

INTERNATIONAL STANDARD

ISO
4121

First edition
1987-12-15



INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE DE NORMALISATION
МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ

Sensory analysis — Methodology — Evaluation of food products by methods using scales

*Analyse sensorielle — Méthodologie — Évaluation des produits alimentaires par des méthodes
utilisant des échelles*

STANDARDSISO.COM : Click to view the full PDF of ISO 4121:1987

Reference number
ISO 4121 : 1987 (E)

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.

Draft International Standards adopted by the technical committees are circulated to the member bodies for approval before their acceptance as International Standards by the ISO Council. They are approved in accordance with ISO procedures requiring at least 75 % approval by the member bodies voting.

International Standard ISO 4121 was prepared by Technical Committee ISO/TC 34, *Agricultural food products*.

Users should note that all International Standards undergo revision from time to time and that any reference made herein to any other International Standard implies its latest edition, unless otherwise stated.

Sensory analysis — Methodology — Evaluation of food products by methods using scales

1 Scope

This International Standard specifies various tests using several types of scales by which the sensory assessment of samples may be carried out.

It is divided into three sections:

section one: Guidance and general conditions for the two types of test;

section two: Tests using interval or ratio scales, for the evaluation of food products by the allocation of scores;

section three: Tests used for evaluating specific food products, consisting of classifying these products as a whole or by their particular attributes, using ordinal scales, by means of a score or expression.

2 Field of application

The tests are applicable to the evaluation of organoleptic attributes (for example, taste, odour, texture) or of a particular attribute (for example, sweetness, peppermint strength, tenderness) of a sample.

The tests may be used in particular to determine variations in raw materials or the effect of methods of production, treatment or storage.

The number of samples and the number of attributes to be evaluated during one session depend on the nature of the products to be examined.

3 References

ISO 5492, *Sensory analysis — Vocabulary*.

ISO 5497, *Sensory analysis — Methodology — Guidelines for the preparation of samples for which direct sensory analysis is not feasible*.

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

ISO 8589, *Sensory analysis — General guidance for the design of test rooms*.¹⁾

4 Definitions

For the purpose of this International Standard, the definitions given in ISO 5492, and the following, apply.

4.1 ordinal scale: Scale where points are arranged according to a pre-established or continuous progression.

4.2 interval scale: Scale where numbers are chosen in such a way that equal numerical intervals are assumed to correspond to equal differences in sensory perception.²⁾

As far as instruments are concerned, the Celsius and Fahrenheit scales for measuring temperature offer two examples of this type of scale.

In the interval scale, the zero is arbitrary and the operation of multiplying by a scalar is of no significance.

4.3 ratio scale: Scale where numbers are chosen in such a way that equal numerical ratios are assumed to correspond to equal sensory perception ratios.

If, for example, for the characteristic "sweetness", sample A scored 6 and sample B scored 3, the ratio 6/3 indicates that sample A is judged to be twice as sweet as B; compared with a sample C which scored 18, A is judged to be three times less sweet than C.

In sensory evaluation, the ratio scales are generally obtained by the so-called magnitude estimation.

The numerical value given to the reference sample may be fixed or left to the choice of the assessor; in the latter case, a subsequent mathematical calculation will be required in order that the results of each assessor may be compared.

1) At present at the stage of draft.

2) However, in sensory analysis this equality is difficult to attain.

Section one: Guidance and general conditions for the two types of test

5 Principle

Classification carried out by placing the products on one or more predetermined ordinal, interval or ratio scales corresponding to each of the attributes assessed.

6 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 standardized apparatus corresponds to the needs of the test, it shall be used.

7 Sampling

Refer to the International Standards relating to sampling, for sensory analysis, of the product or products to be examined.

The method of sampling shall take account of the test objective and, if there is no International Standard for the product concerned, shall be agreed between the contracting parties.

8 General test requirements

8.1 Room

For the conditions in the room in which tests are to be conducted, see ISO 8589.

8.2 Assessors

8.2.1 Qualification

For the conditions which the assessors shall fulfil, see ISO 6658.

All the assessors shall have the same level of qualification (assessor, qualified assessor or expert), which is chosen according to the purpose of the test.

8.2.2 Number

The number of assessors carrying out the tests and the number of repetitions per assessor depend on the aim of the evaluation and the required accuracy. No general rule can be made. Some indications are, however, given in ISO 6658.

8.3 Preliminary discussion

It may be advisable for a preliminary discussion to take place between the assessors and the test supervisor on the problem raised and the nature of the samples, provided that this discussion does not bias their future judgements.

With this type of test, the supervisor may in particular inform the assessors of the meaning of each of the steps on the scale.

Several typical samples of the series for analysis may be presented and discussed. The number of these shall allow the evaluation of all the attributes of the stimulus.

If the test in progress relates to the detection of foreign flavours, this preliminary test shall include the examination of a sample free from any foreign flavour, or, alternatively, if this is possible, the examination of the foreign flavour to be detected.

