INTERNATIONAL STANDARD

ISO 8586-1

First edition
1993-03-15

Sensory analysis — General guidance for the selection, training and monitoring of assessors —

Part 1:
Selected assessors

Analyse sensorielle — Guide général pour la sélection, l'entraînement et le contrôle des sujets —

Partie 1: Sujets qualifiés

Reference number
ISO 8586-1:1993(E)

Contents

International Organization for Standardization
Case Postale 56 • CH-1211 Genève 20 • Switzerland
Printed in Switzerland

Annexes

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.

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.

International Standard ISO 8586-1 was prepared by Technical Committee ISO/TC 34, *Agricultural food products*, Sub-Committee SC 12, *Sensory analysis*.

ISO 8586 consists of the following parts, under the general title *Sensory analysis — General guidance for the selection, training and monitoring of assessors*:

— *Part 1: Selected assessors*

— *Part 2: Experts*

Annex A forms an integral part of this part of ISO 8586. Annexes B and C are for information only.

Introduction

A sensory analysis panel constitutes a true "measuring instrument", and consequently the results of the analyses conducted will depend on its members.

The recruitment of persons willing to participate in a panel therefore needs to be carried out with care and to be considered as a real investment, both in time and financially.

Sensory assessment may be made by three types of assessor:

— "assessors",

— "selected assessors", and

— "experts".

Assessors can be "naïve assessors" who do not have to meet any precise criterion, or "initiated assessors" who have already participated in sensory tests.

"Selected assessors" are assessors who have been selected and trained.

"Experts" can be "expert assessors" who have already demonstrated particular acuity in panel work and have developed a good long-term memory, or "specialized expert assessors" who draw on additional knowledge gained in particular fields.

This part of ISO 8586 concerns only the recruitment, selection, training and monitoring of candidates intended to become selected assessors. The recruitment, selection, training and monitoring of candidates intended to become experts will form the subject of ISO 8586-2.

The selection and training methods to be employed depend on the tasks which one intends to give the selected assessors. It should be pointed out that these methods sometimes only constitute a way of choosing the better candidates among those who are available, rather than those who satisfy predetermined criteria. This is particularly the case when it is necessary to constitute internal panels.

A preliminary selection of candidates has to be undertaken at the recruitment stage, in order to eliminate those who would be unsuited for sensory analysis. However, the final selection can only be made after training and the completion of the envisaged tasks.

The recommended procedure involves

a) recruitment and preliminary screening of naïve assessors;

b) training of naïve assessors who will become initiated assessors;

c) selection of initiated assessors according to ability to perform particular tests; they will then become selected assessors;

d) selection following the performance of an actual sensory assessment (useful in the case of descriptive analyses);

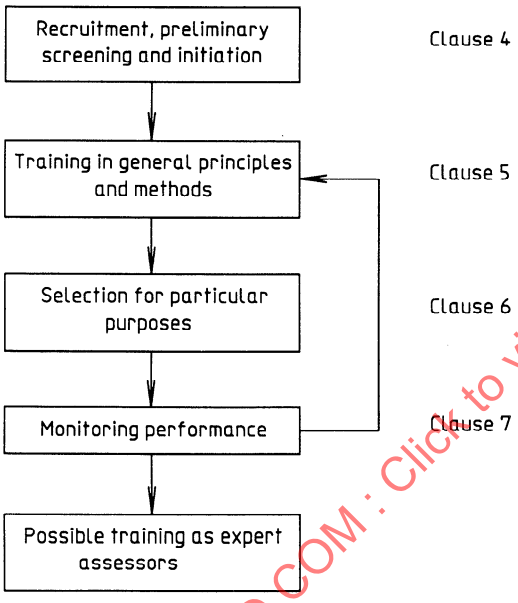
e) possible training of selected assessors to become expert assessors.

In certain cases (especially for descriptive sensory analysis), the panel may be divided into specialized sub-groups.

The exact procedures covered by a) and b) and the nature of the tests performed in c) and d) depend on the tasks which the panel is to perform.

The performance of selected assessors should be monitored regularly to ensure that the criteria by which they were initially selected continue to be met.

The entire process is illustrated in the following diagram.



Sensory analysis — General guidance for the selection, training and monitoring of assessors —

Part 1:
Selected assessors

1 Scope

This part of ISO 8586 specifies criteria for the selection and procedures for the training and monitoring of selected assessors. It supplements the information given in ISO 6658.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this part of ISO 8586. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this part of ISO 8586 are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 4120:1983, *Sensory analysis — Methodology — Triangular test.*

ISO 4121:1987, *Sensory analysis — Methodology — Evaluation of food products by methods using scales.*

ISO 5492:1992, *Sensory analysis — Vocabulary.*

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

3 Definitions

For the purposes of this part of ISO 8586, the definitions given in ISO 5492 apply.

4 Recruitment, preliminary screening and initiation

Recruitment is an important starting point in forming a panel of selected assessors. Different recruitment methods and criteria are available and there are various tests that can be used for screening candidates for suitability for further training.

