AES20-1996

AES recommended practice for professional audio — Subjective evaluation of loudspeakers

Published by
Audio Engineering Society, Inc.

Abstract

This standard is a set of recommendations for the subjective evaluation of high-performance loudspeaker systems. It is believed that, for certain audio components, including loudspeakers, subjective evaluation is a necessary adjunct to objective measurements. The strong influence of listening conditions, program material, and individual evaluators is recognized. This document seeks, therefore, to assist in avoiding testing errors rather than to attempt to establish a correct procedure.

Contents

Foreword

[This foreword is not a part of *AES recommended practice for professional audio — Subjective evaluation of loudspeakers*, AES20-1996.]

The scope of this document identifies its application for “loudspeakers in domestic listening environments and in professional environments of similar acoustics.” However, there are several levels of interest in subjective assessments of these loudspeakers, and it is in this diversity that there is a lack of focus. At one extreme are the experimenters who probe the limits of human perception and measurement capability. These researchers constantly push the state of the art in experimental and statistical procedures. Their needs are so variable, however, that it is unlikely that any standard could be adequately embracing. Closer to the mainstream are some consumer-product testing groups who seek a standardized approach to follow and to point to in any challenge of their findings. At the opposite extreme are persons, including some product reviewers, who depend on their perceptual insights to transcend experimental controls, hearing imperfections, and room acoustics.

However, the vast majority of listening tests are motivated by immediate needs and are performed in spite of limited facilities and significant financial and time constraints. Intentions are good, but all too often the results contain biases or errors. Such errors may be serious when they cause a manufacturer to launch a new product, only to have it falter in the marketplace, or when biased opinions reach the public through a product review in a magazine.

In the evaluation of loudspeakers, the science of listening tests has made considerable progress. There are definite signs of order, and certain generalizations about loudspeaker performance seem to be safe, but listening rooms remain a significant factor in what is heard by listeners, and important aspects of sound quality and stereo imaging will, with certainty, relate only to the specific circumstances of the test. In such situations, experimental blinds and statistical analyses cannot substantially improve the utility or meaning of the results.

For these reasons, the working group decided to focus its efforts on the areas where the largest problems exist and the most substantial improvements can be anticipated. Mainly, these relate to the room, the program material, and the basic test procedure. The initial objective, therefore, has been not to develop a rigid standard, in the conventional sense, but to issue a set of procedures and guidelines by which persons conducting listening tests can assess the circumstances of their tests with a view to detecting sources of obvious bias or error.

The writing group consisted of Peter Aczel, Paul Barton, Søren Bech, Marshall D. Buck, David L. Clark, Laurence R. Fincham, M. Raymond Jason, D. B. Keele, Jr., Peter W. Mitchell, David Moulton, Thomas A. Nousaine, Sean E. Olive, Daniel Queen, and Floyd E. Toole.

David L. Clark, Chairman
Floyd E. Toole, Vice-chairman
AESSC WG-07 Working Group on Listening Tests
September 1994

AES20-1996

AES recommended practice for professional audio — Subjective evaluation of loudspeakers

1 Scope

The recommendations of this document apply to the subjective evaluation of sound reproduction of loudspeakers in domestic listening environments and in professional environments of similar acoustics. The recommendations apply most directly to user-installed free-standing and in-wall loudspeakers. Specifically excluded are

1) custom-built and equalized professional monitors;
2) loudspeakers used at very short listening distances;
3) loudspeakers for computer workstations.

For broadcast monitoring applications, attention is drawn to ITU and EBU standards documents.

1.1 Object

This document gives recommendations for test procedures, data acquisition and analysis, and interpretations of subjective evaluations of reproduced sound. Consideration is given to tests designed to reveal the presence of differences between devices under test, as well as tests intended to yield subjectively scaled ratings according to any of several possible criteria.

The recommendations include practical measurements and experimental procedures by which the subjective influences of certain physical, psychological, and experimental variables can be identified, isolated, and controlled. The objective is to minimize or control the biases and variations in listeners' judgments that are attributable to factors other than the devices under test.

1.2 Application

This standard applies to the evaluation of high-performance loudspeakers by loudspeaker manufacturers, professional users, product testing organizations, and consumers.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this document. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this document are encouraged to investigate the possibility of applying the most recent editions of the indicated standards.

IEC 268-3, *Sound system equipment — Part 3: Amplifiers.* Geneva, Switzerland: International Electrotechnical Commission, 1988.

IEC 268-5, *Sound system equipment — Part 5: Loudspeakers.* Geneva, Switzerland: International Electrotechnical Commission, 1989.

IEC 651, *Sound level meters.* Geneva, Switzerland: International Electrotechnical Commission, 1979.

IEC 804, *Integrating-averaging sound level meters.* Geneva, Switzerland: International Electrotechnical Commission, 1985.

ISO 1996-1, *Acoustics — Description and measurement of environmental noise — Part 1: Basic quantities and procedures.* Geneva, Switzerland: International Organization for Standardization, 1982.

ISO 2204, *Acoustics — Guide to International Standards on the measurement of airborne acoustical noise and the evaluation of its effects on human beings.* Geneva, Switzerland: International Organization for Standardization, 1979.

ISO 3382, *Acoustics — Measurement of reverberation time in auditoria.* Geneva, Switzerland: International Organization for Standardization, 1975.

ISO 7029, *Acoustics — Threshold of hearing by air conduction as a function of age and sex for otologically normal persons.* Geneva, Switzerland: International Organization for Standardization, 1983.

3 Terminology

3.1 acoustical recording: Recording made with microphones, of musical instruments that do not rely on electroacoustics for their sound emission.

