TECHNICAL REPORT

ISO/TR 9272

Second edition
2005-07-15

Rubber and rubber products — Determination of precision for test method standards

Caoutchouc et produits en caoutchouc — Évaluation de la fidélité des méthodes d'essai normalisées

Reference number
ISO/TR 9272:2005(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

In exceptional circumstances, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example), it may decide by a simple majority vote of its participating members to publish a Technical Report. A Technical Report is entirely informative in nature and does not have to be reviewed until the data it provides are considered to be no longer valid or useful.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO/TR 9272 was prepared by Technical Committee ISO/TC 45, *Rubber and rubber products*, Subcommittee SC 2, *Testing and analysis*.

This second edition cancels and replaces the first edition (ISO/TR 9272:1986), which has been technically revised.

Introduction

The primary precision standard for ISO test method standards is ISO 5725, a generic standard that presents the fundamental statistical approach and calculation algorithms for determining repeatability and reproducibility precision as well as accuracy and a concept related to bias called trueness. However there are certain parts of ISO 5725 that are not compatible with precision determination in the rubber manufacturing and carbon black industries over the past four decades.

two major problems exist:

a) strict adherence to ISO 5725 conflicts with the operational procedures and the past history of testing as conducted in these two industries and

b) ISO 5725 does not address certain requirements that are unique to rubber and carbon black testing.

Thus although ISO 5725 is necessary as a foundation document for this Technical Report and is used as such, it is not sufficient for the needs of TC 45.

This Technical Report replaces ISO/TR 9272, an interim document that has been used for guidance on precision determination since 1986. This new edition of the Technical Report has a more comprehensive approach to the overriding issue with precision determination over the past several decades — the discovery that the reproducibility (between-laboratory variation) of many test methods is quite large. The existence of very poor between-laboratory agreement for many fundamental test methods in the industry has been the subject of much discussion and consternation. Experience has shown that poor reproducibility is most often caused by only a small number (percentage) of the laboratories that may be designated outlier laboratories. This new edition of ISO/TR 9272 describes a "robust" analysis approach that eliminates or substantially reduces the influence of outliers. See Annex E for a more detailed discussion of these issues and additional background on ISO 5725.

Five annexes are presented. These serve as supplements to the main body of the Technical Report. They are in addition to the terminology section proper.

— Annex A defines the Mandel h and k statistics, illustrates how they are calculated and gives tables of critical h and k values.

— Annex B lists the calculation formulae for repeatability and reproducibility. It also describes how to generate and use six tables that are required for a spreadsheet precision analysis.

— Annex C outlines the procedure for calculating replacement values for outliers that have been rejected by h and k value analysis. Outlier replacement rather than deletion is an option that may be used for precision determination with a minimum number of laboratories and/or materials.

— Annex D is an example of a typical general precision determination programme for Mooney viscosity testing. It shows how a precision database is reviewed for outliers, using both the h and the k statistics, and illustrates some of the problems with outlier identification and removal as described in ISO 5725-2.

— Annex E presents some background on ISO 5725, robust analysis and other issues related to precision determination.

Annex E is given mainly as background information that is important for a full understanding of precision determination. Annexes A, B, and C contain detailed instructions and procedures needed to perform the operations called for in various parts of this Technical Report. The use of these annexes in this capacity avoids long sections of involved instruction in the main body of the Technical Report, thus allowing better understanding of the concepts involved in the determination of precision.

Rubber and rubber products — Determination of precision for test method standards

1 Scope

This Technical Report presents guidelines for determining, by means of interlaboratory test programmes (ITPs), precision for test method standards used in the rubber manufacturing and the carbon black industries. It uses the basic one-way analysis of variance calculation algorithms of ISO 5725 and as many of the terms and definitions of ISO 5725 as possible that do not conflict with the past history and procedures for precision determination in these two industries. Although bias is not determined in this Technical Report, it is an essential concept in understanding precision determination. The ISO 5725 concepts of accuracy and trueness are not determined in this Technical Report.

