



**Recommendation ITU-R BS.1284-2**  
(01/2019)

**General methods for the subjective  
assessment of sound quality**

**BS Series**  
**Broadcasting service (sound)**



## 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 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 Recommendations

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

| Series     | Title                                                                                |
|------------|--------------------------------------------------------------------------------------|
| <b>BO</b>  | Satellite delivery                                                                   |
| <b>BR</b>  | Recording for production, archival and play-out; film for television                 |
| <b>BS</b>  | Broadcasting service (sound)                                                         |
| <b>BT</b>  | Broadcasting service (television)                                                    |
| <b>F</b>   | Fixed service                                                                        |
| <b>M</b>   | Mobile, radiodetermination, amateur and related satellite services                   |
| <b>P</b>   | Radiowave propagation                                                                |
| <b>RA</b>  | Radio astronomy                                                                      |
| <b>RS</b>  | Remote sensing systems                                                               |
| <b>S</b>   | Fixed-satellite service                                                              |
| <b>SA</b>  | Space applications and meteorology                                                   |
| <b>SF</b>  | Frequency sharing and coordination between fixed-satellite and fixed service systems |
| <b>SM</b>  | Spectrum management                                                                  |
| <b>SNG</b> | Satellite news gathering                                                             |
| <b>TF</b>  | Time signals and frequency standards emissions                                       |
| <b>V</b>   | Vocabulary and related subjects                                                      |

**Note:** This ITU-R Recommendation was approved in English under the procedure detailed in Resolution ITU-R 1.

Electronic Publication  
Geneva, 2019

© ITU 2019

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

## RECOMMENDATION ITU-R BS.1284-2

**General methods for the subjective assessment of sound quality**

(Question ITU-R 19-1/6)

(1997-2003-2019)

**Scope**

This Recommendation describes general methods of the subjective assessment for sound quality with or without a reference signal.

**Keywords**

Subjective assessment, audio quality, stereophonic image quality, front image quality, timbral quality, sound colour, sound homogeneity, localization quality, environment quality

The ITU Radiocommunication Assembly,

*considering*

- a) that the introduction of new kinds of sound signal processing, such as digital coding and bit rate reduction, new kinds of television signals using time multiplexed components and new services such as enhanced television, high definition television (HDTV), ultra-high definition television (UHDTV), high dynamic range television (HDR-TV) and advanced immersive audio-visual (AIAV) system, may require new or amended methods of subjective sound quality assessment;
- b) that these techniques entail their own specific signal impairments;
- c) that subjective listening tests permit assessment of the degree of annoyance caused to the listener by any impairment of the wanted signal during its transmission between the source and the listener;
- d) that many different methods of subjective testing are possible;
- e) that it is highly desirable to standardise the methods of subjective testing and the interpretation of the results, so that the best possible comparisons may be made between results obtained at different times and/or in different places;
- f) that it is highly desirable that the grading scales which are used to describe the subjective quality of sound should permit more consistent statistical processing methods, independent from the language used to express the opinions;
- g) that it would be desirable for a single assessment scale to be available for both sound and television programmes;
- h) that the geometric and acoustic properties of control rooms and listening rooms can have a considerable influence on audition, and therefore listening conditions should be closely specified,

*recommends*

that the testing and evaluation procedures given in Annex 1 to this Recommendation should be used for the subjective assessment of the quality of reproduced sound.

## Annex 1

### 1 General

This Annex is divided into the following sections, giving detailed requirements for various aspects of the tests:

- 1 General
- 2 Experimental design
- 3 Selection of the listening panel
- 4 Test method
- 5 Attributes
- 6 Programme material
- 7 Reproduction devices
- 8 Listening conditions
- 9 Statistical treatment of data
- 10 Presentation of results
- 11 Contents of test reports

### 2 References

This Recommendation is intended as a guide to the general assessment of sound quality. It is based on Recommendation ITU-R BS.1116 – Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems. However, the requirements of Recommendation ITU-R BS.1116 are stringent, being intended for the assessment of small impairments. More general assessments usually involve larger differences and therefore do not usually need such close control of the test parameters. Recommendation ITU-R BS.1116 contains a glossary of terms, some of which are used in the present Recommendation.

