Designation: ~~C670 – 10~~ C670 – 13

Standard Practice for Preparing Precision and Bias Statements for Test Methods for Construction Materials[1]

This standard is issued under the fixed designation C670; 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.

This standard has been approved for use by agencies of the Department of Defense.

1. Scope*

1.1 The ~~*This practice supplements Practice*~~*Form and Style for* ~~*E177, in*~~*ASTM Standards* ~~order to provide guidance in preparing precision and bias statements for ASTM test methods pertaining to certain construction materials~~ (requires that all test methods contain statements on precision and bias. Further, the precision statement is required to contain a~~Note 1~~). ~~Recommended forms for precision and bias statements are included. A discussion~~ statement on single-operator precision (repeatability) and a statement on multilaboratory precision (reproducibility). This practice provides guidance for preparing precision and bias statements that comply with these requirements. Discussion of the purpose and significance of ~~these~~ precision and bias statements for ~~the~~ users of ~~those~~ test methods is also provided.

~~Note 1—Although under the jurisdiction of Committee C09, this practice was developed jointly by Committees C01, D04, and C09, and has been endorsed by all three committees. It has subsequently been adopted for use by Committee D18.~~Examples of precision statements that conform to this practice are included in Appendix X1. This practice supplements Practice E177 and has been developed to meet the needs of ASTM Committees dealing with construction materials.

Note 1—Although this practice is under the jurisdiction of Committee C09, the current version was developed jointly by Committees C01 and C09 and has subsequently been adopted for use by other committees dealing with construction materials.

1.2 This practice assumes that an interlaboratory study (ILS) has been completed in accordance with Practice C802 or Practice E691. The interlaboratory study provides the necessary statistical values to write the precision and bias statements.

1.3 The system of units for this practice is not specified. Dimensional quantities in the practice are presented only in examples of precision and bias statements.

1.4 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety 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)~~
C802 Practice for Conducting an Interlaboratory Test Program to Determine the Precision of Test Methods for Construction Materials
C1067 Practice for Conducting a Ruggedness Evaluation or Screening Program for Test Methods for Construction Materials
D6607 Practice for Inclusion of Precision Statement Variation in Specification Limits
E177 Practice for Use of the Terms Precision and Bias in ASTM Test Methods
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 of Terms Specific to This Standard:*~~

[1] This practice is under the jurisdiction of ASTM Committee C09 on Concrete and Concrete Aggregates and is the direct responsibility of Subcommittee C09.94 on Evaluation of Data (Joint C09 and C01).

Current edition approved ~~Oct. 1, 2010~~July 1, 2013. Published ~~November 2010~~July 2013. Originally approved in 1971. Last previous edition approved in ~~2003~~2010 as C670-~~03.~~10. DOI: ~~10.1520/C0670-10.~~10.1520/C0670-13.

[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**

3.1 *one-sigma limit (1s)—Definitions:* the fundamental statistic underlying all indexes of precision is the standard deviation of the population of measurements characteristic of the test method when the latter is applied under specifically prescribed conditions (a given system of causes). The terminology "one-sigma limit" (abbreviated (1s)) is used in Practice E177 to denote the estimate of the standard deviation or sigma that is characteristic of the total statistical population. The one-sigma limit is an indication of the variability (as measured by the deviations above and below the average) of a large group of individual test results obtained under similar conditions.

3.1.1 *single-operator one-sigma limit*—the one-sigma limit for single-operator precision is a quantitative estimate of the variability of a large group of individual test results when the tests have been made on the same material by a single operator using the same apparatus in the same laboratory over a relatively short period of time. This statistic is the basic one used to calculateFor definitions of general statistical terms, refer to Terminology E456the single-operator index of precision given in the precision statement for guidance of the operator..

3.2.2 *multilaboratory one-sigma limit*—the one-sigma limit for multilaboratory precision is a quantitative estimate of the variability of a large group of individual test results when each test has been made in a different laboratory and every effort has been made to make the test portions of the material as nearly identical as possible. Under normal circumstances the estimates of one-sigma limit for multilaboratory precision are larger than those for single-operator precision, because different operators and different apparatus are being used in different laboratories for which the environment may be different.

