INTERNATIONAL STANDARD

ISO 13299

Second edition
2016-03-15

Sensory analysis — Methodology — General guidance for establishing a sensory profile

Analyse sensorielle — Méthodologie — Directives générales pour l'établissement d'un profil sensoriel

Reference number
ISO 13299:2016(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 on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword - Supplementary information.

The committee responsible for this document is ISO/TC 34, *Food products*, Subcommittee SC 12, *Sensory analysis*.

This second edition cancels and replaces the first edition (ISO 13299:2003), which has been technically revised by presenting the principles and methods in general, including some new ones, which are developed in the annexes.

Introduction

The purpose of this International Standard is to serve as guidance for establishing sensory profiles performed by trained assessors.

A sensory profile is the result of a descriptive analysis of a sample by a panel of assessors. The sample may be for example food, beverage, tobacco product, cosmetic, textile, paper, packaging, sample of air or water, etc. Profiling can be carried out in a number of ways. Over the years, a few of these have been formalized and codified as descriptive procedures by professional societies or by groups of producers and users for the aim of improving communication between themselves.

The purpose of this International Standard is to provide agreed guidelines for descriptive sensory procedures.

Sensory profiling is the description of sensory properties of a sample, usually consisting in the evaluation of sensory attributes with assignment of an intensity value for each attribute. The attributes are generally evaluated in the order of perception. Some sensory profiles take a view across all of the senses; others (partial profiles) concentrate in detail on particular senses.

Quality of results depends on the number of assessors and their ability to describe their perceptions. Training and development of a common language help to improve these abilities. Some methods have been used with untrained assessors, but it is out of the scope of this International Standard. Quality of results can also depend on the number of replications by an assessor.

Sensory analysis — Methodology — General guidance for establishing a sensory profile

1 Scope

This International Standard gives guidelines for the overall process for establishing a sensory profile. Sensory profiles can be established for all products or samples which can be evaluated by the senses of sight, odour, taste, touch, or hearing (e.g. food, beverage, tobacco product, cosmetic, textile, paper, packaging, sample of air or water). This International Standard can also be useful in studies of human cognition and behaviour.

Some applications of sensory profiling are as follows:

— to develop or change a product;

— to define a product, production standard, or trading standard in terms of its sensory attributes;

— to define a reference "fresh" product for shelf-life testing;

— to study and improve shelf-life of a product;

— to compare a product with a reference product or with other similar products on the market or under development;

— to map a product's perceived attributes for the purpose of relating them to factors such as instrumental, chemical or physical properties, and/or to consumer acceptability;

— to characterize by type and intensity the off-odours or off-tastes in a sample (e.g. in pollution studies).

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 5492, *Sensory analysis — Vocabulary*

ISO 5496, *Sensory analysis — Methodology — Initiation and training of assessors in the detection and recognition of odours*

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*

ISO 11035, *Sensory analysis — Identification and selection of descriptors for establishing a sensory profile by a multidimensional approach*

ISO 11136, *Sensory analysis — Methodology — General guidance for conducting hedonic tests with consumers in a controlled area*

3 Terms and definitions

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

3.1
attribute
perceptible characteristic attached to a product

[SOURCE: ISO 5492:2008, modified]

3.2
sensory profile
description of the sensory properties of a sample by means of sensory *attributes* (3.1), most often with their intensity values

3.3
partial sensory profile
profile comprising certain selected sensory *attributes* (3.1), most often with their intensity values

EXAMPLE Odour profile, flavour profile, texture profile.

3.4
quantitative descriptive profile
description of a sample consisting of both *attributes* (3.1) and their intensity values

[SOURCE: ISO 5492, modified]

3.5
qualitative sensory profile
description of the sensory attributes of a sample without intensity values

[SOURCE: ISO 5492, modified]

3.6
consensus sensory profile
profile derived from agreement after discussion in a group of assessors, who evaluated the product on various *attributes* (3.1)

3.7
deviation from reference method
relative-to-reference rating
procedure of quantitative descriptive *sensory profile* (3.2) in which all samples are evaluated against a reference sample

3.8
free-choice sensory profile
procedure in which each assessor chooses and scores his/her own *attributes* (3.1) to describe a group of samples

3.9
flash profile
procedure for characterizing products by having assessors choose their own descriptive terms and rank the products on each term

Note 1 to entry: This is a variant of sensory free-choice profiling distinguished by the use of ranking rather than rating.

[SOURCE: ISO 5492:2008/Amd.1:—[1]]

3.10
temporal dominance of sensations
TDS
procedure in which each assessor is asked to successively indicate the dominant sensation over the time the product is being assessed

1) To be published.

3.11
sensory panel
group of assessors participating in a sensory test

[SOURCE: ISO 5492:2008, modified]

3.12
panel leader
person whose primary duties are to manage panel activities and recruit, train, and monitor the assessors

Note 1 to entry: This person may also design and conduct sensory tests, and analyse and interpret data.

[SOURCE: ISO 13300 (all parts), modified]

3.13
selected assessor
assessors chosen for their ability to perform a sensory test

[SOURCE: ISO 5492:2008, 1.6]

3.14
expert sensory assessor
selected assessors (3.13) with a demonstrated sensory sensitivity and with considerable training and experience in sensory testing, who are able to make consistent and repeatable sensory assessments of various products

[SOURCE: ISO 5492:2008, 1.8]

4 General test conditions

4.1 Equipment and test room

The laboratory shall have the appropriate equipment for sample preparation as specified in ISO 6658.

Sensory profiling shall be performed under the conditions specified in ISO 8589. When a discussion is needed (e.g. about results, products, reference substances, etc.), the room should be arranged in a manner that allows communication between assessors and the panel leader still ensuring appropriate conditions for evaluating products (for example, appropriate lights).

A panel leader shall be designated to perform sensory profiling. The panel leader shall

— train assessors,

— maintain the panel, and

— execute tests.

The panel leader should meet the required qualifications (e.g. steps for recruitment and training) as described in ISO 13300-1 and ISO 13300-2.

4.2 Assessors

This International Standard applies to profiling methods performed by either selected or expert assessors. Requirements for the selection, training, and monitoring of assessors can be found in ISO 8586.

The number of assessors and their training shall be adapted to the profiling method. Repeatability and reproducibility are improved with the selectivity level of the assessors and with training time. The interpretation of results and the highlighted differences between products are also dependent on the number of assessors and their training.

Candidates shall be recruited through talks, circulars, or personal contact. Two to three times the number of assessors required shall be interviewed and screened. The following characteristics shall be considered especially important:

— health that is compatible with product testing;

— interest and motivation;

— engagement for the agreed duration and availability for panel sessions;

— promptness;

— capacity to concentrate;

— ability to memorize;

— ability to honestly communicate and report sensations;

— ability to discriminate between the studied characteristics;

— ability to work in a group setting.

Sensory acuity can be balanced by establishing panels of 10 or more assessors.

4.3 Products

The products of the study and their conditions of preparation shall be defined.

EXAMPLE Soluble coffee prepared with water or milk, with or without sugar.

4.4 Samples

For the preparation and presentation of product samples, ISO 6658 shall apply. Special care shall be taken to ensure that assessors cannot draw conclusions about the nature of samples from the way they are presented. For example, coloured testing glasses or coloured lights shall be used to mask differences in appearance, if needed.

The preparation and distribution of samples at uniform temperature shall be standardized. Samples shall be coded with three-digit random numbers and the order of presentation shall be defined using an appropriate design.

To increase the reliability and validity of results, any sample or sample group shall be presented two or three times or more, if possible on different days. The choice of the number of replications shall be guided by the precision required, by the observed dispersion of results, and by any specific trend towards improved discrimination as the assessors become familiar with the samples. Replication provides an estimation of the experimental error. Repeating the assessment of a product from the same batch shows the dispersion of scores given by one assessor, whereas repeating the assessment of a product from different batches also reflects variations within the product. The protocol shall define which sample(s) is/are duplicated and under which conditions they are prepared and assessed.

The identity of the samples shall not be disclosed until the assessors have completed all the assessments.

4.5 Preliminary discussion

It shall be ascertained that the assessors are fully familiar with any particular characteristic to be studied and with the mechanics of the test as specified in ISO 6658. If necessary, a preliminary general discussion concerning the test problem and the nature of the samples shall be arranged. A few samples typical of the product category shall be presented and discussed. The panel leader shall make sure that the discussion does not bias future assessments.

5 Descriptive methods: principle and main characteristics

5.1 Consensus profile

In the consensus profiling, the assessors share their individual views to achieve a consensus on the different attributes, their order of appearance, and their intensity.

Usually, the scale is limited to a few marks. Results shall consist in a single score (the agreed one) for each attribute. It is possible for an assessor to disagree with the group: this shall be recorded in the report.

5.2 Deviation from reference profile (relative-to-reference scaling)

The products shall be presented in pairs. For each attribute of a common list, the two products shall be compared to one another, either directly by the assessors or, *a posteriori,* from the scores given to each product of the pair. If more than two products need to be compared, each product should be compared to the reference product under the same conditions.

Data analysis is performed on the differences between the samples and reference.

5.3 Free-choice profile

In the free-choice profile, each assessor shall use his/her individual list of terms instead of a common list.

The results shall be interpreted with an appropriate multidimensional analysis such as generalized procrustes analysis. The output shall be displayed in the form of a map.

5.4 Flash profile

The flash profile is a variant of the free-choice profile, with a simultaneous presentation of the whole sample set and comparative evaluation of the samples via ranking.

The results are interpreted with an appropriate multidimensional analysis such as generalized procrustes analysis. The output is always in the form of a map.

5.5 Quantitative descriptive profile

In the quantitative descriptive profiling, the assessors evaluate samples on a common list of attributes and score their intensity.

There are several methods for establishing a quantitative descriptive sensory profile, among which some techniques have been trademarked[2)]. Results shall consist of intensity scores for each attribute that can be submitted to univariate or multivariate analyses.

5.6 Qualitative sensory profile

In the qualitative sensory profile, the assessors shall evaluate only the presence or absence of the attributes from a common list of terms without indicating their perceived intensity.

The list of attributes is larger and less product-dependent than for a quantitative descriptive sensory profile. The training of the panel shall be focused on the recognition and memorization of numerous references. References that are stable and do not change over time are necessary for memorization.

The number of assessors and/or replicates should be higher than for quantitative descriptive sensory profile.

2) Methods QDA®, Spectrum™ are examples of suitable procedures available commercially. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of these procedures.

Results shall be expressed as frequency of quotation of each attribute.

5.7 Temporal Dominance of Sensations (TDS)

TDS is a temporal profiling technique in which each assessor is asked to successively indicate the dominant sensation over time while the product is being assessed.

The dominant sensations are chosen from a common list of attributes. As an option, the intensity of the dominant attribute chosen can also be scored.

Data shall consist of the proportion of each attribute chosen as dominant at each moment. Data are usually converted into curves where time defines the *x*-axis. The curves of the different attributes for a given product shall be pooled into a chart.

