INTERNATIONAL STANDARD

ISO 8588

Second edition
2017-07

Sensory analysis — Methodology — "A" – "not A" test

Analyse sensorielle — Méthodologie — Essai "A" – "non A"

Reference number
ISO 8588:2017(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 second edition cancels and replaces the first edition (ISO 8588:1987), which has been technically revised. The following changes have been made:

— more detailed explanations of all aspects of the test method have been added;

— the option of testing more than one "not A" sample in a single test has been added;

— statistical calculations are presented in detail for all examples;

— an alternative data analysis procedure that deals directly with the one-sided nature of the "A" – "not A" test has been added.

Sensory analysis — Methodology — "A" – "not A" test

1 Scope

This document specifies a procedure for determining whether a perceptible sensory difference exists between samples of two products. The method applies whether a difference exists in a single sensory attribute or in several.

The "A" – "not A" test can be used in sensory analysis in the following ways:

a) as a difference test, particularly for evaluating samples having variations, for example, in appearance (making it difficult to obtain strictly identical repeat samples) or in aftertaste (making direct comparison difficult);

b) as a recognition test, particularly for determining whether an assessor or group of assessors identifies a new stimulus in relation to a known stimulus (for example, recognition of the quality of the sweet taste of a new sweetener);

c) as a perception test, to determine the ability of an assessor to discriminate stimuli.

The "A" – "not A" test is not appropriate for assessing if two products are sufficiently similar to be used interchangeably (i.e. for similarity testing) because the "A" – "not A" test inherently involves replicate evaluations of the same products by all assessors. These replicate evaluations violate the basic assumptions for similarity tests to be statistically valid.

Examples of its application are given in Annex B.

NOTE Bi and Ennis[1] point out that the estimate of the discriminal distance, d', between the "A" and "not A" samples is the same regardless of the nature of the replicated evaluations performed in the test but that the estimate of the variance of d' does depend on how the replicate evaluations were performed. As such, no general discussion of a Thurstonian analysis of the "A" – "not A" method, nor of the power of the test is undertaken in this document. Interested readers are referred to Reference [1] for a detailed discussion of the topic.

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 3534-1, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*

ISO 5492, *Sensory analysis — Vocabulary*

ISO 8586:2012, *Sensory analysis — General guidelines for the selection, training and monitoring of selected assessors and expert sensory assessors*

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

3 Terms and definitions

For the purpose of this document, the terms and definitions given in ISO 5492, for terms concerning sensory analysis, and ISO 3534-1, for statistical terms, 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 http://www.iso.org/obp

4 Principle

An assessor is presented with a series of samples, some of which are composed of the "A" product while others are composed of one or more "not A" products. For each sample, the assessor indicates whether the sample is an "A" product or is a "not A" product. This test requires the assessor to be familiar with product "A", possibly through exposure to known samples of product "A", prior to exposure to the test samples.

5 Apparatus

The apparatus shall be selected by the test supervisor according to the nature of the product to be analysed, the number of samples, etc. and shall in no way affect the test results.

If standard apparatus corresponds to the needs of the test, it shall be used.

6 Sampling

Refer to sampling standards for the sensory analysis of the product or products being tested.

In the absence of such standards, agreement shall be sought among the parties concerned.

7 General test conditions

7.1 Clearly define the test objective in writing.

7.2 Carry out each session of the test under conditions that prevent communication among assessors until all evaluations have been completed.

7.3 The facilities in which the tests are conducted shall comply with ISO 8589.

7.4 Assessors shall not be able to identify the samples from the way in which they are presented. For example, in a taste test, one should avoid any differences in temperature or appearance. Mask any irrelevant colour differences using, for example, light filters, subdued lighting or opaque serving vessels.

7.5 Code the vessels that contain the test samples in a uniform manner, using 3-digit numbers chosen at random for each sample. Each test sample in a set shall have a different code. The same two codes (one for the "A" sample and one for the "not A" sample) can be used for all assessors within a test session provided different codes are used from one session to another, if multiple sessions are required to complete the test.

7.6 The quantity or volume of product served shall be identical for all test samples. In a taste test, the quantity or volume to be placed in the mouth can be specified. If it is not, assessors shall be instructed to evaluate the same quantity or volume of each test sample.

7.7 The temperatures of the test samples shall be identical, preferably at the temperature at which the product is generally consumed.

7.8 Occupational safety of assessors shall be taken into account. The assessors shall be instructed as to how they should assess the test samples. For example, the assessors shall be instructed whether or not they are to swallow the test samples or whether they are free to do as they please. In the latter case, the assessors shall be instructed to proceed in the same manner for all test samples.

7.9 During the test session, avoid giving information about product identity, expected treatment effects or individual performance until after all testing is complete.

8 Assessors

8.1 Qualification, selection, arrangement

All assessors shall possess the same level of qualification, this being chosen on the basis of the test objective in accordance with ISO 8586:2012. Depending on the objective of the test, assessors may be completely naïve or highly trained. However, within a test, all assessors shall be equally qualified. For example, if the test is being conducted because there is a suspicion that the "not A" product may exhibit a particular taint, assessors with a history of being highly sensitive to the taint may be selected. 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 "A" – "not A" test (the format, the task and the procedure of evaluation).

8.2 Numbers of assessors and assessments

The number of assessors to be used depends on the objective of the test and on the required significance level. Between 10 and 50 assessors who are familiar with the "A" product shall be used in the test. The number of replicate evaluations performed by each assessor shall be determined based on how fatiguing the product is. The total number of evaluations performed in an "A" – "not A" test typically falls between 20 and 100 evaluations.

