

International Telecommunication Union

ITU-R
Radiocommunication Sector of ITU

Report ITU-R BS.2399-0
(03/2017)

**Methods for selecting and describing
attributes and terms, in the preparation of
subjective tests**

BS Series
Broadcasting service (sound)



International
Telecommunication
Union

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Annex 1 of Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from <http://www.itu.int/ITU-R/go/patents/en> where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Reports

(Also available online at <http://www.itu.int/publ/R-REP/en>)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	Broadcasting service (sound)
BT	Broadcasting service (television)
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	Spectrum management

Note: This ITU-R Report was approved in English by the Study Group under the procedure detailed in Resolution ITU-R 1.

Electronic Publication
Geneva, 2017

© ITU 2017

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without written permission of ITU.

REPORT ITU-R BS.2399-0

**Methods for selecting and describing attributes and terms,
in the preparation of subjective tests**

(2017)

1 Introduction

The ITU-R has published several Recommendations for subjective test methods for assessing audio quality. The best known of these are Recommendations ITU-R BS.1116 and ITU-R BS.1534. Whilst there is a very large amount of information in these Recommendations, they do not specify exactly what is to be evaluated in a test, although they do list examples of attributes that might be used.

When planning a subjective test, the experimenter needs to decide what is to be evaluated. Then the subject needs to understand the attributes of the stimuli that are to be given a grade.

This Report describes the process by which attributes to be assessed might be chosen and described.

2 Overview

Consensus vocabularies (CV) are extensively employed in the field of sensory science and are thoroughly described in textbooks [1],[2],[3]. Fundamentally, CV approaches seek to define attributes or descriptors used by assessors to characterize the perceptual differences between products in a test. The aim is to develop, or re-use, an attribute list that is consensually agreed upon across the panel and that allows a detailed characterization of the perceived differences between the systems under study.

One of the purposes of labelling and defining consensus attributes is to be able to communicate perceived sound characteristics in an objective and consistent manner. Furthermore, the attributes shall be suitable for reliable assessments of perceptual product characteristics. This leads to the following list of desired attribute characteristics [4]:

- 1 Powerful: Good discrimination power among stimuli (i.e. a large range of mean values and small confidence intervals).
- 2 Good for individual usage: Assessors are able to give reproducible and discriminative assessments.
- 3 Good consensus characteristics: Good agreement among assessors, unambiguous attributes to all subjects.
- 4 Independent: Low redundancy and correlation with other attributes, little or no overlap with others terms.
- 5 Specifiable attribute scales: Using text labels and reference sound samples to illustrate the meaning and polarity of the attribute scale.
- 6 Relation to reality: Attributes should be practically useable to assess differences between products.
- 7 Relation to Preference: Should be related to concepts that influence consumer's preference, through further analysis (e.g. preference mapping).
- 8 Relates to metrics: Relates to physical measures defining the stimuli. Such metrics may not yet exist.

Points 1-4 should be fulfilled for good consensus attributes. The degree of fulfilment can be demonstrated by results from listening tests. Whether points 5-8 are important depends on the purpose.

The first step in any attribute selection is to decide whether an appropriate lexicon is available for the given product category, or if the product category cannot be described by a lexicon that already exists.

3 Procedure for development of attributes

The traditional way of working with consensus attributes in a project is starting with the word elicitation. This is normally done by gathering a group of experienced assessors, expose them to a selection of stimuli representing the products for testing and then discuss and agree on which characteristics that best are describing and discriminating the products and how these characteristics (attributes) should be named and defined. This is done prior to the training of assessors and performing the test of the actual products.

