

## RECOMMENDATION ITU-R BS.1679

**Subjective assessment of the quality of audio in large screen digital imagery applications intended for presentation in a theatrical environment**

(Question ITU-R 15/6)

(2004)

The ITU Radiocommunication Assembly,

*considering*

- a) that it will be necessary to verify the suitability of the technical solutions considered for members of that family of large screen digital imagery (LSDI) applications;
- b) that such verification will also need to include, when necessary, subjective assessment tests performed under rigorous scientific conditions;
- c) that Recommendation ITU-R BS.1284 specifies general requirements applicable to the subjective assessment of the quality or impairment of program audio;
- d) that LSDI programs intended for presentation in a theatrical environment will be generally accompanied by multichannel audio, thus requiring a subjective assessment procedure designed for multichannel audio;
- e) that Recommendation ITU-R BS.775 covers multichannel stereophonic audio signals with and without accompanying picture;
- f) that the subjective assessment of the quality of audio in LSDI applications intended for presentation in a theatrical environment requires a procedure in which the quality of the audio is assessed in the presence of the image component of the LSDI program, since perceptual interaction between audio and picture can affect the assessment of audio quality;
- g) that Recommendation ITU-R BS.1286 covers methods for the subjective assessment of audio systems with accompanying picture;
- h) that, the source-coding (if any) used for the delivery of LSDI program audio for presentation in a theatrical environment should be transparent or virtually transparent to the audio quality present on the program master, and the subjective assessment of source-coding transparency requires a procedure designed to assess small audio impairments;
- j) that Recommendation ITU-R BS.1116 covers methods for the subjective assessment of small impairments in audio systems including multichannel audio systems,

*recommends*

**1** that the subjective assessment of audio quality or audio impairments in LSDI applications designed for program presentation in a theatrical environment should be based on a choice among the specifications contained in Recommendations ITU-R BS.1284, ITU-R BS.1286 and ITU-R BS.1116;

2 that the listening environment used for those subjective assessments should be based on the universal multichannel stereophonic audio system specified in Recommendation ITU-R BS.775. If the subjective assessment uses a loudspeaker arrangement different from the reference one indicated in Recommendation ITU-R BS.775, the assessment report should describe it in detail;

3 that reference should be made to Annex 1 for a summary indication of the provisions to be selected in the four Recommendations above, for implementation in the subjective assessment of audio of LSDI applications for presentation in a theatrical environment, and reference should be made to the four Recommendations themselves for full details of the selected provisions.

## **Annex 1**

### **Summary of provisions for the subjective assessment of LSDI audio quality**

#### **1 Introduction**

This Annex provides a summary of the provisions that should be implemented when performing subjective assessment tests of audio quality or audio impairment for LSDI applications designed for program presentation in a theatrical environment.

These provisions have been taken from those contained in Recommendations ITU-R BS.775, ITU-R BS.1116, ITU-R BS.1284 and ITU-R BS.1286. They apply to the case of the indicated LSDI applications, which can be characterized as follows:

- The audio to be assessed is a multichannel program audio.
- The audio accompanies program images presented on a large screen in a theatrical environment.
- The expected impairment is small with respect to the subjective audio quality present on the program master.

Reference should be made to the Recommendations listed above for full details of the selected provisions.

#### **2 General provisions related to the assessment of program audio**

Recommendation ITU-R BS.1284 specifies general requirements for the subjective assessment of audio quality. Several provisions apply to the specific case of the subjective assessment of small impairments in multichannel program audio with accompanying picture. This particularly applies to the following elements.

## Listening panel

Expert listeners are 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.

## Grading scales

The following five-grade scale is recommended for the subjective assessment of “basic audio quality”<sup>1</sup>. Due to the fact that LSDI applications focus on high quality the five-grade quality scale is not appropriate.

	<b>Impairment</b>
5	Imperceptible
4	Perceptible, but not annoying
3	Slightly annoying
2	Annoying
1	Very annoying

For comparison tests, either a method based on the following 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. Taking into account that LSDI is focusing on high quality comparison tests in general are not adequate.