6 Procedure for establishing a sensory profile

6.1 General

This Clause presents the steps common to sensory profiling methods. For a detailed description of each method, refer to the corresponding annex.

6.2 Prepare the test

6.2.1 Select products for training

Products for training shall be selected as specified in ISO 8586.

6.2.2 Select assessors

Assessors shall be selected as specified in ISO 8586 and ISO 5496.

6.2.3 Choose the optimal attributes

The purpose is to identify and select a set of non-overlapping, singular, objective, unambiguous, and referenced attributes that, as far as possible, permit a complete descriptive analysis of the samples under study. This important step can be done individually or collectively and depends on the sensory profiling method. If a common list is needed, the panel leader may use one of the three approaches set out in Table 1 or any combination.

Table 1 — Procedures for choosing optimal attributes

No.	Principle	Method	Advantages	Disadvantages
1	Use existing terminology and reference standards.	Consult the literature and experts to make an appropriate selection. Acquire the prescribed standards and use these to teach the assessors the quality of each descriptor and, if needed, an intensity scale for that descriptor.	The accumulated experience of the experts is utilized. Profiles may be interpreted by other groups and compared to other research.	Existing terminology or reference standards may include choices that are imprecise or inappropriate for a particular set of samples. Attributes may be missed that could have been discovered during the development of new terminology.

Table 1 *(continued)*

No.	Principle	Method	Advantages	Disadvantages
2	Use the panel in special sessions to develop the terminology it will use.	Use a panel of selected assessors; develop the terminology in round table discussion under the direction of an experienced panel leader. Reference standards are used which may be provided by the panel leader or the test requester, or by an assessor during the session. May be combined with method 1.	The process of terminology development is less time-consuming than method 3.	Profiles obtained are unique to a given panel and set of samples. They cannot be interpreted by other groups if no reference standards are given.
3	Use the panel in special sessions to generate the terminology it will use.	Consult ISO 11035 which describes a recommended method of identifying and selecting discriminating terms using a set of prepared training samples; then reduce the number of terms by stepwise elimination using statistical techniques.	A fully objective process of selection and elimination is used, thus terms based on traditional misconceptions or preconceived notions are minimized, and the selected terms will give optimal coverage of the qualities which the assessors perceive in the samples.	Profiles obtained are unique to a given panel and set of samples. They cannot be interpreted by other groups if no reference standards are given. The process is relatively time-consuming and requires a certain level of experience, especially in data analysis.

As a part of a profiling session, it is possible to ask assessors to provide one or more overall evaluations. Examples include

— overall fruitiness or spiciness, and

— overall flavour intensity.

CAUTION — An overall evaluation by trained sensory panels shall not be hedonic because they can be biased: sensory assessors have been trained to be objective in describing products and may, consciously or unconsciously, adopt a different evaluation strategy from a representative consumer target. If hedonic evaluations are used to guide new product development, *proceed as specified in* ISO 11136.

6.2.4 Determine the order of evaluation

In some products (e.g. beverages), the order of perception of certain attributes is a characteristic of the product profile. In other products, the order of perception can change during the evaluation, for example in accordance with textural or physical properties.

EXAMPLE A piece of chocolate which melts; a facial tissue which is moistened.

Usually, the order of perception should determine the order in which attributes are evaluated. Aftertaste or afterfeel should be evaluated last; these are attributes which can still be perceived after the sample has been consumed or used. In oral texture, as in skinfeel and fabric feel applications, the order of perception can be predetermined by the way the product is handled. The panel leader shall control manipulation (one chew, one manual squeeze) in order to determine which attributes are perceived first.

6.2.5 Select an appropriate response scale

When the profiling method consists of indicating the intensity of each attribute present in a given sample, the response scale should be selected carefully. Response scales used in sensory profiling may be numerical or semantic, continuous or discontinuous, unipolar or bipolar, as presented in ISO 4121.

NOTE Data obtained with response scales may suffer from so-called “end effects”: assessors are reluctant to use numbers near either end of the scale for fear that later samples may be more extreme. Continuous response scales, e.g. line scales, are considered as less prone to produce end effects.

6.2.6 Train the assessors

Assessors shall be trained as specified in ISO 8586. It is recommended that feedback be given to assessors.

6.3 Conduct the test

6.3.1 Scoresheets

Pre-printed scoresheets containing instructions regarding the scale to be used or computer screens or tablet digitizers to record the verdicts shall be used. A blank space shall be left on scoresheets and assessors’ comments or suggestions for additional attributes shall be requested.

NOTE Several ratings placed on a single sheet or screen may mutually influence or distort each other (i.e. halo effect: a positive or negative evaluation may carry over from one attribute to the following).

6.3.2 Evaluate the samples

It shall be arranged for assessors to work alone in a booth.

Depending on the method, samples shall be presented either monadically (in succession, one-by-one) or simultaneously. Simultaneous evaluation facilitates the comparison of samples but profiles become interdependent. In both cases, the order of samples from one assessor to another should be balanced or randomly assigned to reduce the impact of order effects.

The number of samples per assessor and per session should be adapted to the length of the session, the nature of the products, the number of attributes, and the expected differences: present a limited number of samples if small differences are expected, and for samples with strong or persistent flavour.

6.4 Statistical interpretation

Investigating and visualizing the raw data prior to further analysis is always recommended in order to

— detect and correct obvious mistakes,

— eliminate outliers from the analysis, if appropriate,

— obtain an initial impression of the main structures of the data set, and

— discover possible tendencies that may be of interest later on in the analysis.

Various techniques are available: mean values accompanied by standard deviations calculated for the complete data set or with focus on either individual assessors or specific attributes, box plots, histograms, line plots.

The statistical treatments depend on the method. In a very general way, the interpretation involves three steps. A first step focuses on the performance of the assessors and checking the data for any experimental errors. A second step focuses on each of the descriptors evaluated and seeks to determine the descriptors that best discriminate the study’s products. This step is often called: univariate analysis (see Annex H). A third step takes into account all descriptors estimated useful in the first stage. This step is often called multivariate analysis. It can be performed after a segmentation of the descriptors, for example visual descriptors, flavour descriptors, taste descriptors, texture descriptors.

NOTE There are various multivariate techniques and they cannot all be presented in this International Standard. However, they are mentioned in the annexes presenting the different profile methods.

6.5 Study report

The study report shall include the following information:

a) name of the company/laboratory which performed the study, names of the panel leader and study supervisor;

b) objective of the study;

c) full identification of sample(s);

d) date(s), time of the test(s), and duration of sessions;

e) operating conditions of the study (including any conditions differing from the recommendations given in this method):

   1) reference to this International Standard, i.e. ISO 13299, and quality control procedures;

   2) detailed design of the study: order and mode of presentation of samples, definition of replicates [same/different day(s), same/different batch(es)];

   3) number and type of assessors (selected assessors or experts);

   4) list of attributes evaluated with their definition, their protocol of evaluation, and the reference substances;

   5) response scales used;

   6) any other information given to the assessors during the study (e.g. information regarding the type or brand of product);

f) results, with statistical analysis and appropriate representations (graphs, diagrams, maps, etc.).

If agreed with the partner/sponsor, the content of the report may be less exhaustive.

Annex A
(informative)

Consensus profile

A.1 Principle

In consensus methods, assessors evaluate products and then discuss the attributes and/or intensities to determine an agreed description of the product in question. Consensus requires that agreement has to be reached rather than averaging data over the group. In consensus methods, the panel has to be extensively trained to determine properties and able to make decisions about the attributes and intensities as they tasted.

Historically, the consensus was a key element of both techniques known as the Flavor Profile Method (FPM) developed by Arthur D. Little, Inc. in the late 1940s and the Texture Profile Method (TPM) developed by the Product Evaluation and Texture Technology groups at General Foods Corp. during the 1960s and which gave birth to the Spectrum™[3] Method during the 1990s. But there is a fundamental difference between these two techniques: in the first FPM, the consensus is required, in the second TPM, the consensus is optional (see Clause 6).

A.2 Sensory attributes

Assessors, with the panel leader's help in providing and maintaining reference samples, develop, and define the common terminology to be used by the entire panel. Reference standards and definitions for each descriptor are created during the training phase. In the FPM, the assessors describe the flavor notes and estimate their intensity, but also evaluate the "general impression" (called amplitude) caused by the product. The TPM uses a standardized terminology to describe the textural characteristics; these are described by both their physical and sensory aspects.

A.3 Scales

Originally, the scale of the FPM had 5 points.

0: the character is not present;

) (: the character is at threshold level;

1: the character is slight;

2: the character is moderate;

3: the character is strong.

But this scale has subsequently been expanded, including the use of arrows, 1/2s, or plus and minus symbols. The original TPM used an expanded 13-point version of the Flavor Profile Method. Today, the most common for TPM is probably a 15-point scale; a 10-cm or 15-cm line scale is also possible.

3) Methods QDA®, Spectrum™ are examples of suitable procedures available commercially. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of these procedures.

A.4 Assessors

Two features characterize the assessors involved in the establishment of a consensus profile.

— There are few of them (4-6 in FPM).

— They are highly trained. In the FPM, assessors are provided with a broad selection of references samples representing the product range as well as examples of ingredient and processing variables for the product type. In the TPM, assessors need to learn to use in a reproducible way different points on the scale.

See Reference [16].

A.5 Procedure

— The assessors individually evaluate one sample at a time and they record the intensities of attributes and their order of appearance.

— At the end of the evaluation of a product by the group members (including the panel leader), the results are collected by the panel leader. He/she leads a general discussion to reach a consensus profile on the attributes, intensities, and any other aspects (such as the order of the attributes). Discussion can take several minutes or several tens of minutes. Reference samples can be provided to enrich the discussion. Once the panel has reached consensus, a final sensory descriptive “profile” is created.

— When there is no agreement, the panel leader may decide to present the product in a new session or to record the differences among assessors.

— This process is repeated until all products have been evaluated.

— The panel leader has to be careful to avoid a senior member or a dominant personality imposing his or her views.

The above procedure is the basic procedure. Many variants exist.

A.6 Statistical analysis and interpretation of the results

There is no statistical analysis in the FPM. The data are generally reported in tabular form, although a graphic representation is possible. In the TPM, the panel verdicts may be derived by group consensus or by statistical analysis of data. For final reports, the data may be displayed in tabular or graphic form.

Various multivariate analyses may be used (such as “tree” map or PCA) to highlight the similarities and differences between products.

Annex B
(informative)

Deviation from reference method (or relative-to-reference rating)

B.1 Principle

Deviation from reference method, also called relative-to-reference rating, was first developed by Larson and Pangborn (1978). This method is a variation of generic descriptive analysis where a reference sample is used to assess all the other samples against it (samples are presented by pairs). This method is particularly suitable when there is an obvious reference or when distinctions are difficult. There are two different ways of performing such a method: either the samples are directly compared by the assessors to a reference sample and scored regarding the reference sample (the reference sample being a midpoint anchor on the scale) or, the samples are compared by computing the difference between the scores given to the reference sample and to the samples. Usually, in this case, the assessors don't know which sample is considered as the reference sample.