A step-wise procedure is suggested:

- 1 Attribute elicitation. A free verbalization interview. The participants should listen to selected relevant domain specific sound examples and describe their auditory impressions. This step resulted in a list of descriptive attributes from each participant. The lists are collected and redundant attributes are deleted.
- 2 Comprehension check. It should be checked whether all participants understands the attributes from the list. Terms are excluded when they were not associated with a meaning for any participant. This is done in a consensus setting.
- 3 Suitability. The participants assess the suitability of the attributes to describe relevant domain specific sound examples. All terms rated less than 75% of the maximum should be excluded. A check-all-that-apply approach may be used.
- 4 Attribute descriptors are developed. This is typically done by the test leader based on previous discussion and input from step 1, 2, and 3. This step may also involve the panel, though this is often very time consuming.
- 5 Redundancy check. A pilot study listening test using the remaining attributes with descriptors is performed. A cluster analysis of the data provided important clues regarding the similarity of the terms. Several different approaches to the clustering may be considered:
 - a) Group work with expert assessors.
 - b) Hierarchical cluster analysis on scaled attribute data.
 - c) Semantic text/data mining.
- 6 Evaluation of repeatability. The repeatability of the participants' judgments should be evaluated.
- 7 Final consensus session. Evaluating and potentially reducing the list of attributes based on results from step 5 and 6.

4 Procedure for development of lexicons/sound wheels

For a certain product category it is meaningful and efficient to have a selection of predefined attributes of known relevance and validity – a lexicon. Different approaches may be used to obtain that.

Lexicons are standardized vocabularies that facilitate communication across diverse audiences. They are meant for trained descriptive panels as well defined and documented tool to consistently and correctly describe products of interest. The following work procedure applies:

- Establish a “frame of reference” by selecting samples from a wide array of products that represent the entire product space in the category.
- Develop and generate terms that describe the products by the help of appropriate panellists “exposed” to representative samples (reference to general semantic lists such as Semantic Space of Sounds [5] may be useful).
- Select reference products that clarify the terms using examples so the panel more fully understands the terms.
- Develop the final list of descriptors with definitions for the lexicon.
- Validation. Using the terms to compare a few products in the category and determine whether the lexicon allows the panel to differentiate and describe the sensory properties.

The procedure can be seen as a generalized and formalized word elicitation process. Examples on such lexicons can be found in [6]. Efforts have been made to apply this approach to auditory stimuli, see e.g. [7].

A sound wheel is a hierarchical visual representation of a perceptual attribute lexicon. Attributes with similarities are placed in the same category. Similar categories are grouped near each other and placed on a doughnut graph. The result is a visual representation of how attributes are related. The wheel may be used to explain and understand descriptive sensory analysis and it may be used for panellists during training.

An example of a wheel is shown in Fig. 1, with associated example descriptions found in Table 1. An alternative representation, as a tree hierarchy, is shown in Fig. 2 from Zentner [8].