Two precision determination methods are given that are described as "robust" statistical procedures that attempt to eliminate or substantially decrease the influence of outliers. The first is a "level 1 precision" procedure intended for all test methods in the rubber manufacturing industry and the second is a specific variation of the general precision procedure, designated "level 2 precision", that applies to carbon black testing. Both of these use the same uniform level experimental design and the Mandel h and k statistics to review the precision database for potential outliers. However, they use slight modifications in the procedure for rejecting incompatible data values as outliers. The "level 2 precision" procedure is specific as to the number of replicates per database cell or material-laboratory combination.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: Probability and general statistical terms*

ISO 5725 (all parts), *Accuracy (trueness and precision) of measurement methods and results*

3 Terms and definitions

3.1 General

For the purposes of this document, the terms and definitions given in 3.3 apply, together with those in ISO 5725 with modifications in 3.2.

Additional terms concerning certain types of precision can be found in 5.3. Better understanding can be gained by giving these definitions, which relate to the nature of the material to be tested, in that subclause.

3.2 ISO 5725 terms

Terms defined in ISO 5725, usually those from ISO 3534-1, are used when:

a) their definition does not conflict with the procedures required for a comprehensive treatment of precision determination for TC 45 test method standards, and

b) when they are adequate to the task of giving definitions that are informative and promote understanding.

In this subclause, some additional notes have been added to the ISO 5725 term definitions to give greater insight into precision determination for TC 45 test methods.

3.2.1
accepted reference value
value that serves as an agreed-upon reference for comparison and which is derived as:

a) a theoretical or established value, based on scientific principles;

b) an assigned or certified value, based on experimental work of some national or international organization;

c) a consensus or certified value, based on collaborative experimental work under the auspices of a scientific or engineering group;

d) when a), b) and c) are not available, the expectation of the (measured) quantity, i.e. the mean of a specified population of measurements.

3.2.2
test result
value of a characteristic obtained by carrying out a specified test method

NOTE The test method should specify that one or a number of individual measurements, determinations or observations be made and their average or another appropriate function (median or other) be reported as the *test result*. It may also require standard corrections to be applied, such as correction of gas volumes, etc.

3.2.3
accuracy
closeness of agreement between a test result and the accepted reference value

NOTE The term accuracy, when applied to a set of test results, involves a combination of random components and a common systematic error or bias component.

3.2.4
bias
difference between the expectation of the test results and an accepted reference value

NOTE Bias is the total systematic error (deviation) as contrasted to random error. There may be one or more systematic error components contributing to bias. A larger systematic difference from the accepted reference value is reflected by a larger bias.

3.2.5
laboratory bias
difference between the expectation of the test results from a particular laboratory and an accepted reference value

3.2.6
precision
closeness of agreement between independent test results obtained under stipulated conditions

NOTE 1 Precision (for within-laboratory conditions or repeatability) depends on the distribution of random errors and does not relate to the true value (accepted reference value) or the specified value. For a global testing domain (between-laboratory conditions), see 3.3.1 below, the between-laboratory precision (reproducibility) is influenced by laboratory bias as well as the random variations inherent in such a global testing domain.

NOTE 2 The measure of precision is usually expressed in terms of the imprecision and computed as a standard deviation of the test results. Less precision is reflected by a larger standard deviation.

NOTE 3 The term "independent test results" is defined as a set of results where the measurement of each value (of the set) has no influence on the magnitude of any other test result in the set.

NOTE 4 Quantitative measures of precision depend critically on the stipulated conditions (the type of test domain). Repeatability and reproducibility conditions are particular sets of extreme conditions.

NOTE 5 Alternatively, precision may be defined as a "figure of merit" concept. It is proportional to the inverse of the dispersion of independent replicate (test or observed) values, as estimated by the standard deviation, for a specified testing domain.

3.2.7
repeatability conditions
conditions where independent test results are obtained with the same method on identical test items (or elements) in the same laboratory by the same operator using the same equipment within short intervals of time

NOTE As defined in 3.3.1, a "local test domain" is the locale or environment (in a particular laboratory) under which repeatability tests are conducted. The word "identical" should be interpreted as "nominally identical", i.e. no intentional differences among the items. The "intervals of time" between repeat measurement of test results may be selected by the consensus of a particular testing community. For TC 45 and the international rubber manufacturing industry, the time interval between repeat tests is of the order of one to seven days.

3.2.8
repeatability
precision under repeatability conditions

NOTE 1 Repeatability, defined by the symbol r, is expressed in terms of an interval or range that is a multiple of the standard deviation; this interval should (on the basis of a 95 % probability) encompass duplicate independent test results obtained under the defined local testing domain.

NOTE 2 Relative repeatability, (r), is expressed in terms of an interval (a multiple of the standard deviation) that is a percentage of the mean level of the measured property; this interval should (on the basis of a 95 % probability) encompass duplicate independent test results (on a percentage basis) obtained for a defined local testing domain.