	<b>Comparison</b>
3	Much better
2	Better
1	Slightly better
0	The same
-1	Slightly worse
-2	Worse
-3	Much worse

NOTE 1 – The scales should be treated as continuous, with a recommended resolution of 1 decimal place.

NOTE 2 – It has been shown that the use of pre-defined intermediate anchor points may introduce bias. It is possible to use the number scales without descriptions of anchor points. In such cases, the intended orientation of the scales must be indicated. This may help to overcome translation problems when comparing the results of tests written in different languages.

---

<sup>1</sup> Basic audio quality is used here in the same way as it is used in Recommendation ITU-R BS.1116.

If intermediate anchor points are not used it is essential that the results for individual subjects are normalized with respect to mean and standard deviation. Recommendation ITU-R BS.1284 provides the normalization algorithm that can be used.

### **Test procedures**

Tests may be of single presentations, paired comparisons (one of which may be the reference) or multiple comparisons, with or without references. The presentations may be repeated as required.

Short-term human memory limitations may dictate that each program excerpt should not last longer than 15 to 20s.; excerpts may be very short (a few seconds) for some tests. In the case where the sequence is a musical item, the phrase should not appear to be interrupted.

When the test sequence is not under the control of the subject, it is necessary to provide a clear indication of the current presentation.

No session with any one listener should last longer than about 15 to 20 min without interruption. If the sessions must be consecutive, they should be separated by rest periods of at least the same duration.

### **Program material**

When the system is intended to carry high quality audio, as it is the case of LSDI applications, the test material should be chosen for its highly critical behavior with respect to the impairments introduced by the system being tested.

To ensure the comparability of test data obtained in different places and/or at different times, some program sequences should be the same in all the tests to be compared. Statistical testing on the common test items must be performed to check whether it is allowed to compare the results of two tests.

In any event, the content of a program sequence in general should be neither so interesting nor so disagreeable or boring that the listener is distracted. However a few program sequences designed to stress the systems under test might also sound unpleasant.

### **Statistical treatment of data**

The subjective scores should be processed to derive the mean values and confidence intervals. This will describe the data and, if the resulting discrimination is inadequate to satisfy the objectives of the test, further processing should be carried out, as detailed in Recommendation ITU-R BS.1116.

The overall value of the test will be enhanced if the data is further analysed to verify the underlying assumptions of the test and to evaluate subject reliability.

### **Presentation of test results**

Specifications for the presentation of the test results are given in Recommendation ITU-R BS.1116.

In general, all aspects of the test should be reported, as per Recommendation ITU-R BS.1116, even if some of the aspects were not implemented or controlled.

### 3 Provisions related to the assessment of multichannel program audio

Recommendation ITU-R BS.775 specifies a reference loudspeaker arrangement for multichannel program audio, and the use of five reference recording/transmission signals for left (L), right (R), centre (C), channels for the front, and left surround (LS) and right surround (RS) channels for the side/rear. Additionally the system may include a low frequency extension signal for a low frequency effects (LFE) channel.

The Figure that details the reference loudspeaker arrangement in Recommendation ITU-R BS.775 is reproduced in Fig. 1 for memory and reference. An example of the loudspeaker arrangement in a typical theatre environment is shown in Fig. 2; in this case (see Note 1), in order to obtain coverage over a larger seating area, the surround channels are reproduced by two arrays of loudspeakers.

Depending on the LSDI application for which the subjective assessment test is designed, the loudspeaker configuration that best fits the investigated application should be chosen.

NOTE 1 – Optionally, there may be an even number of more than two rear/side loudspeakers which may provide a larger optimum listening area and greater envelopment.

NOTE 2 – Optimum audio reproduction requires use of wide angular spacing between the left and right loudspeakers of two or three front loudspeaker channel stereophonic systems (see Fig. 1).

