TECHNICAL REPORT

IEC TR 61305-6

First edition
2005-05

Household high-fidelity audio equipment and systems– Methods of measuring and specifying the performance –

Part 6: Listening tests on loudspeakers – Single stimulus ratings and paired comparisons

Reference number
IEC/TR 61305-6:2005(E)

Publication numbering

As from 1 January 1997 all IEC publications are issued with a designation in the 60000 series. For example, IEC 34-1 is now referred to as IEC 60034-1.

Consolidated editions

The IEC is now publishing consolidated versions of its publications. For example, edition numbers 1.0, 1.1 and 1.2 refer, respectively, to the base publication, the base publication incorporating amendment 1 and the base publication incorporating amendments 1 and 2.

Further information on IEC publications

The technical content of IEC publications is kept under constant review by the IEC, thus ensuring that the content reflects current technology. Information relating to this publication, including its validity, is available in the IEC Catalogue of publications (see below) in addition to new editions, amendments and corrigenda. Information on the subjects under consideration and work in progress undertaken by the technical committee which has prepared this publication, as well as the list of publications issued, is also available from the following:

- **IEC Web Site (www.iec.ch)**

- **Catalogue of IEC publications**

  The on-line catalogue on the IEC web site (www.iec.ch/searchpub) enables you to search by a variety of criteria including text searches, technical committees and date of publication. On-line information is also available on recently issued publications, withdrawn and replaced publications, as well as corrigenda.

- **IEC Just Published**

  This summary of recently issued publications (www.iec.ch/online_news/ justpub) is also available by email. Please contact the Customer Service Centre (see below) for further information.

- **Customer Service Centre**

  If you have any questions regarding this publication or need further assistance, please contact the Customer Service Centre:

  Email: custserv@iec.ch
  Tel: +41 22 919 02 11
  Fax: +41 22 919 03 00

TECHNICAL REPORT

IEC TR 61305-6

First edition
2005-05

Household high-fidelity audio equipment and systems– Methods of measuring and specifying the performance –

Part 6: Listening tests on loudspeakers – Single stimulus ratings and paired comparisons

International Electrotechnical Commission, 3, rue de Varembé, PO Box 131, CH-1211 Geneva 20, Switzerland
Telephone: +41 22 919 02 11 Telefax: +41 22 919 03 00 E-mail: inmail@iec.ch Web: www.iec.ch

Commission Electrotechnique Internationale
International Electrotechnical Commission
Международная Электротехническая Комиссия

PRICE CODE R

For price, see current catalogue

CONTENTS

INTERNATIONAL ELECTROTECHNICAL COMMISSION

HOUSEHOLD HIGH-FIDELITY AUDIO EQUIPMENT AND SYSTEMS – METHODS OF MEASURING AND SPECIFYING THE PERFORMANCE –

Part 6: Listening tests on loudspeakers – Single stimulus ratings and paired comparisons

FOREWORD

1) The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides (hereafter referred to as "IEC Publication(s)"). Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.
2) The formal decisions or agreements of IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC National Committees.
3) IEC Publications have the form of recommendations for international use and are accepted by IEC National Committees in that sense. While all reasonable efforts are made to ensure that the technical content of IEC Publications is accurate, IEC cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.
4) In order to promote international uniformity, IEC National Committees undertake to apply IEC Publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC Publication and the corresponding national or regional publication shall be clearly indicated in the latter.
5) IEC provides no marking procedure to indicate its approval and cannot be rendered responsible for any equipment declared to be in conformity with an IEC Publication.
6) All users should ensure that they have the latest edition of this publication.
7) No liability shall attach to IEC or its directors, employees, servants or agents including individual experts and members of its technical committees and IEC National Committees for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this IEC Publication or any other IEC Publications.
8) Attention is drawn to the Normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.
9) Attention is drawn to the possibility that some of the elements of this IEC Publication may be the subject of patent rights. IEC shall not be held responsible for identifying any or all such patent rights.

The main task of IEC technical committees is to prepare International Standards. However, a technical committee may propose the publication of a technical report when it has collected data of a different kind from that which is normally published as an International Standard, for example "state of the art".

