INTERNATIONAL STANDARD

ISO 13301

Second edition
2018-04

Sensory analysis — Methodology — General guidance for measuring odour, flavour and taste detection thresholds by a three-alternative forced-choice (3-AFC) procedure

Analyse sensorielle — Méthodologie — Recommandations générales pour le mesurage des seuils de détection d'odeur, de flaveur et de goût par une technique à choix forcé de 1 parmi 3 (3-AFC)

Reference number
ISO 13301:2018(E)

© ISO 2018



COPYRIGHT PROTECTED DOCUMENT

© ISO 2018

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

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

 © ISO 2018 – All rights reserved

Contents

Page

Foreword ... **iv**

Introduction ... **v**

1 Scope ... **1**

2 Normative references ... **1**

3 Terms and definitions ... **2**

4 Principles ... **3**
4.1 Experimental procedures ... 3
4.2 Data processing ... 3

5 Experimental procedures ... **4**
5.1 Preparation of samples ... 4
5.1.1 General precautions ... 4
5.1.2 Gases ... 4
5.1.3 Liquids ... 4
5.1.4 Solids ... 4
5.2 Selection of concentrations of the stimulus ... 4
5.3 Presentation of samples ... 5
5.3.1 Preparation ... 5
5.3.2 Gases ... 5
5.3.3 Liquids ... 5
5.3.4 Solids ... 5
5.4 Training of assessors ... 5
5.5 Selection of assessors ... 6
5.5.1 General ... 6
5.5.2 Individual threshold ... 6
5.5.3 Distribution of thresholds ... 6
5.5.4 Measurement of thresholds of stimuli ... 6
5.6 Design of the experiment ... 6
5.6.1 Individual threshold ... 6
5.6.2 Distribution of thresholds ... 7
5.6.3 Measurement of the threshold of a stimulus for a group of assessors ... 7

6 Data processing ... **9**
6.1 Mathematical and statistical models ... 9
6.2 Preliminary inspection of data ... 9
6.2.1 Preparation ... 9
6.2.2 Preliminary estimation of threshold and slope using the logit transformation ... 10
6.3 Maximum likelihood procedure for fitting the data to a logistic model and estimating error bounds ... 10
6.3.1 General ... 10
6.3.2 The parameter, b ... 10
6.3.3 Confidence intervals for estimated parameters ... 11
6.4 Interpretation of results ... 11
6.5 Selecting a detection rate target (p_d) other than 0,5 ... 12
6.6 Estimation of the best estimate threshold (BET) ... 12
6.7 Presentation of results ... 12

Annex A (informative) **Estimated number of assessors required for a given degree of precision** ... **14**

Annex B (informative) **Examples** ... **15**

Bibliography ... **28**

© ISO 2018 – All rights reserved

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 13301:2002), which has been technically revised. The following changes have been made:

— the bibliographic references have been updated;

— Additional instructions regarding the use of spreadsheet programs have been added;

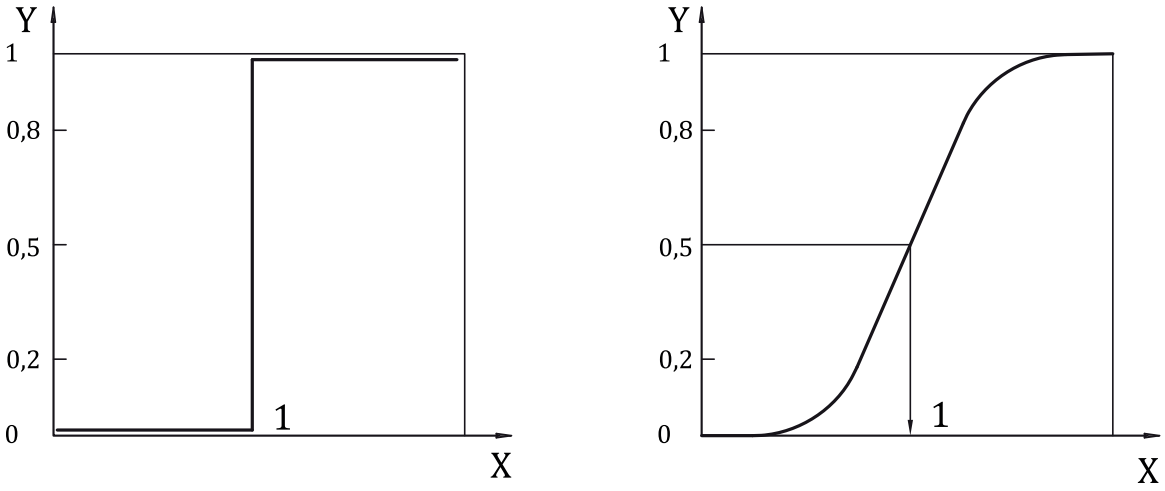
— typographical errors in tables and examples have been corrected.

