INTERNATIONAL STANDARD

ISO 4120

Third edition
2021-03

Sensory analysis — Methodology — Triangle test

Analyse sensorielle — Méthodologie — Essai triangulaire

Reference number
ISO 4120:2021(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

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 of 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 www.iso.org/iso/foreword.html.

This document was prepared by ISO/TC 34, *Food products*, Subcommittee SC 12, *Sensory analysis,* in collaboration with the European Committee for Standardization (CEN) Technical Committee CEN/SS C01, *Food Products,* in accordance with the Agreement on technical cooperation between ISO and CEN (Vienna Agreement).

This third edition cancels and replaces the second edition (ISO 4120:2004), which has been technically revised. The main changes compared with the previous edition are as follows:

— the document has been generalized beyond food and beverage applications;

— guidance on how to use the Thurstonian model in addition to the previously emphasized guessing model has been added.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Sensory analysis — Methodology — Triangle 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 more efficient than the duo-trio test (described in ISO 10399), but has limited use with products that exhibit strong carryover and/or lingering flavours.

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 homogeneous.

The method is effective for:

a) determining that:

   1) either a perceptible difference results (triangle testing for difference);

   2) a perceptible difference does not result (triangle testing for similarity),

   when, for example, a change is made in ingredients, processing, packaging, handling or storage;

b) selecting, training and monitoring assessors.

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:

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

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

3.1
alpha-risk
***α*-risk**
probability of concluding that a perceptible difference 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 meaningful difference 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 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 prepared, presented and evaluated in the test

3.6
sensitivity
statistical parameters that measure the performance characteristics of the test

Note 1 to entry: 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 between the samples are so small that the products can be used interchangeably

3.8
triad
those three samples given to an assessor in the triangle test

Note 1 to entry: In the triangle test, each sample is marked with a different code. Two of the samples are alike (i.e. from one product) and one is different (i.e. from the other product).

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) and are informed that two of the samples are the same and that one is different. The assessors report which sample they believe to be different, even if the selection is based only on a guess.

The number of correct responses is counted, and the significance is determined by reference to a statistical table or an applicable computer program or app.

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 to ISO 8589.

5.3 Prepare the samples out of sight of the assessors and in an identical manner (e.g. same apparatus, same vessels, same quantities 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 test samples in a uniform manner, preferably using three-digit numbers, chosen at random for each test. Each triad is composed of three samples, each with a different code. Preferably, different codes should be used for each assessor during a session. However, the same three 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 triangle tests on different products are being conducted in the same session).

5.6 It is preferable to present the samples under the conditions at which the product is generally used (e.g. in a taste test, present the samples at the temperature at which the product is generally consumed). The serving conditions of the three samples in each triad shall be identical (e.g. in a taste test, the three samples shall be served at the same temperature), just as that of all the other samples in a series of tests on a given type of product.

5.7 The size, quantity or volume presented 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 size, quantity or volume to be evaluated may be imposed. If it is not, the assessors should be told to take sizes, quantities or volumes that are always similar whatever the sample.

5.8 In a taste test, 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 the samples.

5.9 During the test sessions, avoid giving information about product identity, expected treatment effects, or individual performance until all testing is completed. The only necessary information for the assessor is the nature of the product to be tested and the task to be performed.

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 triangle test (i.e. 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 the 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 is often determined by material conditions (e.g. duration of the experiment, number of available assessors, quantity of product). When testing for a difference, typical numbers of assessors are between 24 and 30. When testing for no meaningful difference (i.e. similarity), twice as many assessors (i.e. approximately 60) 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 10 assessors are available, have each assessor evaluate 3 triads to obtain a total of 30 evaluations.

NOTE Treating three evaluations performed by 10 assessors as 30 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 (see References [8] and [9]). References [4], [5] and [6] on replicated discrimination tests suggest alternative approaches for analysing replicated evaluations in discrimination tests.

7 Procedure

7.1 Prepare worksheets and scoresheets (see B.1 and B.2) in advance of the test so as to utilize an equal number of the six possible sequences of two products, A and B:

ABB	AAB	ABA
BAA	BBA	BAB

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

NOTE The worksheets also can be prepared digitally and made available for the evaluations.