B.2 Sensory attributes

The list of sensory attributes is common to all the assessors and developed like for quantitative descriptive analysis. The development of the list is conducted during the training period, either starting from a "closed list" with predefined terms or entirely generated by the panel itself.

B.3 Scales

Response scales may be structured or unstructured.

When the samples are directly evaluated relative to the reference sample, the scale is bipolar with the reference as a midpoint anchor. In this case, the assessors score the sample as more or less than the reference sample, for a given attribute (Figure B.1).

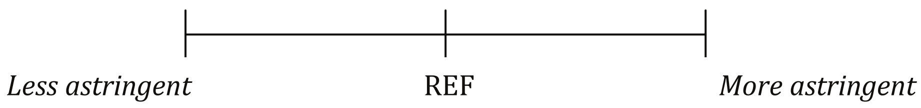

Figure B.1 — Example of unstructured scale used when the reference sample is used as a midpoint anchor

When the samples are not directly evaluated relative to the reference sample (the difference is computed a posteriori), the scales are the same as the ones used for quantitative descriptive analysis but the two scales, for a given attribute, are close to one another (Figure B.2).

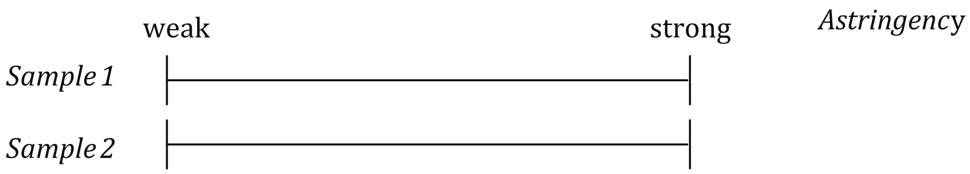

Figure B.2 — Example of unstructured scale used when the difference between reference sample and sample is computed

B.4 Assessors

Usually, the assessors are selected and trained on the products to be tested as in the quantitative descriptive profile. However, it can be assumed that the use of a common reference sample in each assessment will facilitate the alignment between assessors as well as the stability of scoring between separate sessions.

B.5 Procedure

Usually, the training procedure is similar to that of the quantitative descriptive profile. It includes sessions to familiarize the assessors with the product category, to develop and master the list of attributes, and to score the intensity of each attribute on the scale.

For the evaluation, the assessors receive a set of two samples. It is recommended to balance the presentation design between assessors and between sessions.

— When the samples are directly compared, one sample is defined as the reference sample and the other sample is assessed relative to it (see Figure B.1). In this case, the reference sample is assessed prior to the other. Usually, the assessors are allowed to come back to the reference sample as often as necessary to assess the second sample.

— When the samples are not scored relative to one another, the assessors score the two samples (see Figure B.2). The assessors are usually allowed to taste the two samples several times. When it is possible, the assessors may simultaneously evaluate the two samples (i.e. cosmetic products simultaneously applied on the right and left sides of the face). The difference within the two samples is computed a posteriori.

In both cases, the reference is often included as a sample (a “blind reference”, not identified as a reference) and used as an internal measurement of assessor reliability.

In certain conditions, it is technically possible to compare more than two samples.

— Each sample should be compared to the reference sample under the same conditions.

— The reference sample should be selected carefully and should be in the middle of the product category.

As in the quantitative descriptive profile, the evaluation is usually replicated.

B.6 Statistical analysis and interpretation of the results

In both procedures, data analysis is performed on the differences between the samples and the reference. Samples that score less than the reference are indicated negatively in comparison to the midpoint anchor, whereas those that score more than the reference are indicated as positive. Classical univariate statistics can be applied. When there are only two samples to compare (one being defined as the reference sample), a two-tailed paired *t-test* would be appropriate. When all the samples have to be compared one to each other as well as to the reference sample, provided that the conditions previously cited are fulfilled, univariate and multivariate analyses can be applied. In this case, for the reference, the values of the blind reference can be used instead of the values of the reference sample evaluated in each pair.

Annex C
(informative)

Free-choice profile

C.1 Principle

The free-choice profile (FCP) is a procedure in which each assessor creates his/her own list of descriptive terms and quantifies the intensity by predetermined or self-selected scales.

FCP is used to compare three or more products with many attributes and to find differences. It is useful to establish a perceptual mapping of product spaces (for example, for market research). It can be also used for generating attributes.

C.2 Sensory attributes

The assessors create the descriptive attributes individually without need for explanation of the meaning of these terms. There may be a discussion between the assessors about the attributes that may help describe the products. This discussion does not aim to get a consensus: each assessor takes his/her own list. The terms can differ individually and be understood only individually. However, each assessor evaluates all products consistently with his/her own individual list of attributes. He/she may choose all or only some properties like shape, colour, taste, smell, etc. Some assessors may use only a few attributes. There is no rule to take special attributes. An assessor may describe using whatever words and/or word-combinations he/she wishes (e.g. “smells like camping”).

C.3 Scales

Assessors quantify the intensity by predetermined or self-selected scales. They are allowed to evaluate the products differently on different scales. However, each assessor should evaluate all products consistently with his/her own scales. There will be no training on intensities or scales.

C.4 Assessors

Assessors should be able to see, smell, taste, touch, and hear. They are not previously trained specifically on a product. It is required to select assessors with good descriptive abilities. They should have a comprehensive vocabulary to describe the products with a copious amount of attributes. No further training is necessary.

A minimum of 10 assessors is recommended.

C.5 Procedure

In a first session, the assessors collect the attributes in individual booths and choose either the scales they wish or the predetermined scales.

In a second session, the assessors use their attributes consistently for every product they have to compare and perform intensity evaluation of the products under standard conditions. The samples are presented in randomized order. There should be two or more replications for each assessor and product.

A method which can be used instead of FCP is the Repertory Grid Method. The method to generate attributes is more structured than FCP. The generation of attributes is improved by choosing simple attributes and by a succession of comparisons. In the first comparison, the attributes are generated by comparing two to four products (usually three). Then, in the next comparison, one of the products of the first group is carried over and compared with two more products. The assessors are asked how the two products are similar or different from the first product. In this way, the tests can go on with further triads.

C.6 Statistical analysis and interpretation of the results

The complete description of products cannot be statistically evaluated by ANOVA, PCA, or linear regression methods, which are not applicable. FCP data are analysed using a method such as General Procrustes Analysis (GPA). This method analyses the data by every evaluation and repetition in matrices for every assessor to get a consensus configuration by transformation (rotation, translation, scaling) of the data.

Annex D
(informative)

Flash profile

D.1 Principle

Flash profiling was first put forward by Sieffermann in 2000. It is a combination of FCP and a ranking method, the assessors generating attributes individually by describing the differences between a set of products. It shows similar groupings of products in a sensory space, by giving sensory mapping. The direct comparison of attribute intensity by ranking stresses relative differences.

D.2 Sensory attributes

Each assessor is requested to describe the differences between the products in the set on his own and individually generate attributes. This is an important requirement for successful profiling. It is not possible to choose hedonic terms.

There may be a discussion between the assessors about the attributes that may help describe the products. This discussion does not aim to get a consensus: each assessor takes his/her own list.

D.3 Scales

There is no specific scale to be used as assessors are asked to rank the different samples on each attribute.

D.4 Assessors

Assessors should be able to see, smell, taste, touch, and hear. They are not previously trained specifically on a product. It is required to select assessors, with good descriptive abilities. They should have a comprehensive vocabulary to describe the products with a copious amount of attributes. No further training is necessary.

A minimum of 10 assessors is recommended.

D.5 Procedure

One to two sessions to instruct the assessors about the method and how to describe and differentiate between the products are first conducted. Usually, in these sessions, each assessor gets a set of products and has to describe the products individually. This can be done by sorting samples to find out the differences between the products. Afterwards, this will be controlled by the panel leader and the assessor to withdraw irrelevant or hedonic attributes and add relevant ones. These are training sessions on description and differentiation.

The next sessions are spent to the measurement. These consist in ranking the products according to their intensity on each attribute. This is performed one attribute after another. It is recommended to replicate the evaluation of some (within the same session) or all samples (in separate sessions).

Unlike FCP, all products are presented simultaneously in a balanced block design or BIB (balanced incomplete block). The design depends on the number of products which are to be assessed. A minimum of five to six samples is recommended. When more than 12 samples have to be evaluated, this should be done in a BIB.

D.6 Statistical analysis and interpretation of the results

Similar to FCP, the statistical evaluation is done with a multidimensional analysis such as general procrustes analysis (GPA). The correlation between the repetitions can be calculated to show the reliability of the assessors (Spearman-rank correlation).

Annex E
(informative)

Qualitative sensory profile

E.1 Principle

Qualitative profiling is a citation frequency-based technique used principally to characterize complex aromatic products such as wine or cheese. A large list of descriptors is provided to assessors, who have to choose the most pertinent ones to describe a given product. Scores are computed from the number of times a term is selected (frequency of citation), for a given product by the assessors. The large number of available terms allows assessors to precisely describe their perception using the most appropriate term for them. The test is conducted by sensory experts (trained assessors).

E.2 Sensory attributes

The list of attributes evaluated is a “closed list” with predefined terms or developed and adapted to the product. The development of the list is conducted during the training period. The number of attributes is larger than for the quantitative profile method: the number of different terms evaluated may reach around 100. The attributes are focused on aromas or odours and correspond to precise aromatic references. Each term has to be associated with a standard aromatic reference (either commercially available references or natural products).

E.3 Scales

This method does not use scale. The evaluation consists only in selection, among the list of terms, of perceived attributes of the product: detected or not detected.

E.4 Assessors

E.4.1 Number of assessors.

The number of assessors is higher than for a quantitative descriptive profile: generally twice the panel size (from 16 to around 30 qualified assessors).

E.4.2 Qualification and training

The training period is generally composed of two phases: general and product specific training. During the general training (5 to 15 sessions or more, depending on the number of properties to memorize), assessors become familiar with the terms of the list, and smell several standard aromatic references to choose from those which best fit their concept for each term.

When the list is developed and adapted to the product, during the general training, assessors discuss the list and modify it by eliminating the terms they consider irrelevant, ambiguous, or redundant and by adding the terms they consider pertinent.

The specific training consists in sessions in which assessors still smell standard aromatic references and become familiar with the products of the study.

E.5 Procedure

Trained assessors receive a whole set of products. The order of tasting of each product is balanced. For each product, each assessor is requested to select the perceived aromas or odours in the list of terms, but they are limited in the number of terms they can select. The maximum number of terms is defined according to (1) the possibility of allowing assessors to express their perceptions with a relatively high number of terms and (2) the limited capacity of humans to discriminate odour qualities in a mixture.

The evaluation is usually duplicated.