3.2.3 *one-sigma limit in percent (1s%)*—in some cases the coefficient of variation is used in place of the standard deviation as the fundamental statistic. This statistic is termed the "one-sigma limit in percent" (abbreviated (1s%)) and is the appropriate standard deviation (1s) divided by the average of the measurements and expressed as a percent. When it is appropriate to use (1s%) in place of (1s) is discussed in Section 6.

3.3 *Acceptable Range of Results:*

3.3.1 *acceptable difference between two results*—the "difference two-sigma limit (d2s)" or "difference two-sigma limit in percent (d2s%)," as defined in Practice E177, has been selected as the appropriate index of precision in most precision statements. These indexes indicate a maximum acceptable difference between two results obtained on test portions of the same material under the applicable system of causes described in 4.1.1 and 4.1.2 (or whatever other system of causes is appropriate). The (d2s) index is the difference between two individual test results that would be equaled or exceeded in the long run in only 1 case in 20 in the normal and correct operation of the method. The (d2s%) index is the difference between two individual test results expressed as a percent of their average that meets the same requirements. These indexes are calculated by multiplying the appropriate standard deviation (1s) or coefficient of variation (1s%) by the factor $2\sqrt{2}$ (equal to 2.83).

3.3.2 *acceptable range of more than two results*—in cases where the test method calls for more than two test results to be obtained, the range (difference between highest and lowest) of the group of test results must be compared to a maximum acceptable range for the applicable system of causes and number of test results. The range for different numbers of test results including two that would be equaled or exceeded in only 1 case in 20 is obtained by multiplying the appropriate standard deviation (1s) or coefficient of variation (1s%) by the appropriate factor from the second column of Table 1 (Note 2).

NOTE 2—It is important to note that when more than two test results are obtained, an index of precision for the difference between two results can not be used as a criterion for judging acceptability of the range of the group or for other pairs of results selected from the group.

3.3.3 *variations for single operators*—the system of causes designated for obtaining the quantitative guide to acceptable performance by an operator as stated in 4.1.1 leads to single-operator precision, using the system of modifiers given in Practice

TABLE 1 Maximum Acceptable Range of Test Results[A]

Number of Test Results	Multiplier of (1s)Standard Deviation or (1s%)Coefficient forof Maximum Acceptable Range Varitation[AB]
2	2.8
3	3.3
4	3.6
5	3.9
6	4.0
7	4.2
8	4.3
9	4.4
10	4.5

[A] A test result can be a single determination or the average of two or more determinations as defined in the test method.

[B] Values were obtained from Table A7 of "Order Statistics and Their Use in Testing and Estimation," Vol 1, by Leon Harter, Aerospace Research Laboratories, United States Air Force.

~~E177 (Note 3). When two results by the same operator differ by more than (d2s) or (d2s%) or the range of more than two results exceeds that obtained by the method described in 3.3.2 there is a significantly large probability that an error has occurred and retests should be made as directed in Note 4.~~

~~NOTE 3—Single-operator precision is often referred to as "repeatability," and multilaboratory precision is often referred to as "reproducibility."~~

~~NOTE 4—It is beyond the scope of this practice to describe in detail what action should be taken in all cases when results occur that differ by more than the (d2s) limits or by more than the maximum allowable range. Such an occurrence is a warning that there may have been some error in the test procedure, or some departure from the prescribed conditions of the test on which the limits appearing in the test method are based; for example, faulty or misadjusted apparatus, improper conditions in the laboratory, etc. In judging whether or not results are in error, information other than the difference between two test results is needed. Often a review of the circumstances under which the test results in question were obtained will reveal some reason for a departure. In this case the data should be discarded and new test results obtained and evaluated separately. If no physical reason for a departure is found, retests should still be made, but the original tests should not be completely ignored. If the second set of results also differs by more than the applicable limit, the evidence is very strong that something is wrong or that a real difference exists between the two samples tested. If the second set produces a result within the limit, it may be taken as a valid test, but the operator or laboratory may then be suspected of producing erratic results, and a closer examination of the procedures would be in order. If knowledge about the test method in question indicates that certain actions may be appropriate in cases where deviant results occur, then such information should be included in the test method, but details of how this should be done will depend upon the particular test method.~~

~~3.3.4 *variations between laboratories*—the system of causes designated for obtaining the quantitative guide for acceptance of results by different laboratories as given in 4.1.2 is multilaboratory precision, using the system of modifiers given in Practice E177 (Note 3). When results differ by more than (d2s) there is a significantly large probability that one or both laboratories are in error or that a difference exists in the portions of material being used for the tests. In such cases, retests should be made. When possible, newly drawn test samples should be used for such retests as directed in Note 4.~~

3.2 *Definitions of Terms Specific to This Standard:*[3]

3.2.1 *test determination, n*—the value of a characteristic of a single test specimen obtained by a specified test method.