3.2 anechoic recording: Recording made with microphones in an acoustical environment having negligible reflections.

3.3 foreground listening: Listening to program material in a situation where the program material is the center of attention.

3.4 front (of room): End of the room that the listener faces.

3.5 front (of stage): End of the stage closest to the listener.

3.6 impaired listener: Listener having a hearing threshold level below that regarded as normal per ISO 7029.

3.7 listening fatigue: Subjective sensation of annoyance and tiring that develops gradually after a long period of continuous listening (see annex C), not to be confused with auditory fatigue, an objective hearing-threshold shift.

3.8 robustness: Stability of performance with normal listener movements and listening locations (see annex C).

4 Room requirements

4.1 Size and shape

The listening room shall accommodate a reasonably proportioned rectangular working area of not less than 20 m^2. A larger size and non-rectangular walls shall be permissible provided the room remains acoustically representative of the space for which the loudspeakers are intended during use. A larger area shall be used for more than three listeners. The minimum ceiling height shall be 2.1 m.

4.2 Acoustics

The listening room acoustics shall be similar to those of a normally furnished domestic listening room. Within this range, any unusual and distinctive acoustic characteristics should be suppressed. Medium-size control rooms usually strive to mimic the domestic acoustical environment.

4.2.1 Reflections and reverberation

4.2.1.1 At midrange frequencies, a reflected sound field shall decay 60 dB in 0.45 ± 0.15 s (see ISO 3382). Both a higher decay time at low frequencies and a decrease at high frequencies are permissible. These decays are based on measurements of a large number of control rooms and domestic listening rooms having volumes ranging from 50 to 120 m^3. Smaller and larger volume rooms may have somewhat shorter and longer reverberation times, respectively. Allowances should be made for the prevailing acoustic conditions of the spaces for which the loudspeakers are intended.

4.2.1.2 For test purposes, an adequately random sound field may be generated in a small room by using a multiplicity of loudspeakers aimed away from the measurement microphone. Upon switching off the signal, a decay of average sound pressure level (SPL) will be observed. After a possible rapid drop in the first 10 ms, an average constant slope should be observed down to near the background noise level. The first 10 to 20 ms after the initial drop may be used to determine the decay slope. A slope that decreases significantly when only 10 to 20 dB down may indicate an acoustical problem within the room or excessive acoustic coupling to another room. Either condition shall be corrected.

4.2.1.3 Satisfying these goals should be possible with ordinary domestic furnishings. The room should be neither too full of padded furniture nor too sparse. Specifically, extensive use of commercial sound absorbers or diffusers is not recommended.

4.2.1.4 The floor shall be carpeted or at least 75% covered with area rugs. For domestic environments, the coverage shall include the area between the loudspeakers and the listeners. The ceiling shall be acoustically reflective. For rooms using suspended ceilings, absorptive tiles or panels in the central area that are in a position to reflect sound from the loudspeakers shall be replaced with wallboard or other reflective material.

4.2.2 Room modes

Efforts shall be made to damp and stagger low-frequency room modes. At a minimum, the evaluator must know the effects of modes for the loudspeaker-listener locations used. A recommended method is to play constant-energy-per-percentage-bandwidth (pink) noise through loudspeakers of known frequency response located where the loudspeakers to be evaluated will be placed. The microphone of a third-octave real-time analyzer should be moved to various possible listening locations while observing the low-frequency response. This test, if performed, shall include the equivalent of left-equals-right signal feed in the low-frequency range because it is typical of most commercial stereo program material. The effect of room modes should be noted and taken into account.

4.2.3 Flutter echoes

Audible flutter echoes excited by the loudspeakers shall be suppressed by diffusion or absorption.

4.2.4 Background noise

The room shall have an A-weighted ambient noise level lower than 35 dB, measured per ISO 1996-1. In addition, the C-weighted noise level shall be less than 50 dB. The slow response of a type 1 or type 0 sound level meter complying with IEC 651 or with IEC 804 shall be used. In addition, no periodic or tonal sounds may dominate the noise (see ISO 2204).

5 Listening arrangement

5.1 Listening distance and loudspeaker separation

5.1.1 In order to explore potential listener locations, a listening area, rather than a few fixed listening locations, shall be available. This area shall include the stereo centerline and shall be measured from the center of the front panels of the left and right loudspeakers. The width of the listening area shall allow listening off the centerline a distance equal to the distance of either loudspeaker from the centerline. The depth of the listening area shall extend from less than 1 m to more than 3 m from the front panels.

5.1.2 The angular separation of the loudspeakers for a listener on the centerline anywhere within the listening area shall be not less than 40˚. Ideally, 60˚ should be available, particularly for evaluating loudspeakers intended to be used as control-room monitors. It shall be possible to move the loudspeakers closer together to achieve less separation, when desired and so reported.

> NOTE — For multichannel playback the same recommendations apply to the left and right loudspeakers, except in dedicated film surround-sound situations, where the left and right loudspeakers may be located for best effect with regard to the visual image, rather than to meet an angular requirement. Whatever is done, the choice shall be justified and reported.

5.1.3 Achieving the required listening distance shall not violate the standards for loudspeaker and listener proximity to room boundaries given in 5.2 to 5.5. Listening distance and loudspeaker separation availability requirements also apply to the outer loudspeakers of front sound-stage multichannel systems.

5.1.4 Manufacturers' recommendations shall be relied upon for distance and separation requirements for the evaluation of side and surround loudspeakers.

5.2 Loudspeaker locations

5.2.1 When available, the manufacturer's recommendations shall be accommodated as a starting point and used unless an improvement can be obtained. Such improvements shall follow the guidelines in annex A.

