

## RECOMMENDATION ITU-R BS.1286\*

**Methods for the subjective assessment of audio systems  
with accompanying picture**

(1997)

The ITU Radiocommunication Assembly,

*considering*

- a) that Recommendations ITU-R BS.1284, ITU-R BS.1116, ITU-R BT.710 and ITU-R BT.500 have established methods for assessing the subjective quality of audio and video systems respectively;
- b) that Recommendation ITU-R BS.775 specifies one universal multichannel stereophonic sound system with and without accompanying pictures;
- c) that perceptual interaction between sound and picture can affect the assessment of sound quality;
- d) that existing methods for subjective assessment of sound quality are sometimes inadequate for sound systems with accompanying pictures;
- e) that the use of standardized methods is important for the comparison, exchange, compatibility and correct interpretation of data, derived from subjective tests carried out at different times and/or places,

*recommends*

**1** that the testing and evaluation procedures given in Annex 1 be used for the subjective assessment of sound systems with an accompanying picture.

## 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

---

\* Radiocommunication Study Group 6 made editorial amendments to this Recommendation in 2002 in accordance with Resolution ITU-R 44.

- 6 Programme material
- 7 Reproduction conditions
- 8 Statistical analysis and presentation of results.

Sound and picture are inseparably combined in TV and movie programmes. Under normal circumstances, the observers should perceive the sound and picture with a sense of unity. Therefore, the presentation of pictures may be indispensable for some subjective assessment of sound quality.

Visual stimuli often affect the perception of sound. For example, the apparent direction of a sound image is frequently shifted to that of the corresponding picture, which is well-known as the “ventriloquism effect”. Also, visual stimuli sometimes make aural impairments less noticeable.

The following fields of assessment require the presentation of a visual image:

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

## 2 Experimental design

In designing the tests, the considerations of Recommendation ITU-R BS.1116, § 2, should be taken into account. The use of a reference is not always necessary.

## 3 Selection of the listening panel

For the selection of the listening panel, the considerations of Recommendation ITU-R BS.1284 should be taken into account. In cases where the impairments are small, the considerations of Recommendation ITU-R BS.1116, § 3, should also be taken into account.

## 4 Test method

To conduct subjective assessments with an accompanying picture, it is necessary to select appropriate methods.

If the subjective differences are expected to be small it is appropriate to use the double blind triple stimulus with hidden reference method as described in Recommendation ITU-R BS.1116, § 4.

If the subjective differences are not small and it is appropriate to use less sensitive methods than those of Recommendation ITU-R BS.1116, the test methods of Recommendation ITU-R BS.1284 should be used.

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

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

It should be noted that the five-interval scale used for video assessment in Recommendation ITU-R BT.500 differs fundamentally from the five-point scales of Recommendations ITU-R BS.1116 and ITU-R BS.1284.

## 5 Attributes

The following attributes may be assessed:

- Front image quality: for definition see Recommendation ITU-R BS.1116, § 5.3.
- Impression of surround quality: for definition see Recommendation ITU-R BS.1116, § 5.3.
- Basic audio quality: for definition see Recommendation ITU-R BS.1284, § 5.1.
- Correlation between sound and picture images.

The last attribute may include the following characteristics:

- correlation of source positions derived from visual and audible cues (including azimuth, elevation and depth);
- correlation of spatial impressions between sound and picture;
- time relationship between audio and video.

## 6 Programme material

The test programme material should be selected to stimulate the attributes of interest. Different attributes may need different types of test programme.

In general, the aspects described in Recommendation ITU-R BS.1116, § 6, are also valid for sound with accompanying picture.

The time difference between sound and picture may be one of the parameters to be assessed. If it is not, then it should be made as small as possible. In some cases it may be necessary to correct the relative delays of the audio and video signals.

NOTE 1 – The permissible time difference between sound and picture is currently defined in Recommendation ITU-R BT.1359.

