



Designation: D4821 – 20

Standard Guide for Carbon Black—Validation of Test Method Precision and Bias¹

This standard is issued under the fixed designation D4821; 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 covers procedures for using the ASTM Standard Reference Blacks² (SRBs) and the HT and INR Iodine Number Standards to continuously monitor the precision of those carbon black test methods for which reference values have been established. It also offers guidelines for troubleshooting various test methods.

1.2 This guide defines the environmental conditions that are required for laboratories that perform carbon black testing activities for those test methods under D24's jurisdiction.

1.3 This guide establishes procedures for the use of x-charts to continuously monitor those tests listed in Section 2 for within-lab precision (repeatability) and between-lab accuracy (reproducibility).

1.4 This guide provides a statistical procedure for improving test reproducibility when a laboratory cannot physically calibrate its apparatus to obtain the reference values of the ASTM reference blacks, within the ranges given in this guide.

1.5 *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:*³

[D1510 Test Method for Carbon Black—Iodine Adsorption Number](#)

¹ This guide is under the jurisdiction of ASTM Committee D24 on Carbon Black and is the direct responsibility of Subcommittee D24.61 on Carbon Black Sampling and Statistical Analysis.

Current edition approved Nov. 1, 2020. Published November 2020. Originally approved in 1988. Last previous edition approved in 2019 as D4821 – 19. DOI: 10.1520/D4821-20.

² Standard Reference Blacks are available from Laboratory Standards & Technologies, Inc., 227 Somerset St., Borger, TX 79007.

³ 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.

- [D1513 Test Method for Carbon Black, Pelleted—Pour Density](#)
- [D1765 Classification System for Carbon Blacks Used in Rubber Products](#)
- [D2414 Test Method for Carbon Black—Oil Absorption Number \(OAN\)](#)
- [D3265 Test Method for Carbon Black—Tint Strength](#)
- [D3324 Practice for Carbon Black—Improving Test Reproducibility Using ASTM Standard Reference Blacks \(Withdrawn 2002\)⁴](#)
- [D3493 Test Method for Carbon Black—Oil Absorption Number of Compressed Sample \(COAN\)](#)
- [D6556 Test Method for Carbon Black—Total and External Surface Area by Nitrogen Adsorption](#)
- [D7854 Test Method for Carbon Black-Void Volume at Mean Pressure](#)
- [E177 Practice for Use of the Terms Precision and Bias in ASTM Test Methods](#)
- [E2282 Guide for Defining the Test Result of a Test Method](#)
- [E2586 Practice for Calculating and Using Basic Statistics](#)

3. Terminology

3.1 *Definitions:*

3.1.1 *accepted reference value, n*—a value that serves as an agreed-upon reference for comparison, and which is derived as: (1) a theoretical or established value, based on scientific principles, (2) an assigned or certified value, based on experimental work of some national or international organization, or (3) a consensus or certified value, based on collaborative experimental work under the auspices of a scientific or engineering group.

3.1.1.1 *Discussion*—A national or international organization, referred to in (2), generally maintains measurement standards to which the reference values obtained are traceable. **E177**

3.1.2 *accuracy, n*—the closeness of agreement between a test result and an accepted reference value.

3.1.2.1 *Discussion*—The term accuracy, when applied to a set of test results, involves a combination of a random

⁴ The last approved version of this historical standard is referenced on www.astm.org.

component and of a common systematic error or bias component. **E177**

3.1.3 *ASTM reference blacks, n*—a set of carbon blacks that span the useful range of the test method for which they are reference materials. **D3324**

3.1.4 *bias, n*—the difference between the expectation of the test results and an accepted reference value.

3.1.4.1 *Discussion*—Bias is the total systematic error as contrasted to random error. There may be one or more systematic error components contributing to the bias. A larger systematic difference from the accepted reference value is reflected by a larger bias value. **E177**

3.1.5 *characteristic, n*—a property of items in a sample or population which, when measured, counted or otherwise observed, helps to distinguish between the items. **E2282**

3.1.6 *coefficient of variation, CV, n*—for a nonnegative characteristic, the ratio of the standard deviation to the mean for a population or sample. **E2586**

3.1.7 *intermediate precision, n*—the closeness of agreement between test results obtained under specified intermediate precision conditions.