5.2.2 Locations that place the center of the loudspeaker front panel at least 1 m away from reflecting surfaces shall be available. This distance is recommended if manufacturers do not recommend to the contrary. In a small room this distance may not be achievable without some compromises. The compromises are

a) to move the loudspeakers away from the walls, thus reducing the width of the stereo sound stage, or
b) to move the loudspeakers closer to the walls, thus affecting the timbre.

Neither compromise may be a perfectly satisfactory solution, but a small room setup is so constrained. The compromise chosen shall be reported.

5.3 Early reflections

Means shall be available to optionally absorb or retain early reflections from surfaces in the vicinity of the loudspeakers.

5.4 Listening locations

Listeners shall be free to move within the listening area described in 5.1.1. The choice of locations within the area should be determined, according to the guidelines in annex A, by the purpose of the test.

5.5 Loudspeakers requiring special mounting

5.5.1 Large control-room monitors intended for mounting within the structure of the room and domestic loudspeakers intended to be recessed into a wall (in-wall loudspeakers) are examples of loudspeakers requiring special mounting.

5.5.2 Manufacturers' recommendations for such loudspeakers must be followed with regard to acoustic intent. For example, mounting an in-wall loudspeaker in a 1.2-m-wide sample of drywall and stud wall construction placed just in front of an existing listening room wall may meet the acoustic intent. Similarly, this mounting can be simulated by filling the space around the loudspeakers with rigid material. Details of compliance with manufacturers' recommendations shall be reported.

5.5.3 Limited-range loudspeakers designed to be augmented by associated loudspeakers shall be used in a system with appropriate filtering and appropriately placed associated loudspeakers.

> NOTE — An example of a limited-range loudspeaker is a center loudspeaker intended to be used with a high-pass filter with the bass directed to flanking loudspeakers, as in certain systems for reproducing film sound. Another example is represented by satellite loudspeakers intended to be used with dedicated or optional woofers or subwoofers. The satellites should be located for optimum imaging. Dedicated subwoofers shall be used with the crossover provisions provided by the manufacturer and assembled as instructed. Optional subwoofers shall be used with caution since many do not have high-pass filtering provisions for the satellite loudspeakers. The choice of low-pass filter characteristic will modify the perceived performance of the satellite loudspeakers. The evaluator is in the position of a system designer and should exhibit candor in reporting the experience.

6 Program material

6.1 Media

Media providing recorded digital audio are recommended as program sources because they maintain their characteristics throughout their service life and may be copied precisely for renewal.

6.2 Core selections

6.2.1 A core set of acoustically recorded selections should be kept unmodified and should always be used principally to verify natural reproduction of timbre-related and spatial aspects of music reproduction. The acoustically recorded selections may be augmented with electronic instrument selections for testing the special requirements of contemporary music. New selections may be used in addition to the core set. If a selection in the core set needs change, it should be phased out over a period during which both it and its replacement are used. Listeners must be familiar with the sound of every selection used.

6.2.2 The core selections shall include acoustically recorded music selections of recognizable instruments and an anechoic recording of one or more examples of the male speaking voice. Solo male and female singing voices as well as choirs are also recommended. These selections alone shall contain enough spectral, temporal, spatial, and dynamic material to completely challenge and test the loudspeaker. Suitable monophonic versions of all acoustically recorded selections shall be available. Monophonic sound can usually be achieved by playing only one recorded channel. Mixing two or more recorded channels is not recommended.

> NOTE — Locating a suitable recording of a male speaking voice may be difficult. Many sources exhibit spectral imbalances and colorations of various kinds related to the microphone, its orientation, and reflections from tables or script holders. Fortunately, a suitable recording can easily be made of a familiar voice by recording at a 0.5-m distance in front of a talker standing outdoors well away from reflecting surfaces other than the ground. An omnidirectional microphone with flat response over the range of 100 Hz to 12 kHz should be used.

6.2.3 A live microphone feed from an acoustic source in another room should be included as a test signal. Accuracy of the loudspeaker can be assessed by comparing the sound of the feed to the acoustic source.

NOTE — The performance of the musician, the microphone, and its location with respect to the live source should be regarded as additional variables that will be introduced into an already complicated task.

6.3 Technical signals

Use of technical signals is optional. They are useful for finding potential problems. Pink noise and sine sweep technical signals are recommended for quickly diagnosing a polarity reversal of one loudspeaker, buzzing or non-functioning drivers, and other improper conditions (see IEC 268-5). Any non-correctable defect found with technical signals must be verified as audible while reproducing the core selections.

NOTE — Faults audible with technical signals will very likely be less audible in music because of the masking effects of the music itself. This uncertainty creates a dilemma for the evaluator, because somewhere, sometime, there may be a musical selection having the temporal and spectral features required to reveal the defect even though normally the fault is inaudible.

6.4 Source of selections

6.4.1 Commercial recordings or the equivalent are suggested. Recordings called “technically correct” are not definable for two-channel recordings designed for loudspeaker playback in rooms having somewhat reverberant surfaces. Subjective judgment is employed to optimize spectral and spatial aspects even in purist recordings by means of microphone selection and placement.

6.4.2 Poor recordings can be used to distinguish between loudspeakers. For example, an unintended electronic transient in a recording may reveal spectral problems in a loudspeaker. Such a flawed recording shall be treated as a technical signal. The decision as to whether a useful recording is to be treated as challenging music or a technical signal should be based on the artistic intent of the sound. Was it intended or was it a mistake?