NOTE 3 Repeatability may be dependent on the magnitude or level of the measured property and is usually reported for particular property levels or materials or element classes (that determine the level).

NOTE 4 Although repeatability as defined above applies to a local testing domain, it can be obtained in two different ways and the term repeatability can be used in two different contexts. It can pertain to a common community value, obtained as an average (or pooled) value from all laboratories in an ITP among N different laboratories. This can be referred to as a universal or *global* repeatability, that applies to a "typical laboratory", that stands as a representative of all laboratories that are part of a global testing domain. It can also pertain to the long-term or established value for a "particular laboratory" as derived from ongoing testing in that laboratory, not related to any ITP. The second use can be referred to as a *local* repeatability, i.e. repeatability obtained in and for one laboratory.

3.2.9
reproducibility conditions
conditions where test results are obtained with the same method on identical test items (or elements) in different laboratories with different operators using different equipment

NOTE 1 Each laboratory (or location) in the global testing domain, see 3.3.1.5, conducts n repeatability tests on a material (target material) and reproducibility is determined based on the mean values (of the n local domain tests) for the N laboratories for that material. Reproducibility may also depend on the level of the measured property or on the materials tested and it is also usually reported for particular levels or materials.

NOTE 2 The term “different equipment” should be interpreted as different realizations of an accepted and standard test device, i.e. all of the test devices are nominally identical but they are located in different laboratories.

3.2.10
reproducibility
precision obtained under reproducibility conditions

NOTE 1 Reproducibility, R, (for a defined global testing domain) is obtained by way of independent tests conducted in N laboratories (with n replicates each) on nominally identical test items or elements, expressed in terms of an interval or range that is a multiple of the standard deviation; this interval should (on basis of a 95 % probability) encompass duplicate test results, each obtained in different laboratories for a defined global testing domain.

NOTE 2 Relative reproducibility, (R), is expressed in terms of an interval (a multiple of the standard deviation) that is a percentage of the mean level of the measured property; this interval should (on the basis of a 95 % probability) encompass duplicate independent test results (on a percentage basis) each obtained in different laboratories for a defined global testing domain.

NOTE 3 Reproducibility may also depend on the level of the measured property or on the materials tested and it is also usually reported for particular levels or materials. Reproducibility usually does not have the dual interpretation or use as discussed above for repeatability, since it is a “group characteristic” that only applies across a number of laboratories in a global testing domain.

NOTE 4 As indicated in Note 1 in the definition of precision above, reproducibility is determined by the magnitude of random variations in the global testing domain as well as the distribution of bias components in this same global domain. Laboratories that have good agreement with either a reference value or an overall mean value for the ITP, have either zero or a very small bias. Laboratories that do not have good mean value agreement have substantial biases and, although the bias magnitude is relatively constant for each laboratory, it differs among the biased laboratories, i.e. it has the characteristics of a distribution.

3.2.11
outlier
member of a set of values which is inconsistent with the other members of that set

NOTE This TC 45 standard defines a “set” as a “class of elements” that are subjected to measurement. See *element* and *element class* defined in 3.3.1 below.

3.3 Required terms not in ISO 5725

A number of specialized terms are defined here in a systematic sequential order, from simple terms to complex terms. This approach allows the simple terms to be used in the definition of the more complex terms; it generates the most succinct and unambiguous definitions.

3.3.1 Basic testing terms

3.3.1.1
element
entity that is tested or observed to determine a property or characteristic; it may be a single object among a group of objects (test pieces, etc.) or an increment or portion of a mass (or volume) of a material

NOTE The generic term *element* has a number of synonyms: item, test piece, test specimen, portion, aliquot part, sub-sample, laboratory sample.

3.3.1.2
element class
class of elements)
category or descriptive name for a group of elements that have a common origin or have nominally identical properties

NOTE The term “nominally identical” implies that the elements come from a source that is as homogeneous as possible with regard to the property being measured.

3.3.1.3
testing domain
location and operational conditions under which a test is conducted; it includes a description of the element preparation (test sample or test piece), the instrument(s) used (calibration, adjustments, settings), the selected test technicians and the surrounding environment

3.3.1.4
local testing domain
domain comprised of one location or laboratory as typically used for quality control and internal development or evaluation programmes

3.3.1.5
global testing domain
domain that encompasses two or more locations or laboratories, domestic or international, typically used for producer-user testing, product acceptance and interlaboratory test programmes

3.3.1.6
balanced uniform level design

plan for an interlaboratory test programme (ITP) for precision, where all laboratories test all the materials selected for the programme and each laboratory conducts the same number of repeated tests, n, on each material.