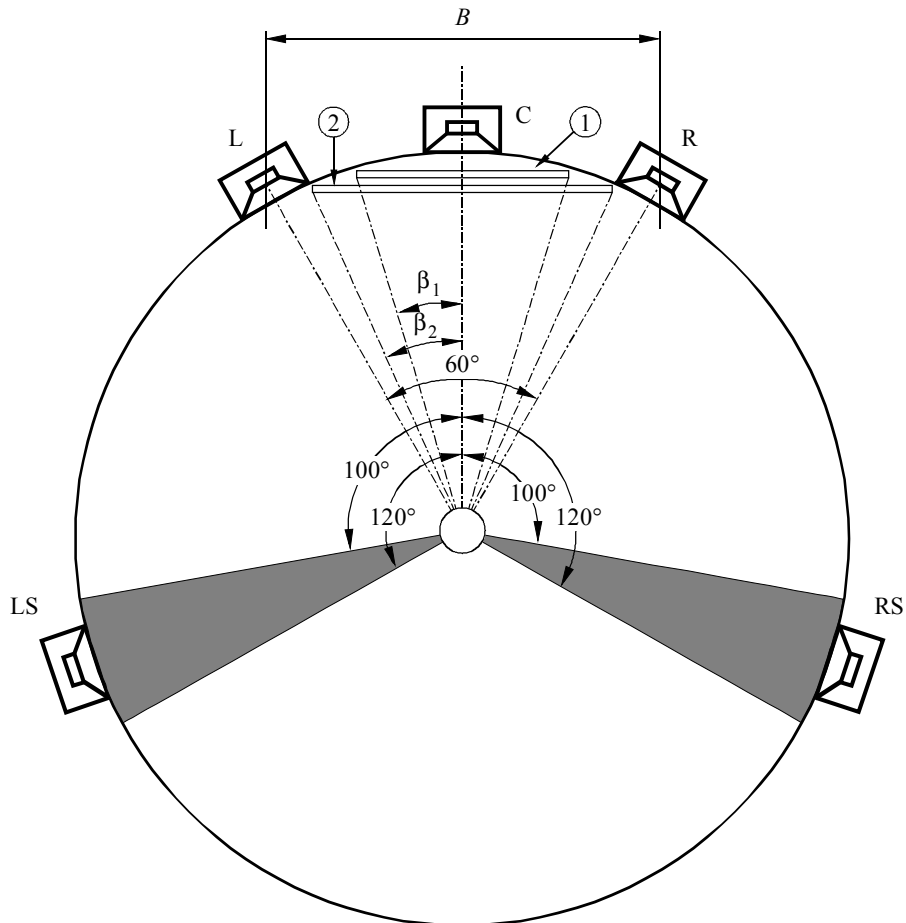
NOTE 3 – The size of the loudspeaker base-width, B (see Fig. 1) is defined for reference listening test conditions in Recommendation ITU-R BS.1116.

NOTE 4 – If more than two rear/side loudspeakers are used, then the loudspeakers should be disposed symmetrically and at equal intervals on the arc from the centre front reference.

NOTE 5 – If more than two rear/side loudspeakers are used, the LS signal should be fed to each of the side/rear loudspeakers on the left side of the room and the RS signal should be fed to each of the side/rear loudspeakers on the right side of the room. In doing so, it will be necessary to reduce the signal gain such that the total power emitted by the loudspeakers carrying the LS (or RS) signal is the same as if that signal had been reproduced over a single loudspeaker. For large room reproduction, it may also be necessary to delay, or otherwise decorrelate, the feeds to some or all of the side/rear loudspeakers.

NOTE 6 – If other audio systems for LSDI application, as for instance 10.2, Wave Field Synthesis or Ambiosonics are under test the loudspeaker arrangement might differ significantly. In that case the test report has to specify the loudspeaker arrangement used in detail.