8.4 Control sample

For ordinal or interval scales, the use of a control sample, to serve as a reference, which will then be reintroduced anonymously into the series of tests, is recommended.

The attributes of this control sample shall be assessed and an agreement shall be reached on the score to be awarded. This may be an overall score or a score for each attribute, as applicable.

8.5 General test instructions

Test instructions relating specifically to the product shall be observed by the assessors, for example stirring the sample before testing the odour. In addition, it may be necessary in certain cases to neutralize gustatory impressions by means of additional substances suitable for the product under test, for example water, weak tea, bread rolls, unsalted crackers, sticks of celery, quartered apples (particularly after tasting fats and oils).

9 Preparation of test samples

(distribution, dilution, cooking, etc.)

9.1 Provide a sufficient quantity of total sample and the required number of individual samples.

9.2 It shall not be possible for the assessors to draw any conclusions on the nature of the samples from the way in which they are presented to them.

The various samples shall be prepared in exactly the same way (identical containers and laboratory ware, same amounts of products).

9.3 The temperature of the samples and also, if possible, that of all the samples within a series of tests shall be the same.

9.4 The vessels containing the test samples shall be coded, preferably using three-digit numbers chosen at random. The coding shall be different for each test.

See ISO 5497 if the samples for examination have an intense flavour or if they cannot be analysed as they are.

The assessors are given the samples either all together or one after the other; the evaluation may be random or follow a defined test plan.

However, it is necessary to take precautions against certain criteria which may produce bias: simultaneous or consecutive presentation of samples, effects of the successive order of presentation of the samples, constitution of the groups of samples presented in a structured order, number of samples to be evaluated in one session, etc.

The means of preparation and presentation shall be carefully defined as shall the procedure, where required, for masking particular properties.

STANDARDSISO.COM : Click to view the full PDF of ISO 4121:1987

Section two: Tests using interval or ratio scales

For general test conditions, see clauses 5 to 9.

10 Procedure

10.1 Establishing scales

The scales used may be interval scales (see 4.2) or ratio scales (see 4.3).

10.1.1 Interval scales are very diverse as the number of steps characterizing each of them may vary. As there is no ideal scale, care shall be taken, whenever a particular scale is constructed or used, that each step corresponds to a fixed level of intensity, satisfying the condition of equality of the intervals. In practice, these levels are defined by reference substances (for example, scale of concentration of the substance under test according to a geometric progression), or by a literal expression or a group of expressions; for this reason, the assessor may give his or her reply either by means of a score, or by means of expressions which would then be converted into scores.

Sometimes, only certain steps are clearly defined. The extreme form of this type of scale is the non-structured scale in which only the end steps are defined.

Some examples of this type of scale are given in 10.2.

10.1.2 Ratio scales in sensory analysis are generally obtained by the so-called magnitude estimation method. See 10.2.7.

10.2 Examples of scales in common use

10.2.1 Six-point scale for assessing intensity¹⁾

Scale with six points	
1	Absent
2	Very slight
3	Slight
4	Clear
5	Pronounced
6	Very pronounced

10.2.2 Seven-point scale for assessing hardness¹⁾

Scale with seven points	
1	Very hard
2	Hard
3	Slightly hard
4	Neither hard nor soft
5	Slightly soft
6	Soft
7	Very soft

10.2.3 Nine-point hedonic scale¹⁾

Scale with nine points	
9	Extremely agreeable
8	Very agreeable
7	Agreeable
6	Slightly agreeable
5	Neither agreeable nor disagreeable
4	Slightly disagreeable
3	Disagreeable
2	Very disagreeable
1	Extremely disagreeable

10.2.4 Incomplete nine-point scale for the assessment of one characteristic (e.g. sweetness)

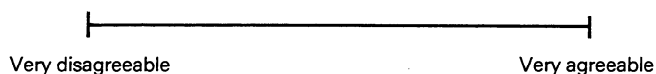
Insufficient								Excessive

10.2.5 Unstructured intensity scale

Slightly aromatic								Very aromatic
-------------------	--	--	--	--	--	--	--	---------------

¹⁾ These scales are only interval scales on the assumption that the intervals are equal. However, if this is not the case, they should be looked upon as ordinal scales and treated as such (see section three, clause 15).

10.2.6 Unstructured hedonic scale



10.2.7 Ratio scale

Several orange juices are presented to the assessor. The assessor is asked to evaluate the acidity of each juice using a number, which reflects the ratio of intensity perceived between this juice and a reference sample R.

For example, if sample R is given a score of 20, and it is felt that the acidity of the juice marked 375 is three times greater than that of sample R, this juice is given a score of 60; if the juice marked 658 seems five times less acid, this juice is given a score of 4.

11 Expression of results

From the scores given, for each characteristic studied, a frequency curve can be constructed and the average and the standard deviation can be calculated.

Other statistical methods, such as variance analysis, regression or correlation analysis, may also be used. If the researcher thinks that the equidistance condition is not achieved, i.e. if the results do not comply with interval or ratio scales, non-parametric methods may prove valuable.