3.1.7.1 *Discussion*—The specific measure and the specific conditions must be specified for each intermediate measure of precision; thus, “standard deviation of test results among operators in a laboratory,” or “day-to-day standard deviation within a laboratory for the same operator.”

3.1.7.2 *Discussion*—Because the training of operators, the agreement of different pieces of equipment in the same laboratory and the variation of environmental conditions with longer time intervals all depend on the degree of within-laboratory control, the intermediate measures of precision are likely to vary appreciably from laboratory to laboratory. Thus, intermediate precisions may be more characteristic of individual laboratories than of the test method. **E177**

3.1.8 *intermediate precision conditions, n*—conditions under which test results are obtained with the same test method using test units or test specimens taken at random from a single quantity of material that is as nearly homogeneous as possible, and with changing conditions such as operator, measuring equipment, location within the laboratory, and time. **E177**

3.1.9 *measured value, n*—an observed test results as opposed to a standard value. **D3324**

3.1.10 *normalization, n*—the practice of applying a statistical correction to test measurements to improve accuracy.

3.1.10.1 *Discussion*—The correction of test data using a straight-line equation (linear regression) where measurements of ASTM reference blacks are analyzed with published accepted reference values to determine a slope and y-intercept. Normalization is a proven technique to improve the accuracy or reproducibility of laboratory data when all other means of calibration do not satisfactorily achieve a desired state of calibration.

3.1.11 *observation, n*—the process of obtaining information regarding the presence or absence of an attribute of a test specimen, or of making a reading on a characteristic or dimension of a test specimen. **E2282**

3.1.12 *observed value, n*—the value obtained by making an observation. **E2282**

3.1.13 *precision, n*—the closeness of agreement between independent test results obtained under stipulated conditions.

3.1.13.1 *Discussion*—Precision depends on random errors and does not relate to the accepted reference value.

3.1.13.2 *Discussion*—The measure of precision usually is expressed in terms of imprecision and computed as a standard deviation of the test results. Less precision is reflected by a larger standard deviation.

3.1.13.3 *Discussion*—“Independent test results” means results obtained in a manner not influenced by any previous result on the same or similar test object. Quantitative measures of precision depend critically on the stipulated conditions. Repeatability and reproducibility conditions are particular sets of extreme stipulated conditions. **E177**

3.1.14 *regression of standard values on measured values, n*—statistical equation derived by the method of least-squares. **D3324**

3.1.15 *repeatability, n*—precision under repeatability conditions.

3.1.15.1 *Discussion*—Repeatability is one of the concepts or categories of the precision of a test method.

3.1.15.2 *Discussion*—Measures of repeatability defined in this compilation are repeatability standard deviation and repeatability limit. **E177**

3.1.16 *repeatability conditions, n*—conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time.

3.1.16.1 *Discussion*—See precision, the “same operator, same equipment” requirement means that for a particular step in the measurement process, the same combination of operator and equipment is used for every test result. Thus, one operator may prepare the test specimens, a second measure the dimensions and a third measure the mass in a test method for determining density.

3.1.16.2 *Discussion*—By “in the shortest practical period of time” is meant that the test results, at least for one material, are obtained in a time period not less than in normal testing and not so long as to permit significant change in test material, equipment or environment. **E177**

3.1.17 *repeatability limit (r), n*—the value below which the absolute difference between two individual test results obtained under repeatability conditions may be expected to occur with a probability of approximately 0.95 (95 %).

3.1.17.1 *Discussion*—The repeatability limit is times the repeatability standard deviation. This multiplier is independent of the size of the interlaboratory study.

3.1.17.2 *Discussion*—The approximation to 0.95 is reasonably good (say 0.90 to 0.98) when many laboratories (30 or more) are involved, but is likely to be poor when fewer than eight laboratories are studied. **E177**

3.1.18 *repeatability standard deviation (sr), n*—the standard deviation of test results obtained under repeatability conditions.

3.1.18.1 *Discussion*—It is a measure of the dispersion of the distribution of test results under repeatability conditions.

3.1.18.2 *Discussion*—Similarly, “repeatability variance” and “repeatability coefficient of variation” could be defined and used as measures of the dispersion of test results under repeatability conditions.—In an interlaboratory study, this is the pooled standard deviation of test results obtained under repeatability conditions.