4.1 Principle

To recruit candidates and to select those most suitable for training as selected assessors.

4.2 Recruitment

The following three questions arise when recruiting persons to form a sensory analysis panel:

— where should one look for the people who will constitute the group?

— how many people shall be selected?

— how shall the people be selected?

4.2.1 Types of recruitment

Two types of recruitment are available to organizations:

— recruit through the personnel department of the organization (internal recruitment), or

— recruit people from outside the organization (external recruitment).

It is possible to constitute a mixed panel made up of both types of recruitment.

4.2.1.1 Internal recruitment

The candidates are recruited from amongst the office, plant or laboratory staff. It is advisable to avoid those persons who are too personally involved with the products being examined, in particular those involved at the technical or commercial level, because they may cause the results to be biased.

In this type of recruitment, it is vital that the organization's general management and hierarchy provide their support and make it known that sensory analysis is considered as forming part of everyone's work. This can be made known at the hiring stage of the personnel.

4.2.1.2 External recruitment

The recruitment is conducted outside the organization.

The most commonly used means for this purpose are:

— recruitment through classified advertisement in the local press, in specialized publications, or in newspapers which are distributed free of charge, etc. (in this case, all types of people will reply and it will be necessary to carry out a selection);

— opinion poll institutes; certain of these institutes can provide the names and addresses of persons likely to be interested;

— in-house "consumer" files, compiled as a result of advertizing campaigns or complaints;

— persons visiting the organization;

— personal acquaintances.

4.2.1.3 Mixed panel

A mixed panel may be formed using internal and external recruitment, in variable proportions.

4.2.2 Advantages and disadvantages of internal and external recruitment

Organizations may wish to use independent internal or external panels for different tasks.

4.2.2.1 Internal recruitment

4.2.2.1.1 Advantages

The advantages are that

— the people are on the spot;

— it is not necessary to make provision for any payment (however, in order to maintain interest, it may be desirable to offer small presents or perks);

— a better confidentiality *vis-à-vis* the results is ensured, which is particularly important if it is a question of research work; and

— there is better stability of the panel with time.

4.2.2.1.2 Disadvantages

The disadvantages are that

— candidates are influenced in their judgements (by knowledge of the products);

— it is difficult to allow for the evolution of the organization's products (people are influenced by their familarization with the organization's products);

— replacement of candidates is more difficult (limited number of persons in small organizations);

— lack of availability.

4.2.2.2 External recruitment

4.2.2.2.1 Advantages

The advantages are that

— there is a wide range of choice;

— there is subsequent supply of new persons by word of mouth;

— there are no problems with hierarchy;

— selection is much easier, without the risk of offending people if they are unsuitable;

— easy availability.

4.2.2.2.2 Disadvantages

The disadvantages are that

— the method is expensive (remuneration, paperwork);

— this method is better suited to urban communities where there is a sufficient number of inhabitants; however, in rural areas, advantage can be taken of co-operatives (e.g. milk, wine);

— since it is necessary that the individuals are available, one sometimes encounters an inordinate number of (old age) pensioners or unemployed women or even students; it is more difficult to recruit working people; and

— after having paid for the selection and training, one risks people leave at a moment's notice.

4.2.3 Number of persons to be selected

Experience has shown that, after the recruitment, the selection procedures eliminate approximately half the people for reasons such as gustative sensitivity, material conditions, etc.

The number of persons to be recruited will vary depending on the following elements:

— the financial means and the requirements of the organization;

— the types and frequency of tests to be conducted;

— whether or not it is necessary to interpret the results statistically.

It is not desirable for a panel to operate with less than 10 selected assessors. It is necessary to recruit at least two to three times the number of persons actually required to constitute the final panel. For example, in order to obtain a panel of 10 persons, it is necessary to recruit 40 and to select 20.

For specialized purposes, a higher level of recruitment will be required.

4.3 Background information

Background information on the candidates may be obtained by submitting them to a combination of clearly understood questionnaires coupled with interviews by persons experienced in sensory analysis. The aspects specified in 4.3.1 to 4.3.8 shall be explored.

4.3.1 Interest and motivation

Candidates who are interested in sensory analysis and the product or products to be investigated are likely to be more motivated and hence are likely to become better assessors than those without such interest and motivation.

4.3.2 Attitudes to foods

Strong dislikes for certain foods and beverages, in particular those which it is proposed to assess, together with any cultural, ethnic or other reasons for not consuming certain foods or beverages, should be determined. Candidates who are venturesome in their eating habits often make good assessors for descriptive analyses.

4.3.3 Knowledge and aptitude

The initial sensory perceptions of the candidates have to be interpreted and expressed, requiring certain physical and intellectual abilities, in particular the capacity to concentrate and to remain unaffected by external influences. If the candidate is then required to evaluate only one type of product, knowledge of all aspects of that product may be beneficial. It is then possible to choose expert assessors from amongst those candidates who have shown an aptitude for sensory analysis of this product.

