INTERNATIONAL STANDARD

ISO 10399

Third edition
2017-12

Sensory analysis — Methodology — Duo-trio test

Analyse sensorielle — Méthodologie — Essai duo-trio

Reference number
ISO 10399:2017(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 34, *Food products*, Subcommittee SC 12, *Sensory analysis*.

This third edition cancels and replaces the second edition (ISO 10399:2004), of which it constitutes a minor revision. The references have been updated, the definition for 3.6 has been replaced and an expression in A.3 has been corrected.

Sensory analysis — Methodology — Duo-trio test

1 Scope

This document specifies a procedure for determining whether a perceptible sensory difference or similarity exists between samples of two products. The method is a forced-choice procedure. The method is applicable whether a difference exists in a single sensory attribute or in several attributes.

The method is statistically less efficient than the triangle test (described in ISO 4120) but is easier to perform by the assessors.

The method is applicable even when the nature of the difference is unknown (i.e. it determines neither the size nor the direction of difference between samples, nor is there any indication of the attribute(s) responsible for the difference). The method is applicable only if the products are fairly homogeneous.

The method is effective for

a) determining that

   1) either a perceptible difference results (duo-trio testing for difference), or

   2) a perceptible difference does not result (duo-trio testing for similarity) when, for example, a change is made in ingredients, processing, packaging, handling or storage, and

b) for selecting, training and monitoring assessors.

Two forms of the method are described:

— the constant-reference technique, used when one product is familiar to the assessors (e.g. a sample from regular production);

— the balanced-reference technique, used when one product is not more familiar than the other.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements 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 5492, *Sensory analysis — Vocabulary*

ISO 8589, *Sensory analysis — General guidance for the design of test rooms*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 5492 and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— IEC Electropedia: available at http://www.electropedia.org/

— ISO Online browsing platform: available at https://www.iso.org/obp

3.1
alpha-risk
α-risk
probability of concluding that a perceptible *difference* (3.3) exists when one does not

Note 1 to entry: This is also known as Type I error, significance level or false positive rate.

3.2
beta-risk
β-risk
probability of concluding that no perceptible *difference* (3.3) exists when one does

Note 1 to entry: This is also known as Type II error or false negative rate.

3.3
difference
situation in which *samples* (3.5) can be distinguished based on their sensory properties

Note 1 to entry: The proportion of assessments in which a perceptible difference is detected between the two products is given the symbol p_d.

3.4
product
material to be evaluated

3.5
sample
unit of *product* (3.4) prepared, presented and evaluated in the test

3.6
sensitivity
< statistic> statistical parameters that measure the performance characteristics of the test

Note 1 to entry: In statistical terms, the sensitivity of the test is defined by the values of α, β and p_d.

3.7
similarity
situation in which any perceptible *differences* (3.3) between the *samples* (3.5) are so small that the *products* (3.4) can be used interchangeably

3.8
triad
three *samples* (3.5) given to an assessor in the duo-trio test

Note 1 to entry: In the duo-trio test, one sample is labelled as the reference, the other two are marked with different codes. One of the coded samples is the same product as the reference; the other coded sample is the other product in the test.

4 Principle

The number of assessors is chosen based on the sensitivity desired for the test (see 6.2 and the discussion in A.3).

Assessors receive a set of three samples (i.e. a triad), one sample of which is labelled as a reference and the other two samples have different codes. The assessors are informed that one of the coded samples is the same as the reference and that one is different. Based on their training and the instructions given prior to the test, the assessors report either which of the coded samples they believe to be same as the reference, or which of the coded samples they believe to be different from the reference.

The number of correct responses is counted and the significance is determined by reference to a statistical table.

5 General test conditions and requirements

5.1 Clearly define the test objective in writing.

5.2 Carry out the test under conditions that prevent communication among assessors until all the evaluations have been completed using facilities and booths that conform with ISO 8589.

5.3 Prepare the samples out of sight of the assessors and in an identical manner (i.e. same apparatus, same vessels, same quantity of product).

5.4 Assessors shall not be able to identify the samples from the way in which they are presented. For example, in a taste test, avoid any differences in appearance. Mask any irrelevant colour differences using light filters and/or subdued illumination.

