SLOVENSKI STANDARD
oSIST prEN ISO 11132:2020

01-marec-2020

Senzorična analiza - Metodologija - Smernice za nadzorovanje izvajanja kvantitativnega senzoričnega panelnega testa (ISO/DIS 11132:2020)

Sensory analysis - Methodology - Guidelines for monitoring the performance of a quantitative sensory panel (ISO/DIS 11132:2020)

Sensorische Analyse - Methodologie - Leitlinien zur Überprüfung der Leistungsfähigkeit eines quantitativen sensorischen Panels (ISO/DIS 11132:2020)

Analyse sensorielle - Méthodologie - Lignes directrices pour le contrôle de la performance d'un jury sensoriel quantitatif (ISO/DIS 11132:2020)

Ta slovenski standard je istoveten z: **prEN ISO 11132**

ICS:

67.240	Senzorična analiza	Sensory analysis

oSIST prEN ISO 11132:2020 **en**

DRAFT INTERNATIONAL STANDARD

ISO/DIS 11132

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

Secretariat: **IRAM**

Voting begins on:
2020-01-30

Voting terminates on:
2020-04-23

Sensory analysis — Methodology — Guidelines for monitoring the performance of a quantitative sensory panel

Analyse sensorielle — Méthodologie — Lignes directrices pour le contrôle de la performance d'un jury sensoriel quantitatif

ICS: 67.240

This document is circulated as received from the committee secretariat.

ISO/CEN PARALLEL PROCESSING

Reference number
ISO/DIS 11132:2020(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 34, Food products, Subcommittee SC 12, Sensory analysis.

This second edition cancels and replaces the first edition (ISO 11132:2012), which has been technically revised.

Sensory analysis — Methodology — Guidelines for monitoring the performance of a quantitative sensory panel

1 Scope

This International Standard gives guidelines for monitoring and assessing the overall performance of a quantitative descriptive panel and the performance of each member.

A panel of assessors can be used as an instrument to identify products' sensory attributes and to assess the magnitude of sensory attributes.

Performance is the measure of the ability of a panel or an assessor to make reliable and valid attribute assessments across the products being evaluated. It can be monitored at a given time point or tracked over time. Performance comprises the ability of a panel to detect, identify, and measure an attribute, use attributes in a similar way to other panels or between assessors within a panel, discriminate between stimuli, use a scale properly, repeat their own results, and reproduce results in comparison to other panels or assessors.

The methods specified allow to monitor and assess the consistency, repeatability, freedom from bias and the ability of discrimination of panels and assessors. Monitoring and assessment can be carried out in one session or over time.

Monitoring performance data enables the panel leader to improve panel and assessor performance, to identify issues and retraining needs or to identify assessors who are not performing well enough to continue participating.

The methods specified in this International Standard can be used, in full or a selection only, by the panel leader to appraise continuously the performance of panels or individual assessors. The methods listed are not exhaustive and other appropriate methods may also be used.

This International Standard applies to individuals or panels in training as well as for established panels.

2 Normative references

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

ISO 5492, *Sensory analysis — Vocabulary*

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

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

3 Terms and definitions

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

3.1 agreement

ability of different panels or assessors to exhibit the same product differences when assigning scores on a given attribute to the same set of products.

3.2 panel drift

phenomenon where a panel, over time, changes in sensitivity or becomes susceptible to biases and as a consequence changes the location on the scale where an attribute is rated for a constant, reference product

3.3 performance

ability of a panel or an assessor to make reliable and valid assessments of stimuli and stimulus attributes

3.4 validation

process of establishing that sensory data correlate with other data on samples of the same product (e.g. laboratory measurements, consumer perception, results from other panels, consumer complaints) or that a panel or assessor is able to meet specified performance criteria

3.5 replicate sessions

sessions in which the assessors, the products, the test conditions, and the task are the same

4 Principle

4.1 Two possible approaches

This International Standard is concerned with sensory panels used to assess the magnitude of one or more sensory attributes in order to make quantitative descriptions or profiles of products. Different methods are appropriate to the assessment and monitoring of the performance of panels used for difference testing.

The performance of a quantitative sensory panel may be evaluated by using assessments already available or from panel sessions conducted specifically for the purpose of obtaining performance data.

4.1.1 Panel validation via a dedicated procedure

A dedicated monitoring procedure at periodic intervals is appropriate for accreditation and other purposes. Figure 1 is a flow chart for this procedure.