© ISO 2018 – All rights reserved

Introduction

The concept of "threshold" has wide use in sensory analysis and is often used in the literature on sensory studies of food and drink. Data on sensory thresholds to stimuli are used in sensory studies in two main ways: as measures of the sensitivity of assessors or groups of assessors to specific stimuli; and as measures of the ability of substances to evoke sensory responses in assessors. In the first, the value of the threshold is taken as a description of an assessor's performance; in the latter, as a measure of a property of the substance.

The term "threshold" was introduced by 19th century psychophysicists and used to denote a stimulus concentration above which the stimulus could be detected, and below which it could not [see Figure 1 a)].



a) Traditional notion of threshold **b) Probabilistic nature of threshold**

Key

X concentration of stimulus, x

Y probability of detection, p_d

1 threshold

Figure 1 — Traditional notion and probabilistic nature of threshold

However, in practice the graph of the probability of detection[1)] against the intensity of the stimulus is always an ogive [see Figure 1 b)], and it is convenient to assume, for purposes of calculation, that the threshold fluctuates so that a particular stimulus concentration exceeds it on some occasions but not on others. The threshold can then be obtained as an estimate of the median of these momentary values, i.e. as the stimulus concentration for which the probability of detection is 0,5. The threshold defined in this way has analogies with median effect levels used in branches of biology such as pharmacology and toxicology, which are concerned with the effect of chemicals on organisms.

Where detection thresholds of a particular substance in air or water have been measured in more than one laboratory, the reported values often span two or three orders of magnitude or more[6][10][14]. This range is greater than can be expected from experimental errors alone or from differences in the processing of data; but it probably can be accounted for by difference in concepts of thresholds between laboratories, and differences in experimental procedure. Reference [6] suggests a procedure for standardizing detection thresholds in air.

1) This document is based on the use of the 3-AFC method of presenting the stimuli, and the probability of detection, p_d, is modelled as $p_d = 1{,}5 \times p_c - 0{,}5$, where p_c is the probability of a correct selection. This is strictly a "guessing model" of the assessor's behaviour. It is not a psychometric model of the assessor's decision process, such as a signal-detection model, which could also be applied, see Reference [13].

© ISO 2018 – All rights reserved

The user needs to be aware that the determination of detection thresholds requires more experimental effort than is at first apparent from this description. Experimental results demonstrate that on repeated testing, the observed individual thresholds tend to decrease, and the difference between individuals likewise tends to decrease. Threshold testing is often an unfamiliar activity, and assessors will improve their sensitivity as they become accustomed to the substance and the mechanics of the test. The 3-AFC procedure requires that assessors can recognize the stimulus. Training programmes require effort but will in turn yield needed information about each assessor's range of partial detection. Results improve as the experimenter learns to tailor the concentrations presented to each assessor's range, see 6.3.

© ISO 2018 – All rights reserved

Sensory analysis — Methodology — General guidance for measuring odour, flavour and taste detection thresholds by a three-alternative forced-choice (3-AFC) procedure

1 Scope

This document gives guidelines for

— obtaining data on the detection of stimuli that evoke responses to odour, flavour and taste by a 3-AFC (three-alternative forced-choice) procedure, and

— the processing of the data to estimate the value of a threshold and its error bounds, and other statistics related to the detection of the stimulus.

Typically, the procedures will be used in one of the following two modes:

— investigation of the sensitivity of assessors to specific stimuli;

— investigation of the ability of a chemical substance to stimulate the chemoreceptive senses.

(Although experiments can encompass both modes.)

Examples of the first mode include studies of the differences among individuals or specified populations of individuals in sensitivities and of the effects of age, gender, physiological condition, disease, administration of drugs and ambient conditions on sensitivity. Examples of the latter mode include

— studies in flavour chemistry and the impact of specified chemicals on the flavour of foods,

— classification of chemicals for their impact on humans, if present in the environment,

— studies on the relationship of molecular structure to capacity of a chemical to act as a stimulant,

— quality assurance of gaseous effluents and of water, foods and beverages, and

— studies in the mechanism of olfaction.

In both modes, the way in which probability of a correct response changes with intensity of stimulus, i.e. the slope of the dose/response curve, could be an important aspect of the study as well as the threshold value, and the data processing procedures described here provide this information.