6.5 Listening sequence

The use of technical signals prior to evaluation will expedite the diagnosis and correction of conditions such as malfunctions, reversed polarity, and buzzing. Male speech and other acoustical program material should be used next to judge timbre-related and spatial qualities. Rock and synthesized music can then be used to test dynamic range and frequency extension qualities. Listening fatigue and robustness may be tested with a variety of musical program material.

7 Test procedure

7.1 Associated equipment

Associated equipment that could affect the audio signal shall have technical performance that is clearly superior to that of the loudspeaker and shall be operated within its limits of linearity. No audible electromagnetic interference (EMI) shall be permitted. Any unusual characteristics that could affect the performance of the loudspeaker, such as an amplifier that presents the loudspeaker with a high source impedance (a low damping factor or high regulation per IEC 268-3), shall be reported.

7.2 Playback level

For foreground listening, the playback level should be set subjectively to match the loudness of the original event for each selection. It should be set lower if the sound is distorted or otherwise annoying. A full orchestra can require 95 dB A-weighted. Some contemporary popular music should be played as loudly as tolerable while testing for an upper limit of loudspeaker capability.

7.3 Loudness matching

Each loudspeaker in an *n*-channel loudspeaker system shall be matched in SPL to that of its counterpart in each multiloudspeaker system to which it is compared. A recommended method is to apply pink noise to each loudspeaker in turn and adjust the A-weighted level of each for equality within ± 0.5 dB. Whatever technical method is used, the matching shall be confirmed subjectively.

7.4 Anchor loudspeakers

Loudspeakers that have been evaluated previously should be inserted into the procedure. A listener's use of rating scales varies with the performance range of the particular group of loudspeakers being evaluated. Anchor loudspeakers can provide reference points to anchor the rating scale. A second use of anchor loudspeakers, important in single-system evaluations, is to prevent the listener from unconsciously tuning out aberrations in the loudspeaker system being evaluated. After listening to the anchor, a switch back often brings the aberrations back to the listener's attention.

7.5 Single-loudspeaker listening

Loudspeakers shall be evaluated as single units in their normal room location for timbre-related accuracy as well as in stereo pairs, as surround sound components, or for other use. Timbre-related accuracy is much more easily heard in single-loudspeaker listening. This procedure is necessary because there are times in almost any program format when a loudspeaker must perform on its own.

7.6 Listening locations

The listener should explore the effect of listening off the axis of a single loudspeaker. The listener should also explore the effect of listening both on and off the centerline of a stereo pair or multiloudspeaker system. When more than one listener is being used, they should exchange places periodically. When a single system — monophonic, stereophonic, or multichannel — is being evaluated for a product review, the reviewer shall explore a range of listener and system room layouts to determine the best possible configuration for performance as well as to identify any unusual sensitivities to location.

7.7 Listeners

7.7.1 Trained, experienced listeners shall be preferred. The best way to determine a listener's training level is to evaluate self-consistency in blind listening tests. That is, does a listener always rank the same system in the same way in different blind tests? They need not give the same ranking as other listeners. Indeed, this would not be expected of an exceptional listener.

7.7.2 Listeners with impaired hearing may perceive differently than do non-impaired listeners. Hearing-impaired persons need not necessarily be excluded, but their test results should be analyzed separately to see whether it is reasonable to pool their data with the data from normal-hearing listeners.

7.8 Critical listening

7.8.1 Judgments are best made by including a critical attitude. That is, the listener should attempt to test a loudspeaker to uncover its faults rather than to dwell on its good points. A full range of critical checks should be made. A reminder list or check sheet is helpful for this. A loudspeaker that performs well in all critical checks is likely to be perceived as excellent in overall fidelity. One that has a single strong shortcoming rarely is.

7.8.2 Listeners should take breaks as needed to maintain the concentration required for critical listening.

7.9 Critical checklist (see annex C)

a) Spectral uniformity
— Frequency extension
— Uniformity of response

b) Sound-stage imaging
— Width
— Depth
— Image specificity
— Image stability

c) Ambience reproduction
— Appropriate amount
— Even distribution around listener
— Spaciousness
— Spectral character

d) Dynamics and distortion
— Loudness capability, frequency dependence
— Transient capability
— Not modulated or compressed
— No timbre change when loud
— Pianissimo clarity
— Problems, buzzes, failures

e) Listening fatigue

f) Robustness
— Effect of listening off centerline
— Effect of standing
— Effect of head rotation

7.10 Blind testing

7.10.1 A listening test shall be blind, preferably double-blind. Blind testing is strongly recommended whenever possible (see Bech, annex D).

7.10.2 Non-blind evaluation is deprecated because it is subject to non-sonic biases which may exceed the real differences between the loudspeakers. An evaluation made with the loudspeakers visible and with the knowledge of which unit under test is playing is a non-blind evaluation and shall be reported as such.

7.10.3 To make the test blind, the loudspeakers shall be hidden from view with a screen. The screen shall be visually opaque and acoustically transparent, two conflicting requirements. A good loudspeaker grille cloth can be used, such as polyester double-knit fabric. Its acoustical transparency should be measured as follows:

1) measure the frequency response of a loudspeaker at a distance of approximately 1 m;

2) without touching anything else in the setup, insert the screen between loudspeaker and microphone and measure again.

The amplitude of the second measurement should exhibit, at most, a smooth rolloff of high frequencies to -1 dB at 10 kHz. Folds and pleats in the screen should be avoided, and any framework associated with its support should not interfere with sounds radiated from the loudspeakers under test. Lights facing down, up, or from the sides can be used to render the screen more opaque. Dark paint or fabric behind the loudspeakers can also help disguise their outlines.

