26
10	–	26.82	26.71
11	24.32	24.53	24.58
12	24.56	24.27	24.15
13	23.84	23.67	23.84

X3.5.2 To illustrate the interpretation of an ANOVA table for a two-stage nested experiment design, an ANOVA was carried out using the data in Table X2.1. The analysis was defined with the factor "batch" nested in the factor "laboratory." The results are shown in Table X3.6. If the ANOVA were conducted incorrectly as a "crossed" experiment without identifying the nesting of "batch," the degrees of freedom for the factor "batch" would be 2 rather than 20.

X3.5.2.1 By substituting the mean squares in Table X3.6 into the applicable equation shown as Eq X3.10 to Eq X3.12, the following values are obtained for the variance components:

$$s_L^2 = \frac{MS_L - MS_B}{n_b n_r} = \frac{220\ 700 - 49\ 874}{3 \times 3} = 18\ 981 \quad \text{(X3.13)}$$

$$s_b^2 = \frac{MS_b - MS_e}{n_r} = \frac{49\ 874 - 4972}{3} = 14\ 967 \quad \text{(X3.14)}$$

$$s_r^2 = MS_e = 4972 \quad \text{(X3.15)}$$

X3.5.2.2 These values are identical (accounting for rounding error) to the values at the bottom of Table X2.1.

TABLE X3.4 ANOVA Table for Data in Table X3.3

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Laboratories	24.72898	12	2.060748	45.81653	3.79E-13	2.203607
Error	1.0345	23	0.044978			
Total	25.76348	35				

TABLE X3.5 Degrees of Freedom and Variance Components for Mean Squares in a Two-Stage Nested ILS Design

Source of Variation	*df*	Variance Components	
Laboratory (MS_L)	p - 1	$MS_L = n_b n_r s_L^2 + n_r s_b^2 + s_r^2 = n_b n_r s_L^2 + MS_b$	(X3.7)
Batch (MS_b)	$p\,(n_b - 1)$	$MS_b = n_r s_b^2 + s_r^2 = n_r s_b^2 + MS_e$	(X3.8)
Error (MS_e)	$p\,n_b\,(n_r - 1)$	$MS_e = s_r^2$	(X3.9)

TABLE X3.6 ANOVA Table for Two-Stage Nested Design Using Data from Table X2.1

Source	df	Sums of Squares	Mean Square	F-ratio	Prob
Constant	1	806.835E6	806.835E6	16177	0.0001
Laboratory	9	1.9863E6	220700	4.4251	0.0027
Batch	20	997490	49874.5	10.031	0.0001
Error	60	298335	4972.26		
Total	89	3.28212E6			

REFERENCES

(1) *ASTM Manual for Conducting an Interlaboratory Study of a Test Method, STP. 335*, ASTM, Philadelphia, PA 1963.

(2) Youden, W. J., *Statistical Techniques for Collaborative Tests*, Association of Official Analytical Chemists, Washington, DC, 1967.

(3) *ASTM Manual on Quality Control of Materials, STP 15C*, ASTM, Philadelphia, PA, 1951, pp. 47-49.

(4) Mandel, J.,"Repeatability and Reproducibility," *Materials Research and Standards*, ASTM, Vol. 11, No. 8, August 1971.

(5) Dixon, W. J., and Massey, F. J., *Introduction to Statistical Analysis*, 4th Ed., McGraw-Hill, New York, NY, 1983.

(6) Box, G.E.P., Hunter, J. S., and Hunter, W.G., *Statistics for Experimenters: Design, Discovery, and Innovation*, 2nd Ed., John Wiley & Sons, Hoboken, NJ, 2005.

(7) Snedecor, G. W. and Cochran, W. G., *Statistical Methods*, 8th Edition, Iowa State University Press, 1989.

(8) Mandel, J and Paule, R. C., "Interlaboratory Evaluation of a Material with Unequal Number of Replicates," *Analytical Chemistry*, Vol. 42, No. 11, September, 1970, pp. 1194-1197.

(9) Paule, R. C. and Mandel, J., "Consensus Values and Weighting Factors," *Journal of Research of the National Bureau of Standards*, Vol. 87, No. 5, Set-Oct. 1982, pp. 377-385.

SUMMARY OF CHANGES

Committee C09 has identified the location of selected changes to this practice since the last issue, C802 – 09a, that may impact the use of this practice. (Approved December 15, 2014.)

(1) This practice has undergone an extensive revision.