INTERNATIONAL STANDARD

ISO 5495

Third edition
2005-11-15

Sensory analysis — Methodology — Paired comparison test

Analyse sensorielle — Méthodologie — Essai de comparaison par paires

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO 5495:2005
https://standards.iteh.ai/catalog/standards/sist/44704a89-240b-4273-9e1d-1d67cab87ce7/iso-5495-2005

Reference number
ISO 5495:2005(E)

© ISO 2005

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO 5495:2005
https://standards.iteh.ai/catalog/standards/sist/44704a89-240b-4273-9e1d-1d67cab87ce7/iso-5495-2005

© ISO 2005

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

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

 © ISO 2005 – All rights reserved

Contents

Page

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO 5495:2005
https://standards.iteh.ai/catalog/standards/sist/44704a89-240b-4273-9e1d-1d67cab87ce7/iso-5495-2005

© ISO 2005 – All rights reserved

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.

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.

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 5495 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 5495:1983), which has been technically revised.

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO 5495:2005
https://standards.iteh.ai/catalog/standards/sist/44704a89-240b-4273-9e1d-1d67cab87ce7/iso-5495-2005

© ISO 2005 – All rights reserved

Sensory analysis — Methodology — Paired comparison test

1 Scope

This International Standard describes a procedure for determining whether there exists a perceptible sensory difference or a similarity between samples of two products concerning the intensity of a sensory attribute. This test is sometimes also referred to as a directional difference test or a 2-AFC test (Alternative Forced Choice). In fact, the paired comparison test is a forced choice test between two alternatives.

NOTE The paired comparison test is the simplest existing classification test since it concerns only two samples.

The method is applicable whether a difference exists in a single sensory attribute or in several, which means that it enables determination of whether there exists a perceptible difference concerning a given attribute, and the specification of the direction of difference, but it does not give any indication of the extent of that difference. The absence of difference for the attribute under study does not signify that there does not exist any difference between the two products.

This method is only applicable if the products are relatively homogeneous.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

The method is effective

a) for determining

ISO 5495:2005
https://standards.iteh.ai/catalog/standards/sist/44704a89-240b-4273-9e1d-1db7cab870c7/iso-5495-2005

— whether a perceptible difference exists (paired difference test), or

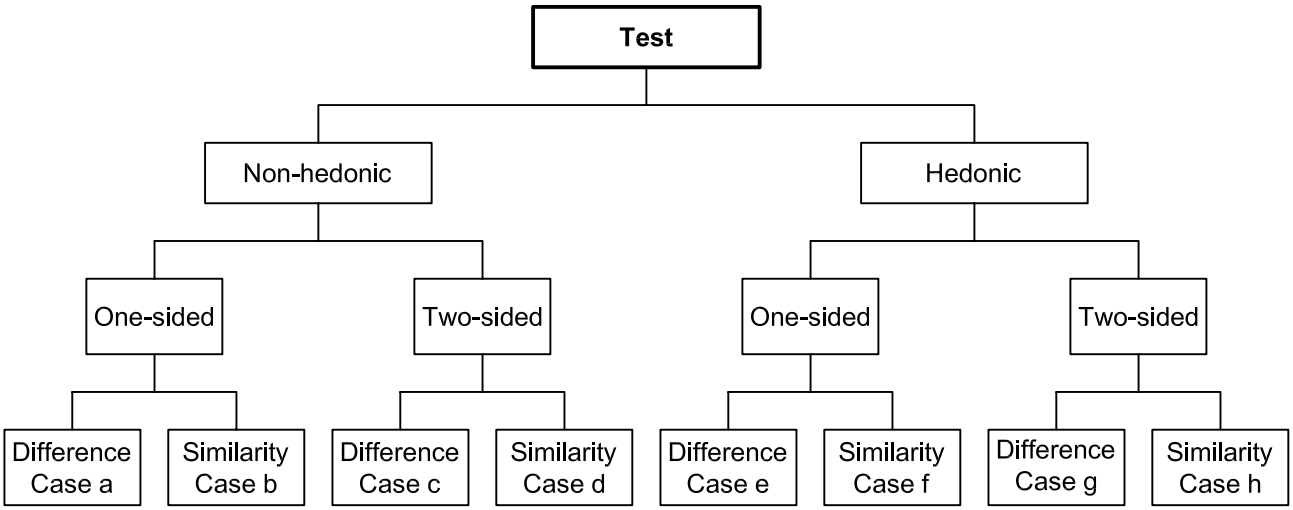
— whether no perceptible difference exists (paired similarity test) when, for example, modifications are made to ingredients, processing, packaging, handling or storage operations, or

b) for selecting, training and monitoring assessors.