When the various attributes of a product are scored separately, it may be interesting to use a graphic representation in which the attributes are, for example, placed along the abscissa and the scores along the ordinate. If the attributes are scored using different scales, weighting is necessary.

12 Test report

The test report shall include the following information:

- a) all necessary information for the complete identification of the sample (or samples):
 - number of samples,
 - whether control samples have been used;
- b) the test parameters that have been used:
 - number of assessors and their level of qualification,
 - test environment,
 - physical conditions;
- c) the results obtained and their statistical interpretation;
- d) a reference to this International Standard;
- e) the deviations from this International Standard;
- f) the name of the person supervising the test;
- g) the date and time of the test.

Section three: Tests used for evaluating specific food products

For general test conditions, see clauses 5 to 9.

13 Procedure

13.1 Establishing scales

The scales used are ordinal scales.

They meet the following criteria:

- the highest steps of the scale are used for designating the highest quality or intensity;
- the scales are only in whole numbers;
- the numbers relate to each class.

The range of the scale depends on the aim of the test, and on the desired accuracy of the results. This range shall be fixed before the test is started.

In general, scales with three to nine points are used.

13.2 Examples of ordinal scales used to evaluate food products

Taking a basic three-point scale, corresponding to the three main qualitative classes

- desirable quality:** Predominantly positive descriptions, i.e. when the standardized quality requirements are met,
- tolerable quality:** Mixture of positive and negative descriptions, i.e. when the standardized quality requirements are met with qualifications,
- undesirable quality:** Predominantly negative descriptions, i.e. when the standardized quality requirements are not met,

the image of the product can be refined by subdividing the scale into six or nine points, in accordance with the principles given in tables 1 and 2.

Table 1 — Example of possible subdivisions of the scales

Quality class	Extent of the scale		
	$k = 3$	$k = 6$	$k = 9$
Desirable	3	6	9
		5	8
			7
Tolerable	2	4	6
		3	5
			4
Undesirable	1	2	3
		1	2
			1

Table 2 — Example of subdivision of scale applied to the overall assessment of a product

9	Perfect, typical	6	Perfect, typical	3	Perfect to minor deviations	
8	Typical, no deviation					5
7	Typical, with minor deviations					
6	Slightly noticeable deviations or minor defects	4	Slightly noticeable deviations or minor defects	2	Slightly noticeable deviations to noticeable defects	
5	Noticeable deviations or minor defects					3
4	Noticeable defects					
3	Strong defects	2	Strong to very strong defects	1	Strong defects to completely changed	
2	Very strong defects					1
1	Completely changed					

Table 3 — Example of a specific scale

This example is applicable to a descriptive scale of colour, odour and texture attributes of green beans.

Score	Attribute		
	Colour	Odour	Texture
9	Strongly green, uniform	Completely characteristic, pure	Very tender, very juicy, uniform
8	Strongly green, nearly uniform	Characteristic, pure	Tender, juicy, nearly uniform
7	Green, still uniform	Still characteristic, pure	Still tender, still juicy, still uniform
6	Slightly too light or too dark green	Slightly flat, slightly pungent, tart	Somewhat soft, hard, mildly dry, floury, watery, non-uniform
5	Noticeably discoloured (olive green, yellowish, brownish, with spots), noticeably non-uniform	Flat, pungent, tart	Soft, hard, dry, floury, fibrous, stringy
4	Markedly discoloured (olive green, yellow, brown, with spots)	Markedly pungent, tart	Markedly soft, hard, dry, floury, fibrous, stringy
3	Strongly discoloured, with many spots	Strongly changed	Mashy, tough, dry, fibrous, stringy
2	Very strongly discoloured	Very strongly changed	Strongly mashy, tough, fibrous, stringy
1	Completely changed	Completely changed	Completely changed

14 Test technique

Once the preliminary discussion has taken place between the assessors and the test supervisor, and the panel has come to an agreement on the qualitative scores assigned to the control sample (overall score or score per attribute, as applicable), the assessors assess the attributes using the whole numbers of the scale proposed to them.

This evaluation is based on a descriptive evaluation scheme.

Depending on the type of sample, a successive assessment is made of the attributes of a single sample, or of the individual attributes of all the samples (for example colour).

If the samples examined are heterogeneous (e.g. cooked meals, where each component is to be evaluated separately), this heterogeneity shall be taken into consideration both when the attributes to be assessed are chosen and when the results are evaluated.

The assessors report their evaluation on an answer form. A specimen answer form is given in table 4.

Table 4 — Specimen answer form

Product:		Date:	
Test:		Assessor:	
Sample No.	Attribute	Score	Additional remarks
	A B C etc.		
	A B C etc.		

15 Expression of results

When the results (scores) from the assessors forming the panel have been confirmed, it is possible to proceed with the statistical analysis of these results by calculating the frequency distributions and the median.

16 Test report

See clause 12.