Designation: E1169 – 21

An American National Standard

Standard Practice for Conducting Ruggedness Tests[1]

This standard is issued under the fixed designation E1169; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This practice covers conducting ruggedness tests. The purpose of a ruggedness test is to identify those factors that strongly influence the measurements provided by a specific test method and to estimate how closely those factors need to be controlled.

1.2 This practice restricts itself to experimental designs with two levels per factor. The designs require the simultaneous change of the levels of all of the factors, thus permitting the determination of the effects of each of the factors on the measured results.

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

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

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:*[2]

E456 Terminology Relating to Quality and Statistics

E1325 Terminology Relating to Design of Experiments

E1488 Guide for Statistical Procedures to Use in Developing and Applying Test Methods

E2282 Guide for Defining the Test Result of a Test Method

F2082 Test Method for Determination of Transformation Temperature of Nickel-Titanium Shape Memory Alloys by Bend and Free Recovery

3. Terminology

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

3.1.1 *factor, n*—independent variable in an experimental design. **E1325**

3.1.1.1 *Discussion*—For experimental purposes, factors must be temporarily controllable. In a ruggedness test, a factor is a test variable that may affect either the result obtained from the use of the test method or the variability of the result.

3.1.2 *fractional factorial design, n*—a factorial experiment in which only an adequately chosen fraction of the treatments required for the complete factorial experiment is selected to be run. **E1325**

3.1.3 *interaction, n*—differences in responses to a factor among levels (versions) of other factors in the experiment. **E1325**

3.1.3.1 *Discussion*—Interaction is the condition where a factor effect changes with the level of other factors in the experiment design.

3.1.4 *level (of a factor), n*—a given value, a specification of procedure or a specific setting of a factor. **E1325**

3.1.5 *main effect, average effect, n*—a term describing a measure for the comparison of the responses at each level (version) of a factor averaged over all levels (versions) of other factors in the experiment. **E1325**

3.1.5.1 *Discussion*—This is also known as a first-order effect. In a ruggedness test, the main effect is the change in the test result due to a change in the level of a factor. This is the difference of the average result at the high level of the factor minus the average result at the low level. There are only two levels in the ruggedness tests considered here.

3.1.6 *Plackett-Burman designs, n*—a set of screening designs using orthogonal arrays that permit evaluation of the linear effects of up to $n = t - 1$ factors in a study of t treatment combinations. **E1325**

[1] This practice 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 June 1, 2021. Published July 2021. Originally approved in 1987. Last previous edition approved in 2020 as E1169 – 20. DOI: 10.1520/E1169-21.

[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.

3.1.7 *ruggedness, n*—insensitivity of a test method to departures from specified test or environmental conditions.

3.1.7.1 *Discussion*—An evaluation of the "ruggedness" of a test method or an empirical model derived from an experiment is useful in determining whether the results or decisions will be relatively invariant over some range of environmental variability under which the test method or the model is likely to be applied.

3.1.8 *ruggedness test, n*—a planned experiment in which environmental factors or test conditions are deliberately varied in order to evaluate the effects of such variation.

3.1.8.1 *Discussion*—Since there usually are many environmental factors that might be considered in a ruggedness test, it is customary to use a "screening" type of experiment design which concentrates on examining many first order effects. The validity of the estimates depends on the assumption that second order effects such as interactions and curvature are relatively negligible. Often in evaluating the ruggedness of a test method, if there is an indication that the results of a test method are highly dependent on the levels of the environmental factors, there is a sufficient indication that certain levels of environmental factors must be included in the specifications for the test method, or even that the test method itself will need further revision. This evaluation may include extra runs in a second experiment.

3.1.9 *screening design, n*—a balanced design, requiring relatively minimal amount of experimentation, to evaluate the lower order effects of a relatively large number of factors in terms of contributions to variability or in terms of estimates of parameters for a model. **E1325**

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

3.1.11 *test unit, n*—the total quantity of material (containing one or more specimens) needed to obtain a test result as specified in the test method. **E2282**

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *foldover, n*—test runs, added to a two-level fractional factorial experiment, generated by duplicating the original design by switching levels of one or more factors in all runs, for the purpose of separating estimates of main effects from two factor interactions.

3.2.1.1 *Discussion*—The most useful type of foldover is with signs of all factors switched. The foldover runs are combined with the initial test results. The combination allows main effects to be separated from interactions of other factors that are aliased in the original design.

3.2.2 *two-factor interaction effect, 2fi, n*—estimate of the condition where a factor effect changes with the level of another factor in the experiment design.

4. Significance and Use

4.1 A ruggedness test is a special application of a statistically designed experiment that makes changes in the test method variables, called factors, and then calculates the subsequent effect of those changes upon the test results. Factors are features of the test method or of the laboratory environment that are known to vary across laboratories and are subject to control by the test method.

4.1.1 Statistical design enables more efficient and cost-effective determination of the factor effects than would be achieved if separate experiments were carried out for each factor. The proposed designs are easy to use in developing the information needed for evaluating quantitative test methods.

4.2 In ruggedness testing, the two *levels* (settings) for each factor are chosen to use moderate separations between the high and low settings. In general, if there is an underlying difference between the levels, then the size of effects will increase with increased separation between the high and low settings of the factors. A *run* is an execution of the test method under prescribed settings of each of the factors under study. A ruggedness test consists of a set of runs.

4.3 A ruggedness test is usually conducted within a single laboratory on uniform material, so that the effects of changing only the factors are measured. The results may then be used to assist in determining the degree of control required of factors described in the test method.

4.4 Ruggedness testing should precede an interlaboratory (round robin) study to correct any deficiencies in the test method and may also be part of the validation phase of developing a standard test method as described in Guide E1488.

4.5 This standard discusses design and analysis of ruggedness testing in Section 5 and contains an example of a basic eight run design. Some caution must be used in interpretation of results, since interaction effects may be present. These effects are present when a factor effect changes with the level of other factors in the experimental design. If it is thought that there may be interaction between variables then additional testing of the basic design is necessary. This is discussed in Section 6. In addition, Annex A3 presents estimates of precision of factor effects when run settings are replicated. An example of a twelve run design is shown in Appendix X1. Annex A1 and Annex A2 provide supplemental information.

5. Basic Ruggedness Test Design and Analysis

5.1 *Design*—A series of fractional factorial designs are recommended for use with ruggedness tests for determining the factor effects on the test results. All designs considered here have only two settings (levels) for each factor, and are known as Plackett-Burman (PB) designs (**1**).[3] These designs occur in multiples of four runs, such as 4, 8, 12, etc., and are listed in Annex A1, Each run conducts the test method at designated levels of the factors to produce a test result as the run response.

5.1.1 Select k factors to investigate. Choose a PB design with at least k+1 runs. Assign each factor to a column in the design table. The unassigned columns in the design are denoted as "dummy" factors, and these may be used to estimate the experimental error (see 5.2.3.2).

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