It is necessary to know, prior to carrying out the test, whether the test is a one-sided test (the test supervisor knows a priori the direction of the difference, and the alternative hypothesis corresponds to the existence of a difference in the expected direction) or a two-sided test (the test supervisor does not have any a priori knowledge concerning the direction of the difference, and the alternative hypothesis corresponds to the existence of a difference in one direction or the other).

The paired test can also be used in order to compare two products in terms of preference. The different cases of use of the paired test are summarized in Figure 1.

© ISO 2005 – All rights reserved



NOTE Only non-hedonic tests are dealt with in this International Standard.

Figure 1 — Possible different cases of use of the paired comparison test

EXAMPLE 1 (Case a) The production of a biscuit has been modified in order to render it more crisp. It is desired to check whether this increase is perceptible. Therefore it is necessary to try to highlight a difference to see whether the new product is perceived as being crispier than the usual product (control).

iTeh STANDARD PREVIEW
(standards.iteh.ai)

EXAMPLE 2 (Case b) A manufacturer knows that the product may contain traces of an ingredient which imparts an off-flavour to the product. He therefore wishes to determine the maximum acceptable quantity so that the flavour difference with a reference product without this ingredient is barely perceptible and therefore without any regrettable consequences.

ISO 5495:2005
https://standards.iteh.ai/catalog/standards/sist/44704a89-240b-4273-9e1d-1467cab8fce7/iso-5495-2005

EXAMPLE 3 (Case c) It is desired to produce a new soup and to compare two ingredients which will provide the salty flavour. For cost-intensive reasons, the ingredient which, at the same concentration, will provide the strongest salty flavour is sought. Therefore it is necessary to try to highlight a difference. It is not known a priori which ingredient will produce the strongest salty flavour.

EXAMPLE 4 (Case d) A manufacturer of plastics used, in particular, by car manufacturers for dashboards is seeking, for economic reasons, to replace the usual lubricant by a new one, but does not wish that the new plastics formula be perceived as presenting less or more surface slip than the usual one. It is a question of determining whether, for a same concentration, the new lubricant provides the same “surface slip” level as the usual product. It is necessary to show that both lubricants are similar in terms of “surface slip”, but it is not known a priori which lubricant can produce the highest surface slip characteristics.

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 5492 1992, *Sensory analysis — Vocabulary*

ISO 6658:1985, *Sensory analysis — Methodology — General guidance*

ISO 8586-1:1993, *Sensory analysis — General guidance for the selection, training and monitoring of assessors — Part 1: Selected assessors*

ISO 8586-2:1994, *Sensory analysis — General guidance for the selection, training and monitoring of assessors — Part 2: Experts*

ISO 8589:1988, *Sensory analysis — General guidance for the design of test rooms*

© ISO 2005 – All rights reserved

3 Terms and definitions

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

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

NOTE This is also called a type I error, significance level or false-positive rate.

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

NOTE This is also called a type II error or false-negative rate.

3.3
difference
situation in which samples can be distinguished based on their sensory attributes

NOTE The proportion of assessments during which a perceptible difference is detected between the products for the sensory attribute under study is given by the symbol p_d.

3.4
one-sided test
test in which the test supervisor has a priori knowledge concerning the direction of difference

iTeh STANDARD PREVIEW
(standards.iteh.ai)

NOTE The null hypothesis is H_0, the products are not different; the proportion of correct responses observed, p, is equal to 1/2. The alternative hypothesis is H_1, $p > 1/2$.

ISO 5495:2005
https://standards.iteh.ai/catalog/standards/sist/44704a89-240b-4273-9e1d-1d67cab87ce7/iso-5495-2005

3.5
two-sided test
test in which the test supervisor does not have any a priori knowledge concerning the direction of difference

NOTE The null hypothesis is H_0, the products are not different; the proportion of responses observed for one of the samples, p, is equal to 1/2. The alternative hypothesis is H_1, $p \neq 1/2$.

3.6
correct responses
expected responses
number of assessors, in the case of a one-sided test, having selected the sample expected by the test supervisor to be the most intense for the sensory attribute under study

3.7
consensual responses
highest value, in the case of a one-sided test, of the number of assessors having selected sample A and those having selected sample B

NOTE This is calculated as above since there are not any correct responses.