Other ITU Recommendations which may be relevant in some special cases, are referred to in Recommendation ITU-R BS.1283 – A guide to ITU-R Recommendations for subjective assessment of sound quality.

### 3 Experimental design

In designing the tests, the considerations of Recommendation ITU-R BS.1116, § 2 should be taken into account. However, because the impairments being tested may not be small, it is not always essential to use a reference. If a reference is used, it need not be unimpaired in an absolute sense.

In general, statistical expertise will be required to design the test. This would include the determination of the number of observations needed, the statistical methods for analysing the data and the correct interpretation of the outcomes of the statistical analysis, including a check of the validity of the model assumptions.

### 4 Selection of the listening panel

Expert listeners are always preferred to non-expert listeners. It has been argued that non-experts may be representative of the general population, and that experts may be excessively critical. However,

with long-term exposure to artefacts, in time some non-experts become experts. Therefore, tests using experts give a better and quicker indication of the likely results in the long term. In cases of doubt, the relationship between expert and non-expert opinion should be investigated.

The minimum number of expert listeners should normally be ten, whilst the minimum number of non-expert listeners should normally be twenty. Whenever the system is intended for high-quality sound broadcasting or reproduction, expert listeners should be used.

Generally, the listeners should undertake training to familiarise themselves with the test procedure, the test materials and the test environment.

## 5 Test method

### 5.1 Grading scales

The following unipolar discrete five-grade scales should be used for the subjective assessment of sound quality or impairment. The nature and purpose of the tests will determine which of the scales is the more appropriate.

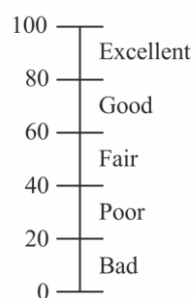
FIGURE 1  
Unipolar discrete five-grade scale

| Quality |           | Impairment |                               |
|---------|-----------|------------|-------------------------------|
| 5       | Excellent | 5          | Imperceptible                 |
| 4       | Good      | 4          | Perceptible, but not annoying |
| 3       | Fair      | 3          | Slightly annoying             |
| 2       | Poor      | 2          | Annoying                      |
| 1       | Bad       | 1          | Very annoying                 |

BS.1284-01

The continuous quality scale (CQS) is also used, for instance, in Recommendations ITU-R BS.1534 and ITU-R BT.500. The CQS consists of a 100-point linear scale which is divided into five equal intervals with five regional verbal anchors used in the five-grade scales.

FIGURE 2  
Unipolar continuous 100-point quality scale with five regional verbal anchors



BS.1284-02

For comparison tests, either a method based on the following bipolar discrete seven-grade comparison scale or one based on numerical differences using the above five-grade scales may be used. In general, these are not equivalent and may not give the same results.

It is essential that the intended direction of the comparison be clearly indicated.

FIGURE 3

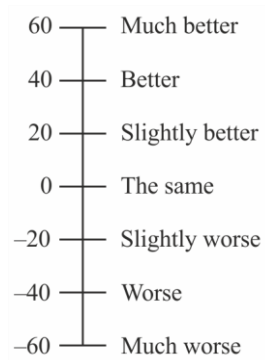
**Bipolar discrete seven-grade scale**

| Comparison |                 |
|------------|-----------------|
| 3          | Much better     |
| 2          | Better          |
| 1          | Slightly better |
| 0          | The same        |
| -1         | Slightly worse  |
| -2         | Worse           |
| -3         | Much worse      |

BS.1284-03

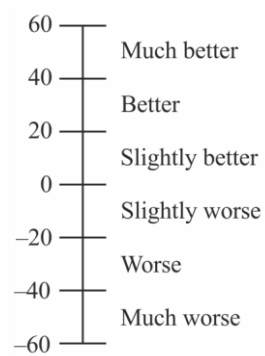
The CQS which consists of a 120-point continuous scale which is divided into six equal intervals is also used.

FIGURE 4

**Bipolar continuous  $\pm 60$ -point comparison scale with seven point verbal anchors**

BS.1284-04

FIGURE 5

**Bipolar continuous  $\pm 60$ -point comparison scale with six regional verbal anchors**

BS.1284-05