## 7 Reproduction conditions

### 7.1 Audio

Unless they are included in the attributes under test, the following parameters should conform to Recommendation ITU-R BS.1284 or Recommendation ITU-R BS.1116, § 7 and 8:

- Reference monitor loudspeakers or headphones
- Reference listening room
- Reference sound field conditions
- Reverberation time
- Operational room response

- Listening level
- Listening arrangements
- Listening position
- Loudspeaker base width
- Background noise level – an effort should be made to minimize the masking effect due to room noise and reproduction device noise. The total noise level, with technical equipment, should not exceed NR30.

## 7.2 Video

A number of ITU-R Recommendations exist for the relationships between picture size and viewing distance as well as loudspeaker base width and listening distance. In general, these are mutually incompatible. Therefore compromises must be made in either audio or video arrangements to permit both listening and viewing to be carried out simultaneously. Which Recommendations take priority will depend upon the objectives of the test. For the assessment of audio attributes, the considerations of Recommendation ITU-R BS.1116 should take priority.

The following considerations should be taken into account when selecting viewing conditions and the size of the video display device:

- Display size and viewing conditions for assessing different TV systems are given in a number of ITU-R Recommendations (see Table 1 of Appendix 1).
- The viewing distance should be similar to the listening distance, though ITU-R gives no recommendations about the combined arrangement of loudspeakers and display device.
- The recommended loudspeaker base width and listening distance for assessing audio systems are given in Recommendation ITU-R BS.1116 (see Table 2 of Appendix 1).

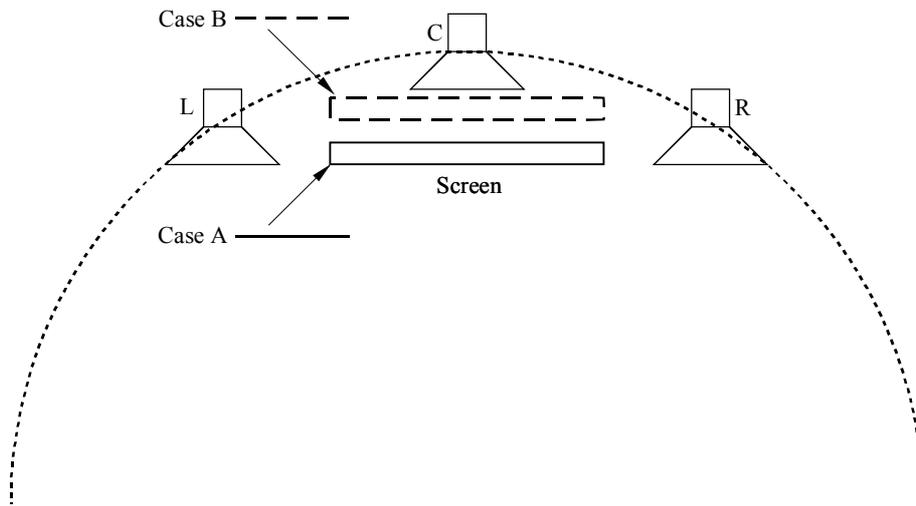
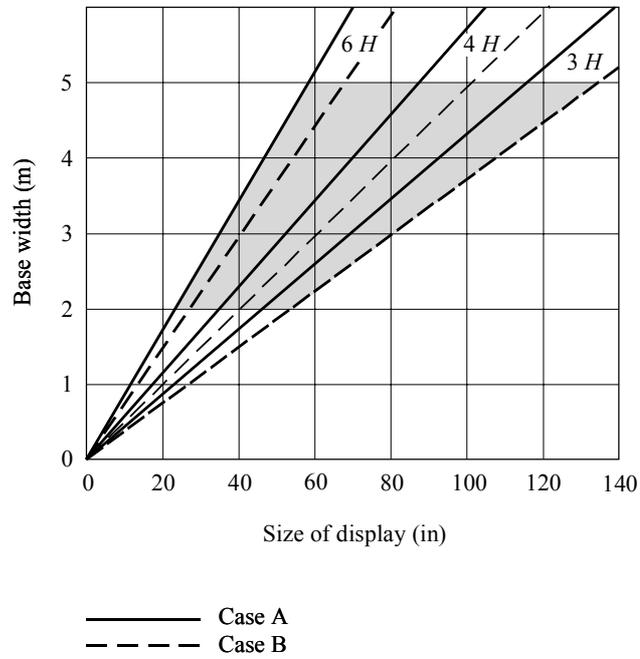
Figures 1 and 2 show the relationship between loudspeaker base widths and display sizes for two aspect ratios according to the considerations mentioned above. These figures indicate the problems which might occur if fixed viewing distances were to be recommended.