FIGURE 1

The Audio Wheel for reproduced sound. Definitions in Table 1

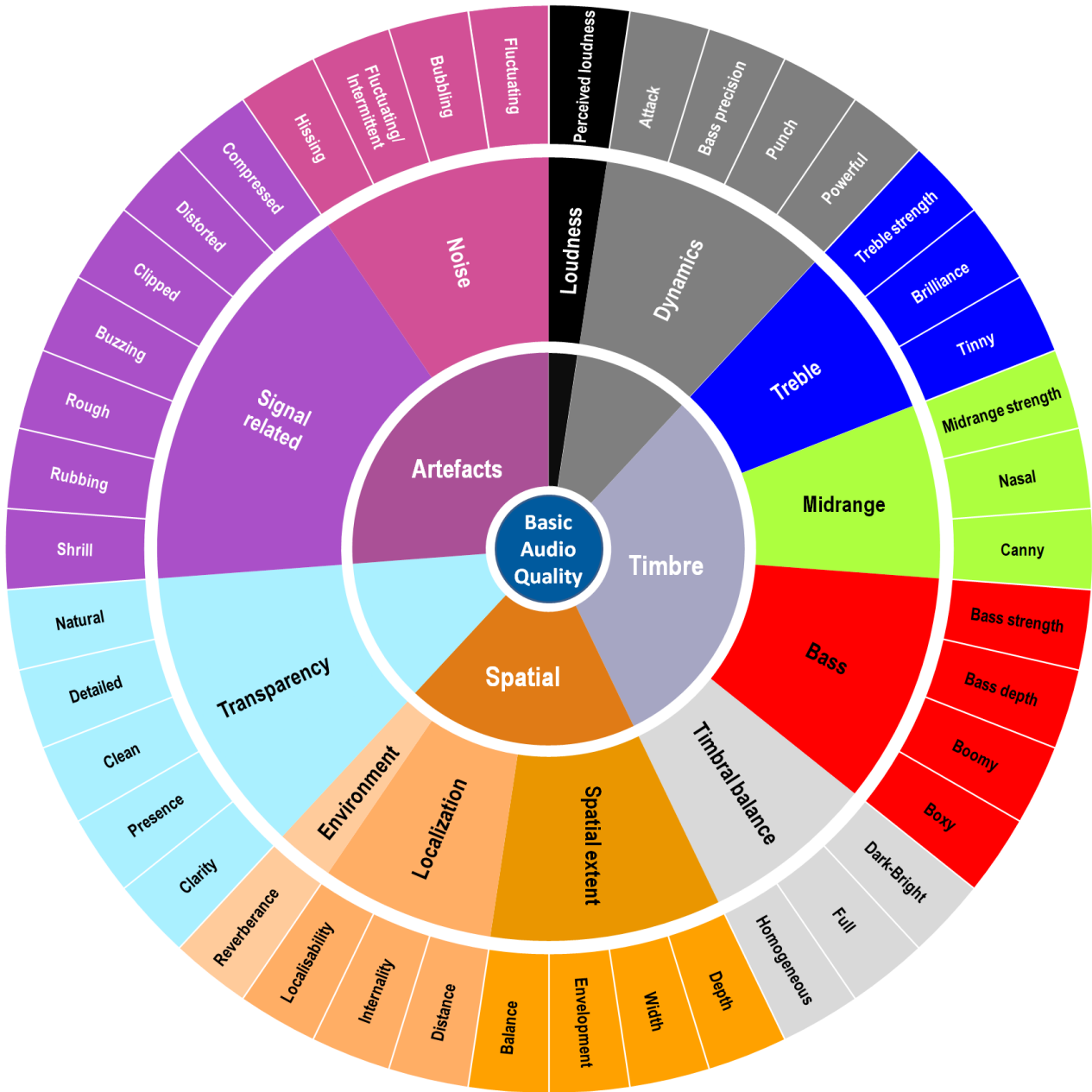


TABLE 1

**Attribute definitions associated with the Audio Wheel (see) for sound reproduction,
drawn from attributes in [9] – [31]**