5.5 Code the vessels containing the samples in a uniform manner, preferably using three-digit numbers, chosen at random for each test. Each triad is composed of three samples, one labelled as the reference and two labelled with different codes. Preferably, different codes should be used for each assessor during a session. However, the same two codes may be used for all assessors within a test, provided that each code is used only once per assessor during a test session (e.g. if several duo-trio tests on different products are being conducted in the same session).

5.6 The quantity or volume served shall be identical for the three samples in each triad, just as that of all the other samples in a series of tests on a given type of product. The quantity or volume to be evaluated may be imposed. If it is not, the assessors should be told to take quantities or volumes that are always similar whatever the sample.

5.7 The temperature of the three samples in each triad shall be identical, just as that of all the other samples in a series of tests on a given type of product. It is preferable to present the samples at the temperature at which the product is generally consumed.

5.8 The assessors shall be told whether or not they are to swallow the samples or whether they are free to do as they please. In this latter case, they shall be requested to proceed in the same manner for all samples.

5.9 During the test sessions, avoid giving information about product identity, expected treatment effects, or individual performance until all testing is completed.

6 Assessors

6.1 Qualification

All assessors should possess the same level of qualification, this level being chosen on the basis of the test objective (see ISO 8586 for guidance). Experience and familiarity with the product may improve the performance of an assessor and, therefore, may increase the likelihood of finding a significant difference. Monitoring the performance of assessors over time may be useful for increased sensitivity.

All assessors shall be familiar with the mechanics of the duo-trio test (i.e. the format, task and evaluation procedure).

6.2 Number of assessors

Choose the number of assessors so as to obtain the sensitivity required for the test (see discussion in A.3). Using large numbers of assessors increases the likelihood of detecting small differences between the products. However, in practice, the number of assessors often is determined by material conditions

(e.g. duration of the experiment, number of available assessors, quantity of product). When testing for a difference, the typical number of assessors is between 32 and 36. When testing for no meaningful difference (i.e. similarity), twice as many assessors (i.e. approximately 72) are needed for equivalent sensitivity.

Avoid replicate evaluations by the same assessor whenever possible. However, if replicate evaluations are needed to produce a sufficient number of total evaluations, every effort should be made to have each assessor perform the same number of replicate evaluations. For example, if only 12 assessors are available, have each assessor evaluate 3 triads to obtain a total of 36 evaluations.

NOTE Treating 3 evaluations performed by 12 assessors as 36 independent evaluations is not valid when testing for similarity using Table A.2. However, the test for difference using Table A.1 is valid even when replicate evaluations are performed[8],[9] Recent publications[6],[7] on replicated discrimination tests suggest alternative approaches for analysing replicated evaluations in discrimination tests.

7 Procedure

7.1 If the product is familiar to the assessors (e.g. a control sample from the production line), use the constant reference technique. If neither product is more familiar than the other, use the balanced-reference technique

a) **Constant-reference technique**: Prepare worksheets and scoresheets (see B.2) in advance of the test so as to utilize an equal number of the two possible sequences of two products, A and B:

   A-REF AB A-REF BA

   Distribute these at random in groups of two among the assessors (i.e. use each sequence once among the first two assessors; use each sequence once again among the next two assessors, etc.) This will minimize the imbalance that results if the total number of assessors is not an even number.

b) **Balanced-reference technique**: Prepare worksheets and scoresheets (see B.1) in advance of the test so as to utilize an equal number of the four possible sequences of two products, A and B:

   A-REF AB A-REF BA

   B-REF AB B-REF BA

   where the first two triads contain product A as the reference (i.e. A-REF) and the last two triads contain product B as the reference (i.e. B-REF). Distribute these at random in groups of four among the assessors (i.e. use each sequence once among the first group of four assessors; use each sequence once again among the next group of four assessors, etc.). This will minimize the imbalance that results if the total number of assessors is not a multiple of four.