E.6 Statistical analysis and interpretation of the results

Once all data are collected, terms present in the list are ranked according to their citation frequency to identify the most relevant descriptors of each product. Only descriptors cited by a minimum proportion of assessors (generally 10 % to 15 % of the panel) in, at least, one product/repetition are considered for subsequent statistical analysis. A contingency table (products in rows and descriptors in columns) containing the citation frequency of the most cited terms is constructed.

Several statistics methods could be used to analyse the results, but the most common are the following:

— chi-squared (χ^2) analysis on the citation frequency (two repetitions) of each term, to find significant differences between the attribute “intensity” in the products;

— multidimensional analysis: correspondence analysis performed on the contingency table containing the citation frequency of terms cited by more than 15 % of the panel in, at least, one product/repetition.

Annex F
(informative)

Quantitative descriptive profile

F.1 Principle

The Quantitative Descriptive Profile (QDP) Method is a type of trained panel procedure in which all the sensory properties of a product are described and their intensities quantified.

The sensory properties, called attributes, are generated by the panel regarding the product category and the training panel is specifically oriented to the product category that will be tested.

Profiles obtained are panel specific and product category specific. They cannot be interpreted by other groups if no reference standards are given.

F.2 Sensory attributes

The Quantitative Descriptive Profile uses trained assessors to choose and define the attributes necessary to differentiate between the samples of the given product category.

Therefore, the first training sessions are usually dedicated to generating sensory characteristics (e.g. appearance, aroma, flavour, texture, and aftertaste attributes for food products) and to verifying the right use of attribute scales for the product range to be tested in the given project.

Often, the attribute generation stage identifies a large number of attributes which is reduced to a limited list, by consensus through discussion or by multivariate analysis (see ISO 11035).

F.3 Scales

Once the relevant attributes have been selected, the next step is to choose an appropriate response scale by which to indicate the intensity of each attribute present in a given sample.

Response scales used in sensory profiling may be continuous (10 cm to 15 cm) or discontinuous. Usually, the extreme limits are nil (or very weak) and intensive (or very strong).

F.4 Assessors

This method uses trained assessors, who after having passed selection tests, are intensively trained on the product category to be tested (see ISO 8586).

F.4.1 Qualification

F.4.1.1 Recruitment and selection

This method requires qualified respondents who are users/acceptors of the product being tested. During selection tests, they have to demonstrate that they can

— discriminate differences among the products being tested, and

— be able to memorize the nature of sensory characteristics.

Assessors also have to commit to attend all sessions required for successful completion of the project

F.4.1.2 Training

The training is product specific.

The selected assessors working under the direction of the panel leader develop a sensory language, or modify an existing one, to describe all of the product's sensory properties.

This task is made by the way of products that illustrate the range of attributes to be encountered. It is very important to include in this step a representative range of the product's category and their deviations (all types of products, including atypical ones, various products…).

The group develops definitions for each one of the attributes as well as a standardized evaluation procedure. A final task is to practice scoring the products.

Reference standards are used which may be provided by the panel leader or the test requester, or by an assessor during the sessions. Usually, the standard references are different from the tested set of samples.

These training activities can require anywhere from 10 h to 20 h or more for one product category. The duration of the training depends on the specificities of the product and on the task that will be assigned to the group.

The panel leader should organize the training, provide the products, introduce references when needed and facilitate the activities, but should not act as a respondent.

F.4.1.3 Performance assessment

At the conclusion of training, monitoring performance data enables the panel leader to improve panel and assessor performance, to identify a lack of training, or to identify those assessors who are not performing well enough to continue.

The panel evaluates the products using a repeated trial design. For most tests, a two or three replicates design is performed.

Statistical analysis is performed in order to check, for example, homogeneity and repeatability of the whole panel and of each assessor.

F.4.2 Number of assessors

A minimum of eight assessors should be used for this method, but 12 to 15 assessors are recommended.

F.5 Procedure

— Usually, the samples are presented using a sequential monadic design.

— The recommended number of replicates is two or three.

F.6 Statistical analysis and interpretation of the results

F.6.1 General

After a visual inspection of data, several statistics methods can be used to analyse the results. The most common one is the Analysis of Variance (ANOVA). Typically, the analysis used is a 2-way ANOVA, taking into account the variability among products and among assessors. However, other sources of variability such as "Replicates" can be considered. These statistical analyses are presented in Annex H.

In addition, multivariate methods may be used to explore the interaction between the attributes. There are many methods that simultaneously analyse the effect of several products on a number of attributes. The PCA (Principal Component Analysis) is probably the most versatile and the most well-known. It is particularly suitable in situations with lots of data and when little prior information is available. PCA can be performed on the variance-covariance matrix or on the correlation matrix. The first one does not give the same weight to all attributes: the weight varies according to the discriminative power of each attribute in the study. The second one gives the same weight to all attributes, regardless of their discriminative power. It should be performed only after elimination of poorly discriminative attributes. The critical probability leading to the rejection of an attribute as poorly discriminative is generally greater than 0,05; it varies between 0,10 and 0,50.

F.6.2 Presentation of results

Results are usually presented in a numerical format and in a graphical presentation (spider web graph, profiles curves, specific charts from multidimensional analysis).

F.7 Variants

F.7.1 Quantitative Descriptive Analysis®[4)]

The Quantitative Descriptive Analysis (QDA)®[3)] Method relies on measurement and quantification. QDA has been described in many publications and books, including References [9], [27], and [28]. From QDA's inception, its developers compiled best practices and knowledge of human sensory measurement and the method has evolved as practices and the information base have grown. The evolution of the original QDA method continues today as Tragon QDA®[3)].

This approach can be used for a wide variety of purposes including understanding the language consumers use to describe and differentiate products, mapping perceived product similarities and differences, ingredient substitution, new product development, competitive assessments, claims substantiation, and advertising, among other uses. When correlated with consumer affective measures (e.g. liking, preference, attitudinal, emotional, packaging, usage), the data can be used to determine key sensory attributes that most impact consumer choice behaviour and to understand the reason why consumers have unique preferences in segmentation. Developers can use the information to formulate products that target specific consumer benefits and needs, and marketing can use the language and sensory properties to help communicate said benefits.

F.7.1.1 Sensory attributes

Language development requires 8 h to 12 h of group discussions. This is an iterative process where each session builds on previous sessions to develop a comprehensive language. Some language development activities may include observational behaviour, in home, and/or extended usage situations. After initial language development sessions and pilot testing, remedial language sessions may be scheduled. It is not unusual to have 30 to 40 or more sensory attributes to fully describe perceptions before, during, and after usage.

F.7.1.2 Scales

QDA uses unstructured or semi-structured line scales for measuring and scaling perceived differences and intensities.

QDA uses an unstructured or semi-structured (6 inch/~15 cm) line scale, anchored 0,5 inch from either end for measuring and scaling perceived differences and intensities. These equal-interval scales are described in psychophysics literature.

---

4) Methods QDA®, Spectrum™ are examples of suitable procedures available commercially. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of these procedures.

F.7.1.3 Assessors

F.7.1.3.1 Number of assessors

Twelve assessors are recommended.

F.7.1.3.2 Qualification and training

F.7.1.3.2.1 Recruitment and selection

Approximately 30 candidates/consumers who are likers and above average users of the product category are recruited, screened, and selected for their sensory acuity and verbal fluency, in addition to availability.

F.7.1.3.2.2 Training

The training is conducted by the panel moderator or panel leader. The panel leader is the group discussion facilitator and provides the schedule of activities and works with the panel to help them develop the common vocabulary to describe the products of interest. The QDA moderator is not in a teaching role but works to observe behaviour and understand consumer perception of the product in typical consumer usage situations.

Sensory acuity screening is based on up to 30 trials with repeated measurement using discrimination method. Tests represent differences expected in the product set and category of interest and range from easy to moderate to difficult differences to detect.

F.7.1.3.2.3 Performance assessment

The QDA method provides for statistical analysis of panel performance including individual assessor performance by attribute, replication, and overall differences observed relative to the panel as a whole. The analysis then focuses on perceived product similarities and differences. The analysis of variance is specified and a detailed description can be found in Reference [5].

F.7.1.4 Procedure

Evaluation procedures are developed to be typical for the category of interest and then standardized. Evaluation procedures are modelled after typical consumer usage behaviour for that category and may require more than one evaluation procedure.

A minimum of three replications are recommended, following a complete balanced block design.

F.7.1.5 Statistical analysis and interpretation of the results

The statistic method used to analyse the results is the Analysis of Variance (ANOVA).

QDA data can be displayed in various ways but the most common are spider charts or radar plots, with means tables and significant differences highlighted.

F.7.2 Spectrum™[5] Method

F.7.2.1 Principle

The Spectrum™[5] Method is based on a descriptive profiling procedure that includes using documented references for both qualitative attributes and intensity scale points. Historically, the

5) Methods QDA®, Spectrum™ are examples of suitable procedures available commercially. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of these procedures.

Spectrum™5) Method was derived from the Flavor Profile and Texture Profile methods by expanding the scale, documenting more attribute references, and expanding the method beyond food descriptive evaluations to include skinfeel and fabric feel procedures. The method has precise steps and procedures at every stage of development from selection of assessors, panel leadership, panel training, validation, and maintenance of panel after training is complete. These practices lead to a descriptive panel that produces reproducible statistically robust data across multiple sessions and categories.

F.7.2.2 Sensory attributes

The Spectrum™5) Method uses an assessor group that has undergone the Spectrum™5) training in the broad category of interest (food, fragrance, skinfeel, fabric, paper, etc.) to generate the sensory attributes for a project or a product. The assessors develop the attributes/lexicon using outlined steps and guidance from a skilled panel leader. These attributes are identified with both physical external references and written definitions. The resulting ballot/lexicon should be able to describe and discriminate among samples in the product category. The attributes that are selected should be perceivable, primary, orthogonal, and singular.

F.7.2.3 Scales

The Spectrum™5) Method scale is based on a 0 to 15 point intensity scale with the ability to rate in tenths for 150 points of discrimination. This scale was expanded from historical scales to give assessors the ability to discriminate using smaller points of difference. The Spectrum™5) scale is a universal scale that covers the entirety of intensities within a scope such as the world of foods or skinfeel experiences. This is achieved by using an absolute intensity scale where 15 is not a true end point to allow for scale expansion as necessary. Even though most products fit within the 0 to 15 points, higher intensity attributes can be rated beyond 15 as necessary (e.g. extreme sour, higher greasiness). The philosophy and training is designed where the intensity points are related across attributes for comparison of intensities within and between samples. For example, a five intensity of sweetness is the same intensity as a five in bitterness or a five in mint aromatic. The Spectrum™5) Method universal scales are based on a very large array of food products that range in intensity from just detectable (threshold) through slight and strong to very strong. The published values for the Spectrum Universal Scale references are the result of replicated ratings from 10 different industry panels. Changes to scale values (because of routine product changes) are reviewed regularly and are published in Reference [4].