IEC 61305-6, which is a technical report, has been prepared by IEC technical committee 100: Audio, video and multimedia systems and equipment

The text of this technical report is based on the following documents:

Enquiry draft	Report on voting
100/855/DTR	100/905/RVC

Full information on the voting for the approval of this technical report can be found in the report on voting indicated in the above table.

The committee has decided that the contents of this publication will remain unchanged until the maintenance result date indicated on the IEC web site under "http://webstore.iec.ch" in the data related to the specific publication. At this date, the publication will be

- reconfirmed,
- withdrawn,
- replaced by a revised edition, or
- amended.

A bilingual version of this publication may be issued at a later date.

INTRODUCTION

This technical report gives recommendations for establishing, conducting and evaluating listening tests.

The tests described in this report are to be performed in a room, the size and acoustical properties are similar to those of an average living room. Specific recommendations about the room size, acoustical properties, arrangement of loudspeakers and listeners, and environmental conditions are given.

This technical report describes experimental procedures, including recommendations on the choice of programme material and the processing and presentation of the final data. It may be useful to consider some of the recommendations in AES 20. It should be understood that the topics of experimental design, execution and statistical analysis are complex, and that only the most general guidelines can be given. It is recommended that professionals with expertise of experimental design and statistics should be consulted.

The use of multichannel formats, principally for domestic presentation of surround audio and cinema is becoming more usual. The procedures described in this report are applicable to any number of channels.

If the number of loudspeakers to be tested is high, the paired comparison listening test is lengthy because each loudspeaker has to be compared with the other. A shorter method is the single stimulus rating. With this method, each object is judged once. The rating is almost independent of the loudspeaker range in test. Each object is rated absolutely, whereas a paired comparison provides a relative ranking of the order of the loudspeakers in test.

Another technical report for listening test has been published as IEC 60268-13 and it is expected that the two technical reports should be combined in the maintenance work

HOUSEHOLD HIGH-FIDELITY AUDIO EQUIPMENT AND SYSTEMS – METHODS OF MEASURING AND SPECIFYING THE PERFORMANCE –

Part 6: Listening tests on loudspeakers – Single stimulus ratings and paired comparisons

1 Scope

This technical report applies to loudspeakers conforming to IEC 61305-5 and intended for home use.

The purpose of this report is, in addition to objective testing according to IEC 60268-5, to establish standards for comparison of the sound characteristics of various loudspeakers with each other.

Two test procedures are described:

– single stimulus ratings;

– paired comparisons.

The procedures described in this report are applicable to any number of channels.

NOTE The test procedures are specified for stereo systems. They can be applied to multichannel systems accordingly.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

IEC 61305-5, *High fidelity audio equipment and systems; Minimum performance requirements – Part 5: Loudspeakers.*

IEC 60268-5, *Sound system equipment – Part 5: Loudspeakers.*

ISO 3382, *Acoustics – Measurement of the reverberation time of rooms with reference to other acoustical parameters.*

3 Test preparation

3.1 Characteristics of the listening room

The volume of the listening room shall be 80 m^3 ± 20 m^3 with a room height of 2,75 m ± 0,25 m. The room should have a rectangular floor plan, whereby the ratio of the sides to each other should not exceed 2:1. A square floor plan is not permitted.