7.11 Testing several loudspeakers at once

Whenever several different loudspeakers are tested together in a single setup, one or more of the loudspeakers will not be positioned optimally. The positioning particularly affects the low-frequency performance of the loudspeaker. Such a setup shall be varied by use of a turntable or a statistical shuffling of the positions of the different loudspeakers. Each loudspeaker in the setup shall be tested at least once in every position.

7.12 Reporting of results

7.12.1 Basic report. A basic report shall include the following items:

a) summary of results;
b) summary of method (blind or non-blind, multi- or single-loudspeaker system);
c) list of non-compliances with AES20;
d) listening room dimensions;
e) loudspeaker and listener locations;
f) description of source selections;
g) listening panel makeup (number, age, sex, training, hearing acuity).

7.12.2 Full report. A full report shall consist of the basic report plus the following items:

a) all listener test data;
b) detail of method, including the statistical method used;
c) room floor plan showing loudspeaker and listener locations;
d) reverberation time at 125 Hz, 500 Hz, 2 kHz, and 8 kHz;
e) indication of effect of low-frequency room modes;
f) temperature, barometric pressure, and humidity;
g) music selections used;
h) means of shuffling loudspeaker locations;
j) identification of all playback components;
k) instructions to listeners;
l) interpretation of results and conclusion.

Annex A
(Normative)

Guidelines for choosing loudspeaker and listener locations

A.1 Loudspeaker locations

A.1.1 The objective is to obtain the best possible performance from the loudspeakers and to report any unusual difficulty in achieving this. The interrelated effects of the locations of both loudspeaker and listener are so strong that they will be a major determinant of performance. When multiple loudspeakers are tested simultaneously, at best, only one set is likely to be in its optimum location. In some cases, decisions about position can determine the outcome of the test. Faced with this responsibility, the evaluator shall exercise care to ensure that the tests are fair.

A.1.2 The effects of positioning may be summarized as follows.

A.1.2.1 Excitation and reception of standing waves or room modes. This effect is a low-frequency sound-pressure variation and transient condition. Room modes will exist in any closed space. To some extent they can be controlled by damping, which reduces the peak-to-trough amplitude ratio and spreads the effect over a wider range of frequencies. Mode frequencies can be staggered rather than made coincident by adjustment of the basic room dimensions. The number of modes below a given frequency can be increased by using a larger listening room. The lowest few modes may be excited and received in a complementary manner by exact placement of loudspeaker and listener. All of these techniques should be employed to partially equalize the effects of room modes for different loudspeakers and listeners.

A.1.2.2 Acoustic load on a loudspeaker due to any proximity of loudspeaker to room boundaries. This effect is a low-frequency sound-pressure variation phenomenon. Most loudspeakers radiate an acoustic-power output roughly proportional to the acoustic load into which they radiate. Monopole radiators exhibit the greatest undesirable variation when spaced equally from two or more boundaries (such as walls, the floor, and the ceiling). This effect should be avoided.

A.1.2.3 Early reflections of sufficient amplitude arriving at the listener in the range of 3.0 ms and less. This effect will cause midrange coloration or other timbre change. In addition, early lateral reflections will cause some degree of image shift or ambiguity and alteration of perceived spaciousness. These reflections may be caused by either loudspeaker or listener proximity to reflecting surfaces. Early reflections, unless intended as part of the loudspeaker's design, shall be controlled so as not to degrade the sound. This control may be accomplished through placement and other acoustical means (see clause 5). It is desirable to retain early ceiling reflections for domestic loudspeaker evaluation because this is often an unavoidable condition of use.

A.2 Listener locations

A.2.1 Listeners shall be free to move within the listening area described in 5.1.1. The choice of locations within the area should be determined by the purpose of the test. For example, some loudspeakers intended for monitoring in control rooms should be evaluated at the appropriate short distance, whereas those intended for leisure listening should be auditioned at a distance of 2 m or more. Assessments should include listeners positioned on the symmetrical axis as well as listeners at positions well off axis.

A.2.2 The most distant listening positions should be at least 1 m from the rear wall. Near the minimum distance from the wall, distortions in imaging or timbre may be noticed. Diffusion, absorption, or redirection available using normal furnishings such as bookcases may not suffice entirely. This distortion can be tested by applying pink noise identically to a pair of loudspeakers and listening to the quality of the virtual image. For a stereo system, this pair should be the left and right loudspeakers. For a multiloudspeaker system, this pair should be the two loudspeakers that might be required to produce a virtual image. A listener equidistant from the two loudspeakers would hear a compact virtual image floating midway between them. This would be true for all equidistant locations from front to rear of the listening area. If the image is perceived to spread, split, or change in timbre as the listener moves to the rear, the evaluator should experiment with some effective absorbent on the rear wall. Acoustically diffusing elements can also be used in this area, but caution is necessary because such elements may be observed to spread the image if the diffusers are too close behind the listeners or cover too large an area behind the listener.

A.3 Loudspeaker measurements

Relevant measurements of the loudspeakers, such as frequency response and directional characteristics, combined with a knowledge of the listening room's characteristics are valuable to optimize the test. Unfortunately data regarding these characteristics may not be available. In any case, the evaluation must deal with the issue of position of listeners and loudspeakers by employing several loudspeaker locations and several listener locations, repeating the tests many times, and analyzing the data. The data can be combined in a global mean to statistically diminish the effect of individual positions, or the individual positional combinations can be examined to determine the most favorable circumstances for each product. In any event such measurement is not a trivial undertaking.