9 Procedure

To ensure familiarity with product "A", assessors may be exposed to a known "A" sample prior to the evaluations of the test samples. Depending on the objective of the test, assessors also may be exposed to known "not A" samples prior to evaluations of the test samples. For example, if the researcher is concerned that one or more of the "not A" products may exhibit a particular fruity note, the assessors may be exposed to a sample that exhibits supra-threshold intensities of the fruity attribute. The assessors shall not have access to any known samples once the evaluations of the test samples have begun. In addition, in the series presented to the assessor, the respective number of "A" and "not A" samples is unknown to the assessor.

Multiple "not A" products may be evaluated in the same test. The number of "not A" products included in a single test shall be limited to avoid sensory fatigue.

The order of presentation of the "A" and "not A" samples shall be random and the order shall be different for each assessor. All assessors shall be presented with the same number of "A" samples and the same number of "not A" samples (these two numbers not necessarily being the same); see A.2. Similarly, if multiple "not A" products are tested, the numbers of each "not A" product need not be the same; see A.3.

According to the nature of the samples and in order to avoid certain interfering effects of sensory adaptation, the same time interval shall be observed between the presentations of any two successive samples.

Specimen answer forms are reproduced in Annex C.

10 Expression of results

Separate analyses are carried out for each “not A” product in the test. For each “not A” product in the test, the analyst obtains a table of three columns and three rows (see Table 1).

Table 1 — Observed numbers

Assessor’s response	Sample presented is “A”	Sample presented is “not A”	Total
Assessor identifies the sample as “A”	n_{11}	n_{12}	$n_{1.}$
Assessor identifies the sample as “not A”	n_{21}	n_{22}	$n_{2.}$
Total	$n_{.1}$	$n_{.2}$	$n_{..}$

where

n_{11} and n_{22} are the numbers of correct “A” and “not A” responses, aggregated across all assessors, respectively;

n_{21} and n_{12} are the numbers of incorrect “A” and “not A”, responses, aggregated across all assessors, respectively;

$n_{1.}$ and $n_{2.}$ are the sums of the responses in rows 1 and 2, respectively;

$n_{.1}$ and $n_{.2}$ are the sums of the responses in columns 1 and 2, respectively;

$n_{..}$ is the total number of responses.

There are two approaches for analysing the data obtained in an “A” – “not A” test.

In the first approach, the interpretation of results is obtained through a two-step process.

a) If the proportion of times the “A” sample is identified as being “not A” ($n_{21}/n_{.1}$) is greater than the proportion of times the “not A” sample is identified as being “not A” ($n_{22}/n_{.2}$), stop and conclude that there is insufficient evidence to conclude that a perceptible difference exists between the products.

b) Otherwise, compute the χ^2 test statistic, T, in Formula (1) and compare it to the 2α critical value of the χ^2 distribution with 1 degree of freedom. If the value of the test statistic exceeds the critical value, conclude that the samples are perceptibly different. Alternatively, compute the p-value associated with the test statistic in Formula (1) and compare it to the 2α level of significance that has been chosen for the test. If the p-value < 2α, conclude that the samples are perceptibly different.

The 2α action standard is used because the χ^2 test is inherently two-sided. It cannot distinguish between the “A” sample receiving too many or too few “not A” responses. Step a) of the two-step process above rules out the possibility of declaring significance due to the “A” sample receiving too many “not A” responses. Since half of the error is associated with this irrelevant alternative, the true Type I error of the two step test procedure is α when using 2α for the χ^2 test in the second step. Formula 1 is shown as:

$$T=\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\left(n_{ij}-E_{ij}\right)^2}{E_{ij}} \qquad (1)$$

Where n_{ij} is the observed number of counts in row i and column j of Table 1 and E_{ij} is the expected number of counts in row i and column j, which is calculated separately for each of the four cells as $E_{ij} = (n_{i.} \times n_{.j})/n_{..}$. The p-value associated with the test statistic, T, in Formula (1) can be calculated using, for example, a spreadsheet function, such as Excel’s[1)] CHISQ.DIST.RT function. The p-value will be

1) Excel is the trade name of a product supplied by Microsoft. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

displayed in the cell that contains "=CHISQ.DIST.RT(T,1)", where T is the value of the test statistic in Formula (1).

In the second approach, the interpretation of results is obtained through a one-step process.

a) Compute the statistic, T_1, in Formula (2) and compare it to the upper-α critical value of the standard normal distribution. If the value of the test statistic exceeds the critical value, conclude that the samples are perceptibly different. Alternatively, compute the p-value associated with the test statistic in Formula (2) and compare it to the α level of significance that has been chosen for the test. If the p-value < α, conclude that the samples are perceptibly different. Formula 2 is shown as:

$$T_1 = \frac{\sqrt{n_{..}}\,(n_{11}n_{22} - n_{12}n_{21})}{\sqrt{n_{1.}n_{2.}n_{.1}n_{.2}}} \tag{2}$$

The p-value associated with the test statistic, T_1, in Formula (2) can be calculated using, for example, a spreadsheet function, such as Excel's[1] NORM.S.DIST function. The p-value will be displayed in the cell that contains "=1 - NORM.S.DIST(T1,TRUE)", where T_1 is the value of the test statistic in Formula (2).

Note that for sample sizes smaller than those typically occurring in sensory tests, data from an "A" – "not A" test also could be analysed using Fisher's Exact Test.

Some examples are in Annex A.