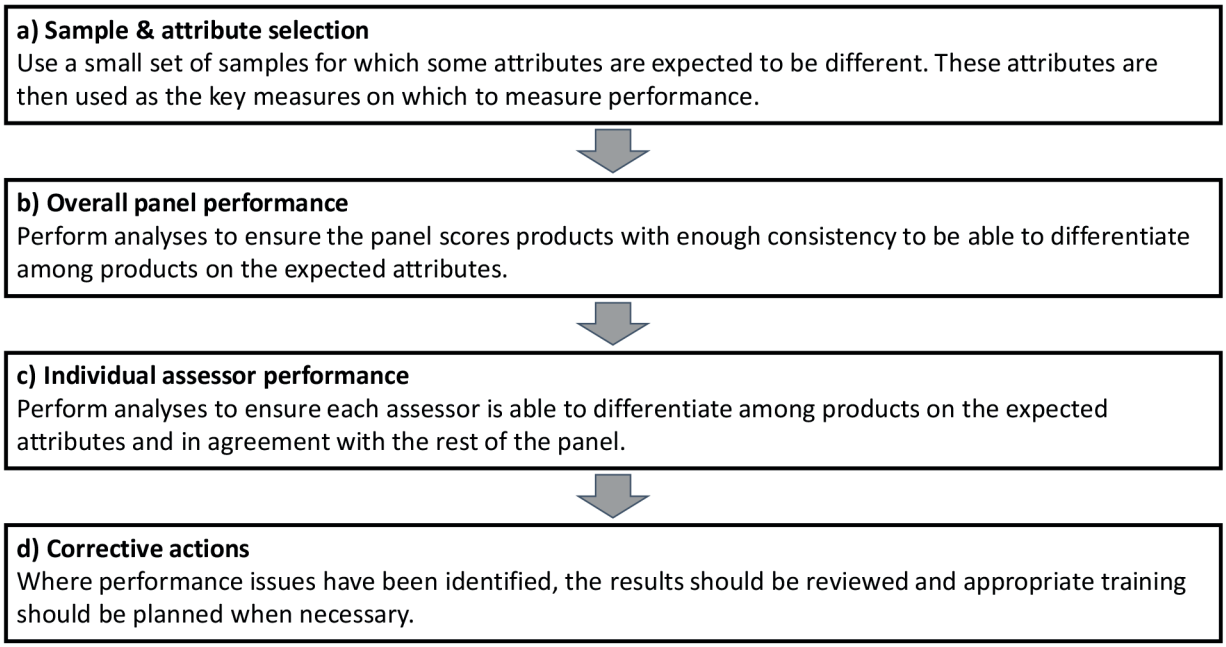

Figure 1 — Process steps for the performance validation

4.1.2 On-going monitoring via routine product profiling

To review ongoing profile data generated by a panel, it can be appropriate to use data that originated from quite different profiling experiments using different product types, product numbers, etc. The procedure is the same as that shown in Figure 1. However, as there are no predefined differences, it is recommended that attributes that are significantly discriminated by the panel as a whole for a given profile be used as the key measures to check the performance of individual panelists. Attributes that result in no significant difference cannot be reliably used to check consistency since the lack of agreement within and between panelists probably means that the products are very similar for those characteristics.

4.2 Indicators of panel or individual assessor performance

In a single session, the following indicators can be determined:

— Discrimination of an assessor, measured as the ability of the assessor to exhibit significant differences among products.

— Discrimination of the panel, measured as the ability of the panel to exhibit significant differences among products.

— Agreement of an assessor, measured as the degree of alignment between the assessor's average product scores and the ones of the panel

— Agreement of the panel, measured as the degree of alignment between the assessors' average product scores

— Repeatability of an assessor, measured as the degree of homogeneity between replicated assessments of the same product.

Across several sessions, the following indicator can be determined:

— Reproducibility of the panel, Repeatability of the panel, measured as the agreement in assessments of the same set of products under similar test conditions by the same or different assessors at different time points.

4.3 Statistical analyses

A single, consistent approach to statistical analysis of the results is described in this document. However, some indicators of panel performance can be assessed by more than one measure. For instance, error mean square and error SD (its square root) both express variability in the evaluation of a product. The measures used should be those that are usual in the field of application.

Other relevant measures of agreement between assessors in the use of the scale for an attribute are the interaction of assessor and product and the coefficient of correlation between an assessor's scores and the panel means. An assessor may have no bias, but may use the scale in a different way. A correlation close to 1, a regression slope close to 1, and a regression intercept close to 0 indicate good agreement between an assessor and the rest of the panel.

With a small number of assessments (fewer than six) the correlation coefficient should be interpreted with caution, as it can be high (up to 0,7), by chance alone.

5 Prerequisites

5.1 Experimental conditions

The test facilities shall be in accordance with ISO 8589.

5.2 Qualification of assessors

The panel shall have the level of qualification and experience of selected assessors (ISO 8586) or better.