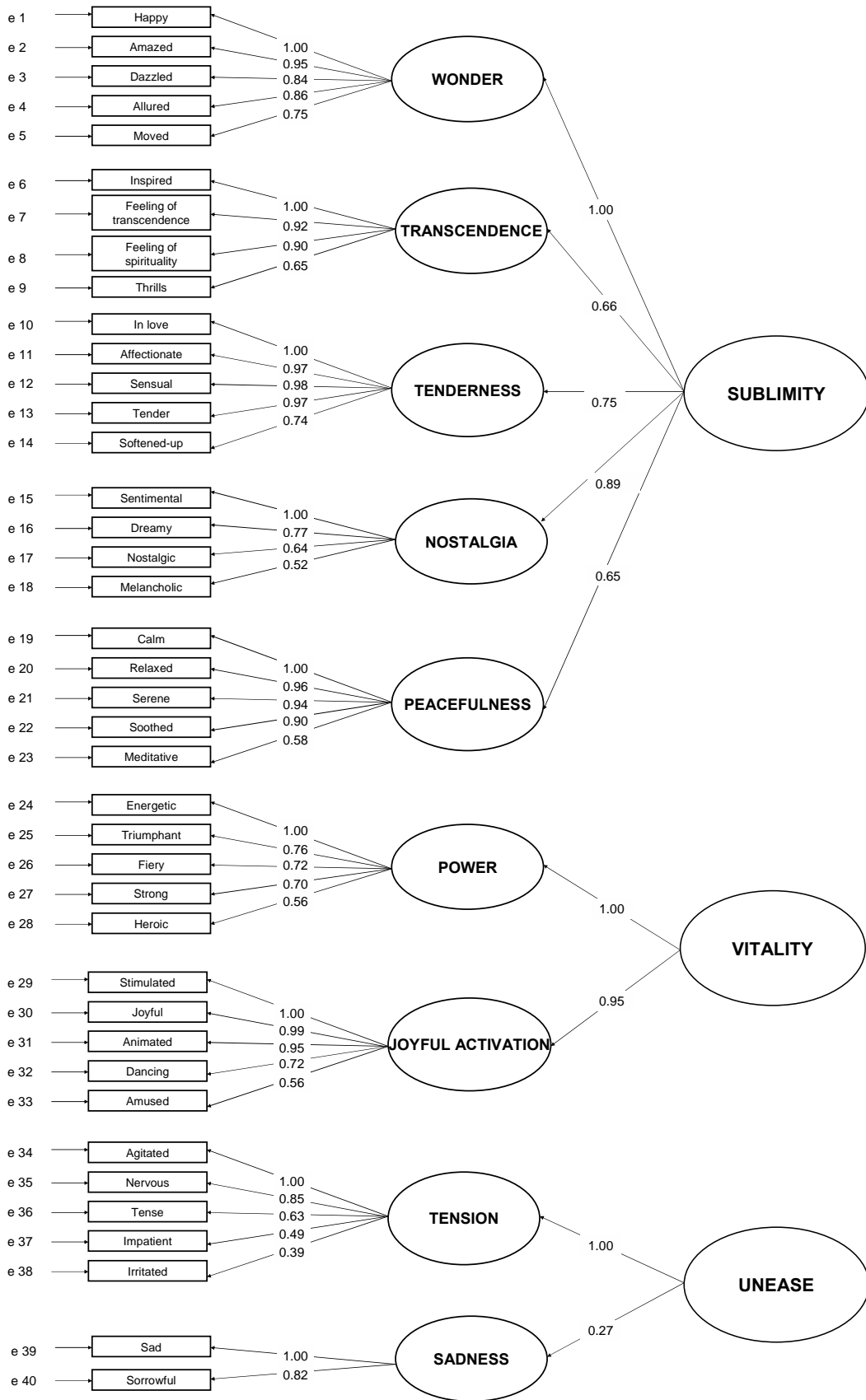
Loudness	Loudness	Perceived loudness	How loud the sound is perceived. Scale: Soft – Loud
	Dynamics		Attack
		Bass Precision	Are instrument impacts from the bass drum and bass precise, crisp and without distortion, are the impacts tight and well defined? Bass precision may be defined as Attack in the bass region. Imprecise means that the attack speeds in time and the peak of the impact is softened. Scale: Imprecise – Precise
		Punch	Specifies whether the strokes on drums and bass are reproduced with clout, almost as if you can feel the blow. The ability to effortlessly handle large volume excursions without compression (compression is heard as level variations that are smaller than one would expect from the perceived original sound). Scale: A little – A lot
		Powerful	The ability to handle high sound levels, especially when striking the drums and bass. Indicates whether the Punch, Attack and Bass precision are maintained at high volume. Scale: A little – A lot
Timbre	Treble	Treble strength	The relative strength of the treble or high frequencies Scale: – Weak: Covered, unsharp. – A little under neutral: A soft sound without being dull. – Neutral: In the middle of the scale, where you can clearly distinguish instruments. – A lot: Treble Raised. Sharp, hard sound.
		Brilliance	Treble or high frequency extension: Scale: – A little: As if you hear music through a door, muffled, blurred or dull. – A lot: Crystal-clear reproduction extended treble range with airy and open treble. Lightness, purity and clarity with space for instruments. Clarity in the upper frequencies without being sharp or shrill and without distortion.
		Tinny	Resonances or narrowband frequency prominence in the treble or high frequencies. Scale: None – Weak – A lot
	Midrange	Midrange strength	The relative level of the midrange, i.e. the middle frequencies, e.g. sopranos, trumpets, violins and xylophones. Not to be confused with Canny which represents prominent narrow frequency ranges (resonances) in the midrange. Scale: Soft – Loud
		Nasal	A closed sound with pronounced midrange. Gives the impression corresponding to vocalists singing through the nose (nasal). Scale: A little – A lot
		Canny	The music sounds like it is being played in a can or tube. The sound is characterized by prominent and narrowband resonances in the midrange. Scale: A little – A lot
	Bass	Bass strength	The relative level of bass, i.e. the low frequencies, for example male voices, bass guitar, bass drum, timpani and tuba. Should not be confused with bass depth that indicates the low frequency bass extension. Scale: Soft – Loud
		Bass depth	Denotes how far the bass extends downwards. If it goes down in the low end of the spectrum, there is great depth. Should not be confused with Bass strength, which indicates the strength of the bass or Boomy which related to resonances in the lower bass region. Scale: A little – A lot
		Boomy	Resonances in the low bass, as sound in a large barrel, which gives a prominent bass resound resounding (reverberating) when bass and bass drums are heard. The representation tends to become muddy and imprecise. Scale: None – Weak – Loud
		Boxy	Boxy denotes a hollow sound, as if the sound was played inside a small box. Represents resonances in the upper bass frequency range. Scale: None – Weak – Loud
Timbral balance	Dark- Bright	Denotes the balance between bass and treble. Scale: – Dark: Excessive bass. Either loud bass or weak treble. – Neutral: Bass and treble are perceived equally loud, there is a balance in the reproduction. This also applies if both bass and treble are equally weak or if the bass and treble are both too loud. If it leads to prominent or soft midrange this is assessed by the Midrange strength. – Bright: Excessive treble. Either loud treble or weak bass. The cause for the sound being dark or light can deduced from the assessments of Bass strength and Treble strength.	
	Full	If both low and high frequencies are well represented with good extension the sound is Full. Scale: A little – A lot	

