This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.

Designation: E3264 – 21

# Standard Guide for Homogeneity of Samples and Reference Materials Used for Inter- and Intra-Laboratory Studies[1]

This standard is issued under the fixed designation E3264; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

## 1. Scope

1.1 This guide presents techniques and guidance for evaluating and assuring homogeneity of individual samples and bulk materials used for interlaboratory and intra-laboratory studies.

1.2 This guide is applicable to samples and reference materials used for proficiency testing programs and for interlaboratory studies to determine precision estimates for test methods. It may also be useful for activities related to quality control of testing within a single laboratory.

1.3 Five techniques are presented for assessing sample homogeneity. The five techniques are not an exhaustive list of available techniques for assessing homogeneity of samples, but the techniques were chosen to cover a range of circumstances (and various degrees of rigor required) for laboratory studies of various types and purposes.

1.4 Each of the first four techniques provides a scheme for testing for homogeneity and a statistical procedure for evaluating the results of the homogeneity testing. The circumstances are described for which each of the techniques is suited.

1.5 For circumstances when homogeneity testing is not possible, the fifth technique provides guidance for producing homogeneous samples.

1.6 The appendixes of this guide provide example spreadsheets for Techniques 1, 2, 3, and 4.

1.7 This guide is not intended for evaluation of certified reference materials (CRMs) or materials used for calibration.

1.8 *Units*—The system of units for this standard is not specified. Dimensional quantities in the standard are presented only as illustrations of calculation methods. The examples are not binding on products or test methods treated.

1.9 *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, health, and environmental practices and determine the applicability of regulatory limitations prior to use.*

1.10 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

## 2. Referenced Documents

2.1 *ASTM Standards:*[2]

C33/C33M Specification for Concrete Aggregates
C1128 Guide for Preparation of Working Reference Materials for Use in Analysis of Nuclear Fuel Cycle Materials
D5956 Guide for Sampling Strategies for Heterogeneous Wastes
D6299 Practice for Applying Statistical Quality Assurance and Control Charting Techniques to Evaluate Analytical Measurement System Performance
D7915 Practice for Application of Generalized Extreme Studentized Deviate (GESD) Technique to Simultaneously Identify Multiple Outliers in a Data Set
E105 Guide for Probability Sampling of Materials
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
E1402 Guide for Sampling Design
E2282 Guide for Defining the Test Result of a Test Method

2.2 *ISO Standards:*[3]

ISO Guide 30 Reference Materials – Selected Terms and Definitions
ISO 13528 Statistical Methods for Use in Proficiency Testing by Interlaboratory Comparison
ISO 3534 Statistics Vocabulary and Symbols – Part 2 Applied Statistics

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

Current edition approved April 1, 2021. Published January 2022. DOI: 10.1520/E3264-21.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[3] Available from International Organization for Standardization (ISO), ISO Central Secretariat, Chemin de Blandonnet 8, CP 401, 1214 Vernier, Geneva, Switzerland, https://www.iso.org.

Copyright © ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959. United States



Documents
Document Sharing Hub
Full document details: ASTM E3264-21

## 3. Terminology

3.1 *Definitions*—Unless otherwise noted in this guide, all terms relating to quality and statistics are defined in E456.

3.1.1 *heterogeneity, n*—the condition of a population under which items of the population are not of uniform structure or composition with respect to the characteristic of interest.

3.1.1.1 *Discussion*—The variation in test results for a property of interest can be used to measure the degree of heterogeneity of a reference material with respect to the specified property.

3.1.2 *homogeneity, n*—condition of being of uniform structure or composition with respect to one or more specified properties. **ISO Guide 30, C1128**

3.1.2.1 *Discussion*—A reference material is said to be homogenous with respect to a specified property if the variation of test results for that property obtained by a specified test method is found to lie within a specified acceptable limit.

3.1.3 *representative sample, n*—a sample collected in such a manner that it reflects one or more characteristics of interest of the population from which it was collected. **D5956**

3.1.3.1 *Discussion*—The term *representative* is often used with other terms to describe when the object of discussion (such as a sample, a group of samples, a test specimen, a portion of a bulk material, a subset of data, or a test unit) reflects characteristics of the population as a whole.

3.1.4 *test result, n*—the value of a characteristic obtained by carrying out a specified test method. **ISO 3534–2, E2282**

3.1.4.1 *Discussion*—The test method specifies that one or a number of individual observations be made, and their average or another appropriate function, (such as the median or the standard deviation), be reported as the test result. It can also require standard corrections to be applied, such as correction of gas volumes to standard temperature and pressure. Thus, a test result can be a result calculated from several observed values. In the simple case, the test result is the observed value itself.

3.1.5 *test specimen, n*—the portion of a test unit needed to obtain a single test determination. **E2282**

3.1.5.1 *Discussion*—When used for a physical test, this is sometimes called "test piece." For a chemical test, it is sometimes called test portion or test sample. For optical and other tests, it is also sometimes called test sample. In interlaboratory evaluation of test methods and other statistical procedures, it is best to reserve the word sample for the whole amount of material involved and not the individual test specimens, pieces or portions being tested.

3.1.6 *test unit, n*—the total quantity of material (containing one or more test specimens) needed to obtain a test result as specified in the test method. (See *test result*.)

3.1.6.1 *Discussion*—In this guide, the term *sample* is used interchangeably with *test unit* and is defined in the context of this guide as a quantity of material, an item, or an artifact provided to a laboratory to be tested as a whole or to be subdivided into multiple test specimens by sub-sampling in a manner consistent with industry protocols most suitable for that type of material, item or artifact. **E2282**

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *between–sample homogeneity, n*—the degree to which the structure or composition with respect to one or more specified properties of each sample is similar to every other sample in a set of samples.

3.2.1.1 *Discussion*—Cans of paint obtained from the same manufacturer and the same batch would be expected to have good between-sample homogeneity and test results on multiple cans would be expected to be similar. The between-sample homogeneity of cans of paint obtained from different batches or from different manufacturers would not be expected to be as good since the ingredients and their proportions may not be the same and the test results from multiple cans would exhibit greater variation.

3.2.2 *between-sample variance,* $s_s^2$, *n*—estimate of the variance of test results due to differences between samples.

3.2.2.1 *Discussion*—The between-sample variance, $s_s^2$, is determined using the variance of the cell averages, $s_{\bar{X}_i}^2$, the within-sample variance, $s_w^2$, and the number of replicates, $k$, as follows:

$$s_s^2 = s_{\bar{X}_i}^2 - \left(s_w^2 / k\right) \tag{1}$$

3.2.3 *cell, i, n*—the set of results obtained from testing performed on a single sample.

3.2.4 *cell average,* $\bar{X}_i$, *n*—the average of the results of replicate tests performed on a single sample.

3.2.5 *cell difference,* $D_i$, *n*—the difference between duplicate results in a single cell.

