



TAPPI/ANSI T 1200 sp-14

TENTATIVE STANDARD – 1961
OFFICIAL STANDARD – 1969
RECOMMENDED PRACTICE – 1986
REAFFIRMED – 1991
REVISED AND COMBINED WITH T 1206 – 2000
CORRECTION – 2006
REVISED – 2007
REVISED – 2014
© 2014 TAPPI

The information and data contained in this document were prepared by a technical committee of the Association. The committee and the Association assume no liability or responsibility in connection with the use of such information or data, including but not limited to any liability under patent, copyright, or trade secret laws. The user is responsible for determining that this document is the most recent edition published.

CAUTION:

This Test Method may include safety precautions which are believed to be appropriate at the time of publication of the method. The intent of these is to alert the user of the method to safety issues related to such use. The user is responsible for determining that the safety precautions are complete and are appropriate to their use of the method, and for ensuring that suitable safety practices have not changed since publication of the method. This method may require the use, disposal, or both, of chemicals which may present serious health hazards to humans. Procedures for the handling of such substances are set forth on Material Safety Data Sheets which must be developed by all manufacturers and importers of potentially hazardous chemicals and maintained by all distributors of potentially hazardous chemicals. Prior to the use of this method, the user must determine whether any of the chemicals to be used or disposed of are potentially hazardous and, if so, must follow strictly the procedures specified by both the manufacturer, as well as local, state, and federal authorities for safe use and disposal of these chemicals.

Interlaboratory evaluation of test methods to determine TAPPI repeatability and reproducibility

1. Scope

1.1 This practice describes techniques for conducting and analyzing the results of intralaboratory and interlaboratory studies. The steps described here will result in a good statistical design that provides sound data for formulating a broadly applicable precision statement regarding the performance of a TAPPI test method.

1.2 Two values are considered: (a) repeatability, which is defined as comparison of test results within a laboratory (same material, operator, apparatus, environmental conditions, making tests in the shortest reasonable timeframe); and (b) reproducibility, which is defined as comparison of test results among laboratories (same material, but different operator, apparatus and perhaps environmental conditions).

1.2.1 In the data chain leading to test results there are many possible sources of variation, and one can conduct studies to isolate these other sources, e.g. same laboratory and operator using different instruments or different laboratories using a shared calibration standard etc. For the purposes of TAPPI test methods, all of these sources of variation are to be captured in a reproducibility value.

2. Summary of practice

2.1 Reliable data may come from any of the following sources: existing data from a known source with sufficient supporting information, intralaboratory study, or interlaboratory study.

2.2 To conduct an interlaboratory study, randomized test specimens of each of several materials are distributed to the participating laboratories, along with appropriate instructions and data sheets. Data returned by the laboratories are analyzed one material at a time. (An assumption that different materials will yield similar precision data is NOT made.)

2.3 Tabular, statistical, and graphical diagnostic procedures are used to estimate averages and variance of the method under question and to show unusual features of the data and patterns among materials that may affect the validity

Approved by the Standard-Specific Interst Group for this Practice



Documents

Document Sharing Hub

or form of the precision statement.

2.4 The information is appropriately summarized in a form from which the precision statement may be readily understood and applied by users of the test method.

3. Significance

3.1 The precision statement gives an indication of the level of variability users of a TAPPI test method can expect. Examples of applications of these values are shown in Appendix A.3.

3.2 *TAPPI Standards: Regulations and Style Guide* requires precision statements in TAPPI Official and Provisional Test Methods. Official Test Methods must have both repeatability and reproducibility estimates. Provisional Test Methods must have repeatability estimates.

4. Definitions

4.1 *Test determination* – the value obtained by carrying out the series of operations specified in the test method whereby readings are made on a test specimen.

4.2 *Test result* – the value obtained for one test unit of the sample. The result is either a single test determination or a specified combination of a number of test determinations (e.g., an average). The method must describe the manner in which each test determination is to be made, the number of test determinations to be made, and how these determinations must be combined to give the test results. For example, the test method for paper thickness, T 551 “Thickness of paper and paperboard (soft platen method)” requires one determination for each of 10 specimen strips; the test result for thickness is defined as the average of these 10 test determinations.

4.3 *Repeatability* – an estimated limit below which the difference between two test results is expected to fall 95% of the time, when the results are obtained at a single laboratory. Repeatability reflects the variability inherent in the experimental system used to determine its value(s).