3.3.2 Material and sampling terms

3.3.2.1
material
specific entity or element class to be tested; it usually exists in bulk form (solid, powder, liquid)

NOTE Material is used as a generic term to describe the “class of elements” that is tested, i.e. a material may be a rubber, a rubber compound, a carbon black, a rubber chemical, etc. A material may or may not be homogeneous. In product testing, the term material may be used to describe the “class of elements” or type of rubber product such as O-rings, hose assemblies, motor mounts, etc. See also the definition of “target material” in 5.3.

3.3.2.2
lot
specified mass or volume of material or number of objects; usually generated by an identifiable process, frequently with a recognized composition or property range

NOTE A lot may be generated by a common production (or natural) process in a restricted time period and usually consists of a finite size or number. A lot may be a fractional part of a population. A recognized property range implies that some rough approximation is available.

3.3.2.3
sample
⟨physical sample⟩ number of elements or the specified mass of a material, selected according to a particular procedure, used to determine material, lot or population characteristics

NOTE The term “sample” should not be used as a synonym for “material”, or “target material”, see 5.3. Ideally several “materials” are tested in any ITP with each material being different (chemically, structurally, property wise). From each material, some number of “samples” (all nominally identical) may be taken for testing.

3.3.2.4
sample
⟨data⟩ number of test or observation values (n = 1, 2, 3, etc.) obtained from (one or more) physical samples by the application of a specific test (observation) method

3.3.2.5
test sample
part of a (physical) sample of any type taken for chemical or other analytical testing, usually with a prescribed blending or other protocol

NOTE A test sample is usually a mass or volume that is some very small fractional part of a bulk material.

3.3.2.6
test piece
object (appropriately shaped and prepared) taken from a sample (or lot) for physical or mechanical testing

NOTE The term "test specimen" is a synonym for test piece.

3.3.3 Additional statistical terms relating to precision

3.3.3.1
replicate
one of a selected number of independent fractional parts or independent number of elements, taken from a sample; each fractional part or element is tested.

NOTE The word replicate as defined above refers to a physical object (element). It can also be used in reference to a data set, where it refers to one of a number of independent data values.

3.3.3.2
true value
measured or observed value for an element, that would be obtained for a testing domain in the absence of errors, deviations or variations of any sort, i.e. where there is no "system-of-causes" variation

NOTE The true value is also defined as the mean that would be obtained by testing all members of any population. Typical "systems of causes" are the unavoidable fluctuations in temperature, humidity, operator technique, fidelity of calibration, etc. in a controlled testing domain.

3.3.3.3
uncertainty
quantity that characterizes, in an inverse manner, the "figure of merit" for a measurement or observation; for a given local domain, it is the magnitude of the difference between the measured element value and an accepted reference value and includes both random and bias deviations

NOTE The definition of "uncertainty" given here attempts to capture the general nature of the concept. It has been defined equivalently, but using different words, by a number of organizations addressing this concept. The word "uncertainty" as defined here is distinguished from the ordinary use of the word. As indicated, "goodness" or "merit" and "uncertainty" (doubt about the measurement) are inversely related. Uncertainty is a characteristic of a *local test domain*; each local domain for any defined test may have a different uncertainty value. Precision as determined by a typical ITP (both repeatability and reproducibility) is a characteristic of a *global test domain*; the precision values obtained in any ITP are intended for universal application, i.e. to a number of laboratories as a group.

4 Field of application

4.1 General background

This Technical Report applies to test methods that have test results expressed in terms of a quantitative continuous variable. It is in general limited to test methods that are fully developed and in routine use in a number of laboratories.

Tests are conducted using standard test methods to generate test data that are used to make technical and other decisions for commercial, technical and scientific purposes. Therefore the precision of a particular test method is an important quality characteristic or figure of merit for a test method and a decision process.

A determination of the precision of a test method is normally conducted with (1) some selected group of materials typically used with that method and (2) with a group of volunteer laboratories that have experience with the test method. The determination represents an "event in time" for the test method for these materials and laboratories. Another ITP precision determination with somewhat different materials or even with the same materials with the same laboratories at a different time may generate precision results that differ from the initial ITP.