F.7.2.4 Assessors

F.7.2.4.1 Qualification

F.7.2.4.1.1 Recruitment and selection

Assessors are selected from a pool of recruited individuals based upon several qualification measures. First during recruitment potential assessors are screened for targeted health concerns, availability, and interest in job function/hours, basic visual scaling, and verbal ability to describe products or scenarios. As a general rule, three times the final desired number of assessors are recruited based upon selected characteristics for the project. Selected assessors that pass basic pre-screening requirements are then interviewed and screened for relevant acuity skills. The resulting group of assessors (recommended 18 to 20) has passed pre-screening and screening criteria and are strong performers in acuity and group interaction dynamics.

F.7.2.4.1.2 Training

The training procedure is completed with a series of intense training weeks each followed by a number of practice sessions to reinforce the material studied during the training phase. An example schedule might be 28 h of training with 45 h of practice time following each training section. The training first consists of introducing the overview philosophy of the Spectrum™5) Method and working through some broad categories related to the panel area of work. This initial training and practice time introduces the overall method, intensity scales, and the protocol for how attributes are selected for a product lexicon/ballot. Later training and practice sessions may focus on specific areas of interest for the panels anticipated work.

F.7.2.4.1.3 Assessing of performance

At the end of the training phase, a validation study determines how the panel is performing in its ability to produce reliable and consistent data. A validation study is used to show overall panel performance, individual assessors' performance and to help identify potential areas for retraining. Data from a SpectrumTM5) trained panel should be both reproducible and able to discriminate among products. The validation study is designed using samples and replicate evaluations to test both of these aspects of performance. It is recommended that performance is measured not just at the end of the training phase but also for functionally working panels (e.g. one time yearly). The concept of panel validation is parallel to the calibration of any other scientific instrument.

F.7.2.4.1.4 Number of assessors

The SpectrumTM5) panel typically uses 10 to 12 assessors. Exceptions for lower numbers (recommended minimum 6 assessors) may be used depending on number of years working as functioning group, validation performance, and data collection methods (e.g. consensus data).

F.7.2.5 Procedure

The SpectrumTM5) Method uses a documented five-step procedure to build a lexicon of attributes for a product space.

a) Frame of Reference step – Assessors experience a broad array of products that define and expand the area of interest. Products may include market products and experimental samples for ideal term generation.

b) Term generation/grouping – Assessors provide their own descriptors for the products based on their own experiences. The panel leader groups these terms based on similarity and with the panel drafts a lexicon.

c) Attribute references – Assessors are exposed to references for the generated terms.

d) Attribute refinement – Discussions around which references are most appropriate and term redundancies are guided by the panel leader. Terminology that is redundant, biasing, or integrated is removed from the lexicon. A final lexicon is generated from these exercises and discussion.

e) Lexicon validated – The generated lexicon is tested using a pair of samples. The lexicon should be able to discriminate between the samples. This step may also lead to lexicon or reference refinement.

F.7.2.6 Statistical analysis and interpretation of the results

Data from a SpectrumTM5) panel may be analysed using parametric statistics. Standard summary statistics such as means, standard deviations, analysis of variance, mean separation procedures, and correlations are appropriate since there are 151 points of differentiation from the intensity scale. Frequencies are used for the nonparametric data such as a present or not present format attributes. Whether means are generated using individual or consensus evaluations, multivariate techniques can be used to summarize product groupings (cluster analysis, principal component analysis/factor analysis, etc.).

Annex G
(informative)

Temporal Dominance of Sensation (TDS)

G.1 Principle

Temporal Dominance of Sensations is a descriptive sensory method, consisting in assessing continuously which sensation is dominant. “Dominant sensation” is defined to the assessors as the sensation that catches his/her attention at a given time, which does not mean that this sensation has to be very or the most intense in the product. Compared to time-intensity, this method takes into account the multidimensionality of the perception over time.

For each run, the method allows to collect a sequence of sensory quoted at different times throughout the tasting.

The assessor can give his answers on a sheet of paper provided he has a stopwatch. But the use of a computer is a much more comfortable technique for the assessor; this does not have to worry about the factor time: he/she may focus exclusively on identification of the dominant attributes.

The collected data can be statistically processed to compute dominance curves for each attribute over time.

G.2 Attributes

One list of attributes is presented to the assessor; he/she has to choose within that list the attributes perceived as dominant throughout the tasting. Therefore, the definition of the list of attributes is a key element as it determines the responses of the assessors. The TDS list of attributes can be built from an attribute generation step, as in Quantitative Descriptive Analysis, followed by a reduction task by consensus or by statistical analysis.

It is recommended to include a maximum of 10 to 12 attributes in that list. If the list is too long, some assessors may not be able to handle the whole list.

The order of attributes in the list tends to influence the selection order of the attributes: attributes at the top of the list may be selected earlier than attributes at the bottom of the list. It is therefore recommended to balance the attribute orders across assessors; However, for a given assessor, the order of attributes should always be the same throughout the experiment (whatever the product and whatever the replication session).

It is possible to combine several attribute types in the same list (for example, flavour attributes and texture attributes). With large differences in levels of intensity versus attribute types, it is recommended, to run at least separate sessions for each type of attribute, and optionally run a session with several types of attributes for complementary information.

G.3 Scales

In its simplest form, the method does not use scale. The evaluation consists only on selection of the dominant attributes among the list of proposed.

G.4 Assessors

G.4.1 Training

Training for TDS studies should be rather different from training for quantitative profiling studies: no training on the intensity is required and the training should be more oriented to the identification of the different sensory attributes to improve the selection of the dominant sensations. The selected assessor, have to develop consensual definitions for each of the TDS attributes. Reference standards can be used to help the assessors to recognize each attribute

A few sessions (one to four depending on the assessors) should be dedicated to familiarize the assessors with the specificities of the TDS method (for example, with the use of the computer system).

Once the assessors familiarized with the TDS method, the number of training sessions required is usually smaller in this method than in Quantitative Descriptive Analysis. Quite often two to four sessions are enough to get used to the list of attributes.

G.4.2 Number of assessors/replicates

For TDS with trained assessors, it is recommended to use slightly more assessors than for quantitative profiling, i.e. about 16 assessors, and two or more replicates to collect at least 30 evaluations of each product.

G.5 Procedure

— Portions of the products, bite-size, are presented monadically to the assessor. He/she puts the entire portion into his/her mouth and simultaneously the run begins (usually by clicking a “start” button).

— Whenever the assessor feels that the dominant sensory perception has changed, he/she has to indicate the new dominant sensation within the list of attributes. The assessor reiterates this procedure until the perception ends; he/she clicks the “stop” button, which stops the chronometer.

— The assessor is free to choose the same attribute for the same product as often as he/she thinks it is necessary, or never to select one or more particular attribute(s) of the list.

— The tasting procedure should clearly indicate when to start and stop the assessment in order to collect reliable data.

G.6 Statistical analysis and interpretation of the results

G.6.1 Treatment of the raw data

The analyst has to determine, during each tasting, the name of the dominant attribute and the time elapsed since the start button. An attribute is considered as dominant from the time it is selected until another attribute is selected.

G.6.2 TDS curves

Computation of TDS curves is illustrated in Figure G.1 for the sweet attribute. The procedure considers each attribute separately. For each point of time, the proportion of runs (assessor × replication) for which the given attribute was assessed as dominant is computed. These proportions, smoothed (non-iterative smoothing spline), are plotted versus time to compute TDS curves. For each product, TDS curves of all the attributes are depicted on the same graph.

In order to get more insight into the TDS graphical display, two lines may be drawn. The first one, called “chance level”, is the dominance rate that an attribute can obtain by chance. Its value, P_0, is equal to $1/p$, p being the number of attributes. The second one, called “significance level”, is the minimum value this proportion should equal to be considered as significantly higher than P_0. It is calculated using the binomial distribution.

NOTE Since the duration into the mouth differs from one assessor to another, the time scales of the sensory perception differ as well. In order to take this phenomenon into account in the computation of the TDS curves, the data from each assessor can be standardized so that the *x*-axis shows values from $x = 0$ (first citation of a dominant attribute) to $x = 100$ (swallowing). Non-standardized curves are interesting to visualize the differences in durations among the products, but they can be difficult to read because all assessments do not end up at the same time, while standardized curves enhance the product properties (higher dominance rates in particular at the very beginning and just before the end).

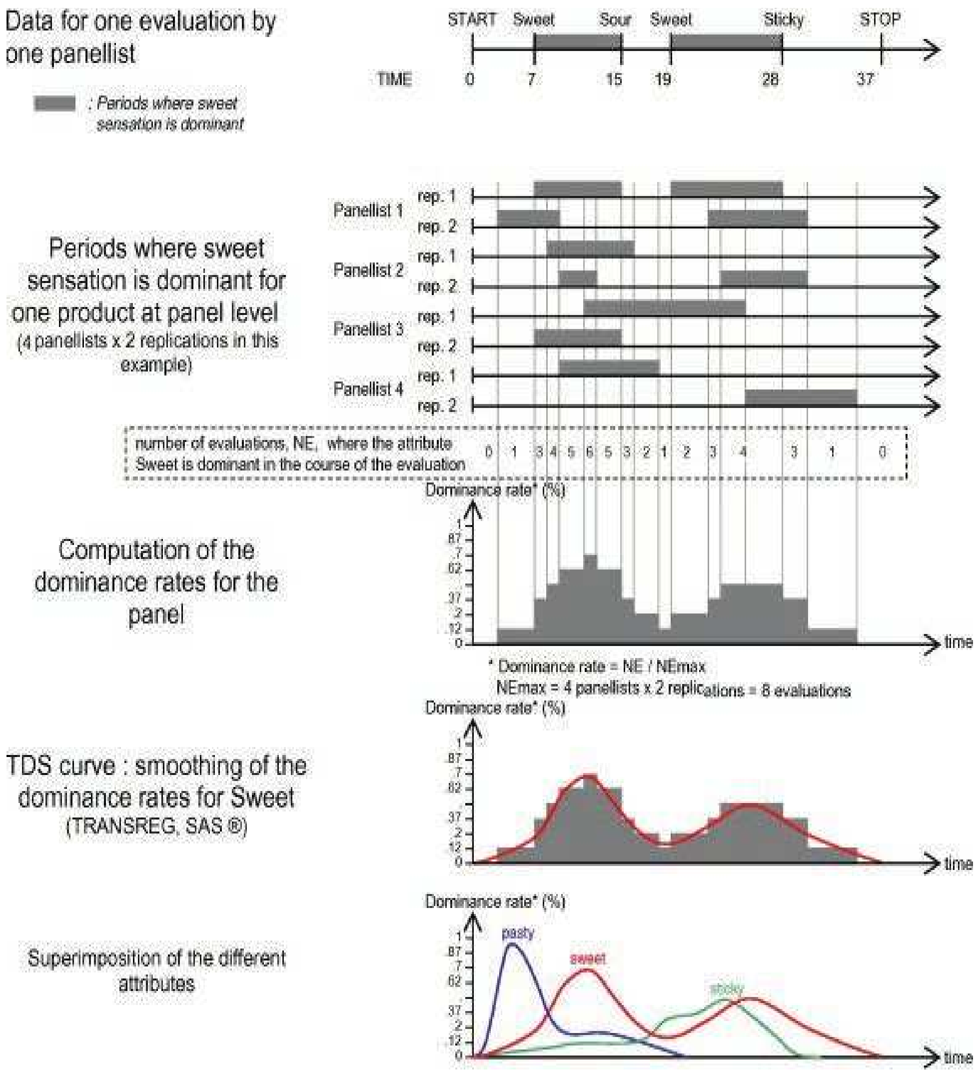

Figure G.1 — Example of calculations for the sweet attribute with 4 assessors and 2 replicates

G.6.3 TDS difference curves

The TDS graph displays the curves of the *p* attributes on the same plot. In order to compare two products, it is possible to display the superimposition of the *p* curves of the differences of the dominance rates. These differences are plotted only when significantly different from zero, highlighting the differences between the products over time. Limit of significance for each curve of difference over time is obtained using the usual test to compare two binomial proportions.

