Designation: ~~E253 – 23~~ E253 – 23a

Standard Terminology Relating to Sensory Evaluation of Materials and Products[1]

This standard is issued under the fixed designation E253; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Referenced Documents

1.1 *ASTM Standards:*[2]
E284 Terminology of Appearance

2. Terminology

absolute judgment, *n*—an evaluation of a stimulus made without direct comparison to other stimuli. (2007)

acceptability/unacceptability, *n*—degree to which a stimulus is judged to be favorable or unfavorable. (2006)

acuity, *n*—the ability to detect or discriminate sensory stimuli. (2007)

adaptation, sensory, *n*—a decrease in sensitivity to a given stimulus which occurs as a result of exposure to that stimulus. (2006)

affective test, *n*—any method to assess acceptance, liking, preference, or emotions for a stimulus or stimuli. (2008)

after effects, *n*—total array of sensations that occur after removal of the stimulus from the sensing field (for example, with foods) or after application of the stimulus (for example, with non-foods). (2008)

after feel, *n*—feel of the skin after application of a sample, with or without touching, usually measured at a specified time point. (2008)

aftertaste, *n*—the oral or nasal sensations that occur after the stimulus has been removed from the oral cavity. (2007) (See also **after effects**.)

aguesia, *n*—lack of sensitivity to taste stimuli. (1996)

α (alpha) risk, *n*—the probability of concluding that a perceptible difference exists when, in reality, one does not. (2014)
DISCUSSION—
α risk also is known as Type 1 Error or significance level.

[1] This terminology is under the jurisdiction of ASTM Committee E18 on Sensory Evaluation and is the direct responsibility of Subcommittee E18.01 on Terminology.
Current edition approved ~~Feb. 1, 2023~~July 15, 2023. Published ~~February 2023~~August 2023. Originally approved in 1965. Last previous edition approved in ~~2022~~2023 as ~~E253 – 22a.~~E253 – 23. DOI: ~~10.1520/E0253-23.~~10.1520/E0253-23A.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

alternative forced choice (AFC), *n*—method in which two, three, or more stimuli are presented, and assessors are given a criterion by which they are required to select one stimulus. (2008)

DISCUSSION—
Typical examples include 2-AFC (directional different test) and 3-AFC (selecting the one stimulus among a set of three that differs in a defined attribute).

alternative forced choice (AFC) test, *n*—method in which two, three, or more stimuli are presented, and assessors are given a criterion by which they are required to select one stimulus. (2008)

DISCUSSION—
Typical examples include 2-AFC (directional difference test) and 3-AFC (selecting the one stimulus among a set of three that differs in a defined attribute).

anchoring point, *n*—a reference point against which other items are judged. (1996)

anosmia, *n*—lack of sensitivity to odor stimuli. (1996)

A-not-A test, *n*—a method of discrimination testing comprised of at least two samples; at least one sample is a previously identified sample ("A") and at least one is a test sample; all samples are presented blindly, and the assessor's task is to assign the label "A" or "not-A" to each of the samples. (2001)

antagonism, *n*—joint action of two or more stimuli whose combination elicits a level of sensation lower than that expected from combining the effects of each stimulus taken separately. (1996)

aroma, *n*—perception resulting from stimulating the olfactory receptors; in a broader sense, the term is sometimes used to refer to the combination of sensations resulting from stimulation of the entire nasal cavity. (1996)

DISCUSSION—
Aroma, odor, and smell have the same basic meaning; however, in common usage they may have different connotations.

aromatic, *n*—perception resulting from stimulating the olfactory receptors retronasally. (2010)

assessor, *n*—a general term for any individual responding to stimuli in a sensory test. (2006)

DISCUSSION—
The terms *assessor*,*judge*,*panelist*,*panel member*, and *respondent* all have the same basic meaning, although sometimes different connotations. Usage of these terms varies with the training and experience of the investigator, habit, tradition, personal preference, and other factors.

assessor drift, *n*—a trend in which an assessor's response to a specific stimulus shifts over time even though the stimulus has not changed. (2018)

DISCUSSION—
Drift happens over a long duration of time, usually weeks or months.

astringency, *n*—the complex of sensations due to shrinking, drawing, or puckering of the epithelium as a result of exposure to substances such as alums or tannins. (1996)

attitude, *n*—a predisposition to respond in a characteristic way toward a class of objects, concepts, or stimuli. (1996)

attitude scale, *n*—a means for eliciting indications of the attitudes or opinions held, usually on a measuring system using marks or value designations. (1996)

attribute, *n*—a perceived characteristic. (1996)

audition, *n*—the sense of hearing. (1996)

aversion, *n*—a predisposition to avoid a stimulus based on a feeling of discomfort or dislike. (2011)