The confidence intervals for the estimated values for repeatability and reproducibility standard deviations is addressed in ISO 5725 and is not part of this Technical Report. The treatment of precision parameter confidence intervals in ISO 5725 assumes the inherent variation in individual values for both repeatability and reproducibility standard deviations (in a long run series of evaluation programmes), is attributable to random test data variations with a normal distribution. Experience as indicated in References [1], [2], [3] and [4] and elsewhere has shown, however, that the poor reproducibility among the laboratories of a typical ITP is due to interlaboratory bias. Certain laboratories are almost always low or high compared to a reference as well as other laboratories in all tests. This offset or bias is typically different for each laboratory that has such a bias. This is in distinction to random deviations compared to a reference as required by a normal distribution. Thus any confidence intervals calculated for the important precision parameter reproducibility, based on a random model, are not valid.

Caution is urged in applying precision results of a particular test method to product testing for consumer-producer product acceptance. Product acceptance procedures should be developed on the basis of precision data obtained in special programmes that are specific to the commercial products and to the laboratories of the interested parties for this type of testing.

An additional concept related to test method technical merit is "test sensitivity". Test sensitivity is defined as the ratio of the test discrimination power for the fundamental property measured, to the property measurement error or standard deviation.

4.2 Defining repeatability and reproducibility

Repeatability and reproducibility are each equal to a range or interval that is a special multiple of the respective standard deviation. The repeatability, designated r, is given by:

$$\text{Repeatability} = r = \phi 2^{1/2} s_r \tag{1}$$

where s_r = the pooled (across all laboratories) "within-laboratory" standard deviation,

and reproducibility, designated R, is given as:

$$\text{Reproducibility} = R = \phi (2)^{1/2} s_R \tag{2}$$

where s_R = the square root (or standard deviation) of the sum of (1) the between-laboratory variance (using the mean of n values for each laboratory for the calculation) and (2) the pooled within-laboratory variance (variance for the n values in each laboratory).

The term $(2)^{1/2}$ is required since r and R are defined as the maximum difference between two single test results that can be expected on the basis of a chance or random occurrence alone at the 5 % probability level or 95 % confidence level. The variance of the difference $(x_1 - x_2)$ for two values taken at random from a population is equal to the sum of the variances for values (of x) taken one at a time from the same population. Since there are two x values, the sum of the variances is simply the variance of x values times two and the square root places this term on a standard deviation basis. In this context each x value represents a "test result" as defined in any particular test method standard.

Thus $(2)^{1/2} s_R$ is the standard deviation of differences. The factor ϕ depends on both the total degrees of freedom in the estimation for either of the standard deviations and on the shape of the distribution of the variable bias terms and the random error terms. The normal assumptions for these are (1) the distributions are unimodal, (2) the number of test results is sufficient (approximately 20) and (3) a probability level of p = 0,05 or confidence level of 95 % is chosen. Under these assumptions the value of ϕ is similar to a t-value or approximately 2,0 and therefore the simplified expressions for r and R are

Repeatability = r = 2,83s_r (3)

Reproducibility = R = 2,83s_R (4)

For more details, see the discussion notes in the definitions for repeatability and reproducibility in 3.1.

5 Precision determination: Level 1 precision and level 2 precision

5.1 Level 1 precision

Two precision categories are described: level 1 precision and level 2 precision. Level 1 precision is discussed first. Level 1 precision determination follows established procedures used in the rubber manufacturing industry on an international basis for the past two decades. The determination is conducted using a balanced uniform level design ITP with three or more materials sent to each of the participating laboratories with tests conducted to generate an independent “test result”, on each of two test days. The ITP database is reviewed for outliers by the Mandel h and k consistency statistics (see Annex A).

a) *Options for outliers* — If no outliers are found, the original database is used to develop a table of precision results. If outliers are identified in any ITP database, there are two options for outlier treatment. Option 1, outlier deletion, is the first choice. Option 2, outlier replacement, is chosen for an ITP with a minimum number of laboratories (ca. six). Issues such as the number of replicate values on each test day and/or the number of technicians or operators used to obtain a test result, which are characteristic of the particular test, are considered on a case-by-case basis by the ITP organizing committee. Outlier treatment is discussed in greater detail in Annexes A, C, D and E.

b) *Types of test method* — Level 1 precision has been successfully used for the broad range of test methods characteristic of the rubber manufacturing industry; from simple “bench type” tests, conducted in few minutes (hardness and pH tests) to a complex multi-step test method, such as an ageing test. Such a test requires preliminary property measurement, a substantial ageing period (days) followed by property measurement after ageing to obtain a final calculated test result or performance index. For such complex tests, any realistic precision determination must include all of the procedural steps in arriving at the test result, the basic datum used in precision analysis and determination. The procedures required for general precision are described in Clauses 8, 9 and 10.