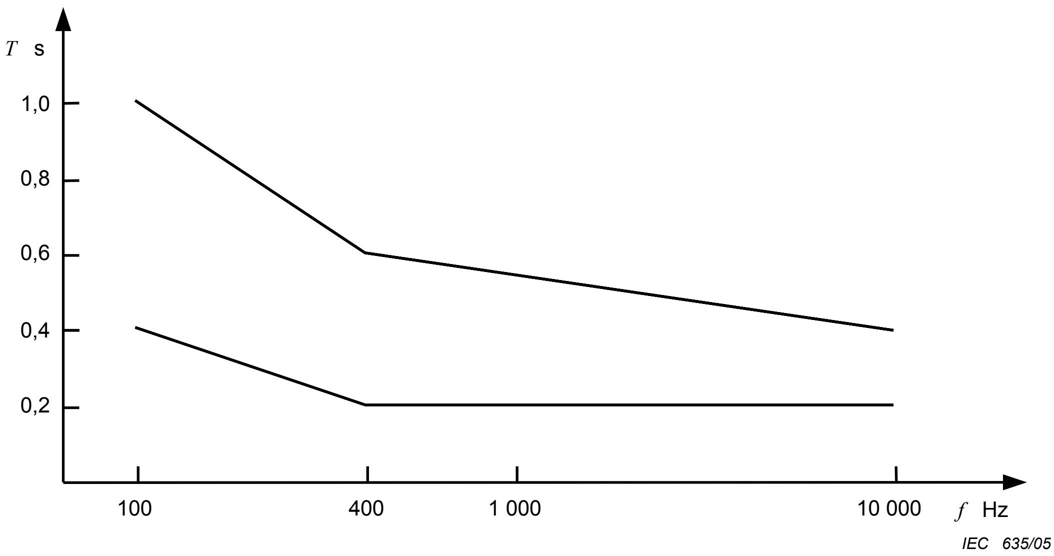

Figure 1 – Range limits for the reverberation time

The reverberation time of the room shall be within the range limits shown in Figure 1. The reverberation time shall not change by more than +25 % or –20 % within one octave in the range from 400 Hz to 8 kHz. The reverberation time is to be measured with noise in 1/3 - octave intervals with listeners present as laid down in ISO 3382. Conversion of the reverberation time measured without listeners to one with listeners is permitted.

Relevant room resonance resulting from furnishings shall be avoided. The distribution of the absorbing material should be as uniform as possible for all frequencies. In particular, parallel walls shall not be allowed to reflect sound to a significantly higher degree than the other walls. This also applies to the floor and the ceiling of the listening room. The AF-rated mean level of noise in the testing room may not exceed a value of L_{AFm} = 25 dB.

3.2 Loudspeaker position and seat arrangement

The loudspeaker shall be hidden behind an acoustically transparent screen to avoid influencing the listener by its design.

The loudspeakers shall be located and listened to in the positions in which they are used normally (as recommended of the manufacturer). This means, for example, loudspeakers designed for placement in shelves shall be placed in shelves or in an acoustically equivalent position. The distances from the floor and the ceiling shall correspond to usual positioning. Unless otherwise specified, the distance from the walls shall be at least 0,5 m.

The loudspeakers shall be tested in the stereo mode. The angle between the imaginary lines joining the loudspeakers and the listener should be not less than 45° for each listener. The loudspeakers shall be positioned so that the stereo base is constant in each case. Only the two sets of stereo loudspeaker (paired comparison) or the set of stereo loudspeakers to be tested (single stimulus rating) may be arranged to prevent feedback from other loudspeakers. When the paired comparison procedure is used, the centre and height of the line between the loudspeakers of each set shall not vary more than is necessary.

If several listeners participate simultaneously in one test, the seats should be arranged so that all listeners can see the loudspeakers directly when the screen is taken away.

3.3 Electrical requirements

The loudspeakers shall be fed by equipment with characteristics that do not affect the test results. The system shall consist of a source, level adjustment, and a power amplifier with a low output impedance and sufficient power (that is, clipping may not occur even at the highest signal levels). The tone control is to be set to a flat position both on the equipment as well as on the loudspeakers. Other settings shall be indicated.

Deviations in the flat frequency response shall be less than 0,5 dB in the range from 20 Hz to 20 kHz as measured at the input of the system and at the output of the power amplifier.

3.4 Level setting

The loudspeakers shall be set to provide the same level of volume as measured at an average listener's position with a test signal similar to a programme. It has been proven good practice to feed pink noise to the loudspeakers and set the A-weighted sound level to 75 dB.

NOTE The problem of setting up different loudspeakers by measurement so that reproduction of speech and music is equal in volume for all loudspeakers cannot be solved completely if the loudspeakers have different frequency responses. It is only possible to find a practical compromise using such a test signal.

3.5 Listening levels for programmes

The listening levels shall be adjusted so that the loudness is the same, as far as possible, as the loudness at an average listening position in the original room where the performance was held.