4.3.4 Health

The candidates shall be in good general health. They shall not suffer from any disabilities which may affect their senses, or from any allergies or illnesses, and shall not take medication which might impair their sensory capacities and thus affect the reliability of their judgements. It may be useful to know whether the candidates have dental protheses, since they can have an influence in certain types of evaluation involving texture or flavour.

Colds or temporary conditions (for instance, pregnancy) should not be a reason for eliminating a candidate.

4.3.5 Ability to communicate

The ability of candidates to communicate and describe the sensations they are perceiving when assessing is particularly important when considering candidates for descriptive analyses. This ability can be determined at the interview and again during screening tests (see 4.4.6).

4.3.6 Availability

Candidates shall be available to attend both training and subsequent assessments. Personnel who travel frequently or have continual heavy work-loads are often unsuited for sensory work.

4.3.7 Personality characteristics

Candidates shall show interest and motivation for the tasks and shall be willing to persevere with tasks demanding prolonged concentration. They shall be punctual in attending sessions and shall be reliable and honest in their approach.

4.3.8 Other factors[1)]

Other information which may be recorded during recruitment are name, age group, sex, nationality, educational background, current occupation and experience in sensory analysis. Information on smoking habits may also be recorded, but candidates who smoke shall not be excluded on these grounds.

1) It is necessary that any files on individual persons comply with the legal requirements of the country concerned.

4.4 Screening

Various tests which may be used for screening purposes are described below.

The choice of the tests and of the materials to be used is conducted on the basis of the envisaged applications and of the properties to be assessed.

4.4.1 Types of screening tests

All the tests described have the dual function of familiarizing the candidates with both the methods and the materials used in sensory analysis. They are divided into three types as follows:

a) those aimed at determining impairment;

b) those aimed at determining sensory acuity;

c) those aimed at evaluating a candidate's potential for describing and communicating sensory perceptions.

Tests, the results of which will be used to aid selection, should only be carried out after previous experience followed by familiarization.

The tests shall be conducted in the actual environment in which products are evaluated and in an appropriate environment in accordance with the recommendations given in ISO 8589. They shall be followed by interviews. Several tests described in this part of ISO 8586 are based on those described in ISO 6658.

The selection of assessors should take into account the intended application, the performance of the candidates at the interviews and their potential rather than their current performance. Candidates with high success rates are to be expected to be more useful than others, but those showing improving results with repetition are likely to respond well to training.

4.4.2 Colour vision

Candidates with abnormal colour vision are unsuitable for tasks involving judgement or matching of colours. Assessment of colour vision can be carried out by a qualified optician or, in the absence of such a person and associated equipment, by using an effective test, for example the Ishihara[2] test.

4.4.3 Ageusia and anosmia

It is desirable that candidates be tested to determine their sensitivity to substances which may be present in small concentrations in products, in order to detect ageusia, anosmia or possible lack of sensitivity (see ISO 3972).

Table 1 — Examples of materials and concentrations for matching tests

Taste or odour	Material	Concentration in water at room temperature g/l	Concentration in ethanol[1] at room temperature g/l
Taste			
Sweet	Sucrose	16	
Acid	Tartaric acid or citric acid	1	
Bitter	Caffeine	0,5	
Salty	Sodium chloride	5	
Astringent	Tannic acid[2]	1	
	or quercitin	0,5	
	or potassium aluminium sulfate (alum)	0,5	
Metallic	Ferrous sulfate, hydrated, $FeSO_4.7H_2O$ [3]	0,01	
Odour			
Lemon, fresh	Citral ($C_{10}H_{16}O$)	—	1×10^{-3}
Vanilla	Vanillin ($C_8H_8O_3$)	—	1×10^{-3}
Thyme	Thymol ($C_{10}H_{14}O$)	—	5×10^{-4}
Floral, lily of the valley, jasmine	Benzyl acetate ($C_8H_{12}O_2$)	—	1×10^{-3}

1) Stock solutions are prepared with ethanol, but the final dilution is made with water and shall not contain more than 2 % of alcohol.

2) This material is not very soluble in water.

3) To avoid the appearance of a yellow coloration due to oxidation, it is necessary to use a solution freshly prepared from neutral or slightly acid water. However, if a yellow coloration occurs, present the solutions in closed opaque containers or under dim or coloured light.

4.4.4 Matching test

Samples of sapid and/or olfactory materials (see table 1) at well above threshold levels are prepared. Each sample is attributed a different, random, three-digit code number. Candidates are presented with one sample of each type and are allowed to familiarize themselves with them (see ISO 6658).

2) See ISHIHARA, S. *Tests for colour blindness*. Kanahara Shuppan Co. Ltd., Tokyo-Kyoto, Japan, 1971.

They are then presented with a series of the same materials labelled with different random numbers. They are asked to match each of them to one of the original set and to describe the sensation they are experiencing.