FIGURE 1  
Reference loudspeaker arrangement as shown in Rec. ITU-R BS.775



Screen 1 HDTV – Reference distance  $= 3 H (2\beta_1 = 33^\circ)$

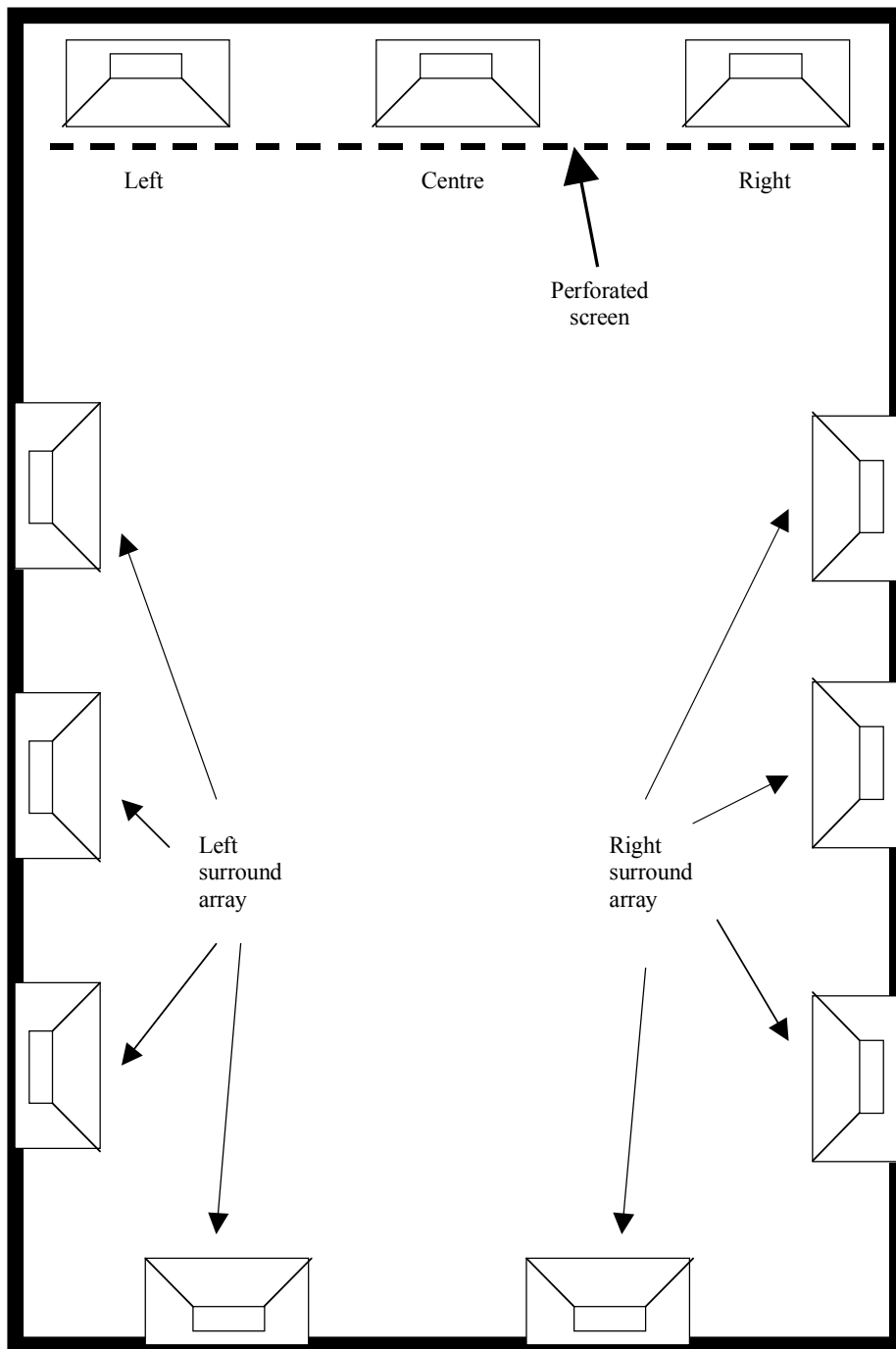
Screen 2  $= 2 H (2\beta_2 = 48^\circ)$

$H$ : height of screen

$B$ : loudspeaker base width

Loudspeaker	Horizontal angle from centre (degrees)	Height (m)	Inclination (degrees)
C	0	1.2	0
L, R	30	1.2	0
LS, RS	100 ... 120	$\geq 1.2$	0 ... 15 down

FIGURE 2  
Example of a loudspeaker arrangement in an LSDI theatrical venue



#### 4 Provisions related to the assessment of program audio with accompanying picture

Recommendation ITU-R BS.1286 specifies methods for the subjective assessment of audio with accompanying picture.