3.6 Programme material

The programme material should include:

a) speech;

b) chamber music;

c) orchestra music (tutti passages of a large orchestra);

d) solo voice and instrumental accompaniment;

e) entertainment music without significant electro acoustical tampering.

The duration of each unit should be between 60 s and 120 s. The sections should correspond to musical phrases. The programme material used shall be listed in detail.

3.7 Qualification and number of listeners

The hearing loss of the listeners shall be less than 20 dB in the frequency range of 125 Hz to 8 kHz [1]. Moreover, they should have some form of training or be interested in music, shall visit concerts regularly and listen to recorded music. The number of listeners shall be not less than 10.

3.8 Test duration

The duration of one test session should not exceed 1,5 h, whereby a suitable break should be made after about 0,75 h. The total test duration should not exceed 3 h per day and per listener.

4 Single stimulus ratings

4.1 Test procedure

The purpose of a listening test is to obtain an evaluation of the overall quality of the loudspeakers. This overall judgement is formed by the various sound characteristics of a loudspeaker. According to studies made by Gabrielson *et al.* [2], three independent characteristics (evaluation aspects) can be found, which can be designated as follows:

- volume;
- definition;
- transparency.

An overall judgement is obtained by combining the evaluations. Since these characteristics can be judged by comparing with an optimum sound obtained from a comparable original performance, these loudspeakers can be judged with a single stimulus rating. The result of the listening test not only provides a relative rating of individual loudspeakers within a set of loudspeakers but an absolute rating, as well.

Since any loudspeaker has to be tested only once for each sample of music, the single stimulus test is essentially a shorter test, especially for a large number of loudspeakers, compared to paired comparison which requires each loudspeaker to be compared with the other. The repeated judgement of loudspeakers required to check the rating reliability of the listeners increases the time insignificantly.

4.2 Questionnaire

In order to obtain ratings for the three evaluation aspects volume, definition and transparency, it is recommended to use a questionnaire listing pairs of terms with opposite meanings. Since the terms do not describe the sensations felt by every test person in exactly the same way it is recommended that several terms are used for each aspect of evaluation. The following questionnaire has proved to be useful in several studies:

Table 1 – Questionnaire for single stimulus rating

		1	2	3	4	5	6	
1	light	---	--	-	-	--	---	dark
2	distinct	---	--	-	-	--	---	indistinct
3	natural	---	--	-	-	--	---	unnatural
4	diffuse	---	--	-	-	--	---	concentrated
5	emphasized bass	---	--	-	-	--	---	poor bass
6	brilliant	---	--	-	-	--	---	faint
7	unpleasant	---	--	-	-	--	---	pleasant
8	round	---	--	-	-	--	---	sharp
9	narrow	---	--	-	-	--	---	wide
10	slim	---	--	-	-	--	---	voluminous
11	subdued	---	--	-	-	--	---	powerful
12	unbalanced	---	--	-	-	--	---	balanced
13	shrill	---	--	-	-	--	---	hollow
14	soft	---	--	-	-	--	---	hard
15	not transparent	---	--	-	-	--	---	transparent
16	dry	---	--	-	-	--	---	echoic
17	rough	---	--	-	-	--	---	clear
18	coloured	---	--	-	-	--	---	uncoloured
19	sharp	---	--	-	-	--	---	dull

Name: Date: Example No.:

The following pairs of terms are relevant for the aspects of judgement:

Aspect of judgement	Pair of opposites
Volume	5, 9, 10, 11
Definition	1, 4, 6, 8, 13, 14, 19
Transparency	2, 3, 15, 16, 17, 18
Overall judgement	3, 7, 12

If other pairs are added, it should be ascertained based on the results of a multifactor variance analysis to be performed subsequently (refer to Bortz, Test von Ratingskalen [3])) that the terms suitably differentiate between loudspeakers and that they are uniformly understood and used by the judges.

NOTE The use of an even number of evaluation steps has proved to be advantageous in helping test persons to make a decision.