3.2.6 *cell standard deviation,* $s_i$, *n*—the square root of the cell variance.

3.2.7 *cell variance,* $s_i^2$, *n*—calculated statistical variance of the replicate test results contained in a single cell.

3.2.8 *replicate test, n*—the process of repeating a test on the same specimen or on sub-samples obtained from a single sample.

3.2.9 *sub-sample, n*—a quantity of material to be used as a test specimen, obtained from a larger sample in a manner consistent with the recommended industry protocols for the material to be tested. (See *test specimen*.)

3.2.10 *variance of the cell averages,* $s_{\bar{X}_i}^2$, *n*—calculated statistical variance of the cell averages for the set of samples selected for homogeneity analysis.

3.2.11 *within-sample homogeneity, n*—the degree to which the structure or composition with respect to one or more specified properties of a single sample is similar throughout.

3.2.11.1 *Discussion*—For example, if a can of paint is allowed to sit for a long period of time, the components of the paint tend to separate into layers. A test specimen from the top of the can would not be expected to have the same test properties as a specimen taken from the bottom of the can and the within-sample homogeneity would be poor. If the paint in the can is stirred before sampling, the within-sample homogeneity would be improved and test specimens from the top and bottom of the can would be expected to have similar test results.

AVA Documents
Document Sharing Hub



Full document details: ASTM E3264-21

3.2.12 *within-sample variance,* $s_w^2$, *n*—average of the cell variances for the set of samples selected for homogeneity analysis.

3.2.12.1 *Discussion*—Assuming that the replicate test specimens obtained from each sample are identical, the within-sample variance is an estimate of the repeatability precision of the test method.

## 4. Summary of Guide

4.1 The guide presents techniques and provides recommendations for evaluating and assuring homogeneity of individual samples or bulk materials and can be used for either interlaboratory or intra-laboratory studies.

4.2 Five techniques are presented covering a range of circumstances encountered with laboratory studies of various types and purposes. Each technique describes the circumstances for which the technique is suited. See Appendix X1, Table X1.1 for an overview of the guide in tabular form.

4.3 Generally speaking, the basic steps for ensuring homogeneous samples are *(1)* use a process that will produce samples that are as homogeneous as possible, *(2)* test a subset of samples (or a portion of the reference material), *(3)* screen results for outliers, and *(4)* estimate the sample variation, and *(5)* decide whether the samples are sufficiently homogeneous for the intended purpose.

4.4 To accomplish the first step, recommended practices for preparing homogeneous samples and bulk materials for all types of laboratory studies are presented in Technique 5. Those practices are also recommended for the preparation of samples used with Techniques 1, 2, 3, and 4.

4.5 For accomplishing the remaining steps, Techniques 1, 2, 3, and 4 are each designed to provide guidance for homogeneity testing and evaluation appropriate to the circumstances described for that technique.

4.6 The practices recommended in Technique 5 for production of individual samples (or units) distributed for use in a laboratory study are also generally applicable to bulk quantities of material. However, for simplicity, Techniques 1, 2, 3, and 4 in this guide do not specifically state how the techniques are to be applied to bulk materials. Since evaluation of the homogeneity of a bulk quantity of material requires testing of multiple representative samples obtained from the bulk quantity, evaluation of the homogeneity of a bulk material is accomplished by applying Techniques 1, 2, 3 or 4 to multiple representative samples obtained from the bulk quantity of material. The results of homogeneity testing of the multiple representative samples can then be used as an indication of the degree of homogeneity of the bulk material.

## 5. Significance and Use

5.1 This guide presents techniques and guidance for evaluating and assuring homogeneity of individual samples or bulk materials and can be used for either interlaboratory or intra-laboratory studies. The types of studies include, but are not limited to, studies to determine precision estimates for test methods, proficiency testing programs, and studies related to quality control of testing within a single laboratory.

5.2 Because the test results of any laboratory study are affected by the quality of the samples tested, producing homogeneous samples and determining the degree of homogeneity is important for interpreting the results of the study.

5.3 Five techniques are presented in this guide to evaluate sample homogeneity for a range of circumstances and degrees of rigor. The circumstances under which the studies are conducted and the degree of rigor required may differ. The user should consider the circumstances listed in each technique to determine which is appropriate for the study at hand.

5.4 Each of the Techniques 1, 2, and 3 provides a procedure for testing and evaluating sample homogeneity when replicate testing of the samples is possible. Technique 4 provides a plan to evaluate sample homogeneity when replicate testing is not possible. Technique 5 recommends practices for producing homogeneous samples for circumstances when homogeneity testing is not possible.

5.5 When the conditions of adequate within-sample homogeneity and between-sample homogeneity are satisfied, any differences in test results on multiple samples can reasonably be attributed to testing variation and not due to sample variation.

5.6 When differences within or between samples are discovered and the samples are deemed insufficiently homogeneous, the sample preparation process can be improved or corrected and a new set of samples can be prepared. Or, in cases where the sample homogeneity cannot be improved or for other reasons when the samples must be used, the method of evaluation for the laboratory study should account for the effect of differences between samples.

5.7 When used in conjunction with studies to develop precision estimates, the guidance in this standard can be used to help quantify sources of test variation (such as effects due to sampling, test method repeatability, and the degree of inhomogeneity) and, therefore, can be useful for determining and stating the conditions under which the precision estimates are valid.

5.8 For proficiency testing programs, the guidance in this standard can provide information to prevent laboratories from being unfairly penalized for testing variation due to inherent differences between samples.

5.9 In a single laboratory, the guidance in this standard could be used to evaluate the homogeneity of samples for studies to measure test variation over time or for studies to compare the results of tests performed by different technicians.

5.10 To minimize the resources required for homogeneity testing, a testing design using a minimum of ten samples with two replicate tests performed on each sample is recommended in Techniques 1, 2, and 3 of this guide. This test design is used in other international standards. See Ref (**1**)[4] and ISO 13528.

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

AVA Documents
Document Sharing Hub



Full document details: ASTM E3264-21

Technique 4, used when replicate testing is not possible, similarly recommends testing a minimum of ten samples. That does not preclude the use of more than ten samples or more than two replicates.

NOTE 1—The spreadsheets provided in this guide for the examples in Techniques 1, 2, and 3 show the calculations when two replicate tests are performed on each sample. The spreadsheets shown for Techniques 1 and 2 may be adjusted using the equations provided in the text when more than two replicate tests are used. Use of Technique 3, as presented in Section 9, is limited to duplicate testing (that is, $k = 2$). To use Technique 3 when $k > 2$, preliminary testing for consistency of replicate results can be performed using the general form of the Cochran's Test as presented in Technique 1, and the homogeneity analysis can be performed as described in the Appendix, X4.3. Also, if desired, the homogeneity criterion in Technique 3 can be used with the calculations using the spreadsheets shown in Technique 2.