Approximately twice as many new samples as original samples shall be presented. None of the samples shall be so intense as to produce strong carry-over effects and hence to influence subsequent tasting. Odourless flavourless water shall be made available for cleansing the palate between samples.

Examples of materials that may be used are given in table 1. For these substances and concentrations, it is generally accepted that candidates who make fewer than 80 % correct matches should not be chosen as selected assessors. A correct description of the sensations produced by the samples is desirable but less important.

4.4.5 Acuity and discriminating ability

The two following tests are recommended.

4.4.5.1 Tests for detection of a stimulus

These tests are based on the triangular test; see ISO 4120.

One material at a time is tested. Two samples of the test material and one sample of water or other neutral medium, or one sample of the test material and two of water or other neutral medium, are presented to each candidate. The concentration of the test material shall be at the supra-threshold level.

The test materials, their concentrations and the neutral medium (if used) shall be chosen by the organizer in relation to the types of assessment for which the candidates will be used. Preferably candidates should have 100 % correct responses.

An inability to detect differences after several repetitions indicates unsuitability for this type of test.

Examples of materials which may be used in detection tests are given in table 2.

Table 2 — Examples of materials which may be used in detection tests

Material	Concentration in water at room temperature
Caffeine	0,27 g/l
Citric acid	0,60 g/l
Sodium chloride	2 g/l
Sucrose	12 g/l
cis-3-Hexen-1-ol	0,4 ml/l

4.4.5.2 Tests for discrimination between levels of intensity of a stimulus

These tests are based on the ranking test described in ISO 8587. The tests are carried out using stimuli for taste, odour (only for very small concentrations), texture (mouth and hand), and colour.

For each test, four samples having different intensities of the property are presented in a random order to the candidates, who are required to put them in order of increasing intensity. This random order shall be the same for all candidates, to ensure that comparisons of their performance are not influenced by the effects of different orders of presentation.

A satisfactory level of success in this task can be specified only in relation to the particular intensities used.

Examples of products that may be used are given in table 3; for these concentrations, candidates who invert the order of more than one adjacent pair of samples shall be considered unsuitable as selected assessors for this type of analysis.

Table 3 — Examples of products which may be used in discrimination tests

Test	Product[1)]	Concentration in water at room temperature
Taste discrimination	Citric acid	0,1 g/l; 0,15 g/l; 0,22 g/l; 0,34 g/l;
Odour discrimination	Isoamyl acetate	5 ppm; 10 ppm; 20 ppm; 40 ppm;
Texture discrimination	To suit the industry concerned (e.g. cream cheese, purée, gelatine)	—
Colour discrimination	Cloth, colour scales, etc.	Intensity of a colour ranging, for example, from dark red to light red

1) Other appropriate products showing a graduation in characteristics may also be used.

4.4.6 Descriptive ability

These tests are aimed at determining a candidate's ability to describe sensory perceptions. Two tests are advocated, one covering odour stimuli and the other textural stimuli. The tests are conducted as combined assessments and interviews.

4.4.6.1 Odour description test

Candidates are presented with between 5 and 10 olfactory stimuli, preferably related to the product or products eventually to be evaluated. The set shall contain some samples which are easy to recognize and others which are less common. The intensity shall be well above the recognition threshold but not greatly above the levels that might be encountered in the products ultimately of interest.

Several methods of sample preparation exist which are either direct or retronasal in nature.

In direct methods, bottles, smelling strips or capsules containing odours are employed.

In retronasal methods, the odours may be evaluated from a gaseous medium, from smelling strips placed inside the mouth or by ingestion of aqueous solutions.

The method most commonly used is still that of the evaluation of odours from bottles. This method is described below.

Samples are absorbed in odourless paraffin wax or cotton wool which is placed in dark coloured, odourless, 50 ml to 100 ml screw cap glass bottles. Sufficient material shall be allowed to evaporate into the headspace of the bottle and the intensity shall be checked before presentation of the bottle to candidates.

Samples may also be presented on smelling strips or pads.

Samples are presented one at a time and the candidate is asked to describe or record what is perceived. Following the initial reaction, the organizer may if desired discuss the sample in order to bring forth further comments and to explore more fully the candidate's capability of discussing the stimuli.

Candidates are graded according to performance on a scale such as the following:

3 points for a correct identification or a description of the most frequent association;

2 points for a description in general terms;

1 point for an identification or description of an appropriate association following discussion;

0 for no response.

A satisfactory level of success in this task can be specified only in relation to the materials used. However, candidates who achieve less than 65 % of the maximum possible score are unsuitable as selected assessors for this type of test.

Examples of olfactory materials which may be used are given in table 4. See also ISO 5496.

Table 4 — Examples of olfactory materials for odour description test

Material	Name most commonly associated with the odour
Benzaldehyde	Bitter almonds, cherry, ...
Octen-3-ol	Mushroom, ...
Phenyl-2-ethyl acetate	Floral, ...
Diallyl sulfide	Garlic, ...
Camphor	Camphor, medicine, ...
Menthol	Peppermint, ...
Eugenol	Clove, ...
Anethol	Aniseed, ...
Vanillin	Vanilla, ...
β-Ionone	Violets, raspberries, ...
Butyric acid	Rancid butter, ...
Acetic acid	Vinegar, ...
Isoamyl acetate	Fruit, acid drops, banana, pear, ...
Dimethylthiophene	Grilled onions, ...