6 Procedure for performance validation

6.1 Sample and attribute selection

At each validation session, the panel of assessors should be presented with a set of samples similar to those the panel are to assess when evaluating products and for which statistically significant differences between at least one pair of the samples are expected for the list of relevant attributes.

As much as possible, it is recommended to use at least eight attributes to encourage panel leaders or sensory managers to identify and select validation samples aligned with the level of difference among products the panel is expected to detect.

These key attributes are used as key measures against which to assess panel performance. The sample set should include replicates. There shall be the same number of replicates of each sample. The number of assessors, samples, and replicates depends on the products, the sensory attributes assessed and the purpose of the procedure. For example, 2 or 3 replicates of three or four samples might be used. Care should be taken to limit the number of assessments required so as to avoid sensory fatigue. The attributes of the samples should be similar to the range of values that the panel assesses when evaluating products.

6.2 Experimental designs

Several types of experimental designs can be used in the performance validation procedure, depending on the most important objective to answer.

6.2.1 Randomized block design

A randomized block experimental design can be used, in which the assessors are the "blocks". This design is appropriate when no carry-over effect from one sample to the next is expected, otherwise a balanced design should be used instead (see 6.2.2.).

6.2.2 Balanced and random designs

If a carry-over effect is expected from one sample to the next, a suitable experimental design is the Williams Latin square, see table 1. The basic design uses four assessors and four samples.

Table 1 — Williams Latin square

Assessor	Session	Order			
		1	2	3	4
1	1	A	B	C	D
2	1	B	D	A	C
3	1	C	A	D	B
4	1	D	C	B	A
1	2	B	D	A	C
2	2	C	A	D	B
3	2	D	C	B	A
4	2	A	B	C	D

In this design, each assessor samples the four products in a different order in a given session and any particular product is followed by a different one for each assessor, for example in session 1, A is followed by B for assessor 1, C for assessor 2, D for assessor 3 and none for assessor 4.

For each replicate of the products' evaluation, it is recommended to use a different product order for each assessor, in order to reduce the order effect and the carry-over effect.

If multiples of four assessors are available, the same design can be repeated for each set of four.

It is also possible to choose a random product order design, i.e. to randomly affect each product to each position in each session.

The advantage of these approaches is to minimize the carry-over effect at panel level and therefore get better estimates of the product means at panel level for performance evaluation. However, if the product order does have an impact, the agreement between assessors will be impacted because each assessor will not experience the same product order. In order to compare the assessors on the exact same task, one can use the same product order for all panellists (see 6.2.3.).

6.2.3 Same order design

In order to focus on individual assessor performance and in an effort to compare assessors under the most similar conditions, an alternative design is proposed, whereby all assessors evaluate the products in the same order, see Table 2.

Table 2 — Same order design

Assessor	Session	Order			
		1	2	3	4
1	1	A	B	C	D
2	1	A	B	C	D
3	1	A	B	C	D
4	1	A	B	C	D
1	2	A	B	C	D
2	2	A	B	C	D
3	2	A	B	C	D
4	2	A	B	C	D

It is worth mentioning that in this case the assessors are not evaluating the products per se but the products at a given position (product and position effects are confounded). This will conduct to a biased estimate of the product effect (biased by the order effect), but to an unbiased estimate of the assessor effect and the product*assessor interaction.

6.3 Statistical analysis of data from performance validation (a single session)

Table 3 illustrates one way to tabulate and summarize the results. Some computer software may require a different organization of the data, for instance with the samples in columns and the assessors in rows.

Table 3 — Results of the assessors

Sample	Assessor								Mean
	1		…	j		…	n_q		
	Scores	Mean		Scores	Mean		Scores	Mean	
1	Y_{111} Y_{112} … Y_{11n_r}	$\bar{Y}_{11.}$		Y_{1j1} Y_{1j2} … Y_{1jn_r}	$\bar{Y}_{1j.}$				$\bar{Y}_{1..}$
…									
i	Y_{i11} Y_{i12} … Y_{i1n_r}	$\bar{Y}_{i1.}$		Y_{ij1} Y_{ij2} … Y_{ijn_r}	$\bar{Y}_{ij.}$				$\bar{Y}_{i..}$
…									
n_p									$\bar{Y}_{n_p..}$ **Error! Not a valid link.**
Mean	$\bar{Y}_{.1.}$			$\bar{Y}_{.j.}$			$\bar{Y}_{.n_q.}$		$\bar{Y}_{...}$

In this table it is assumed that:

$n_p \equiv$ number of samples ($i = 1,2 \ldots n_p$);

$n_q \equiv$ number of assessors ($j = 1,2 \ldots n_q$);

$n_r \equiv$ number of replicates per sample ($k = 1,2 \ldots n_r$).