5.11 This guide is not sufficient for evaluation of certified reference materials (CRMs) or materials used for calibration. Even though homogeneity is required for CRMs, CRMs and calibration materials are typically subject to additional requirements (such as traceability and estimates of uncertainty) that are not addressed in this guide.

## 6. General Considerations for Sampling and Evaluation of Homogeneity

6.1 Effective assurance and evaluation of sample homogeneity requires good sample production, good sampling and sub-sampling to obtain representative specimens for homogeneity testing, and good testing of the representative specimens used to evaluate the homogeneity.

6.1.1 In this guide, the term "homogeneity" applies in two ways. The first is within-sample homogeneity. The second is between-sample homogeneity.

6.2 As stated previously, the basic steps for ensuring homogeneous samples are *(1)* use a process that will produce samples that are as homogeneous as possible, *(2)* test a subset of samples (or a portion of the reference material), *(3)* screen results for outliers, and *(4)* estimate the sample variation, and *(5)* decide whether the samples are sufficiently homogeneous for the intended purpose. To assure a suitable outcome, care and critical consideration are needed at each step.

6.3 Problems that affect initial sample quality can occur during sample production. Examples are such things as materials from a low-grade source, improper handling or storage of materials or components, poor control of the production process, improper or poorly maintained equipment, inadequate blending, sloppy fabrication, use of an ineffective sampling procedure, segregation or separation of components, and degradation. To address some of these issues, the practices recommended in Technique 5 provide guidance for producing good samples. Strict compliance with the practices in Technique 5, or established sampling practices from other sources, appropriate for the material under study are recommended. Those same practices, from Technique 5 or other sources, are also recommended for the preparation of samples used with Techniques 1, 2, 3, and 4.

6.4 Regardless of the care taken in sample production, there will always be some degree of sample variation and some degree of testing variation when testing for homogeneity. It is important to identify, quantify, and account for those variations when determining if the samples are suitable for the laboratory study at hand.

6.4.1 Sample homogeneity is evaluated for both, within-sample variation and between-sample variation.

6.5 Effective homogeneity testing includes ways to detect extreme test results and samples that are out of the ordinary; and ways to treat those instances where extreme test results or out of the ordinary samples are encountered. Techniques 1, 2, 3, and 4 present different outlier routines to identify inconsistent test results or samples.

6.5.1 The outlier routines were selected in an attempt to have calculations that match the homogeneity testing design and the calculations used in the second step of the technique.

6.5.2 Although the outlier routines presented for identifying inconsistent test results in Techniques 1, 2, and 3 are not equivalent, they can generally be considered as alternatives. See Note 2. Other suitable outlier routines may be substituted.

NOTE 2—Cochran's C-test and Mandel's K-test are neither equivalent nor interchangeable as presented, but since $C = K^2/n$, the two tests can be made equivalent with the right choice of critical values. For example, if there are 10 samples, Cochran's C-test with 95 % "confidence" is equivalent to Mandel's K-test with 0.5 % significance, except for the rounding of the tabulated critical values. Both critical values in this case are derived from the 99.5th percentile of the same F-distribution. In general, if the critical value for K is calculated as described in E691, Annex A1, using the $1 - \alpha/n$ fractile of F, and if the K-test is applied in the same manner as Cochran's C-test, the same outcome is achieved as Cochran's C with significance level α. Cochran's C-test is designed so that if it were applied to every cell, it would give an expected number of false rejections. The K-test is designed to give an expected proportion of false rejections, or a specified probability of false rejection for each cell.

6.6 Techniques 1, 2, and 3 each analyze agreement of replicate results for within-cell consistency. Failure to eliminate extreme replicate results that truly represent either poor sub-sampling or a testing error will tend to inflate the apparent within-sample variability and cause an overestimate of the within-sample variance, $s_w^2$. Conversely, falsely eliminating extreme results that are truly representative of expected testing variation has the opposite effect. Both types of error can affect the final evaluation of sufficient between-sample homogeneity for the laboratory study at hand. Therefore, extreme results should be investigated to confirm whether the data are valid. Valid data are those data that would be reported as resulting from the normal performance of laboratory testing of representative samples. Possible causes of invalid data include deviations from the test method, instrument malfunctions, unexpected occurrences during testing, arithmetic errors, and typographical errors. After investigation, data deemed to be invalid should be excluded from further analysis. See Ref (**2**).

6.7 Techniques 1, 2, 3, and 4 also provide guidance for eliminating extreme samples (or cells) before calculating the between-sample variance, $s_s^2$.

6.8 As with the analysis of replicate testing, failure to eliminate the extreme results of a sample that truly represent a testing or production error will tend to inflate the apparent between-sample variability and cause an overestimate of the between-sample variance, $s_s^2$. Conversely, falsely eliminating the results of an extreme sample that is truly representative of

AVA Documents
Document Sharing Hub



Full document details: ASTM E3264-21

expected variation between samples has the opposite effect. Again, both types of errors can affect the final evaluation of sufficient between-sample homogeneity for the laboratory study at hand.

6.9 When evaluating extreme values (whether analyzing consistency of replicate test results or consistency between samples), it can be difficult to distinguish between a bad test result and a bad sample. So, it is important to have as much confidence as possible in any sampling performed to obtain homogeneity test specimens and in the competence of the analysis testing performed on the homogeneity test specimens. Therefore, use of competent testing agencies along with well-established sampling practices and test methods are recommended for evaluating sample homogeneity.

6.10 Generally, the test method used for homogeneity testing is the test under study. However, special consideration should be given when each sample will be subjected to a battery of different tests. It is possible that homogeneity may be sufficient for one test property, but not sufficient for another test property determined using a more discriminating test method. Ideally the homogeneity of the samples would be evaluated for each test method to which the samples are to be subjected. If this is not practical, a single test method may be used to evaluate homogeneity. In this case the test method used for homogeneity testing (it need not necessarily be one of the test methods included in the battery of tests used in the study) should be chosen carefully to be one that is practical (in terms of time and expense) and can most likely be able to discriminate differences between samples that would affect the results of the various tests included in the battery of tests.

6.11 In some cases, it may be desirable to consider the type of statistical distribution that best describes the test results. The evaluation techniques used in this guide are based on a normal distribution and assume that the normal (Gaussian) model is adequate for the evaluation of sample homogeneity. If the homogeneity test results are not normally distributed, the homogeneity evaluation and the decision regarding whether the samples are sufficiently homogeneous can be affected.

NOTE 3—An assumption of a normal distribution is usually adequate. However, if there is concern regarding the type of distribution that best describes the results, the results can be subjected to investigation for normality. The normality of the test results can be visually evaluated using a normal probability plot. If further investigation is desired, an example of a statistical normality test is the Anderson-Darling test for normality. A description of the Anderson-Darling test for normality is not included in this guide, but the test is available in some statistical software packages. The issue of normality is discussed in ASTM Practice D6299. The construction of a normal probability plot and the Anderson-Darling test are described in the Annex of D6299.