4.4 *Repeatability conditions* – conditions that reflect the most controlled experimental circumstances possible. Major measurement biasing factors should remain constant including: equipment, operator, test procedure, environmental conditions, sample preparation and in a single testing interval using presumably homogeneous samples.

4.5 *Reproducibility* – an estimated limit below which the difference between two test results is expected to fall 95% of the time, when the results are obtained in different qualified laboratories for the same homogeneous source of material. While each laboratory tests samples taken at random from a single source of material, all of the other factors influencing measurements not considered for repeatability contribute to the variability in the results obtained in different laboratories.

4.6 *Reproducibility conditions* – conditions that reflect only the most basic control over measurement-influencing factors. Reproducibility conditions should include qualified personnel in qualified laboratories carrying out the same procedure (test method) on different samples of the same presumably homogeneous material. Qualified laboratories are those having the required properly calibrated apparatus and environmental conditions and trained personnel who conscientiously follow the prescribed test method (see Section 5.3 for further information). Many of the influences that are considered constant in repeatability are not controlled for reproducibility conditions including: equipment, operator, environmental conditions etc.

4.7 The 95% probability level for both repeatability and reproducibility (4.4 and 4.6) is generally accepted for precision estimations and is also chosen by other standard writing organizations (e.g. ASTM and ISO). Thus TAPPI repeatability and reproducibility can be considered compatible with the precision statistics in methods currently published by ASTM and ISO. Repeatability and reproducibility are different from, but related to, the 95% confidence interval for a particular test result. Confidence intervals are often estimated by performing replicates and correcting for the number of replicates using Student’s *t* test. For more information see:

<http://www.itl.nist.gov/div898/handbook/prc/section1/prc14.htm>.

4.8 *Intralaboratory study* – a study designed to estimate the precision (repeatability) of a test method that meets the requirements of Section 5 and is conducted in a single laboratory.

4.9 *Interlaboratory study* – a study designed to estimate the precision (repeatability and reproducibility) of a test method that meets the requirements of Section 5 and is conducted at multiple laboratories.

5. Planning the intra- or interlaboratory study

5.1 *Basic design.* Keep the design as simple as possible in order to obtain estimates that are an accurate reflection of the method in common usage. The basic design is represented by a two-way classification table in which the rows represent the laboratories, the columns represent the materials, and each cell (i.e., the intersection of a row with a column) contains the measurements made by a particular laboratory on a particular material (Table 1 and Table 3).

5.2 *Test method.*

5.2.1 The test method should be reviewed to assure that it has been well-written and subjected to careful investigation at one or more competent laboratories to determine clarity and stability. Furthermore, it should be easy to follow by any properly equipped laboratory with competent personnel having knowledge of the materials and of the property to be tested.

5.2.2 The method should have a clearly defined numerical result which can be derived from a single determination or a combination of multiple determinations. The results are assumed to follow a normal distribution.

5.2.3 The method should specify the care of test specimens, including prominent instructions for preconditioning and conditioning when required (see Section 6.2.4).

5.2.4 The method should specify the need for and frequency of adjustment and calibration of the test apparatus.

5.3 *Laboratories.*

5.3.1 Any laboratory that would be considered qualified to run the test in routine testing situations is permitted and encouraged to participate in the interlaboratory study, but the number of laboratories may be restricted if there is a large amount of work involved in preparing samples or there is a significant increase in sampling variability due to the larger amount of material required.

5.3.2 Laboratories must be properly equipped to follow all details of the procedure and be willing to assign the work to a skilled operator on a timely basis. The decision of a laboratory to participate should be based on information provided by the working group, including information as to the required time for calibrating the apparatus and for testing all of the materials.

5.3.3 Subject to the conditions in 5.3.1 and 5.3.2, use as many laboratories as practical, preferably 20 to 30. Because of the possible loss of data from some laboratories for some materials, it is highly desirable to include at least eight laboratories to obtain a usable estimate of the reproducibility of the test method. In consideration of the referee nature of TAPPI test methods, more than a mere usable estimate of the reproducibility between laboratories is highly desirable. Because the data required from each laboratory are held to a minimum in this practice, to encourage the participation of an adequate number of laboratories. Under no circumstances should a precision statement be based on an interlaboratory trial with fewer than five laboratories.