3.1.18.3 *Discussion*—The repeatability standard deviation, usually considered a property of the test method, will generally be smaller than the within-laboratory standard deviation. (See **within-laboratory standard deviation**.) **E177**

3.1.19 *reproducibility, n*—precision under reproducibility conditions. **E177**

3.1.20 *reproducibility conditions, n*—conditions where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment.

3.1.20.1 *Discussion*—Identical material means either the same test units or test specimens are tested by all the laboratories as for a nondestructive test or test units or test specimens are taken at random from a single quantity of material that is as nearly homogeneous as possible. A different laboratory of necessity means a different operator, different equipment, and different location and under different supervisory control. **E177**

3.1.21 *reproducibility limit (R), n*—the value below which the absolute difference between two test results obtained under reproducibility conditions may be expected to occur with a probability of approximately 0.95 (95 %).

3.1.21.1 *Discussion*—The reproducibility limit is times the reproducibility standard deviation. The multiplier is independent of the size of the interlaboratory study (that is, of the number of laboratories participating).

3.1.21.2 *Discussion*—The approximation to 0.95 is reasonably good (say 0.90 to 0.98) when many laboratories (30 or more) are involved but is likely to be poor when fewer than eight laboratories are studied. **E177**

3.1.22 *reproducibility standard deviation (sR), n*—the standard deviation of test results obtained under reproducibility conditions.

3.1.22.1 *Discussion*—Other measures of the dispersion of test results obtained under reproducibility conditions are the “reproducibility variance” and the “reproducibility coefficient of variation.”

3.1.22.2 *Discussion*—The reproducibility standard deviation includes, in addition to between laboratory variability, the repeatability standard deviation and a contribution from the interaction of laboratory factors (that is, differences between operators, equipment and environments) with material factors (that is, the differences between properties of the materials other than that property of interest). **E177**

3.1.23 *standard deviation, n—of a population, σ* , the square root of the average or expected value of the squared deviation of a variable from its mean; *of a sample, s*, the square root of the sum of the squared deviations of the observed values in the sample divided by the sample size minus one. **E2586**

3.1.24 *standard value, n*—the value assigned to a reference black by ASTM Committee D24 on Carbon Black.

3.1.24.1 *Discussion*—Usually this value is calculated as the average test result of an interlaboratory testing program. **D3324**

3.1.25 *test determination, n*—the value of a characteristic or dimension of a single test specimen derived from one or more observed values. **E2282**

3.1.26 *test method, n*—a definitive procedure that produces a test result. **E2282**

3.1.27 *test result, n*—the value of a characteristic obtained by carrying out a specified test method. **E2282**

3.1.28 *test sample, 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**. **E2282**

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

3.1.30 *trueness, n*—the closeness of agreement between the population mean of the measurements or test results and the accepted reference value.

3.1.30.1 *Discussion*—“Population mean” is, conceptually, the average value of an indefinitely large number of test results. **E177**

3.1.31 *variance, σ^2, s^2, n* —square of the standard deviation of the population or sample. **E2586**

3.1.32 *within-laboratory standard deviation, n*—the standard deviation of test results obtained within a laboratory for a single material under conditions that may include such elements as different operators, equipment, and longer time intervals.

3.1.32.1 *Discussion*—Because the training of operators, the agreement of different pieces of equipment in the same laboratory and the variation of environmental conditions with longer time intervals depend on the degree of within-laboratory control, the within-laboratory standard deviation is likely to vary appreciably from laboratory to laboratory. **E177**

4. Significance and Use

4.1 This guide provides insight into the environmental conditions required for operation of laboratories or to aid in the design of new laboratories that perform carbon black testing activities using those test methods that are under D24’s jurisdiction. This guide does not supersede any specific requirements a laboratory may choose to establish.

4.2 This guide recommends the use of statistical x-charts to graphically monitor test data determined for the ASTM reference blacks for those test methods given in Section 2. All laboratories are encouraged to utilize statistical x-charts and ASTM reference blacks because this enables a comparison of testing precision within and between laboratories. The guide describes practices for the use of repeatability and reproducibility limits and x-charts.

4.3 In addition to the calibration of a test method by physicochemical means, a statistical method for achieving calibration of a test method is presented (that is, *normalization*).