## 7. Technique 1 for Evaluating Homogeneity

7.1 *Objective*—The objective of homogeneity testing using this technique is to verify that the variance of test results between samples (estimated by the homogeneity testing protocol proposed in this technique) is not significantly larger than the variance of test results of replicate testing on individual samples. See the criterion in 7.7.

7.2 *Circumstances*—This technique can be used to determine if the homogeneity between samples is acceptable for the intended purpose when all of the following circumstances apply:

7.2.1 An expected target standard deviation ($\sigma_{ET}$) for the laboratory study in which the samples will ultimately be used is unknown or cannot be specified prior to the study.

7.2.1.1 An example is when the precision of the test method used in the laboratory study is not known. This may be the case during the development of a new test method.

7.2.1.2 Another example would be when using an existing test method on a new material type.

7.2.2 Homogeneity testing of the entire set of samples is performed under repeatability conditions.

7.2.3 Since replicate (or duplicate) testing is necessary for homogeneity testing using this technique, the samples to be used in the laboratory study must be such that replicate testing can be performed on each sample. Replicate testing may be performed by retesting the entire sample or by testing replicate specimens obtained as representative sub-samples of the original sample. The technique would not be appropriate for a destructive test method where the entire sample must be used for the test specimen.

7.2.4 During homogeneity testing, it can be assumed that a high degree of homogeneity has been achieved within each individual sample and the sampling variance associated with obtaining replicate specimens is small.

7.2.5 The replicate specimens (obtained by sub-sampling) from each sample are essentially identical and the within-sample standard deviation is a good estimate of the repeatability of the test method.

7.2.6 The test method used for homogeneity testing is capable of discerning differences between the samples to the degree necessary for purposes of the study. This can usually be assumed to be true if the test method used for homogeneity testing is the same test method as will be used in the laboratory study.

7.2.6.1 In some cases, a proxy test method, of equal or better discrimination power for the property of interest, can be used for the homogeneity testing. That is a test method other than the one to be used in the laboratory study can be used for testing sample homogeneity. This may be done in order to better discriminate differences between samples, or for practical reasons such as time or expense. For example, it may be desirable to use a test for the particle size distribution of soil samples as a substitute to evaluate the homogeneity of samples used in a study of a different physical property of the soil such as compaction density.

7.3 *Design*—Select $n \geq 10$ random samples from the group of samples to be distributed for testing. Perform $k \geq 2$ replicate tests on each sample. Replicate testing may be performed by retesting the entire sample or by testing replicate specimens obtained as representative sub-samples of the original sample. Perform testing of the entire set of samples under repeatability conditions. To reduce possible effects of testing trends, the replicate tests must be performed in random order. For each sample, list the replicate test results as $X_{i,1}, X_{i,2}, \ldots, X_{i,k}$ where $i$ corresponds to the cell (or row) number, $i = 1, 2, \ldots, n$.

AVA Documents
Document Sharing Hub



Full document details: ASTM E3264-21

**TABLE 1 Technique 1 Example of Preliminary Test for Consistency of Replicate Results Using Cochran's Test**

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Sample | "*i*" refers to the Cell (that is, row) | Replicate 1 | Replicate 2 | Cell Average (or Mean) of the Replicates[A] $(X_{i,1} + X_{i,2})/k$ | Cell Variance[A] $\frac{(X_{i,1} - \bar{X}_i)^2 + (X_{i,2} - \bar{X}_i)^2}{k-1}$ |
| ID/No. | Cell, *i* | $X_{i,1}$ | $X_{i,2}$ | $\bar{X}_i$ | $s_i^2$ |
| FM1 | 1 | 3.0762 | 3.0491 | 3.06265 | 0.0003672 |
| FM2 | 2 | 3.0799 | 3.0646 | 3.07225 | 0.0001170 |
| FM3 | 3 | 3.0588 | 3.0589 | 3.05885 | 0.0000000 |
| FM4 | 4 | 3.0502 | 3.0621 | 3.05615 | 0.0000708 |
| FM5 | 5 | 3.0506 | 3.0750 | 3.06280 | 0.0002977 |
| FM6 | 6 | 3.0761 | 3.0627 | 3.06940 | 0.0000898 |
| FM7 | 7 | 3.0797 | 3.0636 | 3.07165 | 0.0001296 |
| FM8 | 8 | 3.0466 | 3.0745 | 3.06055 | 0.0003892 |
| FM9 | 9 | 3.0571 | 3.0541 | 3.05560 | 0.0000045 |
| FM10 | 10 | 3.0576 | 3.0573 | 3.05745 | 0.0000000 |
| FM11 | 11 | 3.0520 | 3.1325 | 3.09225 | 0.0032401 |

[A] The formulas in Table 1 are simplified for $k = 2$.

7.3.1 An example data set is shown in Columns 1, 2, 3, and 4 of Table 1 where $n = 11$ and $k = 2$. The replicate results listed in Columns 3 and 4 of Table 1, and displayed in Fig. 1, are for the fineness modulus of a fine aggregate material.

7.4 *Preliminary Test for Consistency of Replicate Results Using the Cochran's Test:*

7.4.1 This step is to assure that the homogeneity tests have all been performed properly and that there are no invalid values in the list of replicate test results. See Cochran's Test in Ref (**3**) and (**4**).

7.4.2 For each sample, calculate the Cell Average of the Replicates, listed as $X_i$ in Column 5 of Table 1, where $k = 2$ for replicate results, $X_{i,1}$ and $X_{i,2}$, and

$$\bar{X}_i = (X_{i,1} + X_{i,2} + \ldots + X_{i,k}) / k \quad (2)$$

7.4.3 For each sample, calculate the Cell Variance, listed as $s_i^2$ in Column 6 of Table 1, where

$$s_i^2 = \left[(X_{i,1} - \bar{X}_i)^2 + (X_{i,2} - \bar{X}_i)^2 + \ldots + (X_{i,k} - \bar{X}_i)^2\right] / (k - 1) \quad (3)$$

7.4.4 Calculate the sum of the Cell Variances, $\Sigma_{i=1}^{n} s_i^2$, where

$$\Sigma_{i=1}^{n} s_i^2 = s_1^2 + s_2^2 + \ldots + s_n^2 \quad (4)$$

For the example, $\Sigma_{i=1}^{n} s_i^2 = 0.004706$.

7.4.5 Identify the largest of the Cell Variances, $s_{\max}^2$, in Table 1, Column 6. In the example, $s_{\max}^2 = 0.0032401$.

7.4.6 Then calculate the Cochran's Test statistic, $C$, where

$$C = s_{\max}^2 / \Sigma_{i=1}^{n} s_i^2 \quad (5)$$

For the example, $C = s_{\max}^2 / \Sigma_{i=1}^{n} s_i^2 = 0.0032401/0.004706 = 0.6885$.