β (beta) risk, *n*—the probability of concluding that no perceptible difference exists when, in reality, one does. (2014)

DISCUSSION—

β risk also is known as Type II Error.

bias, *n*—a systematic error manifested as a persistent positive or negative deviation of the method average from its accepted true value. (2013)

DISCUSSION—

Bias is a general term for any systematic deviation in a method's average from its accepted true value. In sensory, there are several commonly used terms that relate to specific types or causes of bias. These include: sensory adaptation, context effect, contrast effect, convergence effect, error of expectation, order effect, position effect, and response bias.

bipolar scale, *n*—scale where the end anchors are semantic opposites and there is an implied or anchored mid-point. (2008)

DISCUSSION—

Examples of semantic opposites are "too thin" to "too thick," "dislike extremely" to "like extremely."

bite, chemical, *n*—stinging experienced primarily in the oral cavity as a result of exposure to substances such as highly carbonated beverages. (1997)

bitter, *adj*—pertaining to the taste produced by substances such as quinine or caffeine when in solution. (2012)

blinded, *adj*—an element of experimental control in which the identity or an aspect of a treatment, condition, or substance is hidden from the participant (single blind) or both the participant and the experimenter (double blind). (2008)

body (food), *n*—the quality of a food or beverage relating either to its consistency, compactness of texture, fullness, flavor, or combination thereof. (1997)

brightness, *n*—see **color (of an object)**. (2001) (For consensus technical definition see **brightness** in Terminology E284.)

burn, chemical, *n*—perception of irritation resulting from exposure to substances such as ethyl alcohol, acetic acid, and benzoate; the sensation tends to persist after the stimulus is removed. (2013)

calibration (trained assessor), *n*—the process of aligning trained assessors' responses to either established or agreed upon qualitative or quantitative references.

DISCUSSION—

Descriptive analysis methodologies may or may not include trained assessor calibrations.

central location test (CLT), *n*—a test where stimuli are evaluated by consumers at a common designated location(s); the stimuli preparation and presentation are controlled. (2019) (See also **home use test**.)

DISCUSSION—

Examples of designated locations include market research facilities, academic laboratories, grocery stores, or hotel conference rooms.

check-all-that-apply (CATA) scale, *n*—a question format in which any number of items from a list may be selected to indicate which items pertain to the stimulus being measured. (2018)

DISCUSSION—

Check-, choose-, and select-all-that-apply are used interchangeably.

chemesthesis, *n*—perception derived from chemical stimulation of the skin or other tissues, for example, menthol cooling, ammonia pungency. (2015)

chroma, *n*—see **color**. (2001) (For consensus technical definition see **chroma** in Terminology E284).

color (of an object), *n*—the appearance of an object dependent upon the spectral composition of radiant and incident light, the spectral reflectance or transmittance of the object, and the psychological response of the observer; the experience may be described in terms of three attributes: brightness, chroma, and hue. (2001) (For consensus technical definition see **color** in Terminology E284 as defined by Committee E12.)

brightness, n—aspect of visual perception whereby an area appears to emit more or less light.

chroma, n—experienced as color purity, attribute of color used to indicate the degree of departure of the color from a gray of the same brightness.

hue, n—attribute of color related to the wavelength of electromagnetic energy and experienced as "red," "green," "blue," and other elements of the visible spectrum.

color blindness, *n*—total or partial inability to differentiate certain hues. (1997)

consumer, *n*—the user or potential user of a product or service, who may participate in research tests to provide opinions of products, concepts or services. (2010)

context effect, *n*—effect upon the perception of a stimulus arising from its interrelationship with other stimuli in a presentation ~~set. (1997)~~set, the test design, or the environment. (2023)

contrast, visual, *n*—the degree of dissimilarity in appearance of two parts of a field of view seen simultaneously or successively. (1998)

contrast effect, *n*—special case of context effect in which the perceived degree of difference between stimuli is exaggerated as a result of their interrelationship. (1997)

convergence, *n*—tendency of a stimulus to be perceived as similar to prior stimulus or stimuli. (1997)

convergence effect, *n*—special case of context effect in which the perceived degree of difference between stimuli is diminished as a result of their interrelationship. (1997)