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 samples in the order in which they were presented. Inform the assessors that two of the samples are the same and that one is different. Each assessor shall then indicate which one of the three samples is different from the other two.

7.4 The triangle 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. The assessor has one chance out of three of giving the correct answer randomly.

7.5 The assessor shall not go back to any samples from previous triads or change the verdict on any previous test. If an assessor is to carry out more than one test in a session, it is imperative that the assessor shall not be able to change their response once given. For example, collect the completed scoresheet and unused samples prior to serving the subsequent triad or do not allow the assessor to return to an earlier answer screen once a response is confirmed.

7.6 Do not ask questions about preference, acceptance or degree of difference after the initial selection of the odd sample. 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 for guidance.) A comment section asking why the choice was made may be included for the assessor's remarks.

8 Analysis and interpretation of results

8.1 When testing for a difference

Use Table A.1 to analyse the data obtained from a triangle 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 triangle 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 are to 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 samples that were tested, the results and the conclusions. Include that the test was conducted in accordance with this document, i.e. ISO 4120:2021, and that the data were analysed according to the methods given in Clause 8. 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. test facility used, simultaneous or sequential presentation, if the identity of samples was disclosed after the test and, if so, in what manner);

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

10 Precision and bias

Because results of sensory discrimination tests are a function 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 their training and with exposure to the product (e.g. the precision of the test conducted using assessors drawn from the population of employees of the company that makes the product would likely be greater than the precision of the test conducted using the same number of assessors drawn from the population of naïve consumers).

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

Annex A
(normative)

Tables

A.1 Values given in Table A.1 are the minimum number of correct responses required for significance at the stated α-risk level (i.e. column) for the corresponding number of assessors, n (i.e. row). Reject the assumption of "no difference" if the number of correct responses is greater than or equal to the value in Table A.1.

Table A.1 — Minimum number of correct responses needed to conclude that a perceptible difference exists based on a triangle test

n	α					n	α				
	0,20	0,10	0,05	0,01	0,001		0,20	0,10	0,05	0,01	0,001
6	4	5	5	6	—	27	12	13	14	16	18
7	4	5	5	6	7	28	12	14	15	16	18
8	5	5	6	7	8	29	13	14	15	17	19
9	5	6	6	7	8	30	13	14	15	17	19
10	6	6	7	8	9						
						31	14	15	16	18	20
11	6	7	7	8	10	32	14	15	16	18	20
12	6	7	8	9	10	33	14	15	17	18	21
13	7	8	8	9	11	34	15	16	17	19	21
14	7	8	9	10	11	35	15	16	17	19	22
15	8	8	9	10	12						
						36	15	17	18	20	22
16	8	9	9	11	12	42	18	19	20	22	25
17	8	9	10	11	13	48	20	21	22	25	27
18	9	10	10	12	13	54	22	23	25	27	30
19	9	10	11	12	14	60	24	26	27	30	33
20	9	10	11	13	14	66	26	28	29	32	35
21	10	11	12	13	15	72	28	30	32	34	38
22	10	11	12	14	15	78	30	32	34	37	40
23	11	12	12	14	16	84	33	35	36	39	43
24	11	12	13	15	16	90	35	37	38	42	45
25	11	12	13	15	17	96	37	39	41	44	48
26	12	13	14	15	17	102	39	41	43	46	50

Values in the table are exact because they are based on the binomial distribution. For values of n not in the table, compute approximate values for the missing entries based on the normal approximation to the binomial distribution as follows. Minimum number of responses (x) = nearest whole number greater than:

$$x = (n/3) + z\sqrt{2n/9}$$

where z varies with the significance level as follows: 0,84 for α = 0,20; 1,28 for α = 0,10; 1,64 for α = 0,05; 2,33 for α = 0,01; 3,09 for α = 0,001.

NOTE 1 Values of n < 18 are usually not recommended for a triangle test for a difference.

NOTE 2 Adapted from Reference [10].