		Homogeneous	Denotes to which degree the different frequency ranges (bass, midrange and treble) are coherent, continuous, and balanced without gaps between them. That there are seamless transitions between the tone ranges. Scale: A little – A lot
Spatial	Spatial extent	Depth	The radial extent of source, scene or ensemble from the listener (in any direction away from the listener). Scale: Shallow- Deep A talker in a dry acoustic environment may be perceived as having no or little depth. A crowd in a stadium has depth. The sounds of the crowd are simultaneously occurring close to and far from the listener.
		Width	The perceived extent of a sound source, ensemble or scene from side to side. A narrow (or a wide) source may be perceived in any direction seen from the listener, i.e. in front of the listener, or to their side, above them or anywhere around them. Scale: Narrow – wide A single talker in a dry environment might be perceived as narrow. An organ in a church may be perceived as wide.
		Envelopment	Degree of being surrounded by a source, scene or ensemble. Typically, envelopment is associated with a scene. Scale: Not enveloping – Completely enveloping Being surrounded by reverberation would be considered highly enveloping. Being surrounded by a large number of dry sources may also be highly enveloping. This may be heard when standing and listening to the rain hitting the pavement. Envelopment may occur with reverberation or other aspects of the scene such as applause in a concert hall, atmosphere or air conditioning (room tone). Holes (an absence of sound from a certain directions) in the reproduction would normally reduce envelopment. Envelopment may be subdivided in horizontal and vertical envelopment
		Balance	Is the soundstage skewed to one side (left or right) or is it centred in the middle?
		Distance	How far a sound source, ensemble or scene is perceived from the listener. Scale: Near – Far Hearing a person talking close to a listener in a park may be perceived as near, whereas a person at a large distance may be perceived as far.
	Localization	Internality	The distinctness to which a sound source, ensemble or scene is perceived in-side the head. Distinctly internal means that a sound is unambiguous perceived within the head. Distinctly external means that a sound is unambiguous perceived outside the head. If it is ambiguous whether the sound is perceived to be inside or outside of the head, the internality is Indistinct. Scale: Distinctly internal- Indistinct- Distinctly external A dry sample played equally to both ears over headphone may be perceived inside the head Naturally occurring sounds are generally perceived externally.
		Localisability	The degree of precision to which the position and extent of a source or ensemble can be identified. This attribute is typically associated with sources or ensembles, rather than scenes. For a spatially imprecise sound the listener may be unable to identify the position (and extend) of the source or ensemble For a spatially precise sound, the listener can confidently state the position and extend of the source or ensemble. Scale: Imprecise – Precise A clap in a dry environment may be spatially precise. Listening to rain fall in a forest maybe spatially imprecise.
	Environment	Reverberance	Reverberance is the percept of reverberation. The sound that persists after the source has stopped. May be caused by environmental reflections and scattering or artificially generated. Reverberance is most noticeable when a short impulsive sounds stop. However this will also be perceived with continuous sources, such as an orchestra playing in a concert hall or a cathedral. Scale: Dry – Highly reverberant Highly reverberant sounds can be experienced in large spaced such as cathedrals, caves, etc. Dry spaces are often small furnished spaces such as living rooms or spaces outdoor without reflecting objects. Reverberance may be subdivided into "Level of reverberance" and "Duration of reverberance" Level of reverberance: The level of the sound that persists after the source has stopped relative to the level of the source. Scale: Low – high Reverberance is most noticeable when a short impulsive sounds stop. This will also be perceived with continuous sources, such as an orchestra playing in a concert hall or a cathedral. If a singer is close to the listener in a church the level of reverberance would be low. If the listener is far from the singer in a church, the level of reverberance would be high. Furthermore, in a dry room the level of reverberance would also be low. Duration of reverberance: Describes the length of time the sounds takes to decay to an imperceptible level, after the source has stopped. Duration of reverberance is most easily judged when a short impulsive sound stops. Duration of reverberance may also be judged with continuous sources, such as an orchestra playing in a concert hall or a cathedral. Scale: Short – long A short reverberation duration may be experienced in a small room. A long reverberation decay can be experienced in a cathedral.
		Clarity	The impression of how clearly different elements in a scene can be spatially distinguished from each other. Scale: Unclear-clear A singer and a piano performing a duet in a dry acoustic, may be perceived as clear. When listening to a choir from the rear of the church, the sound of the individual signers maybe unclear.
	Transparency	Presence	Does it sound as if the sound sources are present and not distant or absent? Scale: A little – A lot
Clean		It is easy to listen into the music, which is timbral clear and distinct. Instruments and vocals are reproduced accurately and distinctly. The opposite of clean: dull, muddy. Scale: A little – A lot	