Annex B
(Informative)

Listening arrangement

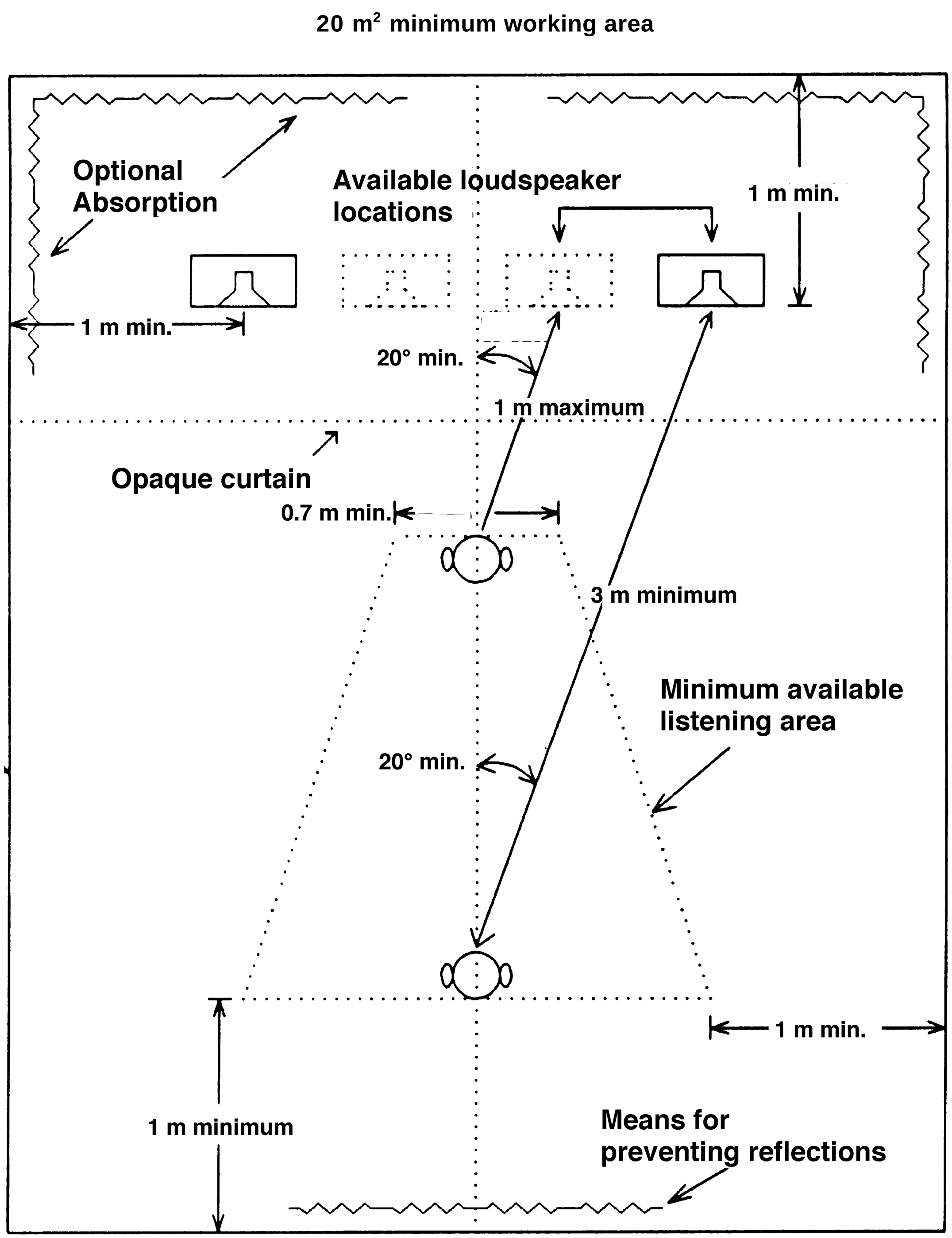

Figure B.1 — Floor plan of listening arrangement

Annex C
(Informative)

Loudspeaker-system evaluation form

C.1 Terminology

C.1.1 spectral uniformity: Desirable loudness balance of the audible range of pitches or frequencies (16 Hz to 20 kHz).

C.1.1.1 coloration: Identifiable tonal quality given to the sound by the reproducing system, for example, a metallic coloration. This quality is undesirable because it distorts the timbre of instruments. Generally, coloration affects narrow ranges of the spectrum (less than a third octave). Coloration can also result from an acoustic condition that adds an identifiable character to the spectrum. Other examples are honky or boxy colorations arising from cavities or reflecting surfaces near the loudspeaker or listener. Coloration modifies the spectrum of the reproduced sound but does not add new sounds.

C.1.1.2 octave balance: Spectral loudness variations covering about one octave (a 2 to 1 frequency ratio). Voices may lack body if the octave containing the fundamentals, around 160 Hz, is suppressed. Clarity can be suppressed or exaggerated to a screechy sound by variations in the octave around 4 kHz. Likewise, perceived effects such as sibilance (10 kHz), boomy bass (100 Hz), and honky (630 Hz) sounds are controlled by octave-wide spectral variations.

C.1.1.3 balance, bass to treble: Emphasis of one half or the other of the spectrum to produce a heavy or thin sound.

C.1.1.4 frequency extension: Quality of reproducing extremely high or low pitches with adequate loudness. The question is, does the response extend far enough? Too much or too little response is better addressed as an octave-balance problem.

C.1.2 sound-stage imaging: Ability to reproduce the sense of directions and distances of the original sound sources. The sound stage is the area containing all of the performers. For reproduced sound, this is the area containing all of the sound images.

C.1.2.1 image: Perceived location of a reproduced sound source. The image may be sharp or broad. A properly reproduced singer may have a sharply defined image, whereas the first violins in an orchestra may be expected to have a broad image.

