



## Standard Guide for Evaluating Test Method Capability and Fitness for Use<sup>1</sup>

This standard is issued under the fixed designation D8146; 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 ( $\epsilon$ ) indicates an editorial change since the last revision or reapproval.

### 1. Scope\*

1.1 This guide covers techniques for evaluating the statistical capability and fitness for use of standard test methods used for measuring properties of petroleum products, liquid fuels, and lubricants. Specifically, this guide provides strategies for evaluating the capability of a test method to provide a sufficiently precise estimate of the intended parameter versus a given level or value of that parameter and for assessing, with sufficient confidence, the fitness for use of a test method for determining the acceptability of products versus specification, regulatory, or manufacturing limits.

1.2 The assessment of capability in this guide is applicable to test methods that always yield non-negative numerical results and have a hard, fixed zero lower limit or fixed upper limit (for example, 100 %). Detailed knowledge of a test method and professional judgement may be required in determining the applicability of this guide to certain test methods.

1.3 In accordance with this guide, the assessment of fitness for use of a test method versus specification, regulatory, or manufacturing limits is generally applicable to test methods that provide numerical results and have applicable reproducibility precision values.

1.4 This guide is intended for use by test method developers, specification setters, and laboratories selecting test methods to evaluate products for specific purposes or versus specific limits.

1.5 This guide is not applicable to test methods that measure temperature to determine properties such as, but not limited to, freeze point, pour point, flash point, and distillations. This guide is not intended for use by laboratories for evaluating laboratory test method execution capability, nor is it intended to evaluate manufacturing process capability.

1.6 The expressions of capability and fitness for use derived from use of this guide should not be the only criteria for selection and use of a given test method. Other factors such as the laboratory's experience with the test method, length of time

to complete an analysis, typical product results relative to the regulatory or manufacturing specification limit, closeness of the results to zero, and cost factors may contribute more significantly to the decision to use a given test method for a specific application.

1.7 This guide draws on statistical approaches covered in other standards supporting petroleum products, liquid fuels, and lubricants, including Practices [D3244](#), [D6259](#), [D6299](#), [D6300](#), [D6792](#) and ISO 4259.

1.8 *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.*

1.9 *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:<sup>2</sup>

- [D3244 Practice for Utilization of Test Data to Determine Conformance with Specifications](#)
- [D6259 Practice for Determination of a Pooled Limit of Quantitation for a Test Method](#)
- [D6299 Practice for Applying Statistical Quality Assurance and Control Charting Techniques to Evaluate Analytical Measurement System Performance](#)
- [D6300 Practice for Determination of Precision and Bias Data for Use in Test Methods for Petroleum Products, Liquid Fuels, and Lubricants](#)
- [D6792 Practice for Quality Management Systems in Petroleum Products, Liquid Fuels, and Lubricants Testing Laboratories](#)

<sup>1</sup> This guide is under the jurisdiction of ASTM Committee [D02](#) on Petroleum Products, Liquid Fuels, and Lubricants and is the direct responsibility of Subcommittee [D02.94](#) on Coordinating Subcommittee on Quality Assurance and Statistics.

Current edition approved Nov. 1, 2022. Published November 2022. Originally approved in 2018. Last previous edition approved in 2018 as D8146 – 18. DOI: 10.1520/D8146-22.

<sup>2</sup> For referenced ASTM standards, visit the ASTM website, [www.astm.org](http://www.astm.org), or contact ASTM Customer Service at [service@astm.org](mailto: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

## 2.2 ISO Standard:<sup>3</sup>

**ISO 4259 Petroleum products—Determination and application of precision data in relation to methods of test**

2.3 ASTM Standards referenced only in **Appendix X2**, Worked Examples, are listed in **X2.3**.

## 3. Terminology

### 3.1 Definitions:

3.1.1 *precision ratio (PR)*, *n*—an estimate of relative magnitude of repeatability and reproducibility. The PR for a given standard test method can provide information on the relative significance between variation caused by different operators and laboratories compared to a single operator in a laboratory performing the standard test method. **D6792**

3.1.1.1 *Discussion*—The PR for a published test method estimates the influence that non-site-specific variations have on the published precision; the PR metric essentially judges the adequacy of between-laboratory standardization, or the agreement between laboratories relative to within-lab repeatability precision.