		Detailed	A well-resolved sound rich in detail. Instruments, voices etc. can easily be separated. The music has many details, details that cannot be measured, details that give the music "soul". It may be small audible nuances: Breathing from a singer, fingers wandering across the guitar strings, the flaps from the clarinet, embouchure sound of the saxophone, the impact from the piano's hammers when they hit the strings. Scale: A little – A lot
		Natural	Sounds reproduced with high fidelity. Acoustic instruments, voices and sounds, sounds like in reality. The sound is similar to the listener's expectation to the original sound without any timbral or spatial coloration or distortion, "Nothing added – nothing missing." The soundstage is clear in space and brings you close to the perceived original sound experience. Scale: A little – A lot
Artefacts	Signal related	Shrill	Treble Distortion. Very sharp s-sounds, cymbals etc. Scale: A little – A lot
		Rubbing	As the sound of something scraping on a (rough) surface. Scale: A little – A lot
		Rough	A hoarse off-sound unintentionally accompanying the reproduced sound. Bass distortion. Scale: A little – A lot
		Buzzing	A zzz-like, undesirable sound typically in the low and midrange frequencies. Scale: A little – A lot
		Clipped	The harmonics are too pronounced and sharp. Scale: A little – A lot
		Distorted	Additional and undesired sounds that add a sharpness to the reproduction. Scale: A little – A lot
		Compressed	Limited dynamic range leading to a lack of natural peaks. Dynamic compression may be heard as a pumping effect. Scale: A little – A lot
	Noise	Fluctuating / Intermittent	Noise with varying loudness and or pauses. Scale: A little – A lot
		Bubbling	Sound or noise with fast (<1 sec.) variations in frequency and/or loudness. Scale: A little – A lot
		Humming	Low frequency noise with tonal components. Scale: A little – A lot
		Hissing	A noise-like sizzling sound, like the sound of bacon in a frying pan. Scale: A little – A lot

FIGURE 2

Example of a hierarchical structure of emotional attributes for music



Future developments

It is anticipated that additions will be made to this report, particularly in the area of spatial audio attributes, adding specific attributes.