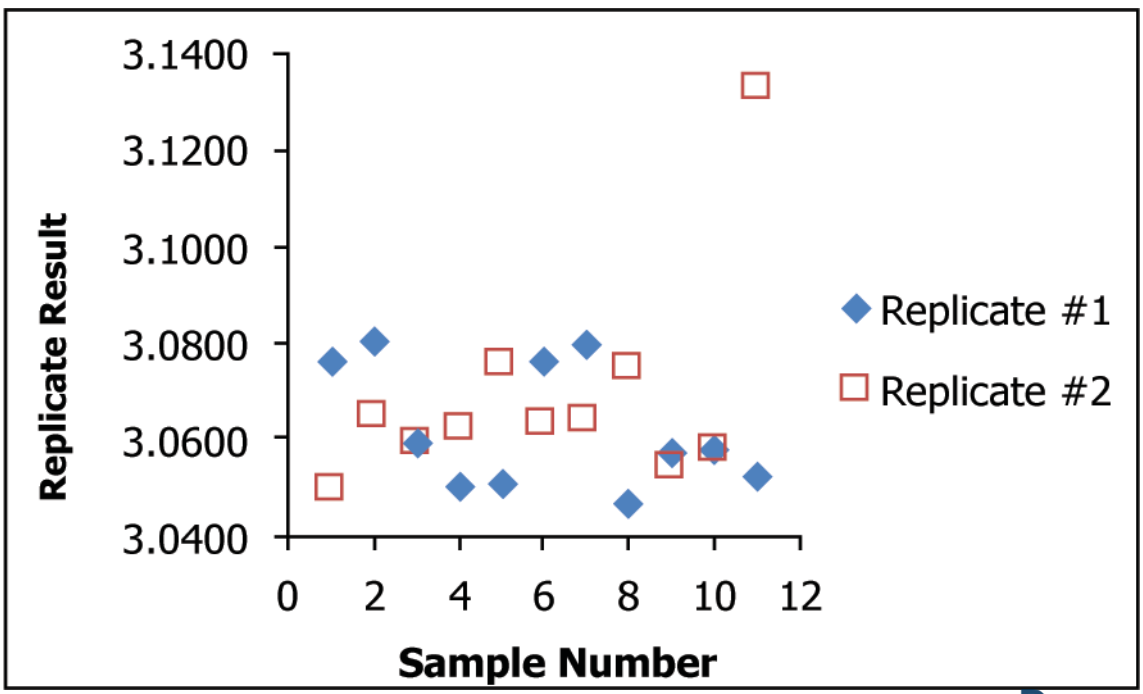


**FIG. 1 All Replicate Results**

Documents
Document Sharing Hub



Full document details: ASTM E3264-21

7.4.7 Determine the Cochran's Test critical value, $C_{crit}$, from Table 2, where $n$ = 11. It is recommended that the Cochran's Test statistic be evaluated at the 95 % or 99 % confidence level. The example uses the 99 % confidence level. For the example, $C_{crit}$ = 0.6852.

7.4.8 If the value of the test statistic, $C$, exceeds its critical value, $C_{crit}$, listed in Table 2, then the data for the sample corresponding to the largest $s_i^2$ (that is, $s_{\max}^2$) are considered inconsistent and should be investigated to confirm whether the data are valid.

7.4.9 In the example, $C$ = 0.6885, is greater than the critical value of 0.6852 for $i$ = 11. Therefore, duplicate results for Sample FM11 are inconsistent with the duplicate results of the other samples and, assuming an investigation has found sufficient reason to believe the data to be invalid, Sample FM11 is eliminated from further analysis. If there were not sufficient reason to believe the data for Sample FM11 were invalid, the data would have been retained for the homogeneity analysis in 7.5.

7.4.10 A column chart of the cell variance for each of the samples is a good graphic for this application. See Fig. 2. In the example, the value of the cell variance for Sample 11 appears to be unusually large, visually verifying the results of the Cochran's Test for variance outliers.

7.5 *Homogeneity Analysis — Estimation of Within-Sample Variation:*

7.5.1 This step uses the replicate test results remaining after cells (that is, samples) containing invalid test results have been eliminated. It is assumed that the replicate test results were obtained under repeatability conditions so that the estimation of the within-sample variation will represent testing precision comparable to the repeatability that can be expected for the material under study. The calculations use the $F$-test in a one-way analysis of variance to determine whether there are any statistically significant differences between the means (that is, within-sample means) of multiple independent groups (that is, cells) and can be found in textbooks.

7.5.2 For the $n_h$ samples remaining after samples with invalid replicate results have been removed, list the replicate test results as $X_{h,1}, X_{h,2}, \ldots, X_{h,k}$ where $h$ corresponds to the new cell (or row) number, as shown in Columns 3 and 4 of Table 3. For each sample, calculate the Cell Average of the replicate test results, $\bar{X}_h$. For the example, see Column 5 of Table 3.

$$\bar{X}_h = (X_{h,1} + X_{h,2} + \ldots + X_{h,k}) / k \tag{6}$$

7.5.3 For the example, the number of replicate results $k$ = 2. Calculate the Overall Average of All Replicate Test Results, $\bar{\bar{X}}_{homog}$, where $n_h$ is the number of samples remaining after cells containing invalid replicate results were eliminated.

$$\bar{\bar{X}}_{homog} = \frac{1}{n_h} \sum_{h=1}^{n_h} \bar{X}_h \tag{7}$$

For the example, the Cochran's Test for variance outliers indicated that replicate results for sample FM11 were inconsistent and, assuming an investigation determined the results to be invalid, sample FM11 was eliminated from further analysis. Therefore $n_h$ = 10 and the value for the Overall Average is $\bar{\bar{X}}_{homog}$ = 3.062735. See Column 6 of Table 3.

7.5.4 For each sample, calculate the Square of the Deviation of each Replicate from the Cell Average, $d_{h,j}^2$.

$$d_{h,j}^2 = \left(X_{h,j} - \bar{X}_h\right)^2 \tag{8}$$

where $h$ = 1, 2, ..., $n_h$ and $j$ = 1, 2, ..., $k$. In the example, $d_{h,j}^2$ values are listed in Columns 7 and 8 of Table 3.

7.5.5 For each sample, calculate the Total of the Squares of the Deviations of the Replicates for each Cell, $d_{h,total}^2$, where

$$d_{h,total}^2 = d_{h,1}^2 + d_{h,2}^2 + \ldots + d_{h,k}^2 \tag{9}$$

See Column 9 of Table 3.

7.5.6 Calculate the Sum of Squares Within Samples (or Cells), $SS_w$, where

$$SS_w = \Sigma_{h=1}^{n_h} d_{h,total}^2 \tag{10}$$

For the example, $SS_w$ = 0.001466, the sum of the values in Column 9 of Table 3.