4.3 Test sequence

When the test persons have been instructed and the selected programme examples have been presented, a few loudspeakers are evaluated by means of the questionnaire as an introduction to the test procedure. The test persons have to complete one questionnaire for each loudspeaker and each programme example. Afterwards, the actual test is started. In order to avoid a chain effect, all loudspeakers and programme examples shall be tested in a random sequence, which will involve changing the loudspeakers very frequently. A useful compromise is obtained by having all programme examples sequentially evaluated for one loudspeaker. In this case, a new questionnaire must be handed out for each programme to prevent the test person being influenced by the judgements of previous programmes. To enable a check of the reliability of the judgements about 1/5 to 1/4 of the loudspeakers shall be evaluated twice without informing the test persons.

4.4 Reliability of judgements

Prior to starting the actual evaluation of data, the reliability of the test persons judgements shall be checked using a simple process. For this purpose, all the questionnaires that were filled out during the double judgement of the loudspeakers are used. The difference between the ratings in the first and in the second pass is determined separately for each test person and for all opposite pairs and programmes. These differences are squared and averaged. The results are standardised on the total variance of the judgements, because the test persons judge differently within the scope of the questionnaire.

Based on the coefficient resulting from this standardisation, the test persons can be compared and evaluated with regard to their judgement reliability. Based on existing knowledge, the coefficient is found with the following equation

$$k = \frac{A_{\mathrm{sd}}}{V_{\mathrm{tot}}} \leq 1{,}5 \tag{1}$$

where

A_{sd} is the average of squared differences;

V_{tot} is the total variance.

The judgements of test persons having a coefficient of more than 1,5 shall not be used for further evaluation.

Notes on calculating the coefficient:

Suppose the ratings of the various test persons for the opposing pairs over all examples and loudspeakers are

$$a_{i,j,k} \quad \text{with} \quad \begin{array}{ll} i = 1 \ldots r & \text{opposing pairs} \\ j = 1 \ldots s & \text{examples} \\ k = 1 \ldots u & \text{loudspeakers} \end{array}$$

and suppose the ratings of repeated measurements for a selected number t of loudspeakers are

$$b_{i,j,k} \quad \text{with} \quad \begin{array}{ll} i = 1 \ldots r & \text{opposing pairs} \\ j = 1 \ldots s & \text{examples} \\ k = 1 \ldots t & \text{loudspeakers} \end{array}$$

where $t \leq u$. Then the average of squared differences A_{sd} is

$$A_{\mathrm{sd}} = \frac{\sum_{ijk} \left(a_{ijk} - b_{ijk} \right)^2}{r\,s\,t} \tag{2}$$

and the total variance is

$$V_{\mathrm{tot}} = \frac{\sum_{ijk} \left(a_{ijk} - \overline{a} \right)^2}{r\,s\,u - 1} \tag{3}$$

with

$$\overline{a} = \frac{\sum_{ijk} a_{ijk}}{r\,s\,u} \tag{4}$$

4.5 Variance analysis

After having eliminated the data of unreliable test persons from the data record, a multifactor variance analysis should be carried out to clarify the variance of variables and to check the significance by means of an F test. The following variables are relevant:

- differences of loudspeakers (desired);
- differences caused by examples (possible but mostly not given);
- differences caused by persons (undesired, i.e. error variance).

If it can be proved that the dissimilarity of loudspeakers is significant with regard to the error variance and that if the differences caused by the examples are low, the judgements can be combined to an average over the examples and persons.

4.6 Factor analysis

As mentioned above, three aspects are relevant for the evaluation. In order to obtain figures for the evaluation aspects for various loudspeakers, the factor analysis is suitable for data compression. (Factor analysis with subsequent varimax rotation followed by factor value estimation). A data record of loudspeaker judgements with regard to the evaluation aspect is obtained as a result.

4.7 Quality judgement

In contrast to the loudspeaker characteristics calculated in 3.6, the values averaged for persons and programmes, for example for the opposing pair "pleasant - unpleasant" shall be used for the quality evaluation. Thus, the original position in the evaluation range is maintained, especially if loudspeakers of a certain class of quality are evaluated in the test.

4.8 Representation of the results

In addition to the representation of loudspeaker characteristics in a table, a graphic representation in the planes of the factor space is useful.

Example: refer to Figures 2 and 3.