5.3.4 Avoid restricting the study to a group of laboratories judged to be exceptionally qualified and equipped. Precision estimates for inclusion in a TAPPI test method should be obtained through the efforts of laboratories and personnel operating under conditions that are representative of the situations in which the test method will be used in practice.

5.4 *Materials.*

5.4.1 *Number and types.* The number and types of materials to be included in the interlaboratory study should (1) span the range of the values of the property being measured, (2) span the number of scales to cover the range if the instrument has more than one scale, and (3) reflect the number of types or classes of materials to which the test method is to be applied. If the precision trial is restricted in any of these areas, the omitted information must be noted in the precision statement.

5.4.2 Without previous knowledge concerning the behavior of different classes of material subject to the method, it should be assumed that different materials will exhibit different performance characteristics. When a method covers different classes of materials, it may be desirable to conduct separate trials to suit the expertise of the laboratories involved. Thus the selected materials within each class differ from each other only in level and should not pose difficulties to the participating laboratories.

5.5 *Replicates.*

5.5.1 If the test method defines a test result as the average of a set number of test determinations, the study should consist of at least three times that number of replicate determinations.

5.5.2 If the test method requires a recalibration or some other special procedure between each test result but not between each test determination, each replicate in the study must be a complete test result, and at least three replicate results should be obtained..

6. Conducting an intra- or interlaboratory study

6.1 The working group should prepare a protocol for the participants in the study, which should include detailed instructions for the following:

6.1.1 Any exception to or selection of the test procedure as specified in the test method. Expected date of the testing, especially if the aging of samples may be important (see 6.2.3);

6.1.3 Any special circumstances that must be addressed in implementing the repeatability conditions (one operator, one apparatus, shortest time span);

6.1.4 The recording of measurements, including the number of decimal places to be used, which, when possible, should be one more than is required to be reported by the test method under study in order to allow reliable calculation of the precision of the method;

6.1.5 The keeping of a log of unusual events that arise during any phase of the testing;

6.1.6 Notifying the study coordinator when an error in test procedure arises so that replacement test specimens may be sent.

6.2 *Sampling and sample preparation and distribution.*

6.2.1 *Sampling of materials.*

6.2.1.1 Sample each material so that the variability among the specimens of that material will be minimized.

6.2.1.2 Prepared in accordance with the sampling section of the method, if applicable.

6.2.1.3 For samples of paper, take all specimens from a small area of a single roll, avoiding the edges of the roll. Usually, specimens taken adjacent to each other in the machine direction are more similar than those adjacent in the cross direction. Similarly, composite specimens, e.g., corrugated board, should be sample from sections of the roll or stack that are expected to be homogenous. As a further refinement for physical tests, each sheet or specimen may be weighed individually to weigh each sheet or specimen and remove any that exhibit weights more than two standard deviations from the mean.. Specimens for nondestructive optical tests may be individually measured in one laboratory. Those specimens with test values that exhibit values more than two standard deviations from the mean can be removed from the study.

6.2.1.4 For a granular solid, liquid or similar material representing any given test level, make the material as homogeneous as possible by mixing or other appropriate means prior to its subdivision into portions or specimens that are allocated to the laboratories.

6.2.2 *Preparation of specimens.*

6.2.2.1 Prepare enough specimens from each sample to provide the required test material for all participating laboratories and a significant number of additional specimens for replacement of any lost or spoiled specimens.. Label each specimen by means of a code symbol and record the identification of the specimens for future reference. Completely randomize the specimens of a particular test material before dividing them into the required number of groups for assignment to the laboratories.

6.2.2.2 The complete randomization of specimens, as specified above, ensures that between-specimen variability is the same within laboratories as between laboratories and greatly simplifies the statistical analysis and interpretation. When each laboratory is given a single sheet of each material from which it is expected to obtain all specimens, intralaboratory variability is confounded with between-sheet variability. On the other hand, when blocks of adjacent specimens (which are presumably more alike) are assigned one from each block to each laboratory, the interlaboratory replication error is confounded with block-to-block variability and cannot be compared with the intralaboratory variability, which does not include the block-to-block variability.

6.2.3 *Aging of samples.* If the samples are such that their properties may significantly change in the course of days or a few weeks, coordinate the dates for the testing among the participating laboratories so that the effect of aging is not confounded with the differences among laboratories.