7.5.7 Calculate the Within-Sample Degrees of Freedom, $d_{fw}$, and the Mean Square Within Samples, $MS_w$, where

$$df_w = n_h(k - 1) \tag{11}$$

and

$$MS_w = SS_w / df_w \tag{12}$$

For the example, $d_{fw}$ = 10 (2 – 1) = 10 and $MS_w$ = 0.001466 / 10 = 0.0001466.

7.6 *Homogeneity Analysis – Estimation of Between-Sample Variation:*

7.6.1 For each sample, calculate the Square of the Deviation of the Cell Average from the Overall Average, $\left(X_h - \bar{X}_{homog}\right)^2$, as listed in Column 10 of Table 3.

7.6.2 Multiply the sum of the values in Column 10, $\Sigma_{h=1}^{n_h}\left(X_h - \bar{X}_{homog}\right)^2$, by the number of replicates, $k$, to determine Sum of Squares Between Samples, $SS_b$, where

$$SS_b = k\Sigma_{h=1}^{n_h}\left(\bar{X}_h - \bar{\bar{X}}_{homog}\right)^2 \tag{13}$$

For the example, $\Sigma_{h=1}^{n_h}\left(\bar{X}_h - \bar{\bar{X}}_{homog}\right)^2$ = 0.0003565 and $SS_b$ = 2(0.0003565) = 0.0007130.

7.6.3 Calculate the Between-Sample Degrees of Freedom, $df_b$, and the Mean Square Between Samples, $MS_b$, where

$$df_b = (n_h - 1) \tag{14}$$

and

$$MS_b = SS_b / df_b \tag{15}$$

**TABLE 2 Cochran's Test Critical Values for $k$ = 2[A]**

| n | 95 % Confidence | 99 % Confidence |
|---|---|---|
| 7 | 0.7271 | 0.8376 |
| 8 | 0.6798 | 0.7945 |
| 9 | 0.6385 | 0.7544 |
| 10 | 0.6020 | 0.7175 |
| 11 | 0.5715 | 0.6852 |
| 12 | 0.541 | 0.6528 |

[A] Values of $C_{crit}$ are from Table 15 of Ref (3).



AVA Documents
Document Sharing Hub
Full document details: ASTM E3264-21

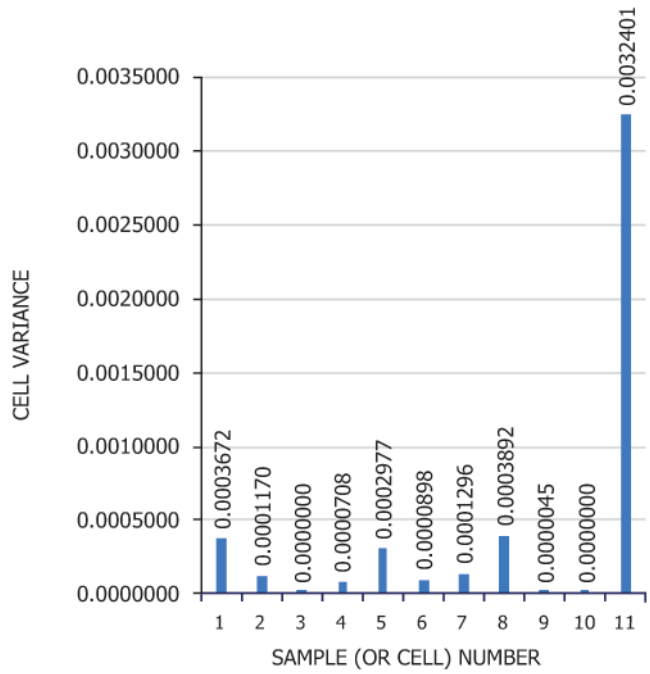


**FIG. 2 Column Chart of Cell Variances, $s_i^2$**

**TABLE 3 Technique 1 Example for Homogeneity Analysis**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Sample | "$h$" refers to the Cell (that is, row) | Replicate 1 | Replicate 2 | Cell Average of Replicates[A] $(X_{h,1} + X_{h,2})/k$ | Overall Average of All Replicate Test Results $\frac{1}{n_h}\sum_{h=1}^{n_h}\bar{X}_h$ | Square of the Deviation of Replicate 1 from the Cell Average $(X_{h,1} - \bar{X}_h)^2$ | Square of the Deviation of Replicate 2 from the Cell Average $(X_{h,2} - \bar{X}_h)^2$ | Total of the Squares of the Deviations of the Replicates for Each Cell[A] $d^2_{h,1} + d^2_{h,2}$ | Square of the Deviation of the Cell Average from the Overall Average |
| ID/No. | Cell, $h$ | $X_{h,1}$ | $X_{h,2}$ | $\bar{X}_h$ | $\bar{\bar{X}}_{homog}$ | $d^2_{h,1}$ | $d^2_{h,2}$ | $d^2_{h,total}$ | $(\bar{X}_h - \bar{\bar{X}}_{homog})^2$ |
| FM1 | 1 | 3.0762 | 3.0491 | 3.06265 | 3.062735 | 0.00018360 | 0.00018360 | 0.000367 | 0.000000007 |
| FM2 | 2 | 3.0799 | 3.0646 | 3.07225 | 3.062735 | 0.00005852 | 0.00005852 | 0.000117 | 0.000090535 |
| FM3 | 3 | 3.0588 | 3.0589 | 3.05885 | 3.062735 | 0.00000000 | 0.00000000 | 0.000000 | 0.000015093 |
| FM4 | 4 | 3.0502 | 3.0621 | 3.05615 | 3.062735 | 0.00003540 | 0.00003540 | 0.000071 | 0.000043362 |
| FM5 | 5 | 3.0506 | 3.0750 | 3.06280 | 3.062735 | 0.00014884 | 0.00014884 | 0.000298 | 0.000000004 |
| FM6 | 6 | 3.0761 | 3.0627 | 3.06940 | 3.062735 | 0.00004489 | 0.00004489 | 0.000090 | 0.000044422 |
| FM7 | 7 | 3.0797 | 3.0636 | 3.07165 | 3.062735 | 0.00006480 | 0.00006480 | 0.000130 | 0.000079477 |
| FM8 | 8 | 3.0466 | 3.0745 | 3.06055 | 3.062735 | 0.00019460 | 0.00019460 | 0.000389 | 0.000004774 |
| FM9 | 9 | 3.0571 | 3.0541 | 3.05560 | 3.062735 | 0.00000225 | 0.00000225 | 0.000004 | 0.000050908 |
| FM10 | 10 | 3.0576 | 3.0573 | 3.05745 | 3.062735 | 0.00000002 | 0.00000002 | 0.000000 | 0.000027931 |

$$ss_w = \sum_{h=1}^{n_h} d^2_{h,total} = 0.001466$$

$$\sum_{hj=1}^{n_h} \left(\bar{X}_h - \bar{\bar{X}}_{homog}\right)^2 = 0.0003565$$

[A] The formulas in Table 3 are simplified for $k$ = 2.