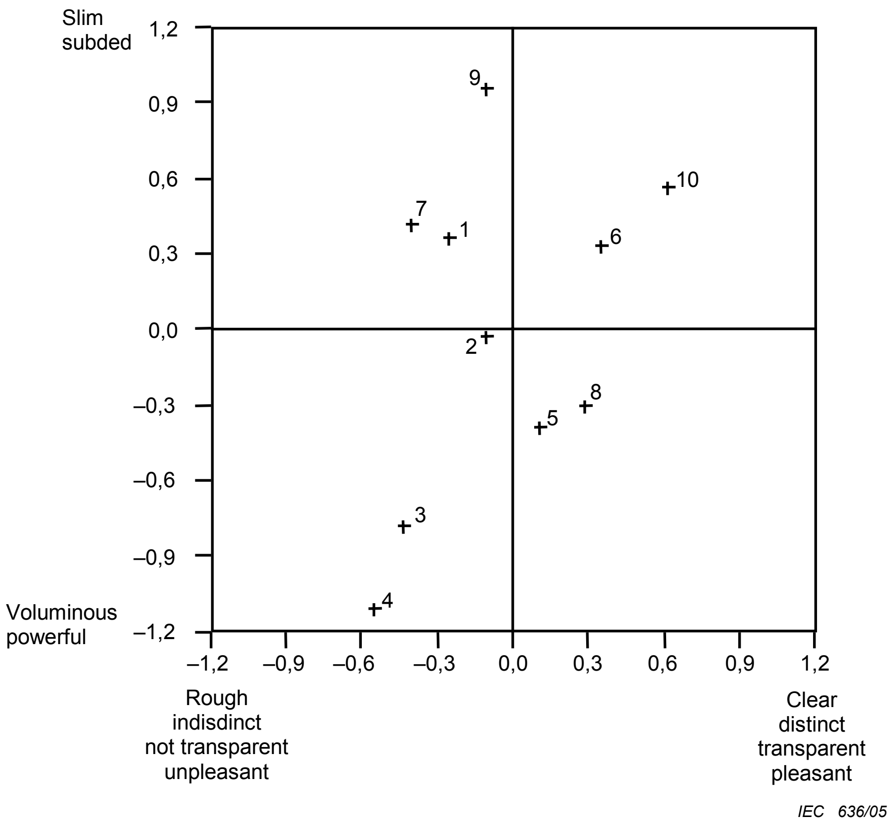

Figure 2 – Representation of results for loudspeakers 1 to 10 in the evaluation aspects transparency/quality and volume