It identifies the four areas of assessment below as requiring the presentation of the visual component of the program, namely:

- correlation between picture and audio images;
- basic audio quality as influenced by the presence of a visual image;
- harmony of spatial impressions of picture and audio;
- assessment of listening and viewing arrangements.

#### Attributes that may be assessed

The following attributes may be assessed:

- front image quality;
- impression of surround quality;
- basic audio quality;
- correlation between audio and picture images, namely:
  - correlation of source positions derived from visual and audible cues<sup>2</sup>;
  - correlation of spatial impressions between audio and picture;
  - temporal relationship between audio and video.

#### Subjective assessment method

Recommendation ITU-R BS.1286 recommends that, if the subjective differences are expected to be small as it is the case of LSDI programs, it is appropriate to use the double-blind triple-stimulus method with hidden reference as described in Recommendation ITU-R BS.1116, § 4.

It should be noted that the reference signal does not need to be unimpaired in an absolute sense.

Subjects should be instructed to assess the audio quality in association with the video presentation, rather than to assess the audio quality alone.

The test program material should be selected to stimulate the attributes of interest. In general a small group of listeners should pre-screen a larger set of program material to find the most critical program material.

Different attributes may need different types of test program.

#### Presentation environment

The presentation environment in the Table below specifies the viewing conditions for the subjective assessment of LSDI program quality.

---

<sup>2</sup> One could devise special matched visual and audible cue test signals, however typical program material such as a conversation among several people seated at random in a room will already provide good cues for the assessment of the degree of agreement between the position of each speaker in the room and the position from which his/her voice is perceived to come.



It should be noted that the audio-image might change in position depending on the position of the viewer-listener with respect to the loudspeakers and the screen. For the purpose of this Recommendation, it is assumed that one viewer-listener is positioned on the perpendicular to the centre of the picture, that the loudspeakers are positioned with respect to him as per Recommendation ITU-R BS.775, and that the picture is centred between the front right and front left loudspeakers. Additional viewer-listener positions should be chosen as per Recommendation ITU-R BS.1116.

To test the coherence of audio and video it is essential that the video being presented corresponds to the audio being tested.

Viewing condition	Setting(s)	
	Minimum	Maximum
Screen width	6 m	16 m
Viewing distance	1.5 H	2 H
Projector luminance (peak white at screen centre)	10 ftL	14 ftL
Screen luminance (projector off)		<1/1 000 of projector luminance

The loudspeakers required to present the multichannel audio component of the LSDI program should be integrated into the presentation environment. Their performance should desirably comply with Recommendation ITU-R BS.1116 that specifies the listening conditions for the subjective assessment of small impairments in audio systems including multichannel audio systems.

For instance, Recommendation ITU-R BS.1116 specifies that the reference (preferred) sound pressure level should be:

$$L_{ref} = 85 - 10 \log n \pm 0.25 \quad \text{dBA}$$

(IEC/A-weighted, slow) where  $n$  is the number of reproduction channels in the total set-up.

This sound pressure should be obtained by adjustment of the channel gain, using an input signal consisting of pink noise with an r.m.s. voltage equal to the “alignment signal level” (0 dB $\mu$ 0s according to Recommendation ITU-R BS.645, or 18 dB below the clipping level of a digital tape recording) fed in turn to the input of each reproduction channel (i.e. a power amplifier and its associated loudspeaker). For alternative loudspeaker arrangements as specified in § 3 Note 1, Note 5 and Note 6 it might be necessary to adjust the sound pressure level manually. To avoid level dependent bias of quality scores the level adjustment might be done in an additional blind-test at the ideal viewer-listener position.

The presentation conditions should be fully described in the test report and they should be kept constant during the test.