For the example, $df_b$ = (10 – 1) = 9, and $MS_b$ = 0.0007130 / 9 = 0.0000792.

7.7 *Homogeneity Analysis — Evaluation and Application of Technique 1 Criterion:*

7.7.1 The homogeneity of the samples can be evaluated using the $F$-test to compare $MS_b$ to $MS_w$, where $df_b = (n_h - 1)$ and $df_w = n_h\ (k - 1)$.

7.7.2 Typically, the samples are deemed sufficiently homogeneous if $F \leq F_{crit}$ for α = 0.05. However, the value of α is somewhat arbitrary.

7.7.3 Calculate the value of the $F$-Statistic where

$$F = MS_b / MS_w \tag{16}$$

For the example, $F$ = 0.0000792 / 0.0001466 = 0.54.



Documents
Document Sharing Hub
Full document details: ASTM E3264-21

7.7.4 Determine the critical value of the $F$-Statistic, $F_{crit}$, from Table 4, where $\alpha = 0.05$, $df_b = 9$ and $df_w = 10$.

For the example, $F_{crit} = 3.02$.

7.7.5 The samples are considered sufficiently homogeneous if the value of the $F$-Statistic does not exceed $F_{crit}$. In the example, $F \leq F_{crit}$ (that is, $0.54 < 3.02$) and the homogeneity of the samples is deemed satisfactory.

## 8. Technique 2 for Evaluating Homogeneity

8.1 *Objective*—The objective of homogeneity testing using this technique is to verify that the between-sample standard deviation ($s_s$) of the samples to be distributed for laboratory testing is sufficiently small as to have little effect on the results of the laboratory study – that is, heterogeneity between samples is deemed sufficiently small in comparison to the expected standard deviation ($\sigma_{ET}$) for the testing to be analyzed in the laboratory study. See the criterion in 8.7.

8.2 *Circumstances*—This technique can be used when all of the following circumstances apply (The circumstances are similar to those of Technique 3.):

8.2.1 The expected target standard deviation in the laboratory study ($\sigma_{ET}$) is known, or can be specified, in advance.

8.2.1.1 For example, for interlaboratory studies of test results from different laboratories, the value of $\sigma_{ET}$ may be based on the published reproducibility standard deviation ($\sigma_R$) for the test method used in the study. For proficiency testing programs, $\sigma_{ET}$ may be the value of the standard deviation used for the evaluation of laboratory testing. For intra-laboratory studies, the value of $\sigma_{ET}$ may be based on the repeatability standard deviation ($\sigma_r$). The value used for $\sigma_{ET}$ may also be derived from the requirements of a specification, industry standard or other source.

8.2.1.2 The value for the estimate for $\sigma_{ET}$ may not be available until after the first round of the laboratory study has been completed. In that case, if there are any available data, even from an internal round robin study, that data could provide a provisional estimate for $\sigma_{ET}$.

8.2.2 Homogeneity testing of the entire set of samples is performed under repeatability conditions.

8.2.3 A test method of sufficient precision is available for homogeneity testing (a poor test method may not detect sample differences).

8.2.3.1 When possible, it is recommended that the test method used for homogeneity testing satisfies the requirement $\sigma_r / \sigma_{ET} < 0.5$, where $\sigma_r$ is the repeatability standard deviation of the test method and $\sigma_{ET}$, as described above, is the expected standard deviation of interest in the study. See Ref (**5**) and (**1**).

8.2.3.2 For intra-laboratory testing, where repeatability ($\sigma_r$) may be the basis for the value of $\sigma_{ET}$, it is obvious that the above requirement cannot be met and another means must be used to assess the test method used for homogeneity testing.

8.2.3.3 Another criterion for assessing the test method used for homogeneity testing is the ratio of $\sigma_r$ / (expected test result). If the repeatability is deemed by the user to be sufficiently small compared to the expected test result for the samples, the test method may be accepted as satisfactory.

8.2.4 Sample preparation and sample homogeneity can be precisely controlled to a degree comparable to the precision of the test method used for the homogeneity evaluation. Little is gained by homogeneity testing in accordance with this technique if sample differences are identified and cannot be eliminated from the sample preparation process.

8.2.5 Since replicate testing is necessary for homogeneity testing using this technique, the samples to be used in the laboratory study must be such that replicate testing can be performed on each sample. Replicate testing may be performed by retesting the entire sample or by testing replicate specimens obtained as representative sub-samples of the original sample.

8.2.6 The sub-sampling variance associated with obtaining replicate specimens for homogeneity testing is small relative to the within-sample (that is, repeatability) testing variance.

8.3 *Design*—Select $n \geq 10$ random samples from the group of samples to be distributed for testing. Perform $k \geq 2$ replicate tests on each sample. Replicate testing may be performed by retesting the entire sample or by testing replicate specimens obtained as representative sub-samples of the original sample. Perform testing of the entire set of samples under repeatability conditions. To reduce possible effects of testing trends, the replicate tests must be performed in random order. For all replicate tests, list the replicate test results as $X_{i,1}, X_{i,2}, \ldots, X_{i,k}$ where $i$ corresponds to the cell (or row) number in Table 5.

8.3.1 An example data set using $n = 11$ and $k = 2$ is shown in Table 5. The replicate results listed in Columns 2 and 3 of Table 5, and displayed in Fig. 3, are for the fineness modulus of a fine aggregate material.

8.4 *Preliminary Test for Within-Cell Consistency Using Mandel's K-Test:*

8.4.1 This step is to assure that the homogeneity tests have all been performed properly and that there are no extreme values in the list of replicate test results. See "Data Consistency" in Practice E691 and Ref (**6**).

8.4.2 Calculate the Cell Averages, listed as $\bar{X}_i$ in Column 5 of Table 5, where

**TABLE 4 *F*-Test Critical Values for $\alpha = 0.05$[A]**

| | | Numerator Degrees of Freedom, $df_b$ | | | | |
|---|---|---|---|---|---|---|
| | | 7 | 8 | 9 | 10 | 11 |
| | 8 | 3.50 | 3.44 | 3.39 | 3.35 | 3.32 |
| Denominator | 9 | 3.29 | 3.23 | 3.18 | 3.14 | 3.11 |
| Degrees of | 10 | 3.14 | 3.07 | 3.02 | 2.98 | 2.95 |
| Freedom, *dfw* | 11 | 3.01 | 2.95 | 2.90 | 2.85 | 2.82 |
| | 12 | 2.91 | 2.85 | 2.80 | 2.75 | 2.72 |

[A] Values of $F_{crit}$ are from Table 3 of Ref (**3**).