The choice of display size is limited; for example, with a base width of 3 m, an aspect ratio of 16:9 and a viewing distance of  $3H$ , the maximum permitted display size is 70 in.

The choice of viewing distance is limited. For displays with built-in loudspeakers the base width and display height are fixed. If the base width was 2 m and the display size 28 in, the viewing distance would necessarily be  $4H$  for a 4:3 aspect ratio display.

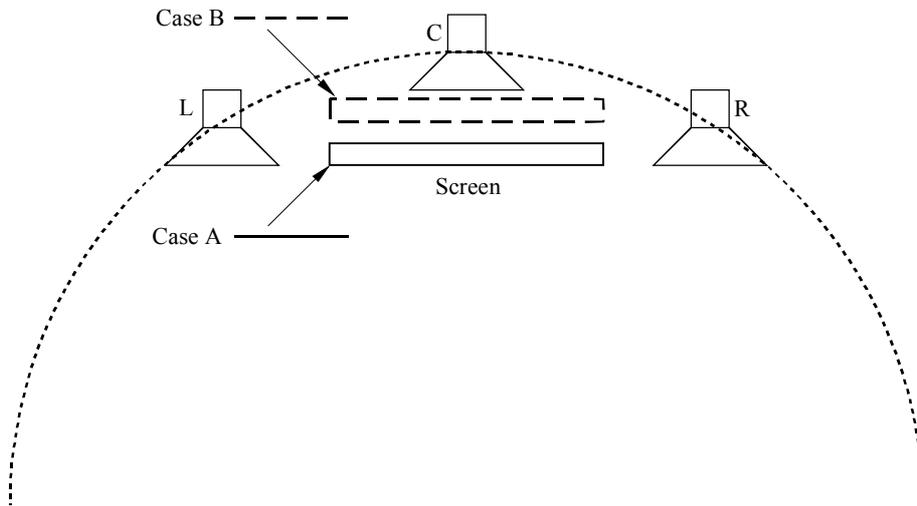
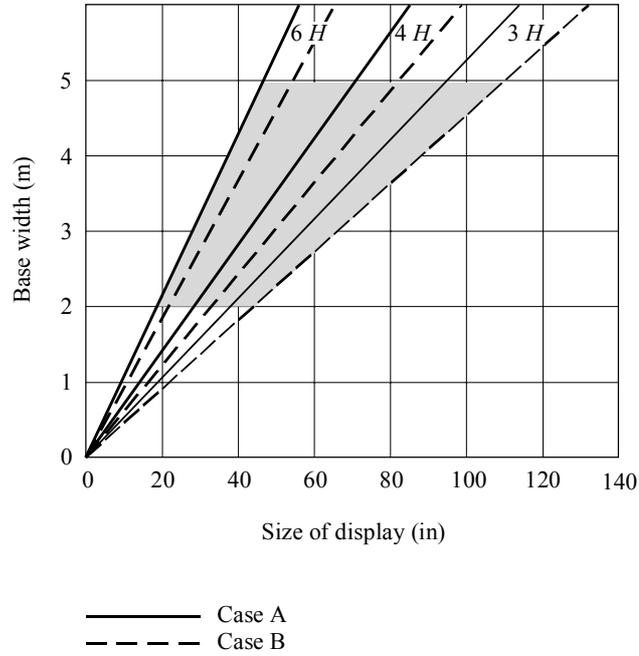
The viewing distance, therefore, cannot be rigidly fixed and can only be recommended with some degree of flexibility. The recommended ranges of viewing distance are  $3H$  to  $4H$  for high definition television (HDTV) and  $4H$  to  $6H$  for conventional TV systems.

