FINAL DRAFT

INTERNATIONAL STANDARD

ISO/FDIS 11056

ISO/TC **34**/SC **12**

Secretariat: IRAM

Voting begins on:
2020-11-05

Voting terminates on:
2020-12-31

Sensory analysis — Methodology — Magnitude estimation method

Analyse sensorielle — Méthodologie — Méthode d'estimation de la grandeur

Reference number
ISO/FDIS 11056:2020(E)

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 Technical Committee ISO/TC 34, *Food products*, Subcommittee SC 12, *Sensory analysis*.

This second edition cancels and replaces the first edition (ISO 11056:1999), which has been technically revised. It also incorporates the amendments ISO 11056:1999/Amd.1:2013 and ISO 11056:1999/Amd.2:2015. The main changes compared with the previous edition concern the statistical treatment of the examples in Annex B:

— the Assessor factor is considered as fixed factor or as random factor (in the previous edition, the Assessor factor was always considered as a fixed factor);

— the R commands used to process the examples and to obtain the different tables are given explicitly (in the previous edition, only the tables of results were given);

— the numerical examples have been preserved without any modification to allow the user to understand the evolution in the processing of the tables;

— a new example has been added as B.2.

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.

Introduction

Magnitude estimation is a psychophysical scaling technique where assessors assign numerical values to the estimated magnitude of an attribute. The only constraint placed upon the assessor is that the values assigned should conform to a ratio principle, i.e. if the attribute appears to be twice as strong in sample B in comparison with sample A, the value assigned to sample B has to be twice that assigned to sample A. Attributes such as intensity, pleasantness or acceptability may be assessed using magnitude estimation.

Magnitude estimation method is often considered as being less susceptible to “end-effects” than the methods which employ an experimenter-defined continuous or discontinuous response scale. These “end-effects” occur when the assessors are unfamiliar with the extent of the sensations elicited by the products. Then assessors can assign one of the initial samples to a category which is too close to one of the ends of the scale. Consequently, they then find themselves short of graduations and are obliged to classify samples perceived as being different into the same category. This should not occur with magnitude estimation since, in theory, there are an infinite number of categories.

Allowing each assessor to start the process at any numerical value, i.e. to use their own scale, gives rise to a particularly important “assessor” effect. However, there are various ways of solving this problem:

— the analysis of variance (ANOVA) allows the “assessor” effect and the interactions to be taken into account;

— the assessors can be forced to a common scale by use of a reference sample to which a value has been assigned;

— the data supplied by each assessor can be reduced to a common scale by applying one rescaling methods.

It is up to the experimenter to choose the most appropriate approach based on the circumstances.

Magnitude estimation is the privileged method to determine the Steven’s equation psychophysical power function. It can also be used to solve concrete problems.

NOTE The magnitude estimation method is not the most efficient technique for determining small differences between stimuli or for conducting assessments in the vicinity of a detection threshold.

EXAMPLE 1 A company produces a moderately successful beverage, but recent products which are sweeter, produced by a competitor, have made inroads into their shares of the market. It is decided to increase the sweetness level by one third in an attempt to recapture some of the market lost. In formulating the new product, knowing the power function of the sweetener will provide an estimation of the amount of sweetener necessary to reach the one third increase in sweetness level.

EXAMPLE 2 In the formulation of the new diet beverage, the intensity of the desired sweetness is known, but it is not yet decided whether to use aspartame or sucrose as a sweetener. Knowing the power functions of each substance, the iso sweetness lines can be plotted to determine the concentrations of each sweetener necessary for the desired sweetness level. This information coupled with cost/volume information can help inform the decision about which sweetener is more cost effective.

The calculations in Annex B were performed using R functions. Access to R packages is free. This information is given for the convenience of users of this document and does not constitute an endorsement or recommendation by ISO of the exclusive use of R packages. Other software may be used to perform the calculations required by this document.

The files are in the ME folder under the USB DISK H (format Text (separator: tabulation)).

The results can sometimes vary due to rounding errors, depending on the software used.

Sensory analysis — Methodology — Magnitude estimation method

1 Scope

This document specifies a method for applying magnitude estimation to the evaluation of sensory attributes. The methodology specified covers the training of assessors, and obtaining magnitude estimations as well as their statistical interpretation.

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 3534-3, *Statistics — Vocabulary and symbols — Part 3: Design of experiments*

ISO 4121, *Sensory analysis — Guidelines for the use of quantitative response scales*

ISO 5492, *Sensory analysis — Vocabulary*

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

ISO 8586, *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 purposes of this document, the terms and definitions given in ISO 3534-1, ISO 3534-3, 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
magnitude estimation
process of assigning values to the intensity of a sensation elicited by a product attribute, or to its hedonic value, so the ratio between the assigned value and the assessor's perception of the attribute remains the same

3.2
external reference
first sample presented as a reference in a sample series in relation to which all subsequent samples are then assessed

3.3
internal reference
first sample introduced in a sample series in relation to which all subsequent samples are then assessed, and presented to the assessor as if it was a test sample

3.4
modulus
numerical value assigned to the *external reference* (3.2), which can be defined by the person conducting the test (fixed modulus) or left to the assessor to choose (non-fixed modulus)

3.5
rescaling
process that consists of multiplying the raw data supplied by each assessor by a factor which reduces the data of all the assessors to a common scale

Note 1 to entry: Adding the logarithm of this factor to the logarithm of the raw data is an equivalent process.