NOTE — It is possible to use food products, spices, extracts, infusions or chemical odorants. Materials chosen shall be suited to local needs and shall be free from other odorous materials.

4.4.6.2 Texture description test

Candidates are provided with a series of products in random order and are asked to describe their textural characteristics.

Solid sample products shall be presented as uniformly sized blocks and liquid sample products shall be presented in opaque vessels.

Candidates are graded according to performance on a scale such as the following:

3 points for a correct identification or a description of the most frequent association;

2 points for a description in general terms;

1 point for an identification or description of an appropriate association following discussion;

0 for no response.

A satisfactory level of success in this task can be specified only in relation to the products used. Candidates who achieve less than 65 % of the maximum

possible score are unsuitable as selected assessors for this type of test.

Examples of products that may be used are given in table 5.

Table 5 — Examples of products for texture description test

Product	Texture most commonly associated with the product
Oranges	Juicy, cellular particles, ...
Breakfast cereals (cornflakes)	Crispy
Pears (Passe Crassane)	Gritty
Granulated sugar	Crystalline, coarse
Marshmallow topping	Sticky, malleable
Chestnut purée	Pasty
Semolina	Grainy
Double cream	Unctuous
Edible gelatine	Gummy
Corn muffin cake	Crumbly
Cream toffee	Tacky
Calamary (squid)	Elastic, springy, rubbery
Celery	Fibrous
Raw carrots	Crunchy, hard

5 Training

5.1 Principle

To provide assessors with rudimentary knowledge of procedures used in sensory analysis and to develop their ability to detect, recognize and describe sensory stimuli. To train assessors to use this expertise so that they may become proficient in the use of such methods with particular products.

5.2 General

A number of assessors one and a half to two times greater than that finally required in the panel shall be trained. To ensure the development of the correct approach to sensory analysis, all training shall be conducted in a suitable environment in accordance with the recommendation given in ISO 6658. It is also useful to train assessors in basic knowledge of the products they assess, e.g. by giving information on the manufacturing process or by organizing visits to plants.

Except in preference tests, assessors shall be instructed at all times to be objective and to disregard their likes and dislikes.

Results shall be discussed and assessors shall be given the opportunity to re-assess samples and to check their replies where disagreement exists.

Assessors shall be instructed not to use perfumed cosmetics prior to or during sessions. They shall also be asked to avoid contact with tobacco or with strong tastes or odours for at least 60 min prior to such sessions. Soap used for washing shall not leave any odour on the hands.

It shall be emphasized to assessors that if they carry any odour into the test room, tests may be invalidated.

5.3 Assessment procedure

At the start of any training programme assessors shall be taught the correct way to assess samples. In all assessments, instructions shall be read thoroughly before any task and adhered to throughout the analysis. The temperature of samples shall be specified. Unless told to concentrate on specific attributes, the assessors should usually examine attributes in the following order:

— colour and appearance;

— odour;

— texture;

— flavour (comprising aroma and taste);

— after-taste.

When assessing odour, assessors shall be taught to take short rather than long sniffs and not to sniff too many times lest they become confused and fatigued.

With both liquid and solid samples, the assessors shall be told in advance the size of the sample (for mouth assessment), the approximate time for which the sample is to be held in the mouth, the number of chews and whether it is to be swallowed or not. The problem of adaptation and the advantages of rinsing the mouth and of standard time intervals between samples shall also be discussed. Any procedure finally agreed upon shall be stated clearly so that all assessors assess products in the same way. The interval between samples shall be sufficient to permit recovery but not so long that assessors lose their ability to discriminate.

5.4 Training in detection and recognition of tastes and odours

Matching, recognition, paired comparison, triangular and duo-trio tests (see ISO 6658 and specific International Standards) shall be used to demonstrate tastes at high and low concentrations and to train assessors to recognize and describe them correctly (see ISO 3972). Identical tests shall be used to de-

velop assessors' acuitiy for odour stimuli (see ISO 5496). Stimuli shall initially be presented singly as an aqueous solution but as experience is gained the base may be replaced by actual foods or beverages. Mixed samples in which the proportions of two or more components vary may also be introduced.

Alteration of the appearance of a sample (for example by the use of coloured light) is particularly useful in demonstrating the need to be objective when trying to detect differences in other sensory characteristics.

Samples used for training and testing shall be characteristic of their origin, style and quality, and representative of the range generally found on the market.

Samples shall be presented in the quantity and at the temperatures generally met with in the trade or in use.

Exceptions may be made when demonstrating excellence, imperfections or faults.

Care shall be taken to ensure that sensory fatigue does not arise owing to the testing of an excessive number of samples.