[3] Terms are listed in order of hierarchy beginning with the basic concept.

3.2.1.1 *Discussion*—

The term "replicate" is often used for a test determination.

3.2.2 *test result, n*—the value of a characteristic of a material obtained by carrying out a specified test method.

3.2.2.1 *Discussion*—

A test result may be a single test determination or the average of a specified number of test determinations, or replicates (see 4.1 for additional discussion).

3.2.3 *identical test specimens, n*—test specimens selected at random and made from a single quantity or batch of material that is as homogeneous as possible.

3.2.3.1 *Discussion*—

In interlaboratory studies of test methods for fresh cementitious mixtures, a practicable approach for obtaining identical tests specimens is to assemble technicians from different laboratories at one location and test specimens are made from the same batch of the fresh mixture. For interlaboratory studies of nondestructive test methods, the same test specimens can be circulated among participating laboratories, provided the characteristic of interest does not change during the time to complete the study.

3.2.4 *single-operator standard deviation, s_r, (or coefficient of variation, CV_r), n*—the standard deviation (or coefficient of variation) of test determinations obtained on identical test specimens by a single operator using the same apparatus in the same laboratory over a relatively short period of time.

3.2.4.1 *Discussion*—

The single-operator standard deviation, or coefficient of variation, is the fundamental statistic underlying the single-operator indexes of precision. The single-operator standard deviation, or coefficient of variation, is an indication of the variability of a large group of test determinations by the same operator on the same material. This value is obtained from an interlaboratory study and is equal to the pooled standard deviation of test determinations obtained by the operators. The coefficient of variation (ratio of standard deviation to the average expressed as a percentage) is used if the standard deviation is proportional to the level of the characteristic being measured. The single-operator standard deviation, usually considered a property of the test method, will generally be lower than the multilaboratory standard deviation. In Practice E177, the single-operator standard deviation is referred to as the *repeatability standard deviation*, and the subscript *r* is used. In previous versions of Practice C670, the terms *one-sigma*

limit (1s) or *one sigma limit in percent (1s%)* were used for the single-operator standard deviation or single-operator coefficient of variation, respectively. In some publications, the term *within-test standard deviation* (or *coefficient of variation*) has been used. The term *within-laboratory standard deviation* (or coefficient of variation) should not be used for this statistic (see 4.2.3).

3.2.5 *multilaboratory standard deviation, s_R (or coefficient of variation, CV_R), n*—the standard deviation or coefficient of variation of test results obtained with the same test method on identical test specimens in different laboratories with different operators using different equipment.

3.2.5.1 *Discussion*—

The multilaboratory standard deviation, or coefficient of variation, is the fundamental statistic underlying the indexes of precision under multilaboratory conditions. The multilaboratory standard deviation is an indication of the variability of a group of test results obtained by different laboratories for identical test specimens. The multilaboratory standard deviation (or coefficient of variation) is usually greater than the single-operator standard deviation (or coefficient of variation), because different operators and different apparatus have been used in different laboratories for which the environments may have differed. In Practice E177, the multilaboratory standard deviation is referred to as the *reproducibility standard deviation* and the subscript *R* is used.

3.2.6 *difference limit (d2s or d2s%), n*—the difference between two test results that is expected to be exceeded with a probability of about 5 % in the normal and correct operation of the test method; used as an index of precision of the test method.

3.2.6.1 *Discussion*—

The difference limit has been selected as the appropriate index of precision in most precision statements. A difference limit (d2s) indicates the maximum acceptable difference between two results obtained on identical test specimens (see 3.2.3.1) under the applicable system of causes (single-operator or multilaboratory conditions). The (d2s%) limit is the maximum acceptable difference between two test results expressed as a percentage of their average. These difference limits are calculated by multiplying the appropriate standard deviation (s_r or s_R) or coefficient of variation (CV_r or CV_R) by the factor 1.96 $\sqrt{2}$, which for the purpose of this Practice is taken to be equal to 2.8. In Practice E177, the terms *repeatability limit* and *reproducibility limit* are used for these difference limits under single-operator and multilaboratory conditions, respectively.