3.6
Steven's equation psychophysical power function
relationship that is expressed as:

$$R = KS^n$$

where

R is the assessor's response (e.g. perceived intensity);

K is a constant which reconciles the units of measurement used for R and S;

S the stimulus (concentration of a chemical substance or physical variable);

n is the exponent of the power function and the slope of the regression curve for R and S when they are expressed in logarithmic units.

Note 1 to entry: In practice, Stevens's equation is generally transformed into natural logarithms:

$$\ln R = \ln K + n \ln S$$

4 Principle

Samples are presented successively to assessors, who are requested to record the intensity of an attribute of each sample by complying with the ratio principle.

The values are assigned by referring to the value of the first sample of the series. For this first sample, either each assessor is free to assign a value to it, or the value is fixed by the person conducting the test. The latter case is called "fixed modulus".

5 General test conditions

For the general test conditions, such as those concerning the facilities, preparation, presentation and coding of samples, International Standards on general methodology shall be followed, in particular ISO 6658 and ISO 8589, as well as those describing the methods using scales and categories, in particular ISO 4121.

6 Selection and training of assessors

6.1 General conditions for selection and training

The general conditions for selection and training shall be in accordance with ISO 8586.

As in all other sensory analysis methods, it is the responsibility of the panel leader to judge the required level of proficiency of the assessors. The objectives of the test, the availability of the assessors, the costs incurred by recruiting additional assessors, as well as their training, shall be taken into account when planning a training programme. Assessors are generally able to use the magnitude estimation methodology after three or four training tests.

6.2 Training specific to the magnitude estimation method

6.2.1 The assessment of surface areas of geometric shapes has been proved to be particularly suited for introducing assessors into the basic concepts of magnitude estimation. A set of 18 shapes (see Table 1) comprising six circles, six equilateral triangles and six squares ranging in size from approximately 2 cm^2 to 200 cm^2 has been used successfully for training assessors.

For the consumer panels, a shorter version may be used. For example, the training can be limited to area estimations.

Table 1 — Dimensions and areas of the training exercise shapes

Circles		Triangles		Squares	
Radius cm	**Surface area** cm^2	**Side** cm	**Surface area** cm^2	**Side** cm	**Surface area** cm^2
1,4	6,2	2,2	2,1	3,2	10,2
2,5	19,6	4,1	7,3	4,2	17,6
3,7	43,0	7,6	25,0	8,5	72,3
5,4	91,6	12,2	64,4	11,1[a]	123,2
6,8	145,3	15,5	104,0	11,1[a]	123,2
8,3	216,4	19,2	159,6	14,2	201,6

[a] Two 11,1 cm squares are introduced into the series in order to be able to evaluate the reproducibility of the assessors.

6.2.2 Prior to presenting the shapes to the assessors, instruct them in the principles of the method. This instruction shall include, but is not necessarily limited to, the following three points:

— the values shall be assigned on a ratio basis: if the attribute is twice as intense, a value twice as high shall be assigned to it;

— there is no upper limit to the scale;

— the value 0 shall be assigned only in the exceptional case where the attribute is not perceived.

Warn the assessors, at the time of training, that the general tendency is often to use round numbers (such as 5, 10, 20, 25, etc.) but that, with this method, all numbers are permitted and may be used.

As assessors are also influenced by the ratios mentioned during training, always take care to suggest to them the use of different ratios, e.g. 3/1, 1/3, 7/5, 5/6 without limiting oneself to 2/1 or 1/2.

6.2.3 Assign codes to the shapes and present the shapes separately by placing them in the centre of a sheet of white paper of approximately A4 size (21 cm × 29,7 cm).

Instruct each assessor to conduct the magnitude estimation, beginning the series with the presentation of the 8,5 cm square (external reference). Record the responses.

Depending on the procedure adopted for the test phase, train the assessors with or without a fixed modulus. With a fixed modulus, the person conducting the test assigns a value between 30 and 100 to the 8,5 cm square.

Without a fixed modulus, leave the assessors to assign the value of their choice to the first figure, but advise them not to choose too small a value.

Randomly present the geometric shapes prior to each test, so that their shapes and dimensions do not form a particular pattern.

6.2.4 After completing the assessment of the set of shapes, allow the assessors to compare their results with the average results of the group. If this is not practical, carry out this comparison with respect to the results obtained by a previous group.

The objective is to provide positive feedback to reassure the panellists that they understand the exercise. Care should be taken not to create the impression that there is a “right” answer. Unless their results are very different, departures from the group results should be explained as order effects; that is, their responses are affected by the order in which they evaluate the samples. They should be reassured that, despite individual order effects, the group’s results will be accurate.

If the results of some assessors are very different, explain once again to these assessors the principles of the method.