Table 6 gives examples of materials which may be used during this phase of training. If possible, stimuli should be chosen to relate to the material or materials it is eventually aimed to assess.

Table 6 — Examples of materials to be used during training in detection and recognition

1	The materials in table 1
2	The products in table 3
3	Saccharin (100 mg/l)
4	Quinine sulfate (0,20 g/l)
5	Grapefruit juice
6	Cider
7	Sloe juice
8	Cold tea
9	Sucrose (10 g/l; 5 g/l; 1 g/l; 0,1 g/l)
10	Hexanol (15 mg/l)
11	Benzyl acetate (10 mg/l)
12	Items 4 to 7 with varying sucrose contents (see item 9)
13	Tartaric acid (0,3 g/l) plus hexanol (30 mg/l); tartaric acid (0,7 g/l) plus hexanol (15 mg/l)
14	Orange-flavoured drink coloured yellow; orange-flavoured drink coloured orange; lemon-flavoured drink coloured yellow
15	Succession of caffeine (0,8 g/l), tartaric acid (0,4 g/l) and sucrose (5 g/l)
16	Succession of caffeine (0,8 g/l), sucrose (5 g/l), caffeine (1,6 g/l) and sucrose (1,5 g/l)

5.5 Training in the use of scales

Assessors shall be introduced to the concepts of rating, classification, interval and ratio scales (see ISO 6658 and ISO 4121) by initially ranking series of single-odour, single-taste and single-textural stimuli with respect to the intensity of a particular characteristic. The various rating procedures are then used to attach meaningful magnitudes to the samples. As indicated in 5.4, the base shall initially be water but actual foods and beverages, with mixed stimuli, both of which may vary independently, may then be introduced.

Table 7 gives examples of materials which may be used during this phase of training.

If possible, stimuli shall be chosen to relate to the product or products it is eventually aimed to assess.

Table 7 — Examples of materials which may be used for training in the use of scales

1	Products in table 3 and in item 9 of table 6				
2	Caffeine	0,15 g/l	0,22 g/l	0,34 g/l	0,51 g/l
3	Tartaric acid	0,05 g/l	0,15 g/l	0,4 g/l	0,7 g/l
	Hexyl acetate	0,5 mg/l	5 mg/l	20 mg/l	50 mg/l
4	Cheeses, e.g. mature hard cheese such as Cheddar or Gruyere, ripened soft cheese such as Camembert				
5	Pectin gels				
6	Lemon juice and diluted lemon juice		10 ml/l	50 ml/l	

5.6 Training in the development and use of descriptors (profiles)

Panelists shall be introduced to the idea of profiling by being presented with a series of simple products and asked to develop vocabularies for describing their sensory characteristics, in particular terms which allow samples to be differentiated. Terms shall be developed individually and then discussed and an agreed list of at least ten devised. This list shall then be used to produce profiles of the products, first by attributing the terms appropriate to each sample and then by scoring their intensities using the various types of scales discussed in 5.5. The organizer shall produce profiles of the products using the results to illustrate the value of descriptive analysis. Examples of products which may be used in this training exercise are given in table 8.

Table 8 — Examples of products which may be used during training in descriptive procedures (profiles)

1	Commercial fruit juice products and blends
2	Breads
3	Cheeses
4	Comminuted fruit or vegetables

See also ISO 6564.

5.7 Practice

The formal training sessions outlined in 5.4 to 5.6 shall be interspersed with exercises to provide assessors with further experience.

5.8 Specific product training

After basic training, assessors may undergo a period of product training, the exact nature of this depending on whether it is intended to use the panel for difference or descriptive testing (visual, odour, textural and flavour evaluations).

5.8.1 Difference assessment

Samples similar to those that will eventually be assessed are presented to the assessors who evaluate them using one of the difference assessment procedures. (See ISO 6658 and the International Standards describing individual difference tests.)

5.8.2 Descriptive assessment

For descriptive analyses which are not intended for a specific product, experience should be gained with a wide range of different products. For assessors who are to assess one specific product type, three samples of this type of product shall be presented in each session, approximately 15 samples being assessed in total.

Descriptors are proposed to describe the various organoleptic characteristics.

The organizer then leads a discussion to help the panel to put similar descriptors into groups and to rationalize the vocabulary by selecting a single descriptor to replace each group of terms. The process is assisted by examining external standards and samples with particular characteristics.

The agreed descriptors are then incorporated into a score sheet. Several further samples are examined and the terminology is further improved. The meaning of intensity scales for each attribute shall be discussed and rationalized by reference to actual samples.

6 Final choice of panels for particular methods

6.1 Principle

Choice of those assessors most appropriate for a given method to make up pools from which panels of assessors for particular tests are taken.

The number of assessors required for each purpose shall be at least the size recommended in ISO 6658. If the number of candidates exceeds only slightly the number required for the panel, it may be necessary to select the best assessors available rather than those meeting predefined criteria.

Candidates selected as appropriate for one purpose are not necessarily appropriate for another, and candidates excluded for one purpose are not necessarily unsuitable for others.