AVA Documents
Document Sharing Hub



Full document details: ASTM E3264-21

**TABLE 5 Technique 2 Example Data and Preliminary Test for Within-Cell Consistency Using the *K*-Statistic**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Sample | "*i*" refers to the Cell (that is, row) | Replicate 1 | Replicate 2 | Cell Average (or Mean) of the Replicates[A] $(X_{i,1} + X_{i,2})/k$ | Cell Variance[A] $\left[\left(X_{i,1} - \bar{X}_i\right)^2 = \left(X_{i,2} - \bar{X}_i\right)^2\right] / (k - 1)$ | Within Cell Standard Deviation $\sqrt{s_i^2}$ | Within-Sample Variance for the Preliminary Test or Average of the Cell Variances $\frac{1}{n}\sum_{i=1}^{n} s_i^2$ | Within-Sample Standard Deviation $\sqrt{s_{wp}^2}$ | Within-Cell Consistency K-Statistic $\frac{s_i}{s_{wp}}$ |
| ID/No. | Cell, *i* | $X_{i,1}$ | $X_{i,2}$ | $\bar{X}_i$ | $s_i^2$ | $s_i$ | $s_{wp}^2$ | $s_{wp}$ | $K_i$ |
| FM1 | 1 | 3.0762 | 3.0491 | 3.06265 | 0.0003672 | 0.019163 | 0.000428 | 0.02068 | 0.93 |
| FM2 | 2 | 3.0799 | 3.0646 | 3.07225 | 0.0001170 | 0.010819 | 0.000428 | 0.02068 | 0.52 |
| FM3 | 3 | 3.0588 | 3.0589 | 3.05885 | 0.0000000 | 0.000071 | 0.000428 | 0.02068 | 0.00 |
| FM4 | 4 | 3.0502 | 3.0621 | 3.05615 | 0.0000708 | 0.008415 | 0.000428 | 0.02068 | 0.41 |
| FM5 | 5 | 3.0506 | 3.0750 | 3.06280 | 0.0002977 | 0.017253 | 0.000428 | 0.02068 | 0.83 |
| FM6 | 6 | 3.0761 | 3.0627 | 3.06940 | 0.0000898 | 0.009475 | 0.000428 | 0.02068 | 0.46 |
| FM7 | 7 | 3.0797 | 3.0636 | 3.07165 | 0.0001296 | 0.011384 | 0.000428 | 0.02068 | 0.55 |
| FM8 | 8 | 3.0466 | 3.0745 | 3.06055 | 0.0003892 | 0.019728 | 0.000428 | 0.02068 | 0.95 |
| FM9 | 9 | 3.0571 | 3.0541 | 3.05560 | 0.0000045 | 0.002121 | 0.000428 | 0.02068 | 0.10 |
| FM10 | 10 | 3.0576 | 3.0573 | 3.05745 | 0.0000000 | 0.000212 | 0.000428 | 0.02068 | 0.01 |
| FM11 | 11 | 3.0520 | 3.1325 | 3.09225 | 0.0032401 | 0.056922 | 0.000428 | 0.02068 | 2.75 |

[A] The formulas in Table 5 are simplified for $k = 2$.

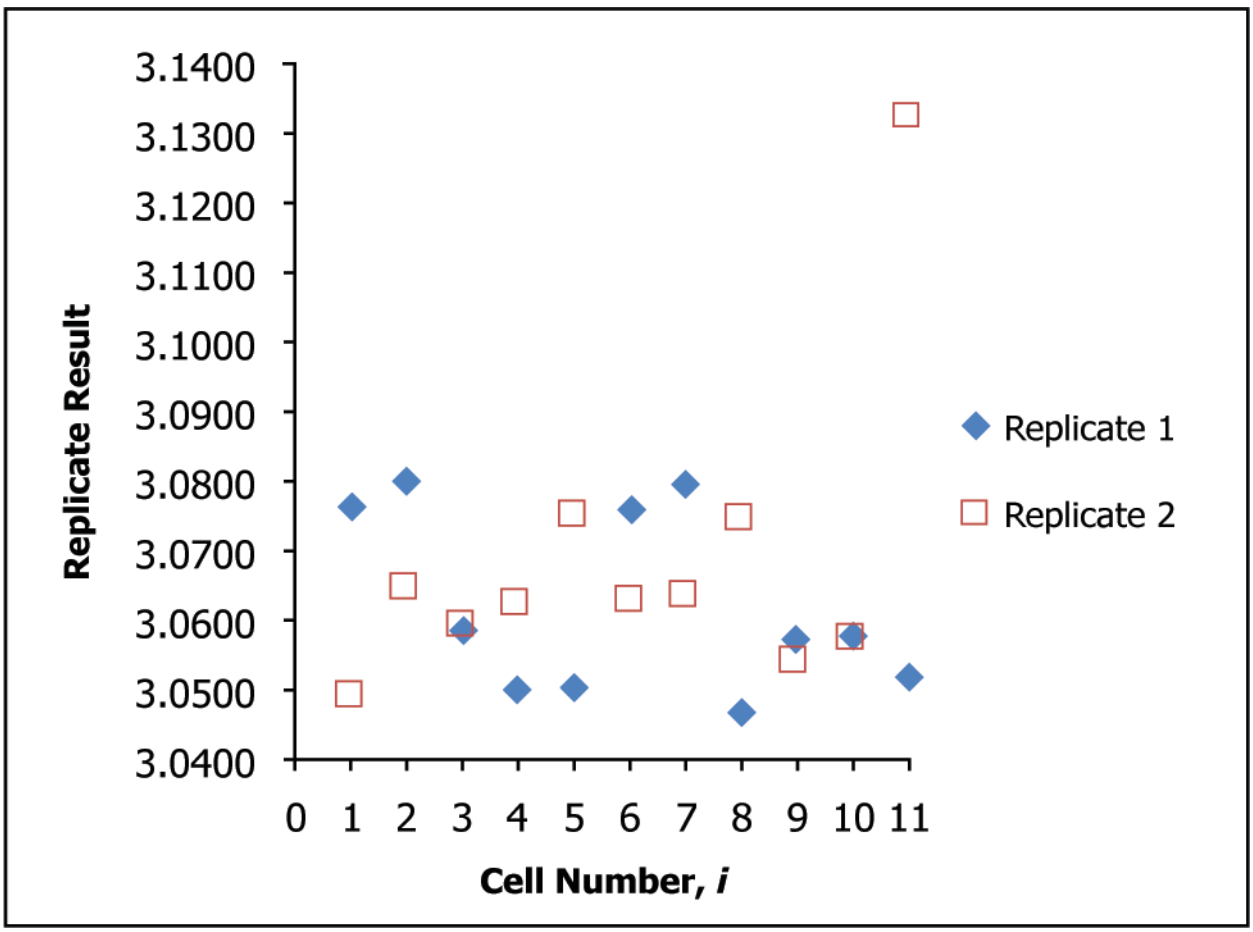


**FIG. 3 Display of Replicate Results**

$$\bar{X}_i = \left(X_{i,1} + X_{i,2} + \ldots + X_{i,k}\right) / k \qquad (17)$$

For the example, $k = 2$.

Documents
Document Sharing Hub



Full document details: ASTM E3264-21