The focus of this document is on data requirements and on computational procedures. Regarding the validity of the data, the text is restricted to general rules and precautions. It does not differentiate between detection and difference thresholds; fundamentally, the procedures measure a difference threshold because a test sample is compared with a reference sample. Typically, the reference sample is not intended to contain the stimulus under investigation, but the guidelines do not exclude experimental design in which the reference could contain the stimulus, or it might not be known if the reference contains the stimulus. The guidelines do not measure a recognition threshold as defined in ISO 5492. They do not address the standardization of methods of determining air quality as discussed in EN 13725.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

© ISO 2018 – All rights reserved

ISO 5492, *Sensory analysis — Vocabulary*

3 Terms and definitions

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

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— IEC Electropedia: available at http://www.electropedia.org/

— ISO Online browsing platform: available at https://www.iso.org/obp

3.1
stimulus
substance that may or may not cause a sensation, detectable by one or more of the senses, depending on the amount present

3.2
medium
material used to dissolve, dilute, disperse or sorb a *stimulus* (3.1) whose threshold is to be measured

3.3
reference sample
quantity of the *medium* (3.2) containing no added *stimulus* (3.1)

3.4
test sample
quantity of the *medium* (3.2) to which a *stimulus* (3.1) has been added at a known concentration

3.5
3-AFC test
three-alternative forced-choice test
test of discrimination in which the assessor is presented with three samples, one of which is a *test sample* (3.4) containing a nominated *stimulus* (3.1) familiar to the assessor, the other two being references, and where the assessor is instructed to indicate the test sample

Note 1 to entry: The standard 3-AFC test is a specified discrimination method. Assessors are instructed as to what attribute to use to make their decision. In threshold testing, the attribute of interest may well be below the recognition thresholds of the assessor, so in threshold tests the assessors conduct an unspecified 3-AFC, in which they select the one sample that is different from the other two, without reference to a specific attribute.

3.6
presentation
set of three samples forming a *3-AFC test* (3.5)

3.7
threshold model
model of sensory detection where a *stimulus* (3.1) presented on a particular trial is either detected (resulting in a correct response) or is not detected (resulting in a response being made at random)

3.8
signal-detection model
model of sensory detection where a *stimulus* (3.1) presented on a particular trial provides some level of evidence of its presence

Note 1 to entry: The evidence contributes to a decision by the assessor about the presence or absence of the stimulus.

© ISO 2018 – All rights reserved

3.9
detection threshold
lowest intensity of a sensory *stimulus* (3.1) that has a probability of detection of 0,5 under the conditions of the test, as calculated from the *threshold model* (3.7)

3.10
individual threshold
detection threshold (3.9) of a single assessor

3.11
average threshold
average (where the type is specified, e.g. arithmetic mean, geometric mean or median) of *individual thresholds* (3.10)

3.12
group threshold from pooled data
estimate obtained by using the sum of outcomes for a particular group of assessors at each concentration of the *stimulus* (3.1) as input when fitting the statistical model

4 Principles

4.1 Experimental procedures

The stimulus is formulated in the medium at a specified concentration and is presented along with a pair of reference samples to the assessor. The assessor selects one of the samples as containing the stimulus or having the stimulus at a greater concentration. The assessor should make a selection. Typically, the stimulus is dissolved in air or water. It is unlikely that a gas other than air will be used as a gaseous medium in tests with human assessors, but solvents other than water, solutions in water or other solvents, or solids, e.g. foods, can be used as liquid or solid medium to dilute the stimulus as the experiment dictates. It is essential that the medium be homogeneous so that the members of the pair of references are identical, and the same in all presentations.

The stimulus is presented at several concentrations. The presentations are replicated, at each concentration, enough times to achieve a desired precision of the threshold and parameters of the mathematical model. The nature of the replications within assessors, across assessors, and combinations of the two are set by the experimental design of the study.

4.2 Data processing

The outcome of a presentation is a binary result; the sample nominated by the assessor is the test sample (a correct selection) or is one of the references (an incorrect selection). The statistical model is that the number of correct selections at a particular concentration comes from a binomial distribution.

For the 3-AFC test, the threshold is the concentration of the stimulus at which the proportion of correct selections is equal to 2/3, which corresponds to the probability of detection = 0,50 [see Formula (1)]. The data, as proportions of correct selections, can simply be inspected and interpolated to derive this point, but a more accurate estimate of the threshold, and its bounds, can be obtained by fitting a mathematical model to the data. A logistic model is used in these guidelines, and the model is fitted by a maximum likelihood procedure, or alternatively, by a least squares procedure. The fitting estimates the two parameters of the model, one a location parameter, the other a shape parameter. The former locates the fitted curve on the stimulus continuum, the latter determines the steepness of the curve. The fitted curve allows estimates of proportions of detection other than 50 % to be derived.

The simplest model to fit is one in which the distribution of proportion of correct selections comes from a single, approximately normal, distribution. This would typically be the case where the data come from replications within a single assessor. A single logistic function can then be adequately fitted, that is, one with a single pair of values for the parameters of the curve. It is not uncommon for the sensitivities to chemicals to be not normally distributed, or even symmetrically distributed, among assessors. For some

© ISO 2018 – All rights reserved