6.2 Difference assessments

Final panel selection is based on repeat examination of actual samples. If the panel is to be used for the detection of a particular characteristic, the ability to detect adulterated samples at decreasing concentrations can also be used as a criterion for selection.

Assessors selected shall perform consistently and be able to differentiate correctly the samples presented. Those who perform this task less well than others shall be rejected.

6.3 Ranking assessment

Final panel selection is based on repeat examination of actual samples. Assessors selected shall perform consistently and be able to rank correctly the samples presented. Those who perform this task less well than others shall be rejected.

See also ISO 8587.

6.4 Rating and scoring

Assessors shall assess approximately six different samples in triplicate, presented in random order, and, if possible, at more than one session. The results should be tabulated as shown in tables A.1 and B.1.

The data should be analysed by analysis of variance as shown in tables A.2 and B.2 to examine the individual results of each assessor.

Assessors who have a high residual standard deviation, indicating inconsistency, or for whom the variation among the samples is not significant, indicating poor discrimination, should be considered for rejection. However, if most of the assessors are poor in one or both of these respects, it may be because the

samples are not sufficiently different to be reliably distinguished.

The combined data should also be analysed by analysis of variance as shown in tables A.3 and B.3. The statistical significance of the variation among assessors, the variation among samples and the assessors/samples interaction should be determined.

Significant variation among assessors indicates the presence of bias, i.e. one or more assessors give(s) scores consistently higher or lower than the others. Significant variation among samples indicates that the assessors as a panel are successfully differentiating among the samples. A significant assessors/samples interaction indicates that two or more of the assessors have a different perception of the dissimilarities between two or more samples. In some cases, an assessors/samples interaction may even reflect a disagreement about the ranking of the samples.

Although analysis of variance is appropriate for scoring, it is not appropriate for some forms of rating. If used for rating, caution is necessary.

6.5 Qualitative descriptive analysis

No additional specific selection procedure is advocated amongst those already outlined.

Assessors are chosen on the basis of their performance in the various exercises, in particular those described in 5.6 and 5.8.2.

6.6 Quantitative descriptive analysis

6.6.1 If controls or reference samples (see ISO 5492) have been provided, candidates shall be tested for their ability to recognize and describe them. Assessors who cannot recognize or adequately describe correctly 70 % of control samples shall be considered unsuitable for this type of test.

6.6.2 Assessors shall assess approximately six samples using the vocabulary and score sheet developed as described in 5.8.2. The samples shall be presented in triplicate in a properly balanced order. Each descriptor for each assessor shall then be subjected to analyses similar to those described in 6.4 and annex A or to other multidimensional methods of analysis.

6.7 Assessors for particular assessments

Despite being selected as the best candidates, selected assessors may fluctuate in their performance. With descriptive analysis, it can often prove an advantage to select the better performers or to divide the assessors into subgroups following a programme of assessments and before any complex statistical examination of the data. For this purpose the procedures used are the same as those given in 6.4.

7 Monitoring of selected assessors

It is necessary to check periodically the effectiveness and performance of selected assessors.

The aim of the check is to examine each individual's performance to determine whether the selected assessor is able to achieve appropriate and reproducible results.

This check may be carried out at the same time as the experiment itself in many cases.

The results of this examination will indicate whether re-training is necessary.

The particular sensory tests to be used depend on the application area of the assessors and the tests are therefore selected by the panel leader. It is recommended that the results be recorded for continuous reference and be used to decide when re-training is necessary.

Annex A
(normative)

Use of analysis of variance in the choice of selected assessors for scoring

A table for the results of the assessors is constructed as shown in table A.1.

In table A.1, Y_{ijk} is the score given by the jth assessor to the kth replicate of sample i, and there are p samples, q assessors and r replicates.

For the more specific case of the final choice of panels for scoring and rating (see 6.4), $p = 6$ and $r = 3$. In this case, the analysis of variance table for the jth assessor is constructed as shown in table A.2.

Table A.1 — Results of the assessors

Sample	Assessor 1		2		j		q		Mean
	Score	Mean	Score	Mean	Score	Mean	Score	Mean	
1									
2									
i					Y_{ijl} Y_{ijk} Y_{ijr}	$\overline{Y}_{ij\cdot}$			$\overline{Y}_{i\,\cdot\cdot}$
p									
Mean					$\overline{Y}_{\cdot j\cdot}$				$\overline{Y}_{\cdot\cdot\cdot}$

Table A.2 — Analysis of variance — Data not combined

Source of variation	Degrees of freedom, v	Sum of squares, SS	Mean square, MS	F
Between samples	$v_1 = p - 1$	$SS_1 = r \sum_{i=1}^{p} (\overline{Y}_{ij\cdot} - \overline{Y}_{\cdot j\cdot})^2$	$MS_1 = SS_1/v_1$	$F = MS_1/MS_2$
Residual	$v_2 = p\,(r-1)$	$SS_2 = \sum_{i=1}^{p} \sum_{k=1}^{r} (Y_{ijk} - \overline{Y}_{ij\cdot})^2$	$MS_2 = SS_2/v_2$	
Total	$v_3 = pr - 1$	$SS_3 = \sum_{i=1}^{p} \sum_{k=1}^{r} (Y_{ijk} - \overline{Y}_{\cdot j\cdot})^2$		