3.8
product
material to be evaluated

3.9
sample
unit of product prepared, presented and evaluated during the course of the test

© ISO 2005 – All rights reserved

3.10
sensitivity
general term employed to summarize the performance characteristics of the test

NOTE In statistical terms, the sensitivity of the test is defined by the values of α, β and p_d.

3.11
similarity
situation in which any perceptible differences between the samples are so small that the products can be used interchangeably

4 Principle

The number of assessors is chosen on the basis of the sensitivity desired for the test (See 6.2 and the footnote that accompanies Tables A.4 and A.5).

The assessors receive a set of two samples (i.e. a pair). They designate the sample which they consider to be the most intense regarding the sensory attribute under consideration, even if this choice is based only on a guess.

NOTE One of the samples may be a control.

The number of times that each sample is selected is counted and the significance is determined by reference to a statistical table, taking into consideration the results obtained for the expected sample (one-sided test) or the highest number of responses obtained for either of the samples (two-sided test).

iTeh STANDARD PREVIEW
(standards.iteh.ai)

5 General test conditions

ISO 5495:2005
https://standards.iteh.ai/catalog/standards/sist/44704a89-240b-4273-9e1d-1d67cab87ce7/iso-5495-2005

5.1 Define the objective of the test in a clear way to determine if the attempt is to be a one-sided or a two-sided test, if it is a difference or similarity test, and which is the most appropriate sensitivity.

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

5.3 Prepare the samples out of sight of the assessors and in an identical manner for each one of them; i.e. same apparatus, same vessels.

5.4 Assessors shall not be able to draw any conclusions regarding the intensity of the attribute from the manner in which the samples are presented to them. For example, for a tactile test, any differences in appearance shall be avoided. Mask all colour differences if the test objective does not concern the colour by using light filters and/or subdued lighting. The samples may also be presented successively and non-simultaneously in the case of slight differences in appearance.

5.5 Code the samples or the vessels containing the samples in a uniform manner, preferably using 3-digit numbers chosen at random for each test. Each pair is composed of two samples, each with a different code. Preferably, different codes should be used for each assessor during a session. However, the two same 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 paired tests on different products are being conducted during the same session).

5.6 The quantity or volume served shall be identical for the two samples constituting each pair, 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 assessed can be imposed. If it is not, it should however be specified to the assessors to take quantities or volumes that are always similar whatever the sample.

5.7 The temperature of the samples constituting each pair 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.

© ISO 2005 – All rights reserved

5.8 The assessors shall be told whether or not they have to follow a special protocol in order to assess the products (e.g. whether or not to swallow the samples for a taste test, or carry out a specific gesture for a tactile test) or whether they are free to do as they please. In this latter case, they should 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 tests are 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 8586-1 and ISO 8586-2). Experience and familiarity with the product can increase the performance of an assessor and can consequently increase the likelihood of finding a significant difference. Monitoring the performance of assessors over time may prove to be useful for increased sensitivity.

All assessors shall be familiar with the mechanisms of the paired test (the scoresheet, the task and the evaluation procedure). In addition, assessors shall be capable of recognising the sensory attribute on which the test is based. This attribute shall be defined verbally, by means of a reference substance or by presenting a few samples having different levels of intensity for the attribute under examination.

6.2 Number of assessors

iTeh STANDARD PREVIEW
(standards.iteh.ai)

Choose the number of assessors so as to obtain the level of sensitivity required for the test (see Table A.4 for a one-sided test and Table A.5 for a two-sided test). The use of a large number 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 conducting a difference test, the number of assessors is typically approximately 24 to 30. When conducting a similarity test, about twice as many assessors (i.e. approximately 60) are required for equivalent sensitivity. When testing for similarity, evaluations should not be replicated by the same assessors. For a difference test, replications may be considered but should still be avoided whenever possible. However, if replicate evaluations are required in order to produce a sufficient total number of 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 perform three paired tests in order to obtain a total of 30 evaluations.

ISO 5495:2005
https://standards.iteh.ai/catalog/standards/sist/44704a89-240b-4273-9e1d-1d67cab87ce7/iso-5495-2005

NOTE Analysing three evaluations performed by 10 assessors as 30 independent evaluations is not valid when testing for similarity using Table A.3. However, the difference test using Tables A.1 and A.2 is valid even when replicate evaluations are performed [5], [6]. Some recent publications [1], [2] on replicated discrimination tests suggest alternative approaches for analysing replicated evaluations.

7 Procedure

7.1 Prepare the worksheets and scoresheets (see Figures B.1, B.2 and B.3) prior to conducting the test so as to use an equal number of the two possible presentation sequences of both products, A and B.