Glossary

Attribute: A property that can be perceived (perceptual, affective or connotative), it may or may not be prominent. Other common terms for attributes are: factors, auditory feature.

Descriptor or technical descriptor: A word or phrase that describes, identifies, or labels an attribute or a characteristic. Often including specialist terms that may be interpreted differently by the layman or researcher from different laboratories – this is to be expected. Descriptors are usually provided with descriptions and example samples to convey the specific meaning to new panels or other laboratories.

Lexicon: Lexicons are standardized vocabularies that objectively describe the sensory properties of consumer products.

Perceptual measurement: An objective quantification of the sensory strength of individual sensory descriptors of a perceived stimulus. Perceptual tests are measurements where humans (usually expert assessors) are used as “measuring instruments”.

Sound character: The overall concept of a set of characteristics that portrays the sound. The “sound character” can be specified by a number of attributes and/or metrics. The sound character is not to be confused with the sound quality, which involves a comparison with some desired or subjectively preferred features.

Sound system characteristics: Sound reproduction systems are required to present acoustic signals to the listeners’ ears in such a way that the intention of the creator is delivered to the listeners’. Any perceived deviation from this is caused by the characteristics of the sound reproduction system.

Affective measurement: Subjective measurements of preference, liking, annoyance or of connotative attributes (ideally made with individuals from the relevant consumer segment or target group). Often associated with personal opinion in the form of a subjective assessment.

Sound Quality: Assessment of quality involves a (conscious or unconscious) comparison with some desired features (a personal “reference”). For that reason quality is a subjective characteristic. The better the characteristics of the sound match the desired features, the higher the quality will be rated. People judge a product to be high quality when their expectations are met or exceed. The sound quality is not to be confused with the sound character. Sound quality is generally viewed as affective in nature, as are many opinions regarding its definition.

References

- [1] Stone, H. and Sidel, J.L., *Sensory evaluation practices*, 2nd ed., Academic Press, San Diego, CA, USA, 1993
- [2] Lawless, H. & Heymann, H., *Sensory evaluation of food – Principles and Practices*, Springer, 2010
- [3] Bech, S. & Zacharov, N., *Perceptual Audio Evaluation*, Wiley, ISBN 0-470-86923-2, 2006
- [4] Pedersen, T. H. & Zacharov, N., *How many psycho-acoustic attributes are needed?* Acoustics08 proceedings, 2008
- [5] Pedersen, T. H., *The semantic space of sound*, Technical report, DELTA, ISBN 978-87-7716-036-3, 2008
- [6] ASTM International, DS72, *Lexicon for Sensory Evaluation: Aroma, Flavor, Texture and Appearance*, ISBN-13: 978-0-8031-7025-4, 2011
- [7] Pedersen, T. H. & Zacharov, N., *The Development of a Sound Wheel for Reproduced Sound*. AES 138, Warsaw Poland
- [8] Zentner, M., Grandjean, D., & Scherer, K. *Emotions evoked by the sound of music: Characterization, classification, and measurement*. *Emotion*, 8, 494-521, 2008
- [9] H. Wilkens, *Mehrdimensionale Beschreibung subjektiver Beurteilungen der Akustik von Konzertsälen*, Ph.D. thesis, TU Berlin, 1975.
- [10] A. Gabrielsson and H. Sjögren, "Perceived sound quality of sound-reproduction systems," *Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 1019–1033, 1979.
- [11] A. Gabrielsson, "Dimension analyses of perceived quality of sound reproduction systems," *Scandinavian Journal of Psychology*, vol. 20, pp. 159–169, 1979.
- [12] Floyd E. Toole, "Subjective measurements of loudspeaker sound quality and listener performance," *J. Audio Eng. Soc.*, vol. 33, no. 1/2, pp. 2–32, 1985.
- [13] C Lavandier, *Validation perceptive d'un modèle objectif de caractérisation de la qualité acoustique des salles*, Ph.D. thesis, Université du Maine, Le Mans, France, June 1989.
- [14] Tomasz Letowski, "Sound quality assessment: Concepts and criteria," in *Audio Engineering Society Convention 87*, Oct 1989.
- [15] G. Kahle, *Validation d'un modèle objectif de la perception de la qualité acoustique dans un ensemble de salles de concerts et d'opéras*, Ph.D. thesis, IRCAM, Paris, France, June 1995.
- [16] Russell Mason and Francis Rumsey, "An assessment of the spatial performance of virtual home theatre algorithms by subjective and objective methods," in *Audio Engineering Society 108th Convention*, Feb 2000.
- [17] N. Zacharov and K. Koivuniemi, "Unravelling the perception of spatial sound reproduction: Language development, verbal protocol analysis and listener training," in *Audio Engineering Society Convention 111*, Nov 2001.
- [18] N. Zacharov and K. Koivuniemi, "Perceptual audio profiling and mapping of spatial sound displays," in *Proceedings of the Int. Conf. on Auditory Display. ICAD*, 2001, pp. 95–104.