6.2.4 *Conditioning of samples.* If required by the test method under investigation, preconditioning (see TAPPI T 402 “Standard Conditioning and Testing Atmospheres for Paper, Board, Pulp Handsheets, and Related Products”) of the samples at low relative humidity (RH) prior to conditioning and testing at 50% RH will avoid confounding hysteresis effects with the differences among laboratories. The preconditioning and conditioning should be left to each participating laboratory, although a preliminary preconditioning could be done while the test specimens are being randomized.

6.2.5 Allocation and shipment of specimens. Allocate and distribute the specimens from a single place. Ensure that the specimens are packaged properly for arriving without meaningful changes in the property under consideration, and addressed as specified by the participant for prompt delivery to the testing laboratory.

6.3 *Report form and questionnaire.* Supply each participating laboratory with a form for reporting the data to ensure that all pertinent data and information are reported in a uniform manner by all participants. In addition, provide

space for other relevant information, such as relative humidity, temperature, type of instrument, and log of unusual events.

7. Calculating only repeatability from an intralaboratory study (for Provisional Methods)

7.1 *Data display.* Place the data to be used for the calculation of repeatability into a table. Display the materials as columns. Display the determination or result values under the respective materials. Place the calculation results at the bottom of the table. See Table 1 for a generalized example, and Table A1 for a specific example. To assist in conducting these calculations an Excel spreadsheet is available for downloading from the general Standards page on the TAPPI website (go to any TAPPI web page, click on the “Standards” button at the top of the page, then scroll down to the “General Information Document list.”).

7.2 *Data inspection.* The working group reviews the data to ensure that sufficient data have been gathered to form a reasonable precision statement. Because the calculation results will not be combined, there is not a requirement that the data be balanced between the materials, provided that sufficient data are present to adequately fill the requirements of Section 5. For example, the cells for Materials A, B and C may consist of 5 test results, whereas the cell for Material D may consist of only 4 test results due to a testing error.

Table 1. General example of the data display for an intralaboratory trial

<i>Material A</i>	<i>Material B</i>	<i>Material C</i>	<i>Material D</i>
Data 1, Data, 2, Data 3, etc.	Data 1, Data, 2, Data 3, etc.	Data 1, Data, 2, Data 3, etc.	Data 1, Data, 2, Data 3, etc.
calculation results	calculation results	calculation results	calculation results

7.3 *Calculations for repeatability.* Prior to performing the repeatability calculations, the working group members should evaluate the quality of the data (see section 9). For each material, use the equations and definitions shown in Table 2 to calculate means, standard deviations, repeatability, and repeatability ratio. Repeatability is related to the repeatability standard deviation (s_r) of results obtained in accordance with the limitations outlined in Section 4.5. TAPPI repeatability (r), is chosen to be 2.77 times the repeatability standard deviation (s_r) and is an estimate of the maximum difference between two test results at a 95% confidence level. The factor $2.77 = 1.96 \times \text{SQRT}(2)$, where 1.96 is the factor for symmetric 95% limits for a two-sided normal distribution and SQRT(2) results from the fact that one is calculating the difference between two normally distributed test results.

8. Calculating repeatability and reproducibility from an interlaboratory study (for Official Methods)

8.1 *Data display.* Place the data to be used for the calculation of repeatability into a table. The materials shall be displayed as columns. The measurements from the laboratories shall be displayed as rows under the respective materials for each laboratory. Group the data by laboratory, and reserve the rows following each laboratory’s data for calculation results.. Place the summary results at the bottom of the table See Table 3 for a generalized example, and Table A2 for a specific example. To assist in conducting these calculations an Excel spreadsheet is available for downloading from the general Standards page on the TAPPI website (go to any TAPPI web page, click on the “Standards” button at the top of the page, then scroll down to the “General Information Document list.”).

8.2 *Data inspection.* The working group must review the data to ensure that sufficient data have been gathered to form a reasonable precision statement. Because the calculation results will not be combined, there is not a requirement that the data be balanced between the materials provided that sufficient data is present to adequately fill the requirements of Section 5. For example, a laboratory’s submitted data for Materials A,B and C may consist of 5 test results, whereas the cell for Material D contains only 4 test results due to a testing error. Another example would be if a laboratory does not have the capability to test a particular material, so that the precision of Materials A, B, D and E are