FIGURE 1  
 Relationship between the loudspeaker base width and the video display screen size for a 16:9 aspect ratio



Base width: 2-5 m  
 Viewing distance: 3 H-6 H  
 Base angle: 60°

FIGURE 2  
 Relationship between the loudspeaker base width and the video display screen size for a 4:3 aspect ratio



Base width: 2-5 m  
 Viewing distance: 3 H-6 H  
 Base angle: 60°

## 8 Statistical analysis and presentation of results

The analysis of the experimental data and the method of presentation of the results should conform to Recommendation ITU-R BS.1284 or ITU-R BS.1116, § 9, 10 and 11 as appropriate.

### APPENDIX 1

#### TO ANNEX 1

TABLE 1

**Viewing distance and display size recommended in ITU-R Recommendations**

System type	Viewing distance and display size	ITU-R Recommendations
Conventional systems	Viewing distance: $6 H$ Display size: $\geq 22$ in	ITU-R BT.1128
Conventional systems under critical conditions	Viewing distance: $4 H$ for 625-line systems and $4 H$ or $5 H$ for 525-line systems	ITU-R BT.1128
Enhanced PAL and SECAM systems	Viewing distance: $4 H$ and $6 H$ for the enhanced picture, $4 H$ and $6 H$ for the compatible picture Display size (16:9): $\geq 28$ in (4:3): $\geq 22$ in	ITU-R BT.811
HDTV	Viewing distance: $3 H$ Display size: 55 in (If not available, $\geq 30$ in)	ITU-R BT.710
Digital television systems at or near the quality of conventional systems	Viewing distance: $4 H$ and $6 H$ Display size: $\geq 22$ in	ITU-R BT.1129
All systems	Maximum observation horizontal angle relative to the normal = $30^\circ$	ITU-R BT.500

TABLE 2

**Listening distance and base width recommended in Recommendation ITU-R BS.1116**

System type	Viewing distance and display size	ITU-R Recommendations
Monophonic	The minimum listening distance: 2 m All listening positions should be within an angle of $\pm 30^\circ$ from the loudspeaker axis.	ITU-R BS.1116
Two-channel stereophonic	Base width: $B = 2$ m to 3 m (Value of $B$ up to 4 m is acceptable in suitably designed rooms.) Limits of listening distance: $D = 2$ m to $1.7 B$ Reference listening distance = $B$ (Reference base angle = $60^\circ$ )	ITU-R BS.1116

The relationship between base width,  $B$ , and display size,  $P$ , is given in Table 3.

TABLE 3

Aspect/screen position	Case A	Case B
16:9	$P = 2.04 H$ $B = D = 2 n H \tan 30 = 0.566 n P = 0.0144 n P'$	$P = 2.04 H$ $B = D = n H = 0.490 n P = 0.0124 n P'$
4:3	$P = 1.67 H$ $B = D = 2 n H \tan 30 = 0.693 n P = 0.0176 n P'$	$P = 1.67 H$ $B = D = n H = 0.600 n P = 0.0152 n P'$

Picture height:  $H$  (m)  
 Display size:  $P$  (m),  $P'$  (in)  
 Viewing distance:  $n H$   
 Loudspeaker base width:  $B$  (m)  
 Listening distance:  $D$  ( $= B$  at the reference listening position).

---