- [19] Francis Rumsey, “Spatial quality evaluation for reproduced sound: Terminology, meaning, and a scene-based paradigm,” *J. Audio Eng. Soc.*, vol. 50, no. 9, pp. 651–666, 2002.
- [20] Jan Berg and Francis Rumsey, “Systematic evaluation of perceived spatial quality,” in *Audio Engineering Society Conference: 24th International Conference: Multichannel Audio, The New Reality*, Jun 2003.
- [21] C. Guastavino and B. F. G. Katz, “Perceptual evaluation of multi-dimensional spatial audio reproduction,” *Journal of the Acoustical Society of America*, vol. 116, no. 2, pp. 1105–1115, 2004.
- [22] Gaëtan Lorho, “Individual vocabulary profiling of spatial enhancement systems for stereo headphone reproduction,” in *Audio Engineering Society Convention 119*, Oct. 2005.
- [23] Jan Berg and Francis Rumsey, “Identification of quality attributes of spatial audio by repertory grid technique,” *J. Audio Eng. Soc.*, vol. 54, no. 5, pp. 365–379, 2006.
- [24] Sylvain Choisel and Florian Wickelmaier, “Extraction of auditory features and elicitation of attributes for the assessment of multichannel re- produced sound,” *J. Audio Eng. Soc.*, vol. 54, no. 9, pp. 815–826, 2006.
- [25] Robert Sazdov, Garth Paine, and Kate Stevens, “Perceptual investigation into envelopment, spatial clarity, and engulfment in reproduced multi-channel audio,” in *Audio Engineering Society Conference: 31st International Conference: New Directions in High Resolution Audio*, Jun 2007.
- [26] Andreas Silzle, “Quality taxonomies for auditory virtual environments,” in *Audio Engineering Society Convention 122*, May 2007.
- [27] Helmut Wittek, *Perceptual differences between wavefield synthesis and stereophony*, Ph.D. thesis, University of Surrey, 2007.
- [28] Tapio Lokki, Jukka Pätynen, Antti Kuusinen, and Sakari Tervo, “Disentangling preference ratings of concert hall acoustics using subjective sensory profiles,” *The Journal of the Acoustical Society of America*, vol. 132, no. 5, pp. 3148 – 3161, 2012.
- [29] Sarah Le Bagousse, Mathieu Paquier, and Catherine Colomes, “Categorization of sound attributes for audio quality assessment a lexical study,” *J. Audio Eng. Soc.*, vol. 62, no. 11, pp. 736–747, 2014.
- [30] Alexander Lindau, Vera Erbes, Steffen Lepa, Hans-Joachim Maempel, Fabian Brinkmann, and Stefan Weinzierl, “A spatial audio quality inventory for virtual acoustic environments (SAQI),” in *Proc. of the EAA Joint Symposium on Auralization and Ambisonics*, 2014.
- [31] Neofytos Kaplanis, Søren Bech, Søren Holdt Jensen, and Toon van Waterschoot, “Perception of reverberation in small rooms: A literature study,” in *Audio Engineering Society Conference: 55th International Conference: Spatial Audio*, Aug 2014.
-