C.1.2.2 stage: Reproduced sound stage, that is, area perceived to contain all the images. It usually has side-to-side and depth dimensions. It may also have a perceived height direction or dimension.

C.1.2.3 front–back stage location: Perception of the stage as behind, all around, or in front of the listener. Front stage location is the live-performance standard and is preferred.

C.1.2.4 up–down stage location: Perception of an unnaturally high or low stage location. If there is no problem, this perception is rated as neutral.

C.1.2.5 left–right stage evenness: Perception of continuous sound (appropriate source material such as a choir is assumed) from the left to the right extremes of the stage.

C.1.2.6 width of stage: Perceived angular width of extreme left to extreme right edges of the stage. This width should be between 30° and 90°. Symmetry with respect to the seating position is not critical as long as it is forward.

C.1.3 localization: Determination by a subject of the apparent direction or distance, or both, of a sound source.

C.1.3.1 left–center–right image localization: Localization of the image at the intended location. In general, left- and right-intended locations are localized properly. Center-intended sounds are a problem for off-centerline listening with stereo. For many stereo systems, the important question is, is the soloist at the center of the stage?

C.1.3.2 depth localization: Ability to hear that performers are appropriately localized from the front to the rear of the sound stage.

C.1.3.3 image separation: Ability to hear images as distinct from one another. Clarity and definition are words used to describe this perception. Black space between instruments is a visual analog that may be helpful. Listen for ensemble inner voices, that is, individual instruments or voices heard while they are playing as part of a group.

C.1.3.4 open or transparent: Quality of an image being clearly localized but having no perceptible relationship to the location of the loudspeakers. This quality is related to left–center–right and depth localization and to spaciousness. It is possible, however, to hear an open, transparent sound from a single loudspeaker with little or no recorded reverberation.

C.1.3.5 split stage or split image: Localization of one instrument in more than one place within the sound stage. High and low frequencies may come from different places. Also, the sound stage may be split left to right or front to back. Split stage is a worst-case condition of poor left–right stage evenness.

C.1.3.6 stage or image stability: Proper localization of the stage or images as stationary when so intended by the recording. Localization of an image may change with changes in pitch, loudness, or timbre. Localization of stage or image, or both, may change with the rotation of the listener's head or movement within the normal seating area. A system could rate highly in all other sound-stage imaging categories but be downgraded because movement of the listener's head destroys the illusion.

C.1.4 ambience reproduction: Ability to convey the acoustically identifiable aspects of the recorded space, whether large or small as well as dead or reverberant.

C.1.4.1 spaciousness: Perceived quality of listening within (not just to) a reverberant environment. The sound is perceived as open, not constrained to the locations of the loudspeakers. The perception is an important part of the “you are there” sensation. Spaciousness is natural compared to phasiness.

C.1.4.2 diffuseness: Property, associated with spaciousness, of perceiving the reverberant sound as coming from all directions.

C.1.4.3 phasiness: Property consisting of confusing and inconsistent directional effects achieved by electronic phase shifting. With phasiness, image shift and spectral changes may result from small listener head movements. It is often generated electrically by shifting the phase of signals that are identical in each channel.

C.1.4.4 direct-to-reverberant ratio rendition (direct–reverb rendition): Ability to make audible the ratio of direct (first arriving sound) to reverberant sound. It is particularly difficult to reproduce accurately a weak reverberation combined with a strong direct sound. This condition exists for an instrument intended to be localized at the front of the sound stage. It should contrast with the lower direct-to-reverberant ratio for instruments further back.

C.1.4.5 size-of-space rendition: Ability to perceive audible differences in the size of recording environments. A small room may be very reverberant, but will still sound small. The sound system should be able to convey this information. This is another important part of the “you are there” sensation. Excessive local acoustics, the acoustics of the listening environment as stimulated by the sound reproduction system, can inhibit size rendition. An excessively strong local acoustical character masks the recorded ambience.

C.1.4.6 spectral uniformity (of reverberation): Perception of appropriate modification of the spectrum of the diffuse reverberation. A common problem is a bass-heavy reverberant sound.

C.1.5 dynamics: Rendition of loudness changes.

C.1.5.1 distortion: Effect on the signal that produces new sounds or timbre change. It is not to be confused with noise that is independent of the signal. Distortion is usually a problem only at high loudness.

C.1.5.2 maximum loudness: Perceived limit on loudspeaker output. Almost without limit, the louder a sound system can play, the better it will be able to cope with percussive and synthesized music.

C.1.5.3 transient impact or punch: Rendition of loud, quickly changing sounds such as percussion. Poor rendition is heard as a softening rather than a distortion of time.

C.1.5.4 compression: Reduction of dynamic extremes without distortion. Electronic signal compression is used intentionally in the amplifiers of some self-powered loudspeakers to prevent clipping. Its operation should be so slight or well managed that it is inaudible as an effect.

C.1.5.5 modulation: Audible effect of compression in which midrange frequency instruments seem to drop in loudness when bass notes are sounded. This can be due to electronic compression, clipping, or overdriven loudspeakers, in which last case it is heard as a “blurring” of the midrange sound.

C.1.5.6 pianissimo clarity: Ability to retain the appropriate timbre during the soft passages without adjusting the volume control from the loud passage setting.

C.1.5.7 strain and glare: Sense of strain or annoyance at maximum loudness. Loud sounds should not become “flattened” or “glaring.” This sense is a rating of loudness purity.

C.1.5.8 timbre change for loud low-frequency sounds: Sharpening or “hardening” in timbre as volume is increased. Low-frequency transient and sustained sounds may sharpen in this undesirable manner. At worst, a kick drum “thud” can turn into a “blat.”