cooling, chemical, *n*—sensation of reduced temperature experienced as a result of exposure to certain substances such as menthol or anise; the sensation usually persists after the stimulus is removed. (1997)

cooling, physical, *n*—sensation of reduced temperature experienced as a result of exposure to thermally cold substances (such as ice), to substances that evaporate rapidly (such as acetone or alcohol), or to substances that have a negative heat of solution (such as crystalline sorbitol); the duration of the sensation is usually limited to the time of direct contact with the stimulus. (1998)

cutaneous sense, *n*—any of the senses whose receptors lie in the skin or immediately beneath it (or in the external mucous membranes), such as contact, pressure, warmth, cold, and pain. (1997)

descriptive analysis, *n*—any method to describe and quantify the sensory characteristics of stimuli by a panel of trained assessors. (1998)

difference limen, *n*—see **threshold**,**difference**. (1997)

directional difference test, *n*—a paired comparison or 2-AFC (Alternative Forced Choice) method in which assessors select the stimulus from a pair of stimuli that is perceived to be higher or lower in intensity of a specified sensory attribute. (2009)

discrimination, *n*—the process of qualitatively or quantitatively differentiating among stimuli. (1998)

discrimination test, *n*—any method to determine if differences among stimuli are perceptible; for example, triangle tests, duo-trio tests, paired comparison tests, and so forth. (1998)

duo-trio test, *n*—a method of discrimination testing comprised of two coded samples and one identified reference; one of the coded samples and the reference are samples of the same stimulus; the other coded sample is a sample of one other stimulus; the assessor is asked to select which of the two coded samples is different from the reference or which of the two coded samples is the same as the reference. (2012)

DISCUSSION—
Stimuli can represent different lots of products, formulations, or processes, and so forth.

end effect, *n*—effect where the end points of a scale are used less frequently than other scale points. (2014)

expectation, error of, *n*—a bias due to preconceived ideas that influences an assessor's judgment. (1999)

expert, *n*—an evaluator with extensive experience and knowledge in a product category who makes judgments about the product's qualities or value. (2020) (See also **assessor**,**expert assessor**.)

DISCUSSION—
Experts often operate alone, not as a member of a sensory panel. Examples of experts related to product assessment are sommeliers and meat graders.

expert assessor, *n*—an assessor with a high degree of sensory acuity who has experience in the test procedure and established ability to make consistent and repeatable sensory assessments; an expert assessor functions as a member of a sensory panel. (1995) (See also **assessor**,**expert**.)

extended use testing, *n*—sensory or consumer testing of a product over a time period that allows for multiple usage occasions.

flavor, *n*—(*1*) perception resulting from stimulating a combination of the taste buds, the olfactory organs, and chemesthetic receptors within the oral cavity; (*2*) the combined effect of taste sensations, aromatics, and chemical feeling factors evoked by a substance in the oral cavity. (2001)

fragrance, *n*—(*1*) see **aroma**; (*2*) an aromatic substance. (2011)

DISCUSSION—
The term fragrance is commonly used in household and personal care industries.

free-choice profiling, *n*—a form of sensory profiling in which each assessor independently generates attributes to evaluate a group of samples; the assessors' attributes may be the same or may differ from sample to sample; the assessors' sensory profiles are combined statistically (for example, by Generalized Procrustes Analysis) to produce a map of the samples. (2000)

gloss, *n*—a shiny appearance resulting from the tendency of a surface to reflect light energy at one angle more than at others. (2000) (See **reflectance, directional**. For the consensus technical definition see **gloss** in Terminology E284.)

gustation, *n*—the sense of taste. (2011)

heat, chemical, *n*—sensation of increased temperature resulting from exposure to substances such as capsaicin or hot peppers; the sensation tends to persist after the stimulus is removed.

heat, physical, *n*—sensation experienced as a result of exposure to thermally hot substances such as water above 120 °F; the duration of the sensation is usually limited to the time of direct contact with the stimulus.

hedonic scale, *n*—a scale on which liking or disliking of a stimulus is expressed. (2000)

home use test (HUT), *n*—a test where stimuli are evaluated by consumers at home or in the environment typical of the actual use situation; the stimuli preparation and presentation are self-administered. (2019) (See also **central location test**.)