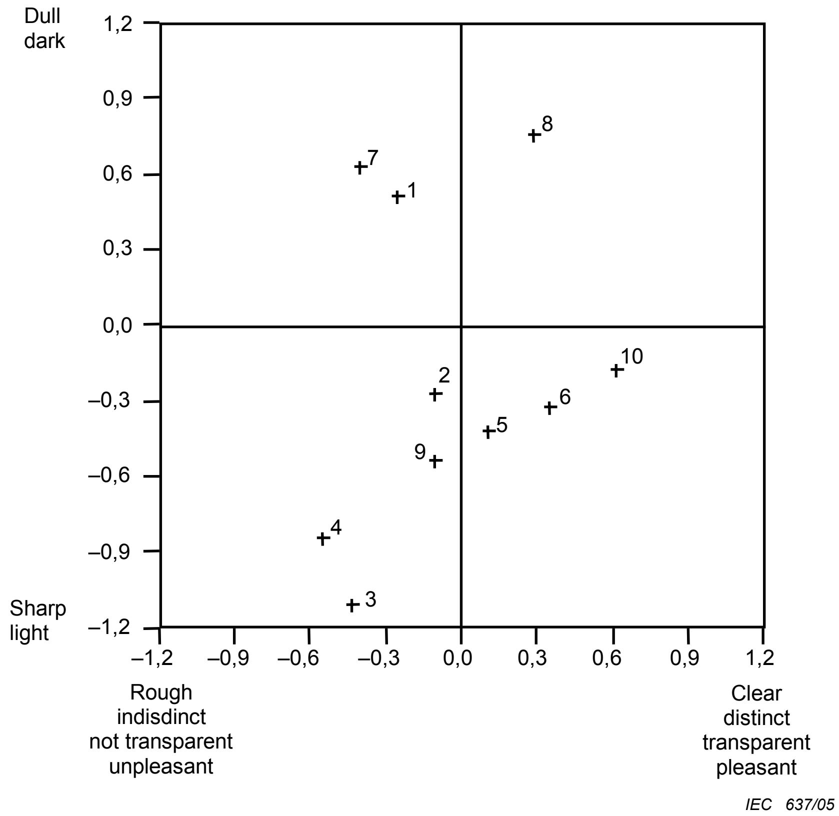

Figure 3 – Representation of results for loudspeakers 1 to 10 in the evaluation aspects transparency/quality and definition

To enable a quality evaluation, the graphic representation should maintain the arrangement of the rating over the original range. Refer to Figure 4.

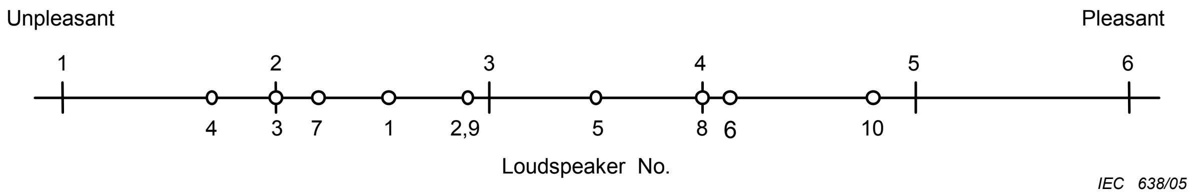

Figure 4 – Rating of loudspeakers 1 to 10 in the opposite pair pleasant-unpleasant

4.9 Computer programs

Computer programs are needed to calculate the various evaluation steps. These are:

- multifactor variance analysis;
- factor analysis with varimax rotation, and
- factor value estimation.

5 Paired comparisons

5.1 Test procedure

The purpose of a listening test is to rate several objects using one criterion of evaluation over a certain scale. As the differences can be determined only by comparing the objects and as this is not possible simultaneously, the paired comparison procedure (AB comparison) can be chosen. In this case, each object is compared with the other and a decision is made as to which object meets the test criterion the best. The following explanations and the sample calculation in Annex A, are based on a two-stage scale (less - more). If a three-stage scale is used (less - equal - more), the equations (A.1), (A.2a) and (A.2b) in Annex A do not apply. Together with the numerous ordinal judgements obtained in this manner, the objects can be arranged on an interval scale using an evaluation model (law of comparative judgement, see for example [4] and [5]).

The use of this test procedure assumes that a suitable selection of test criteria is available. These must be selected so that only the different magnitude of the test criterion in question forms the basis for judgement for all comparisons; the test criteria must therefore be one-dimensional. If this is not the case, inconsistent sequences (circular triads) may result if different aspects of one test criterion form the basis for judgement in different comparisons, despite of the same question (multidimensional test criteria).

If circular triads occur despite correctly selected test criteria, this may be due to uncertainty in the judgement. A figure for the uncertainty of the judgement can be found by comparing the circular triads that actually occur and the maximum possible number [4] or the number occurring by chance [6]. This should also be examined for each test.

After all comparisons have been carried out, the results can be compiled in a so-called dominance matrix. Since the occurring differences do not correspond to the subjective distances, transformation of the values to a normal distribution (z-conversion [4]) is necessary.

Moreover, the ranking order of various objects can depend on the type of programme material; for this reason, all judgements shall be evaluated separately first for each programme. If the ranking order deviates only slightly, the judgements can be combined.

The ranking order found is not generally valid either in terms of its zero point or its spread; it applies primarily for the selected object field.

5.2 Test criteria

The paired comparison shall be carried out for a series of one-dimensional criteria allowing an evaluation of the sound characteristics of the loudspeakers. The following criteria have proven to be usable for judgements that do not contradict and should be used for evaluation of the sound characteristics:

a) bright – dark
b) soft treble – hard treble
c) voluminous – slender
d) powerful bass – weak bass
e) clear – unclear
f) transparent – not transparent
g) coloured – not coloured
h) balanced – unbalanced

The criteria used shall be indicated.