G.6.4 Other statistical analysis

Several other statistical methods can be used. For example: ANOVA of the “stop” times to study the overall product length, PCA or Canonical Variate Analysis (CVA) of dominance durations over time to give an overview of the product differences averaged over time, PCA on the product trajectories over time to highlight the product differences in their temporal evolution.

G.7 Similar and method variations

Usually, the TDS method is performed with selected and trained assessors. It may be also performed with “naive consumers” using the TDS computer system. The list of attributes has to be simple enough to be understood by naive consumers. A few minutes at the beginning of the session should be spent to introduce the consumers to the notion of temporality of sensations, give them a clear definition of the attributes, and give them instructions to use the TDS computer system. It is necessary to give the consumers several warm-up samples in order to familiarize them with the method and with the attributes. For TDS with naive consumers, no replicate is usually performed. Therefore, it is recommended to use at least 30 respondents.

The assessor can optionally be asked to score the intensity of the selected attributes as described in Reference [28]; this request results in a significant increase in the difficulty of the task, requiring a consequent training. In order to not mix up two different cognitive processes: the selection of a dominant attribute (qualitative task) and the rating of the intensity (quantitative task), it is recommended to use buttons rather than intensity scales for indicate the intensity. Figure G.2 gives an example of response for a biscuit.

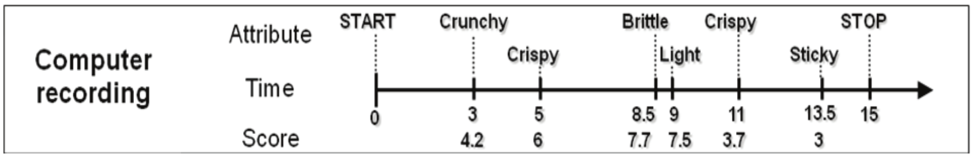

Figure G.2 — Example of computer recording for a TDS run.

The crispy attribute is dominant two first during 3,5 s (between time 5 and time 8,5, Score: 4,2), then a second time during 2,5 s (between time 11 and time 13,5, Score: 3,7). The total duration of crispy dominance in this TDS run is thus 3,5 s + 2,5 s = 6 s. Regarding the crispy intensity score, it would be computed as a weighted (by duration) average over the two scores: 4,2 * 3,5/6 + 3,7 * 2,5/6 = 3,99 ≈ 4.

Annex H
(informative)

Univariate analysis when one attribute is quantified by all the assessors of a panel

H.1 General

The methods described hereafter under are particularly suitable for quantitative descriptive profile, Spectrum™ and QDA®[6] methods where the assessors rate products. They provide a way to determine the relative importance of attributes for the comparison among products.

H.2 Data analysis in a 2-way ANOVA (product and assessor factors) without replicates

H.2.1 ANOVA model

Usually, the basic model is

$$y_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} \quad \text{(H.1)}$$

where

y_{ij} is the score given by the i^{th} assessor (i: 1 to a) to the j^{th} product j (j: 1 to p);

μ is the mean of scores;

α_i is the main assessor effect; it is assumed random;

β_j is the main product effect; it is assumed fixed;

$\alpha\beta_{ij}$ is the interaction assessor*product.

The experimental design is often named *Randomized Complete Block Design* (each assessor is one block). It can also refer to as a *within-assessors design* or a *repeated measures design*.

H.2.2 Numeric example

Ten assessors evaluated three products. Results are given in Table H.1. Table H.1 leads to the 2-way ANOVA given in Table H.2.

6) Methods QDA®, Spectrum™ are examples of suitable procedures available commercially. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of these procedures.

Table H.1 — Example of results in a study without replicates

Assessor	Product 1	Product 2	Product 3
1	2	3	3
2	8	6	7
3	5	4	5
4	4	3	5
5	5	3	5
6	4	4	8
7	4	5	5
8	6	2	4
9	4	3	8
10	5	6	5
Mean	4,7	3,9	5,5

Table H.2 — 2-way ANOVA of Table H.1

Source of variance	Status of factor	Number of degrees of freedom (*df*)	Sum of square (*SS*)	Mean squares (*MS*)	*F* value	*p* value
Assessor	Random	$a - 1 = 9$	34,300	MS_a = **3,811**		
Product	Fixed	$p - 1 = 2$	12,800	MS_p = **6,400**	MS_p/MS_{a*p} = **3,95**	0,038
Assessor*Product	Random	$(a - 1) * (p - 1) = 18$	29,200	MS_{a*p} = **1,622**		

The product factor is tested against the assessor*product interaction. It is not significant for the descriptor of Table H.1.

NOTE When the assessor factor is fixed, for example to study the performance of the assessors during the training of a group, none of the two fixed factors cannot be tested unless it is assumed that the assessor*product interaction estimates the experimental error. With this assumption, both fixed factors are tested against the error, i.e. the assessor*product interaction.

H.3 Data analysis in a 2-way ANOVA (product and assessor factors) with replicates

H.3.1 ANOVA model

Usually, the basic model is

$$y_{ijr} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijr} \quad \text{(H.2)}$$

where

y_{ijr} is the score given by the i^{th} assessor (i: 1 to a) to the j^{th} product j (j: 1 to p) at the r^{th} replicate;

(r: 1 to r); the number of replicates is assumed to be equal for all products;

μ is the mean of scores;

α_i is the main assessor effect; it is assumed random;

β_j is the main product effect; it is assumed fixed;

$\alpha\beta_{ij}$ is the interaction assessor*product;

ε_{ijr} is the random replicate error; it provides an estimate of the experimental error.

This design is often named *Randomized Complete Block Design with more than one observation per experimental unit,* an experimental unit being formed by the intersection between a product and an assessor. In this model, the replicates have no systematic structure among them. This model is relevant, for example, in a study where the replicates are performed during the same session and on samples from the same batch of product. Indeed, in such a model, there is no reason to associate the score given to product P1 at the replicate 1 with the score given to product P2 at the replicate 1 or with the score given to product P3 at the replicate 1. An example of this type of structure is when 2 single carrots from each 4 varieties are served to each of 10 assessors in a single tasting session. At each replicate, one carrot is given to each assessor, so 20 carrots of each variety are used, and each assessor evaluates 4*2 = 8 carrots.

H.3.2 Numeric example

The assessors of Table H.1 rated the products not just once, but twice. Table H.1 gave only the results of the replicate 1. Table H.3 gives the results for the two replicates. The 2-way ANOVA is given in Table H.4.

Table H.3 — Example of results in a study with two replicates

Assessor	Product 1		Product 2		Product 3	
	Replicate 1	Replicate 2	Replicate 1	Replicate 2	Replicate 1	Replicate 2
1	2	3	3	3	3	5
2	8	7	6	6	7	7
3	5	6	4	6	5	7
4	4	5	3	5	5	8
5	5	7	3	5	5	6
6	4	6	4	2	8	8
7	4	5	5	8	5	7
8	6	8	2	5	4	6
9	4	3	3	2	8	7
10	5	7	6	4	5	8

Table H.4 — 2-way ANOVA of Table H.3

Source of variance	Status of factor	Number of degrees of freedom (*df*)	Sum of square (*SS*)	Mean squares (*MS*)	*F* value	*p* value
Assessor	Random	a - 1 = 9	48,350	MS_a = 5,372	MS_a/MS_e = 3,322	0,006 3
Product	Fixed	p - 1 = 2	38,033	MS_p = 19,017	MS_p/MS_{a*p} = 6,673	0,006 8
Assessor*Product	Random	(a - 1) * (p - 1) = 18	51,300	MS_{a*p} = 2,850	MS_{a*p}/MS_e = 1,763	0,082 4
Error		(r - 1) * a * p = 30	48,500	MS_e = 1,617		
Total		a*p*r – 1 = 59	186,183			

The product factor is tested against the product*assessor interaction; it is significant for α-risk = 0,05.

The assessor factor and the product*assessor interaction are tested against the error. The assessor factor is significant, the product*assessor interaction is not.

NOTE When the assessor factor is fixed, for example to study the performance of the assessors during the training of a group, the three factors of Table H.4 are all tested against the error. The F product is equal to 11,763 (p value = 0,000 17); the other two F values are not changed.

H.4 Data analysis in a 3-way ANOVA (product, assessor, and replicate factors)

H.4.1 The replicate factor is crossed with the other two factors

H.4.1.1 ANOVA model

It has 8 components:

$$y_{ijr} = \mu + \alpha_i + \beta_j + \delta_r + \alpha\beta_{ij} + \alpha\delta_{ir} + \beta\delta_{jr} + \alpha\beta\delta_{ijr} \quad \text{(H.3)}$$

where

y_{ijr} is the score given by the i^{th} assessor (i: 1 to a) to the j^{th} product j (j: 1 to p) at the r^{th} replicate (r: 1 to r);

α_i is the main assessor effect; it is considered as random;

β_j is the main product effect; it is considered as fixed;

δ_r is the main replicate effect; it is considered as a fixed factor or a random; probably the most relevant factor is the random factor;

$\alpha\beta_{ij}$ is the assessor*product interaction;

$\alpha\delta_{ir}$ is the product*replicate interaction;

$\beta\delta_{jr}$ is the product*replicate interaction;

$\alpha\beta\delta_{ijr}$ is the assessor*product*replicate interaction.

This design is often referred as a *two-factor completely repeated measures design* or a *three-crossed factors design*.

When there is more than one observation per experimental unit, i.e. per triplet Assessor-Product-Replicate, the design has a 9th component providing an estimate of the experimental error.

Usually, for a profile study, the replicate factor is the factor *session*. If in the example Table H.3 replicate 1 was performed on a Tuesday and replicate 2 Thursday of the same week, replicate 1 becomes session 1 and replicate 2 becomes session 2.

H.4.1.2 Numeric example

When the assessor factor is random and the product and session factors are fixed, the analysis of Table H.3 leads to the 3-way ANOVA given in Table H.5.