DISCUSSION—
HUT and in-home use test (iHUT) are used interchangeably.

hue, *n*—see **color (of an object)**. (2001) (For consensus technical definition see **hue** in Terminology E284.)

intensity, *n*—the perceived magnitude of a stimulus. (2000)

interval data, *n*—data obtained from a scale for which numerically identical differences on any part of the scale correspond to the same magnitude of difference of the characteristic being measured. (2018)

DISCUSSION—
The numerical values of the scale are arbitrary as long as the intervals remain the same. An example of interval data is the numerals 1 to 9, which can be rescaled to –4 to +4 without changing the information content of the data.

judge, *n*—see **assessor**. (2000)

just-about-right scale, *n*—bipolar scale used to measure the level of an attribute relative to an assessor's ideal level, having a midpoint labeled "just about right" or "just right." (2007)

just noticeable difference, *n*—see **threshold difference**. (2000)

kinesthesis, *n*—perception of bodily movement or position. (2015)

labeled affective magnitude scale (LAM), *n*—a type of labeled magnitude scale, with verbal labels related to liking and disliking; there is a neutral point in the center of the line scale and the opposing end anchors are "greatest imaginable like" and "greatest imaginable dislike;" see Fig. 1. (2015)

FIG. 1 Labeled Affective Magnitude Scale

DISCUSSION—
The remaining verbal anchors are equivalent to the anchors used with the well-known, nine-point hedonic scale, from "like extremely" to "dislike extremely." Positions of the verbal anchors were determined by magnitude estimation to be (in terms of % of the scale):[3]

Greatest Imaginable Like	100.0
Like Extremely	87.1

[3] Cardello and Schutz, "Numerical Scale-Point Location for Constructing the LAM (Labeled Affective Magnitude) Scale," *Journal of Sensory Studies*, Vol 19, 2004, pp. 341–346.

Like Very Much	78.1
Like Moderately	68.1
Like Slightly	55.6
Neither Like nor Dislike	50.0
Dislike Slightly	44.7
Dislike Moderately	34.1
Dislike Very Much	22.3
Dislike Extremely	12.3
Greatest Imaginable Dislike	0.00

labeled magnitude scale (LMS), *n*—a semantic scale of perceptual intensity characterized by approximately logarithmic spacing of verbal labels along a line scale. See Fig. 2. (2020)

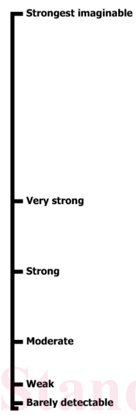

FIG. 2 Labeled Magnitude Scale

Discussion—
The verbal anchors are spaced on the LMS based on calibration using ratio-scaling. It is critical that the spacing be maintained in order to accurately reflect the nonlinear relationship between stimulus and sensation. Positions of the verbal anchors were determined by magnitude estimation to be (in terms of % of the scale): "barely detectable," 1.37; "weak," 5.8; "moderate," 16.2; "strong," 33; "very strong," 50; and "strongest imaginable," 96.[4]

lexicon (sensory), *n*—a set of defined terms that describe the sensory characteristics of stimuli. (2017)

Likert scale, *n*—as originally described, a five or seven point bipolar scale that allows the assessors to express how much they agree or disagree with a particular statement.[5,6] (2016)

Discussion—
The original Likert scale used the following response categories: strongly agree, agree, undecided (or neither agree nor disagree), disagree, and strongly disagree. Modifications to the Likert scale can be five, seven, or nine points, with or without a neutral midpoint, and measure attitudes or opinions to responses such as agreement, frequency, likelihood, and importance.

magnitude estimation, *n*—process of assigning values to the intensities of an attribute in such a way that the ratios between pairs of assigned values are the same as between the magnitudes of the perceptions to which they correspond. (2003)

malodor, *n*—an odor that is unpleasant or disliked when perceived in a specific context. (2012)

Discussion—
For example, "sweaty" is considered a malodor by many people when it is present on the human body. However, "sweaty" notes are common in many cheeses and may not be considered malodor in that context.

masking, *n*—the phenomenon where one quality within a mixture obscures one or several other qualities present. (2001)

---

[4] Green, B. G., Shaffer, G. S., Gilmore, M. M., "Derivation and Evaluation of a Semantic Scale of Oral Sensation Magnitude with Apparent Ratio Properties," *Chemical Senses*, Vol 18, No. 6, 1993, pp. 683–702.

[5] Likert, R., "A Technique for the Measurement of Attitudes," *Archives of Psychology*, No. 140, 1932, pp. 1–55.

[6] McLeod, S. A., "Likert Scale Definition, Examples and Analysis," 2008, www.simplypsychology.org/likert-scale.html.