5.2 Level 2 precision

The carbon black industry has adopted a slightly revised precision determination procedure designated “level 2 precision”. The number of replicates in each cell of a uniform level design ITP is specified as four, two by each of two test technicians. The outliers are reviewed by a special procedure that depends on the number of laboratories in the ITP and the precision, absolute or relative, is expressed by a specified procedure. The procedures for this precision are listed in Clause 11.

5.3 Types of level 1 and level 2 precision

In addition to the ageing tests cited above, other tests also require a more complex total sequence of operations to generate a final test result. One important test of this type is a “performance-in-rubber” test; the evaluation of various rubbers, reinforcement fillers or other compounding materials in standardized formulations. The typical stress-strain evaluation of a lot of a specified rubber will require:

a) a representative sample of the rubber;

b) a standardized formulation and mixing operation to prepare a compound using standard materials;

c) processing of this compound to prepare cured moulded sheets for a selected time and temperature;

d) cutting and gauging of dumbbell (or other) test pieces;

e) the testing of these to obtain the final test results for modulus, elongation and tensile strength properties.

To permit realistic precision determination for performance-in-rubber testing, it is necessary that *all the steps* in the operation be replicated, starting from the raw materials to the final test result. Each of these steps has a potential component of variance and the sum of all variance components establishes the overall test variance and standard deviation. To address this, two types of precision are defined. The two types are characterized by the relationship between the material (or element class) *tested* and the material *directly* evaluated for precision. To explain this, it is necessary to introduce and define a new term:

— *target material*: the material (or class of elements) that is the primary *focus of attention* for a precision determination programme; however it may not be tested in its usual or ordinary physical state.

Using the term “target material”, two types of precision may be defined:

— *Type 1 precision* — A precision determined *directly* on, a target material; prepared test pieces or test portions of the target material (class of elements) drawn from a homogeneous source are tested, with no processing or other operations required prior to testing.

NOTE 1 An example is a lot comprised of died-out, gauged dumbbells for stress-strain testing.

— *Type 2 precision* — A precision determined *indirectly* for a target material; the target material is usually combined with a number of homogeneous ancillary materials to form a composite material and testing is conducted on samples of this and the property response of the target material is determined.

NOTE 2 The properties of the composite material are directly related to the quality or properties of the target material. An example: To determine the quality of a grade of SBR, a sample of the rubber, plus curatives, fillers, antioxidants, etc., are mixed and cured, test pieces are prepared and the resulting compound tested for specified quality properties.

NOTE 3 It is possible that a type 1 precision programme might be conducted on test pieces or portions that require some minimum processing or other simple operations prior to actual testing. This is, in a strict sense, an intermediate level of precision. However, to avoid unnecessary complications, this will be designated a type 1 precision.

6 Steps in organizing an interlaboratory test programme

The steps required to organize an ITP, with a discussion for each procedural step, are as follows:

a) *Organization committee* — An organization committee or task group and a programme co-ordinator should be selected. One member of the committee or group should be a statistician familiar with the technology of the test method as well as the content of this Technical Report. Most ITPs are organized on the basis of a *balanced uniform level design* for the precision programme. For more advanced designs, see ISO 5725.

b) *Category and type of precision* — For all programmes except for carbon black testing, a level 1 precision ITP is organized. For carbon black testing a level 2 precision ITP is organized. The type of precision to be determined shall be selected (see 5.3). Type 1 precision is the most frequently determined. For some test methods, such as rubber or polymer or other performance-in-rubber evaluations using standard formulations, a type 2 precision is required.

c) *Test operator or technician selection* — For simple level 1 precision testing requiring only one operator or technician, all replicate tests should be conducted by the same technician unless the effect of different technicians is part of the intended programme. For more complex tests where several operators or technicians are required to perform a sequence of different steps to arrive at a test result, the same “operator team” should conduct testing for all replicates. For level 2 precision testing, follow the procedure of using two technicians on each of two test days (see Clause 11).