3.2.7 *acceptable range, n*—the difference between the largest and smallest of three or more test determinations or test results that is expected to be exceeded with a probability of about 5 % in the normal and correct operation of the test method; used as an index of precision of the test method, if applicable.

3.2.7.1 *Discussion*—

This index is usually reported in precision statements of test methods that define a test result as the average of three or more determinations. Otherwise, the difference limit (d2s or d2s%) is used. See 4.3 for additional discussion on how to determine this index.

~~3.4 *Number of Tests:*~~

~~3.4.1 *single test results*—the number of tests run must be taken into account when evaluating testing variations. Usually, the statistics used in evaluating precision and the indexes of precision based on them are based on the population distribution of single test results. When this is the case, the index of precision may be used in comparing single tests results only, not averages of two or more tests.~~

~~3.4.2 *test results based on averages*—if the precision statement is based on test results that are averages of two or more measurements, then the number of measurements averaged must be stated, and in using the index of precision, averages of exactly that number of measurements must be used. In some cases a test result is defined in the method as the average of two or more individual measurements. In such cases the index of precision for a test result applies to a test result as so defined, although indexes of precision for ranges of individual measurements within a laboratory may also be included as described in 3.3.3.~~

~~3.4.3 *precision of individual measurements averaged to obtain a test result*—when two or more measurements are averaged to obtain a test result, the range of the individual measurements may be examined to determine whether the latter meet the criterion of being valid individual measurements under the conditions of the test method. The maximum acceptable range for individual measurements is obtained by multiplying the appropriate standard deviation (1s) or, coefficient of variation (1s%) obtained from averages by the appropriate factor from the second column of Table 2 (Note 5). The maximum acceptable range for individual measurements obtained by this method may be included in the precision statement as an index of precision for individual measurements in the same laboratory as described in Example 8.~~

~~NOTE 5—This procedure is only valid if the individual measurements are subject to the same sources of variation as the test result. For example, the single-operator precision of Test Method C109/C109M mortar cubes is calculated from test results that include a contribution from variation among batches of mortar. Variation among individual cubes from a single batch does not contain this component of variation. Therefore, differences among~~

individual cubes from a single batch cannot be inferred from the single-operator standard deviation given in Test Method C109/C109M and the values in Table 2.

3.4.4 *multilaboratory precision expressed as a maximum allowable difference between two averages*—when the test method calls for the reporting of more than one test result, multi-laboratory precision may be expressed as a maximum allowable difference between averages of such groups, one from each laboratory, and both the (d2s) or (d2s%) limit for individual results and this maximum allowable difference of two averages may be included in the multilaboratory precision statement (Note 6). The maximum allowable difference for averages of a given number of test results, *n*, is obtained by dividing the appropriate (d2s) or (d2s%) limit by the square root of *n*.

NOTE 6—Note that this is not the same as the situation where a test result is defined as the average of two or more individual measurements. A given test method may include both features. It is important to bear in mind, however, that when more than one result is obtained in one or both laboratories, the (d2s) or (d2s%) limit may not be used as a criterion for judging the differences between selected pairs of results from the two laboratories.

3.5 *field versus laboratory tests*—precision indexes for ASTM test methods are normally based on results obtained in laboratories by competent operators using well-controlled equipment on test portions of materials for which precautions have been taken to ensure that they are as nearly alike as possible. Such precautions and the same level of competence may not be practicable for the usual quality control or routine acceptance testing. Therefore, the normal testing variation among laboratories engaged in quality control and acceptance testing of commercial materials may be larger than indicated by the relationship derived from the one-sigma limit for multilaboratory precision. In this case it is recommended that studies be made to determine the one-sigma limit for tests made under field conditions and realistic adjustments in specification tolerances be made accordingly.

4. General Concepts

4.1 *Test Result*—The result of a test method may be a single test determination or the average of two or more test determinations (or replicates). The precision statement of a test method applies to a test result as defined in the test method and should state clearly this fact.