Table H.5 — 3-way ANOVA of Table H.3

Source of variance	Status of factor	Number of degrees of freedom (df)	Sum of squares (SS)	Mean squares (MS)	F value	p value
Assessor	Random	a - 1 = 9	48,350	MS_a = 5,372		
Product	Fixed	p - 1 = 2	38,033	MS_p = 19,017	MS_p/MS_{a*p} = 6,673	0,006 7
Session	Fixed	r - 1 = 1	16,017	MS_s = 16,017	MSs/MS_{a*s} = 8,405	0,017 6
Assessor*Product	Random	(a - 1) * (p - 1) = 18	51,300	MS_{a*p} = 2,850		
Product*Session	Fixed	(p - 1) * (s - 1) = 2	1,233	MS_{p*s} = 0,617	MS_{p*s}/MS_{a*p*s} = 0,787	0,47
Assessor*Session	Random	(a - 1) * (s - 1) = 9	17,150	MS_{a*s} = 1,906		
Ass.*Prod.*Ses.	Random	(a - 1) * (p - 1) * (s - 1) = 18	14,100	MS_{a*p*s} = 0,783		
Total		$a*p*s$ – 1 = 59				

The product factor is tested against the assessor*product interaction; it is significant. The session factor is tested against the assessor*session interaction; it is significant. The product*session interaction is tested against the assessor*product*session interaction; it is not significant (see Reference [10], pp. 361-367).

All other factors cannot be tested. However, if the assessor*product*session interaction is assumed negligible, this component can be regarded as an estimate of the experimental error; under this hypothesis, the assessor factor, the assessor*product interaction and the assessor*session interaction can be tested against the assessor*product*session interaction

NOTE 1 When the session factor is random, the design has two random factors (the assessor and session factors) and a single fixed factor (the product factor).

— The assessor factor is tested against the assessor*session interaction; F value is equal to 5,372/1,906 = 2,818 (p value = 0,069).

— The session factor is tested against the assessor*session interaction; F value is equal to 16,017/1,906 = 8,403 (p value = 0,017).

— The assessor*product interaction is tested against the assessor*product*session, F value is equal to 2,850/0,783 = 3,640 (p value = 0,004 4).

— The product*session interaction is tested against the assessor*product*session interaction as in Table H.5 (p value = 0,47).

— The product factor and the assessor*session interaction cannot be tested. That is very annoying for the product factor. But this defect can be overcome in calculating a Quasi F symbolized by F' (see Reference [10], pp. 368-370):

$F' = MS_p/(MS_{a*p} + MS_{p*s} - MS_{a*p*s})$ = 19,017/(2,850 + 0,617 – 0,783) = 19,017/2,684 = 7,085

The number of degrees of freedom of the denominator of F' is given by the so-called Satterthwaite's approximation:

$$\frac{\left(MS_{a*p} + MS_{p*s} + MS_{a*p*s}\right)^2}{\frac{MS^2{}_{a*p}}{df_{a*p}} + \frac{MS^2{}_{p*s}}{df_{p*s}} + \frac{MS^2{}_{a*p*s}}{df_{a*p*s}}}$$

For this example:

$df = (2{,}684^2)/(0{,}451 + 0{,}190 + 0{,}034) = 7{,}204/\ 0{,}675 = 10{,}672$

For F' = 7,085 with 2 *df* for the numerator and 11 *df* (the nearest integer value of 10,67) for the denominator, the probability is equal to 0,0 027. The product factor is significant.

NOTE 2 When the three factors are fixed, no factor can be tested unless the assessor*product*session interaction is assumed to be negligible. Under this assumption, all the six factors (three main factors and three interactions) are tested against the assessor*product*session interaction.

H.4.2 The replicate factor is nested within the product factor

To be identified, this type of design requires a careful analysis of the experimental conditions. Here are three examples. First example: it concerns a sensory analysis of cooked meat roast where the focus is on pork breed. The study has five breeds with four animals per breed. The five animals (the replicates) are said to be nested within each breed (main factor). Second example: a brewer is producing a new brand of beer with five alternatives formulations. For each formulation, it has three batches. The batch factor is nested within the formulation factor. Third example: a researcher studies the differences between six varieties of potatoes. Each of six varieties is grown in the three rows, randomly assigned in a field. The potatoes are harvested and the potatoes from each row are kept in separate containers. The row factor is nested in the variety factor.

This design is often referred as a *nested factorial design.*

The name of the nested factor varies according the assay. In the present subclause, it is named *batch*.

H.4.2.1 ANOVA model

It has five components:

$$y_{ijr} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \delta_r + \beta\delta_{ir} \qquad \text{(H.4)}$$

where

- y_{ijr} is the score given by the i^{th} assessor (i: 1 to a) to the j^{th} product j (j: 1 to p) at the r^{th} batch (r: 1 to r; the number of batches is assumed to be identical for all products);
- α_i is the assessor effect; usually it is a random factor;
- β_j is the product effect; usually it is a fixed factor;
- $\alpha\beta_{ij}$ is the assessor*product interaction;
- δ_r is the batch effect; usually it is a random factor; but it may be a fixed factor;
- $\beta\delta_{jr}$ is the interaction.

When both factors product and batch are fixed and the only random factor is the product factor, the analysis of Table H.3 is simple; it leads to Table H.6.

Table H.6 — Nested ANOVA of Table H.3 (the product and batch factors are fixed)

Source of variance	Status of factor	Number of degrees of freedom (df)	Sum of square (SS)	Mean squares (MS)	F value	p value
Assessor	Random	a - 1 = 9	48,350	MS_a = 5,372		
Product	Fixed	p - 1 = 2	38,033	MS_p = 19,017	MS_p/MS_{a*p} = 6,673	0,006 8
Assessor*Product	Random	(a - 1) * (p - 1) = 18	51,300	MS_{a*p} = 2,850		
Batch (Prod.)	Fixed	(b - 1) * p = 3	17,250	MS_b = 5,750	MS_b/MS_{ab} = 4,970	0,007 1
Ass.*Batch(Prod.)	Random	(a - 1) * (b - 1) * p = 27	31,250	MS_{a*b} = 1,157		
Total		a*p*b – 1 = 59				

The Product factor is tested against the assessor*product and the batch factor is tested against the assessor*batch nested in the product factor. Both factors are significant.

NOTE 1 When the batch factor is random, the design has two random factors (the assessor and session factors) and a single fixed factor (the product factor). The assessor factor, the assessor*product interaction and the batch factor are tested against the assessor*batch(product) interaction. The F assessor is equal to 5,372/1,157 = 4,643 (p value = 0,096) and the F assessor*product interaction is equal to 2,850/1,157 = 2,463 (p value = 0,016 6). The F batch has already been calculated in Table H.6.

The product factor is tested with a Quasi F or F'.

First step: to find the test mean square. It is obtained as

$$MS_{test.P} = MS_{b(p)} + MS_{a*p} - MS_{a*b} = 5{,}750 + 2{,}850 - 1{,}157 = 7{,}443$$

Second step: to find the number of df of the $MS_{test.P}$. It is given by:

$$Den = \frac{MS^2_{test.P}}{\frac{MS_{b(p)}^2}{df_{p(b)}} + \frac{MS_{a*p}^2}{df_{a*p}} + \frac{MS_{a*b}^2}{df_{a*b}}} = \frac{7{,}443^2}{\frac{5{,}75^2}{3} + \frac{2{,}85^2}{18} + \frac{1{,}157^2}{27}} = 4{,}808$$

Third step: to compute the Quasi F:

$$F' = \frac{MS_b}{MS_{test.P}} = \frac{19{,}017}{7{,}443} = 2{,}555$$

Last step: to find the probability associated, under H0, to F' with df = 2 for the numerator and df = 5 (the nearest integer value of 4,808) for the denominator. It is equal to 0,172. The product factor is not significant.

NOTE 2 When the three factors are fixed, all the tests are performed against the assessor*batch(product) interaction. The F assessor is equal to 4,64 (p-value = 0,000 9); the F product is equal to 16,43 (p-value = 0,000 02) and the assessor*product interaction is equal to 2,46 (p-value = 0,017); the F batch had already been calculated in Table H.6.

H.5 Data analysis of assessments over time

H.5.1 General

In H.4, the design had three factors: assessor, product and session (or batch). It is the same in this subclause where the time factor takes the place of the session factor. But while the session factor was somehow a secondary factor compared to the product factor, in this subclause the time factor is as important factor as the product factor because the products change over time.

An example is given in Table H.7; it is borrowed from a study where fresh character of four toothpastes was assessed four times by 10 assessors on a scale 0-20. Sums are given in Table H.8

Table H.7 — Intensity of fresh character of four toothpastes assessed 4 times over time

P1				P2				P3				P4			
T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
8	12	9	8	11	10	2	2	9	10	7	7	6	3	3	1
14	17	15	15	18	17	14	14	16	18	14	15	15	18	17	14
16	17	12	12	15	14	10	9	14	10	8	9	13	15	9	6
15	19	14	14	13	17	10	10	5	15	10	10	10	13	8	7
9	8	6	6	7	9	5	5	7	9	3	4	6	7	4	2
20	20	16	16	19	19	16	16	19	17	15	15	19	17	15	12
11	14	11	10	10	10	7	8	11	14	9	8	10	15	9	8
15	19	13	13	15	17	7	6	12	15	9	10	13	14	10	7
12	12	5	4	10	10	4	3	10	10	2	2	9	9	3	8
10	10	7	7	13	8	3	2	13	11	5	6	11	9	5	2

Table H.8 — Sums of Table H.7

	Time 1	Time 2	Time 3	Time 4	
Product 1	130	148	108	105	491
Product 2	131	131	78	75	415
Product 3	116	129	82	86	413
Product 4	112	120	83	67	382
	489	528	351	333	1 701

H.5.2 All assessors rate the four products: the three factors product, time, and assessor are crossed

In Table H.7, the line 1 gives the data of assessor 1; the line 2 gives the data of assessor 2, etc.

H.5.2.1 Analysis with a 3-way ANOVA

Results are given in Table H.9. The presentation is slightly different from that of Table H.5. But the tests are performed as in Table H.5. The Product factor (fixed factor) is tested against the product*assessor interaction. The time factor (fixed factor) is tested against the time*assessor interaction and the time*product interaction is tested against the error (i.e. the product*time*assessor interaction).