C.1.6 listening fatigue: Subjective sensation of annoyance and tiring that develops gradually after a long period of continuous listening. You want to turn it off. This effect may be due to spectral peaks, unstable imaging, or a number of other problems, alone or in combination. Listening fatigue is to be distinguished from immediate dislike of a sound system on first hearing. Listening fatigue due to the loudspeaker must also be differentiated from listening fatigue caused by the source material. A good loudspeaker system playing good recordings at a natural loudness will not cause fatigue for many hours, or ever.

C.1.6.1 time without fatigue: Period from outset of testing to onset of listening fatigue. Fifteen minutes to the beginning of listening fatigue is poor performance. Several hours or more without fatigue is excellent.

C.1.6.2 selection-dependent fatigue: Listener fatigue occurring only on a certain selection. The system should rate highly if most selections are free of fatigue after several hours.

C.1.7 robustness: Stability of performance with normal listener movements and listening locations. Spectral uniformity, sound-stage imaging, and ambience in particular may be adversely affected.

C.1.7.1 good off centerline: Imaging performance not dependent on sitting exactly (± 0.5 m) on the stereo centerline. This and other aspects should remain good within the listening area.

C.1.7.2 good seated or standing: Listening height not too critical. Only a slight diminishing of high frequencies should be heard at a standing position.

C.1.7.3 good with head rotation: Imaging and ambience performance not diminished by a 90° rotation of the listener's head.

C.2 Recommended form for listeners' evaluations

The form shown in table C.1 is suggested for listeners' evaluations of loudspeakers during listening tests. Point 5 on the scale represents the highest performance and 1 the worst. It is suggested that listeners be instructed to use a resolution no smaller than one-half point in their final judgment in each performance category. This resolution results in a nine-point rating scale. Listeners should be asked to consider each of the issues within a performance category, paying particular attention to any poor performance that would pull the category rating down strongly. Four music types are suggested on the form. A new form can be used to rate the loudspeakers in each of the music types. From these data, performance and application scores may be obtained separately or combined into an average. Consideration of unequal weightings of importance is suggested for different performance and different music categories. Experience with this form has confirmed that reasonably quick ratings can be achieved by experienced listeners provided that challenging, short (30 s to 60 s) music segments are used.

Table C.1 — AES20 loudspeaker evaluation form

LOUDSPEAKER SYSTEM EVALUATION FORM

Music Category (Circle one): Acoustical (unamplified) music
Pop, rock, jazz
Easy listening
Sonic spectacular

Listener name______ _____ _____ _____ Date _ _____ _____ _____ _____ _____

Loudspeaker system_ _____ _____ _____ Test number _ _____ _____ _____ _____

Test location Seat location

1 Spectral uniformity	**1**	**2**	**3**	**4**	**5**	
1.1 Uncolored						
1.2 Octave balance, clarity						
1.3 Balance, bass to treble						
1.4 High-frequency extension						
1.5 Low-frequency extension						
2 Sound-stage imaging	**1**	**2**	**3**	**4**	**5**	
2.1 Stage location front–back, up–down						
2.2 Stage evenness left–right						
2.3 Stage width						
2.4 Image left–center–right localization						
2.5 Image depth localization						
2.6 Separation of images						
2.7 Open, transparent						
2.8 Split stage or image						
2.9 Stability of stage or image						
3 Ambience reproduction	**1**	**2**	**3**	**4**	**5**	
3.1 Spacious, diffuse						
3.2 Direct–reverb rendition						
3.3 Size-of-space rendition						
3.4 Spectral uniformity						
4 Dynamics and distortion	**1**	**2**	**3**	**4**	**5**	
4.1 Maximum loudness						
4.2 Transient impact or punch						
4.3 Not modulated, compressed						
4.4 Pianissimo clarity						
4.5 Loud, no strain or glare						
4.6 Loud, low frequencies, no timbre change						
5 Listening fatigue	**1**	**2**	**3**	**4**	**5**	
5.1 Long time without fatigue						
5.2 Fatigue only on some selections						
6 Robustness	**1**	**2**	**3**	**4**	**5**	
6.1 Good off centerline						
6.2 Good seated to standing						
6.3 Good with head rotation						

Annex D
(Informative)

Informative references and bibliography

BECH, S. Selection and Training of Subjects for Listening Tests on Sound-Reproducing Equipment. *J. Audio Eng. Soc.*, July–August 1992, vol. 40, no. 7/8 p. 590–610.

CLARK, DL. High-Resolution Subjective Testing Using a Double-Blind Comparator. *J. Audio Eng. Soc.*, Engineering Reports, May 1982, vol. 30, no. 5, p. 330–338.

GREEN, DM. and SWETS, JA. *Signal Detection Theory and Psychophysics.* New York: Krieger, 1966.

IEC 268-5, *Sound system equipment — Part 13: Listening tests on loudspeakers.* Geneva, Switzerland: International Electrotechnical Commission, 1985.

The Proceedings of the AES 8th International Conference: The Sound Of Audio. Washington, DC, USA, 1990-05-03 to 1990-05-06. New York: Audio Engineering Society, 1990.

The Proceedings of the AES 12th International Conference: Perception Of Reproduced Sound. Copenhagen, Denmark, 1993-06-28 to 1993-06-30. New York: Audio Engineering Society, 1993.

TOOLE, FE., Subjective Measurements of Loudspeaker Sound Quality and Listener Performance. *J. Audio Eng. Soc.* January–February 1985, vol. 33, no. 1/2, p. 2–32.