4.1.1 *Number of Test Determinations*—The number of test determinations required to obtain a test result by a test method must be taken into account when evaluating testing variations. The statistic used in evaluating single-operator precision is based usually on the standard deviation (or coefficient of variation) of single test determinations. The single-operator standard deviation (or coefficient of variation) may be used in evaluating the acceptable range of test determinations.

4.1.2 *Test Result Based on Averages of Determinations*—For test methods that define a test result as the average of two or more test determinations (or replicates), the fundamental statistic is still the standard deviation (or coefficient of variation) of single test determinations. The report of the analysis of the interlaboratory study (see 5.2) must include this statistic. The single-operator standard deviation of test determinations can be used to calculate the standard deviation of a test result that is the average of multiple determinations and thereby define the maximum acceptable difference between two test results obtained by the same operator on identical test specimens. The precision statement may also include the maximum acceptable range of individual determinations that comprise the test result (see 4.3).

4.1.3 *Standard Deviation of an Average*—The standard deviation of the average of *n* test determinations obtained from identical specimens taken from the same population is equal to the standard deviation of the individual determinations divided by the square root of *n*. This relationship is valid, however, only if the determinations are obtained using identical specimens. It is not applicable to averages obtained on specimens made from different batches of cementitious mixtures as discussed in 4.2.3.

4.2 *Types of Precision*—A precision statement meeting the requirements of this practice normally contains two main elements described elements: (*1* as follows:) single-operator precision, and (*2*) multilaboratory precision. For test methods that require test results on specimens made from more than one batch, the single-operator, multi-batch precision is also included.

4.2.1 *Single-Operator Precision*—A measure of the The pooled, single-operator standard deviation (or coefficient of variation) of test determinations obtained from the interlaboratory study is the underlying statistic of the test method. This is used to calculate the greatest difference between two results or more determinations that would be considered acceptable when properly conducted repetitive determinations are made on the same material by a competent operator. As discussed in 4.1.2, the single-operator standard deviation (or coefficient of variation) of test determinations is also used to calculate the greatest acceptable difference between test results defined as the average of two or more determinations. The single-operator precision provides a quantitative guide to acceptable performance by an operator. If two determinations or test results by the same operator differ by more than the difference limit, (d2s) or (d2s%), or if the range of more than two determinations or test results exceeds the values defined in 4.3, there is a high probability that an error has occurred and retests should be made.

NOTE 2—It is beyond the scope of this practice to describe in detail what action should be taken in all cases if two test results differ by more than the (d2s) or (d2s%) limits or the range of more than two determinations exceeds the maximum expected range. Such an occurrence is a warning that there may have been some error in the test procedure, or some departure from the prescribed conditions of the test on which the limits appearing in the test method are based; for example, faulty or misadjusted apparatus or improper conditions in the laboratory. In judging whether or not results are in error, information other than the difference between two test results is needed. Often a review of the circumstances under which the test results in question were obtained will reveal some reason for a departure. In this case, the data should be discarded and new test results obtained and evaluated separately. If no physical reason for a departure is found, retests should still be made, but the original tests should not be ignored. If the second set of results also differs by more than the applicable limit, the evidence is very strong that something is wrong or that a real difference exists between the specimens tested. If the second set produces a result within the limit, it may be taken as a valid test, but the operator or laboratory may then be suspected of producing erratic

results, and a closer examination of the procedures would be in order. If knowledge about the test method in question indicates that certain actions may be appropriate in cases where deviant results occur, then such information should be included in the test method, but details of how this should be done will depend upon the particular test method.

4.2.2 *Multilaboratory Precision*—~~A~~ The multilaboratory standard deviation (or coefficient of variation) obtained from the interlaboratory study provides a measure of the greatest difference between two test ~~results~~determinations that would be considered acceptable when properly conducted ~~determinations~~tests are made by two different operators in different laboratories on portions of a material that are intended to be identical, or as nearly identical as possible. If results differ by more than the difference limit (d2s) or (d2s%), there is a high probability that one or both laboratories are in error or that a difference exists in the characteristics of the test specimens used for the tests. In such cases, retests should be made. If possible, newly drawn test specimens should be used for such retests.