In table A.2, the mean for sample i is given by

$$\bar{Y}_{ij\cdot} = \frac{\sum_{k=1}^{r} Y_{ijk}}{r}$$

and the overall mean is given by

$$\bar{Y}_{\cdot j\cdot} = \frac{\sum_{i=1}^{p}\sum_{k=1}^{r} Y_{ijk}}{pr}$$

The residual standard deviation is calculated as follows:

$$\sqrt{MS_2}$$

For the combined data, the analysis of variance table is constructed as shown in table A.3.

In table A.3, the mean for sample i is given by

$$\bar{Y}_{i\cdot\cdot} = \frac{\sum_{j=1}^{q}\sum_{k=1}^{r} Y_{ijk}}{qr}$$

and the mean for assessor j is given by

$$\bar{Y}_{\cdot j\cdot} = \frac{\sum_{i=1}^{p}\sum_{k=1}^{r} Y_{ijk}}{pr}$$

The mean of scores given by assessor j to sample i is

$$\bar{Y}_{ij\cdot} = \frac{\sum_{k=1}^{r} Y_{ijk}}{r}$$

and the overall mean is

$$\bar{Y}_{\cdots} = \frac{\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r} Y_{ijk}}{pqr}$$

The statistical significance of interaction between assessors and samples is determined by testing the ratio MS_6/MS_7 against statistical values in tables of the F distribution with v_6 and v_7 degrees of freedom.

If the interaction is not statistically significant at the level $\alpha = 0{,}05$, the statistical significance of the variation between assessors is determined by testing the ratio MS_5/MS_7 against statistical values in tables of the F distribution with v_5 and v_7 degrees of freedom.

Table A.3 — Analysis of variance — Combined data

Source of variation	Degrees of freedom, v	Sum of squares, SS	Mean square, MS
Between samples	$v_4 = p - 1$	$SS_4 = qr \sum_{i=1}^{p} (\bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdots})^2$	$MS_4 = SS_4/v_4$
Between assessors	$v_5 = q - 1$	$SS_5 = pr \sum_{i=1}^{q} (\bar{Y}_{\cdot j\cdot} - \bar{Y}_{\cdots})^2$	$MS_5 = SS_5/v_5$
Interaction	$v_6 = (p - 1)(q - 1)$	$SS_6 = r \sum_{i=1}^{p}\sum_{j=1}^{q} (\bar{Y}_{ij\cdot} - \bar{Y}_{\cdots})^2 - SS_4 - SS_5$	$MS_6 = SS_6/v_6$
Residual	$v_7 = pq(r - 1)$	$SS_7 = \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r} (Y_{ijk} - \bar{Y}_{ij\cdot})^2$	$MS_7 = SS_7/v_7$
Total	$v_8 = pqr - 1$	$SS_8 = \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r} (Y_{ijk} - \bar{Y}_{\cdots})^2$	

Annex B
(informative)

Example of practical application

An assessment by each assessor of three samples from six batches of fish stored in ice for different lengths of time gave the results shown in table B.1 using a 10-point scoring system (individual scores and means).

The analysis of variance table is then constructed as shown in table B.2.

The overall analysis of variance is then calculated as shown in table B.3.

It would be concluded that assessors 1 and 4, having low residual standard deviations and statistically significant variation between the samples, were suitable. Assessor 2, having very high residual standard deviation and consequently no significant variation between the samples, would not be suitable, nor would assessor 3 who had no significant variation between the samples.

Variation between assessors is significant, and it can be seen that assessors 2 and 3 give lower scores than do assessors 1 and 4. On the other hand, the assessors/samples interaction is not significant, and it is not possible to assert that the assessors have disagreements about the ranking of the samples.

Table B.1 — Assessor's scores

Sample	Assessor 1		2		3		4		Mean
	Score	Mean	Score	Mean	Score	Mean	Score	Mean	
1	8 8 9	8,3	5 8 9	7,3	6 7 5	6,0	9 8 8	8,3	7,50
2	6 8 7	7,0	6 7 4	5,7	5 4 7	5,3	7 7 6	6,7	6,17
3	4 5 5	4,7	5 2 3	3,3	4 3 5	4,0	5 5 5	5,0	4,25
4	6 6 5	5,7	6 4 6	5,3	4 2 4	3,3	6 5 5	5,3	4,92
5	4 5 3	4,0	3 2 4	3,0	4 4 5	4,3	4 5 4	4,3	3,92
6	5 6 6	5,7	4 2 7	4,3	5 4 6	5,0	7 5 7	6,3	5,33
Mean	5,89		4,83		4,67		6,00		5,35