7.2 Present the three samples of each triad simultaneously if possible, following the same spatial arrangement for each assessor (e.g. on a line to be sampled always from left to right, in a triangular array). Within the triad, assessors are generally allowed to make repeated evaluations of each sample as desired (if, of course, the nature of the product allows for repeated evaluations).

7.3 Instruct the assessors to evaluate the reference sample first, then evaluate the two coded samples in the order in which they were presented. Inform the assessors that one of the coded samples is the same as the reference and that one is different from the reference. Instruct the assessors to indicate either which of the two coded samples is the same as the reference, or which of the two coded samples is different from the reference.

NOTE When deciding whether to instruct the assessors to select the sample that is the same as the reference or to select the sample that is different from the reference, consideration is given to whether or not the panel routinely uses other discrimination test methods. Many discrimination test methods like the triangle test, for example, focus on identifying the "odd" or "different" sample in the test. Instructing the assessors to identify the "different" sample in one method and to identify the "same" sample in another method can cause confusion and lead to higher levels of incorrect responses.

7.4 Each scoresheet should provide for a single triad of samples. If an assessor is to carry out more than one test in a session, collect the completed scoresheet and unused samples prior to serving the subsequent triad. The assessor shall not go back to any of the previous samples or change the verdict on any previous test.

7.5 Do not ask questions about preference, acceptance or degree of difference after the assessor has made a selection. The selection the assessor has just made may bias the reply to any additional questions. Responses to such questions may be obtained through separate tests for preference, acceptance, degree of difference, etc., see ISO 6658. A comment section asking why the choice was made may be included for the assessor's remarks.

7.6 The duo-trio test is a forced-choice procedure; assessors are not allowed the option of reporting "no difference". An assessor who detects no difference between the samples should be instructed to randomly select one of the samples and to indicate that the selection was only a guess in the comments section of the scoresheet.

8 Analysis and interpretation of results

8.1 When testing for a difference

Use Table A.1 to analyse the data obtained from a duo-trio test. If the number of correct responses is greater than or equal to the number given in Table A.1 (corresponding to the number of assessors and the α-risk level chosen for the test), conclude that a perceptible difference exists between the samples (see B.1).

If desired, calculate a confidence interval on the proportion of the population that can distinguish the samples. The method is described in B.3.

8.2 When testing for similarity

NOTE In this document, "similar" does not mean "identical". Rather, "similar" means that the two products are sufficiently alike to be used interchangeably. It is not possible to prove that two products are identical. However, it can be demonstrated that any difference that does exist between two products is so small as to have no practical significance.

Use Table A.2 to analyse the data obtained from a duo-trio test. If the number of correct responses is less than or equal to the number given in Table A.2 (corresponding to the number of assessors, the β-risk level and the value of p_d chosen for the test), conclude that no meaningful difference exists between the samples (see B.2). If results will be compared from one test to another, then the same value of p_d should be chosen for all tests.

If desired, calculate a confidence interval on the proportion of the population that can distinguish the samples. The method is described in B.3.

9 Test report

Report the test objective, the results and the conclusions. The following additional information is recommended:

— the purpose of the test and the nature of the treatment studied;

— full identification of the samples (i.e. origin, method of preparation, quantity, shape, storage prior to testing, serving size, temperature); sample information should communicate that all storage, handling and preparation was done in such a way as to yield samples that differ only due to the variable of interest, if at all;

— the number of assessors, the number of correct responses and the result of the statistical evaluation (including the values of α, β and p_d used for the test);

— assessors: experience (in sensory testing, with the product, with the samples in the test), age and gender (see ISO 8586 for guidance);

— any information and any specific recommendations given to the assessors in connection with the test;

— the test environment (i.e. the test facility used, simultaneous or sequential presentation, if the identity of samples was disclosed after the test, if so, in what manner);

— the location, date of the test and name of the panel leader.

10 Precision and bias

Because the results of sensory discrimination tests are functions of individual sensitivities, a general statement regarding the reproducibility of results that is applicable to all populations of assessors cannot be made. Precision regarding a particular population of assessors increases as the size of the panel increases and also with training and with exposure to the product.

As a forced-choice procedure is used, results obtained by this method are bias-free, provided that the precautions in Clause 7 are fully observed.