4.2.2.1 If the test method calls for reporting the average of more than one test determination, multilaboratory precision is expressed as a maximum allowable difference between averages of such groups obtained by two laboratories (Note 3). In this case, the multilaboratory standard deviation derived from the interlaboratory study is based on the number of replications required to obtain a test result as defined by the test method.

NOTE 3—Example 5 in Appendix X1 shows an example of this situation. If a test result is based on tests of specimens made from different batches of the cementitious mixture, the consideration in 4.2.3 apply, and Example 6 provides an example of this situation.

4.2.3 *Single-Operator, Multi-Batch Precision*—Some test methods require reporting the averages of two or more determinations obtained on specimens from two or more batches made using the same materials. The single-operator, multi-batch standard deviation is a measure of the variation of the averages among the batches. This standard deviation will usually be greater than the value obtained by dividing the single-operator standard deviation by the square root of the number of determinations used to obtain the average test result for each batch. This is because the single-operator, multi-batch standard deviation includes the batch-to-batch variability. The precision statement for this type of test method will include three indexes of precision: (*1*) the single-operator precision, (*2*) the single-operator, multi-batch precision, and (*3*) the multilaboratory precision. In some test methods, the term *within-laboratory precision* has been used. The preferred term, however, is *single-operator, multi-batch precision* because this is more descriptive of the conditions under which results are obtained. The single-operator, multi-batch precision statement would indicate the acceptable range (or difference limit, if only two batches are involved) among batch averages. The advice of a statistical consultant should be sought in planning the interlaboratory study for this type of test method so that the necessary statistics can be determined.

4.2.4 *Other Measures of Precision*—The elements described in 4.2.1, 4.2.2, and 4.2.3 involve the main systems of causes that are of interest to users of test methods involving construction materials. In cases where other systems of causes apply (for example, *single-operator-apparatus*,*multi-day precision*; or *multi-operator, single-day-apparatus precision*), the appropriate statistics for those systems of causes need to be developed and the appropriate combination of modifiers given in Practice E177 should be used to describe those statistics. These should not, however, be taken as the fundamental precision parameters for the test method. The advice of a statistical consultant should be sought in planning the interlaboratory study so that the correct statistics can be determined.

~~4.2 *Other Measures of Precision*—The two elements described in 4.1.1 and 4.1.2 involve the main systems of causes of interest to users of test methods involving construction materials. In cases where other systems of causes apply, the appropriate statistics for those systems should be used and the appropriate combination of modifiers given in Practice E177 should be used to describe those statistics.~~

4.3 ~~*Use of Indexes of Precision in Specifications*—~~*Acceptable Range Among Results*—~~The indexes of precision described in this practice are to be used as guides to determine (with a prescribed degree of certainty) whether a given series of results can be considered as valid tests under the conditions assumed in the test method. Comparisons of test results with specification limits should be made only after~~If the test method requires more than two test results, as so defined in the method, the difference between highest and lowest test results in the group must be compared to the maximum acceptable range for the applicable system of causes. The range among different numbers of test results in the group, including two, that would be expected to be exceeded with no more than about 5 % probability is obtained by multiplying the appropriate standard deviation or coefficient of variation by the corresponding factor from the second column of Table 1~~there is reasonable assurance that the determinations are adequate. Writers of specifications have the responsibility of recognizing the variability of results characteristic of a given test method in setting specification limits, but indexes of precision of the test method should never be added to specification limits by the users of those specifications for the purpose of judging acceptance or rejection of materials.~~. If more than two test results are obtained, the index of precision for the difference between two results cannot be used as a criterion for judging acceptability of the differences between pairs of results selected from the group.

4.4 ~~*Use*~~*Uses of Indexes of Precision ~~for Qualifying an Operator~~*—~~Indexes of single-operator precision are sometimes used as a basis for qualifying an operator. The assumption is that results that do not differ by more than the stated index are indicative of proper performance of the test. However, this assumption is not necessarily correct. Uniform misunderstanding of instructions or maladjustments of equipment may produce consistent but erroneous test results. Thus, tests conducted for the purpose of qualifying an operator should be made on materials for which the measured characteristic is known, whenever possible, so that accuracy as well as precision can be evaluated. (See Practice E177 for a discussion of the terms precision and accuracy.)~~