3.1.2 *repeatability (r)*, *n*—the quantitative expression for the random error associated with the difference between two independent results obtained under repeatability conditions that would be exceeded with an approximate probability of 5 % (1 case in 20 in the long run) in the normal and correct operation of the test method. **D6300**

3.1.2.1 *Discussion*—This definition is applicable to precision values obtained following the practices described in **D6300**. Repeatability determined using different or modified practices may not have the same statistical significance as that inferred when following **D6300**.

3.1.3 *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. **D6300**

3.1.4 *reproducibility (R)*, *n*—a quantitative expression for the random error associated with the difference between two independent results obtained under reproducibility conditions would be exceeded with an approximate probability of 5 % (1 case in 20 in the long run) in the normal and correct operation of the test method. **D6300**

3.1.4.1 *Discussion*—This definition is applicable to precision values obtained following the practices described in **D6300**. Repeatability determined using different or modified practices may not have the same statistical significance as that inferred when following **D6300**.

3.1.5 *reproducibility conditions*, *n*—conditions where independent test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment. Note that different laboratory by necessity means a different operator, different equipment, and different location and under different supervisory control. **D6300**

3.1.6 *site precision (R')*, *n*—the value below which the absolute difference between two individual test results obtained under site precision conditions may be expected to occur with a probability of approximately 0.95 (95 %). It is calculated as 2.77 times the standard deviation ( $\sigma_R$ ) of results obtained under site precision conditions. **D6299**

3.1.6.1 *Discussion*—Site precision is often estimated in the setup of a quality control chart, and the TPI can be used to evaluate the performance of the test method in the laboratory as described in **D6792**.

3.1.7 *site precision conditions*, *n*—for a single analytical measurement system, conditions under which test results are obtained by one or more operators in a single site location practicing the same test method on a single measurement system, using test specimens taken at random from the same sample of material over an extended period of time spanning at least a 20 day interval.

3.1.7.1 *Discussion*—For an in-depth discussion, see Practice **D6299**. **D6299**

3.1.8 *test performance index (TPI)*, *n*—an approximate measure of a laboratory's testing capability, defined as the ratio of test method reproducibility to site precision. **D6792**

3.1.8.1 *Discussion*—The term capability as used in this definition refers to ability of the laboratory performing the test method relative to the published reproducibility and not to the capability of the test method with respect to application at a given level or value.

3.1.8.2 *Discussion*—The TPI term is similar to precision ratio ( $PR = R/r$ ), but the use of site precision instead of repeatability provides a direct evaluation of the laboratory's performance of the test method versus reproducibility. Since there is a general expectation that  $R > R' > r$ , then generally it would be expected that  $TPI < PR$ .

### 3.2 Definitions of Terms Specific to This Standard:

3.2.1 *analytical performance value (APV)*, *n*—a metric that assesses the precision of a test method relative to applicable specification limits. APV is the ratio of the repeatability or reproducibility of a test method to specific limits, expressed as a percent.

3.2.1.1 *Discussion*—The APV is the percent relative repeatability or reproducibility.

3.2.2 *capability*, *n*—as applied to this guide, an estimate of the ability of a test method to provide a precise estimate of the intended parameter at a given level or value of that parameter.

3.2.2.1 *Discussion*—APV is the metric used to express capability.

3.2.3 *fit for use*, *n*—as applied to this guide, an expression of the adequacy of a test method to provide a precise estimate of the intended parameter with a desired level of confidence at the level of a corresponding specification, regulatory, or manufacturing limit to support compliance decision.

3.2.3.1 *Discussion*—Fit for use is the effectiveness of an analytical measurement system to adequately meet the precision requirements for the intended application relative to specific limits.

3.2.3.2 *Discussion*—Also known as *fitness for use*.

<sup>3</sup> Available from American National Standards Institute (ANSI), 25 W. 43rd St., 4th Floor, New York, NY 10036, <http://www.ansi.org>.