Table H.9 — 3-way ANOVA of Table H.7

Source of variance	Status of factor	Number of degrees of freedom (df)	Sum of square (SS)	Mean squares (MS)	F value	p value
Ass.	Random	$a - 1 = 9$	2 056,18	$MS_a = 228,46$		
Prod.	Fixed	$p - 1 = 3$	161,22	$MS_p = 53,74$	$MS_p/MS_{a*p} = 8,842$	0,000 3
Ass.*Prod.	Random	$(a - 1) * (p - 1) = 27$	164,09	$MS_{a*p} = 6,08$		
Time	Fixed	$(t - 1) = 3$	716,12	$MS_t = 238,71$	$MS_t/MS_{t*a} = 38,85$	0,000 0
Time*Prod.	Fixed	$(t - 1) * (p - 1) = 9$	45,21	$MS_{t*p} = 5,02$	$MS_{t*p}/MS_{error} = 2,43$	0,016 8
Time*Ass.	Random	$(t - 1) * (a - 1) = 27$	167,19	$MS_{t*a} = 6,19$		
Error		$(a - 1) * (p - 1) * (t - 1) = 81$	167,23	$MS_{error} = 2,06$		
Total		159	3 477,24			

H.5.2.2 Analysis with a MANOVA on repeated measures

The 3-way ANOVA requires the assumption about the variance–covariance known as *sphericity*. This assumption is less restrictive if we perform a multivariate analysis of variance (MANOVA) with the Time factor (T1, T2, T3, T4) as dependent variable. For the interpretation of Table H.10, the reader is invited to consult a book of statistics.

Table H.10 — MANOVA of Table H.7 when all assessors rate the four products

	Test	Value	F	Effect df	Error df	p-value
Time	Wilk	0,087 10	87,345 27	3	25,000 00	0,000 000
	Pillai	0,912 90	87,345 27	3	25,000 00	0,000 000
	Hotelling	10,481 43	87,345 27	3	25,000 00	0,000 000
	Roy	10,481 43	87,345 27	3	25,000 00	0,000 000
Time*Product	Wilk	0,491 16	2,299 36	9	60,994 04	0,026 935
	Pillai	0,597 36	2,237 62	9	81,000 00	0,027 499
	Hotelling	0,855 90	2,250 70	9	71,000 00	0,028 152
	Roy	0,486 99	4,382 87	3	27,000 00	0,012 273
Time*Assessor	Wilk	0,175 62	2,220 86	27	73,655 21	0,003 738
	Pillai	1,202 13	2,005 92	27	81,000 00	0,008 874
	Hotelling	2,698 18	2,365 07	27	71,000 00	0,002 065
	Roy	1,621 9	4,866 57	9	27,000 00	0,000 640

With the MANOVA, the time*product interaction is significant for the four tests.

NOTE It is also possible to perform another MANOVA by creating new variables that represent the amount of change from time: $Diff_1$ = Time 2 – Time 1, $Diff_2$ = Time 3 – Time 2 and $Diff_3$ = Time 4 – Time 3. Each of these new change or difference variables measures how much change has taken place at each interval (see Reference [8]).

H.5.3 Each assessor rates one of four products: the assessor factor is nested in the product factor which is crossed with the time factor

Consequently, Table H.7 involves 40 assessors: 10 for the Product 1, 10 for the Product 2, 10 for the Product 3 and 10 for the Product 4.

The data of Table H.7 lead to Table H.11 and Table H.12.

Table H.11 — Nested ANOVA of Table H.7

Source of variance	Status of factor	Number of degrees of freedom (df)	Sum of square (SS)	Mean squares (MS)	F value	p value
Product	Fixed	p - 1 = 3	161,22	MS_p = 53,74	MS_p/MS_{a*p} = 8,84	0,000 3
Assessor (Product)	Random	(a - 1) * p = 36	2 220,28	MS_{a*p} = 61,67		
Time	Fixed	t - 1 = 3	716,12	MS_t = 238,71	MS_t/MS_{error} = 77,09	0,000 00
Time*Product	Fixed	(t - 1) * (p - 1) = 9	45,21	MS_{t*p} = 5,02	MS_{t*p}/MS_{error} = 1,62	0,12
Error		108	334,43	MS_{error} = 3,10		

In Table H.6, the factor nested in the product was a fixed factor: the batch factor. Here, the factor nested in the product is a random factor: the assessor factor. The tests of time factor and time*product are different.

Table H.12 — MANOVA of Table H.8 when assessor is nested in product

	Test	Value	F	Effect df	Error df	p-value
Time	Wilk	0,149 451	64,499 77	3	34,000 0	0,000 000
	Pillai	0,850 549	64,499 77	3	34,000 0	0,000 000
	Hotelling	5,691 156	64,499 77	3	34,000 0	0,000 000
	Roy	5,691 156	64,499 77	3	34,000 0	0,000 000
Time*Product	Wilk	0,636 200	1,880 93	9	82,897 7	0,066 100
	Pillai	0,401 348	1,853 34	9	108,000 0	0,066 810
	Hotelling	0,512 838	1,861 41	9	98,000 0	0,066 763
	Roy	0,339 206	4,070 47	3	36,000 0	0,013 742
NOTE This MANOVA is performed with the Time factor (T1, T2, T3, T4) as dependent variable.						

The time*product is significant for Roy's test. The p-value is close to the significant value for the other three tests.

Bibliography

General

[1] ISO 704, *Terminology work — Principles and methods*

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

[3] ISO 13300-1, *Sensory analysis — General guidance for the staff of a sensory evaluation laboratory — Part 1: Staff responsibilities*

[4] ISO 13300-2, *Sensory analysis — General guidance for the staff of a sensory evaluation laboratory — Part 2: Recruitment and training of panel leaders*

[5] GACULA Jr. M.C., 1997. Descriptive Sensory Analysis in Practice. Food & Nutrition Press, 712 pp.

[6] LAWLESS H.T.. Quantitative Sensory Analysis. Psychophysics, Models and Intelligent Design. Wiley & Sons, 2013, p.

[7] LAWLESS H.T.. HEYMANN H., 2010, Sensory Evaluation of Food. Principles and Practices. 2nd ed., Springer, 596 pp.

[8] MEILGAARD M.C.. CIVILLE G.V., CARR B.T., 2007. Sensory Evaluation Techniques. 4th ed., CRC Press, 448 pp.

[9] STONE H.. BLEIBAUM R.N., THOMAS H.T., 2012. Sensory Evaluation Practices, 4th ed., Academic Press, 438 pp.

Statistical Analysis

[10] ABDI H.. EDELMAN B., VALENTIN D., DOWLING W.G., 2009. Experimental Design and Analysis for Psychology, Oxford University Press, 538 pp.

[11] GACULA M.. SINGH J., BI J., ALTAN S. Statistical Methods in Food and Consumer Science, Second Edition, 2009, p.

[12] HOWELL D.. Statistical Methods for Psychology. Duxbury, Fifth Edition, 2002, p.

[13] NAES T.. BROCKHOFF P.B., TOMIC O., 2010. Statistics for Sensory and Consumer Science, Wiley & Sons, 287 pp.

Consensus profile

[14] BRANDT M.A.. SKINNER E.Z., COLEMAN J.A. Texture Profile Method. J. Food Sci. 1963, **28** (4) pp. 404–409

[15] CAIRNCROSS S. E.. and SJOSTROM L. B. Flavor Profiles - A New Approach to Flavor Problems. Food Technol. 1950, **4** pp. 308–311

[16] CHAMBERS D.H.. ESTEVE E., RETIVEAU A. Effect of milk pasteurization on flavor properties of seven commercially available French cheese types. J. Sens. Stud. 2010, **25** (4) pp. 494–511

Deviation from Reference profile

[17] LARSON-POWERS N.M.. PANGBORN R.M. Descriptive analysis of the sensory properties of beverage s and gelatine containing sucrose and synthetic sweeteners. J. Food Sci. 1978, **43** (11) pp. 47–51

[18] BOYLSTON T.D.. REITMEIER C., MOY J., MOSHER G.A., TALADRIZ L. Sensory quality and nutrient composition of three Hawaiian fruits treated by X-irradiation. J. Food Qual. 2002, **25** (5) pp. 419–433

Flash profile

[19] SIEFFERMANN J.-M.. 2000. Le profil Flash. Un outil rapide et innovant d'évaluation sensorielle descriptive, AGORAL 2000, XIIèmes rencontres, Montpellier, France In : Tec. & Doc., L'innovation: de l'idée au succès, 335-340.

[20] DELARUE J.. SIEFFERMANN J-M., 2004. Sensory mapping using Flash profile - Comparison with a conventional descriptive method for the evaluation of the flavour of fruit dairy products. Food Quality and Preference, 15 (4), 383-392.

Free-choice profile

[21] WILLIAMS A. A.. LANGRON S. P. The use of free-choice profiling for the evaluation of commercial ports. J. Sci. Food Agric. 1984, **35** (5) pp. 558–568

[22] GUÀRDIA M. D.. AGUIAR A. P. S., CLARET A., ARNAU J. and GUERRERO L. Sensory characterization of dry-cured ham using free-choice profiling. Food Qual. Prefer. 2010, **21** (1) pp. 148–155

[23] PIGGOT J.R.. WATSON M.P. A comparison of free-choice profiling and the repertory grid method in the flavor profiling of cider. J. Sens. Stud. 1992, **7** (1) pp. 133–145

Qualitative sensory profile

[24] CAMPO E.. BALLESTER J., LANGLOIS J., DACREMONT C., VALENTIN D. Comparison of conventional descriptive analysis and a citation frequency-based descriptive method for odor profiling: An application to Burgundy Pinot noir wines. Food Qual. Prefer. 2010, **21** (1) pp. 44–55

Quantitative descriptive profile

[25] FINDLAY C. J.. CASTURA J. C., and LESSCHAEVE I. Feedback calibration: A training method for descriptive panels. Food Qual. Prefer. 2007, **18** (2) pp. 321–328

[26] LAWLESS L. J. R., CIVILLE G. V., 2013. Developing Lexicons:Review A. J. Sens. Stud., **28** (4) pp. 270–281

Quantitative descriptive analysis

[27] STONE H.. SIDEL J., OLIVER R.S., WOOLSEY A. and SINGLETON R.C. Sensory evaluation by quantitative descriptive analysis. Food Technol. 1974, **28** (1) pp. 24–34

[28] See also STONE et al. 2012, pp. 250-274 (op cit.)

Spectrum

[29] CIVILLE G. V.. DUS C. A. Development of terminology to describe the handfeel properties of paper and fabrics. J. Sens. Stud. 1990, **5** (1) pp. 19–32

[30] MUNOZ A. M. and CIVILLE G. V.. Universal, product and attribute specific scaling and the development of common lexicons in descriptive analysis. J. Sens. Stud. 1998, **13** (1) pp. 57–75

[31] See also MEILGAARD et al. 2007, pp. 189-253 (op cit)

Temporal dominance of sensations

[32] PINEAU N., CORDELLE S., IMBERT A., ROGEAUX, M., SCHLICH P., 2003. Dominance temporelle des sensations – Codage et analyse d'un nouveau type de données sensorielles. XXXVèmes Journées de Statistiques, Lyon, 2–6 Juin, 777-780.

[33] PINEAU N.. SCHLICH P., CORDELLE S., MATHONNIERE C., ISSANCHOU S., IMBERT A., ROGEAUX M., ETIEVANT P., KOSTER P. Temporal dominance of sensations. Construction of the TDS curves and comparison with time-intensity. Food Qual. Prefer. 2009, **20** (6) pp. 450–455

ICS 67.240

Price based on 41 pages