7.2 Present the two samples constituting a pair successively or simultaneously (see 5.4). In the case of simultaneous presentation, arrange the two samples in the same manner for each assessor (in line from left to right, in line from the bottom up, etc.). The assessors shall examine the two samples constituting the pair in the order indicated in the scoresheet, but assessors are generally authorized to make repeated evaluations of each sample if so wished (if, of course, the nature of the product allows for repeated evaluations).

7.3 Provision should be made for one scoresheet per pair of samples. If an assessor is to perform more than one test during the course of a session, collect the completed scoresheet and the unused samples prior

© ISO 2005 – All rights reserved

to serving the subsequent pair. The assessor can neither go back to any of the previous samples, nor modify his/her verdict concerning any of the previous tests.

7.4 Do not ask any questions about preference, acceptance or degree of difference following the selection of the most intense sample. The selection the assessor has just made may bias the response to any additional questions. Responses to such questions may be obtained through separate tests concerning preference, acceptance, degree of difference, etc. (see ISO 6658). A “Comments” section requesting the reasons for the choice may be included for the assessors' remarks.

7.5 The paired test is a “forced choice” procedure; assessors are not allowed to choose the “no difference” option. An assessor who detects no difference between the samples should be instructed to 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

8.1.1 Case of a one-sided test

Use Table A.1 to analyse the data obtained from a paired 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 to 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 able to distinguish the samples. This method is described in B.5.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

No conclusion should be drawn for maximum numbers of correct responses under $n/2$.

ISO 5495:2005
https://standards.iteh.ai/catalog/standards/sist/44704a89-240b-4273-9e1d-1d67cab87ce7/iso-5495-2005

8.1.2 Case of a two-sided test

Use Table A.2 to analyse the data obtained from a paired test. If the number of consensual responses is greater than or equal to the number given in Table A.2 (corresponding to the number of assessors and to the α-risk level chosen for the test), conclude that a perceptible difference exists between the samples (see B.3).

If desired, calculate a confidence interval on the proportion of the population able to distinguish the samples. This method is described in B.5.

8.2 When testing for similarity [1)]

8.2.1 Case of a one-sided test

Use Table A.3 to analyse the data obtained from a paired test. If the number of correct responses is less than or equal to the number given in Table A.3 (corresponding to the number of assessors, to the β-risk level and to the value of p_d chosen for the test), conclude that no meaningful difference exists between the samples (see B.2). If the 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 able to distinguish the samples. This method is described in B.5.

No conclusion should be drawn for maximum numbers of correct responses under $n/2$.

---

1) In this International Standard, “similar” does not mean “identical”. This term signifies rather that the two products are sufficiently alike to be used interchangeably. It is impossible to prove that two products are identical. However, it can be demonstrated that any difference that does exist between two products is so minor as to have no practical significance.

© ISO 2005 – All rights reserved

8.2.2 Case of a two-sided test

Use Table A.3 to analyse the data obtained from a paired test. If the number of consensual responses is less than or equal to the number given in Table A.3 (corresponding to the number of assessors, to the β-risk level and to the value of p_d chosen for the test), conclude that no meaningful difference exists between the samples (see B.4). If the 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 able to distinguish the samples. This method is described in B.5.

9 Report

Report the test objective, the results and the conclusions. It is recommended to add the following additional information:

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

— the complete identification of the samples: origin, method of preparation, quantity, shape, storage prior to testing, quantity served, temperature (the information concerning the sample should indicate that all storage, handling and preparation operations have been carried out so as to yield samples that differ only due to the variable of interest, if any);

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

iTeh STANDARD PREVIEW
(standards.iteh.ai)

— the assessors: experience (in sensory testing, with the product, with the test samples), age and gender (see ISO 8586-1 and ISO 8586-2);

ISO 5495:2005

— any specific information and recommendations given to the assessors in connection with the test, in particular in the case where a precise definition and reference samples illustrating the attribute under test and/or a test protocol have been indicated to the assessors;

— the test environment: test facility used, simultaneous or sequential presentation, whether the identity of the samples was disclosed after the test and if so, in what manner;

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

10 Precision and bias

Because the results of sensory discrimination tests are dependent on individual sensitivities, it is impossible to make a general statement regarding the reproducibility of the results that is applicable to all populations of assessors. The precision regarding a particular population of assessors increases as the size of the panel increases, as well as with the training and exposure to the product.

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

© ISO 2005 – All rights reserved