6.2.5 When an assessor has successfully completed the area estimation exercise, further training should be given based on the product or type of substance that will be assessed in the actual test. This gives the assessor experience in applying magnitude estimation to attributes characterizing the test substance. The panel leader may need to design exercises for training panellists to identify correctly the attributes to be evaluated. This training may be drawn up using the general guidelines given in ISO 8586.

7 Number of assessors

7.1 General

As for the other methods employing scales, the required number of assessors depends on:

— how close together the various test products are in the attribute being evaluated;

— the training received by the assessors;

— the importance that will be attached to the decision following the test results (see ISO 8586);

— the objectives that can be identified in terms of statistical power.

In the absence of such distinctly identifiable objectives, refer to the recommendations given in 7.2 and 7.3.

7.2 Analytical and research panels

The panels shall be made up as given in Table 2.

Issues of statistical power need to be resolved on the variance of individual evaluations and the magnitude of the differences which need to be detected.

Table 2 — Formation of panels

Types of assessors	Minimum number of assessors	Recommended number
Experienced assessors, highly trained in the product and in the assessment of the attribute being studied	5	10
Experienced assessors, trained in the product and in the assessment of the attribute being studied	15	20 to 25
Newly trained assessors	20	20 and over

7.3 Consumer panels

The magnitude estimation method can also be used with consumer panels or for conducting market research studies. The number of persons to be selected shall then be determined on the basis of the population sampling requirements connected with these types of tests. The use of the magnitude estimation method does not offer any particular advantage in terms of the number of assessors required, and this number shall be the same as for a typical consumer type test, namely at least 60 persons and often much more.

8 Procedure

8.1 Presentation of samples

All the samples shall be presented in an identical manner (i.e. identical serving vessels and the same quantity of product).

The vessels containing the samples shall be coded, preferably using randomly selected three-digit numbers.

8.2 External reference sample

It is desirable that the reference sample has, for the attribute being studied, an intensity that is close to that of the geometric mean of all of the products under test.

NOTE A reference that can present an extreme value for the attribute would introduce a distortion.

One or more randomly coded reference samples may be included in the test series without the assessors being informed. This allows assessment of the repeatability of the assessor within the session.

8.3 Order of presentation of samples

The samples shall be presented all at once or in a sequential way to the assessors. The assessors shall follow the order indicated. As in all sensory tests, this order differs from one assessor to the other, the ideal situation being that the orders of the samples are balanced.

The panel leader can refer to tables proposed in Reference [3], which uses Latin squares to balance the design for order and carryover effects. Where this is not possible, use random order.

8.4 Magnitude estimations

8.4.1 General

Carry out the test in accordance with one of the techniques described in 8.4.2 to 8.4.4.

Questionnaire models for the reference sample are given in Annex A.

8.4.2 Without fixed modulus for the external reference

Each assessor evaluates the reference and assigns a value to it. Advise the assessors not to choose too small a value.

The assessor then evaluates each subsequent coded sample, comparing it with the reference, and assigns to it a value in relation to that which he/she has previously assigned to the reference.

8.4.3 With fixed modulus for the external reference

The panel leader specifies to the assessor that the reference sample has a value of, for example, 30, 50, 100, or whatever seems appropriate to the panel leader.

The leader instructs the assessor to make his or her subsequent judgements relative to the value assigned to the reference (fixed modulus).

8.4.4 Without external reference

It is possible to use magnitude estimation methodology without using any external reference sample. Due to the limits of the sensory systems (memory), it may be difficult for the assessors to refer systematically to the first sample. There are two possible cases, as follows.

a) The assessors are not forced to re-evaluate the first sample prior to evaluating each of the subsequent samples.

   It is then advisable to encourage the assessors to memorize the degree of the attribute being studied for this reference sample and to re-evaluate this reference if it appears necessary to them.

   It is therefore possible:

   1) prior to the test: to choose the presentation design in such a way that the first sample is not the same for each of the assessors; ideally, each sample should be used as a reference for an equal number of assessors; the variances of the mean differences between samples will therefore be equal;

   2) at the time of analysis: to apply an arbitrarily very high (theoretically infinite) weighting to the evaluations of the first sample of each assessor, so that the variances of the differences are correctly estimated.

b) The assessors are asked to evaluate each sample by comparing it to the immediately preceding sample.

   NOTE The problem that then arises is that, for each assessor, the evaluation errors are autocorrelated, and the variance of the difference of two successive samples will be smaller than the variance of the difference of two non-successive samples.

   It is therefore possible:

   1) prior to the test: to choose the presentation design in such a way that all the possible permutations of the samples are presented to an equal number of assessors; if this is not possible, try to propose orders which approximate best this ideal model; the variances of the mean differences between samples will then be equal or, at least, fairly close;

   2) at the time of analysis: to employ autocorrelated error models, the methodology of which is, however, slightly more complicated.

It is to be noted that even if one proceeds as proposed in case a) (systematic comparison with the first sample in order to carry out the evaluation), an autocorrelation term, linked to the evaluation of the preceding sample, however small it may be, very probably remains (this is also true, incidentally, for the tests with reference described in 8.4.2 and 8.4.3). The advice given earlier that the orders of the samples are balanced is therefore valid in all cases.