5.3 Test sequence

Each set of stereo loudspeakers shall be compared with the other by switching the test programmes from one set to the other. The switching sequence to the loudspeakers shall be random; however, the sequences must exclude the possibility of one set of loudspeakers being used in more than two sequential paired comparisons.

The method of switching from one set of loudspeakers to the other shall not give away the anonymity of the objects. The tests may be carried out in two different ways:

a) each of the two sets of stereo loudspeakers should be activated for approximately the same time; the intervals should be 2 s to 10 s and should not exceed 20 s under any circumstances. The programmes should be long enough to allow the loudspeakers to be switched back and forth a number of times during one programme. Neither the test supervisor nor the listener shall switch the loudspeakers;

b) the comparison can also be carried out by listening to the complete programme with one set of stereo loudspeakers and then with the second. The length of the programme should be less than 60 s if this method is used.

With both methods, an judgement shall be made after each programme using all test criteria. For each test criterion the listener shall decide which of the two sets of loudspeakers meets the criterion the best. The judgement shall be kept confidential and registered without any previous discussion.

5.4 Number of contradictory judgements

The judgements of the test persons shall then be evaluated for each criterion separately according to the programme and the test person in order to determine the number of circular triads. This shall be compared with the maximum possible number of circular triads or the number that could probably occur. Then the judgements shall be compiled according to the criterion and the programme. If the judgements deviate only slightly with regard to the programmes then the judgements for all the programmes can be summarized.

5.5 Determination of the scale values

The scale values shall be determined from the individual judgements as shown in the example in Annex A. If the differences in the scale values shall correspond to the differences in the perceptions, they shall be z-standardised as well [4].

5.6 Variance analysis

After eliminating the data of obviously unreliable persons from the data record, a multifactor variance analysis should be carried out to clarify the variance of variables and to check their significance by means of an F test. The following variables are relevant:

- differences of loudspeakers (desired)
- differences caused by examples (possible but mostly not given)
- differences caused by persons (undesired, i.e. error variance)

If it can be proved that the dissimilarity of loudspeakers is significant with regard to the error variance and that the differences caused by the examples are low, the judgements can be combined to give an average for the examples and persons.

5.7 Factor analysis

As mentioned above, three aspects are relevant for the evaluation. In order to obtain figures for the evaluation aspects for various loudspeakers, the factor analysis is suitable for data compression. (Factor analysis with subsequent varimax rotation followed by factor value estimation). A data record of loudspeaker judgements with regard to the evaluation aspect is obtained as a result.

Annex A
(informative)

Example calculation for tests using paired comparison

Annex A provides an example calculation for listening tests using paired comparison. This example includes the judgements of five test persons listening with respect to one criterion to five sets of stereo loudspeakers (referred to in the following as "loudspeakers").

A.1 Judgement reliability check

The 10 judgements of listener 1 for one complete set of paired comparisons are entered in a dominance matrix:

	Listener 1				
No	1	2	3	4	5
1		0	1	1	0
2	<u>1</u>		1	1	1
3	<u>0</u>	<u>0</u>		1	0
4	<u>0</u>	<u>0</u>	<u>0</u>		0
5	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	
a_j	2	0	3	4	1
a_j^2	4	0	9	16	1

Numbers underlined correspond to the answers in the questionnaire. A "1" is always awarded when one loudspeaker in the head of a column is preferred to another in the head of a line with regard to a criterion. The upper half is filled in by supplementing a "1" for equal pairs in the reverse order.

The column totals a_j indicates how often a certain loudspeaker was preferred to another during the course of the paired comparisons. Using a_j^2 it is possible to calculate how many circular triads occurred. This number is

$$d = n(n-1)(2n-1)/12 - \frac{1}{2}\sum_{j=1}^{n} a_j^2 \tag{A.1}$$

where n is the quantity of objects and a_j^2 is the square of the column totals for the individual object. Therefore, listener 1 produced the following number of circular triads:

$$d = 5 \cdot 4 \cdot 9/12 - \frac{1}{2}(4+0+9+16+1) = 0$$

The dominance matrices for the 4 other listeners could be:

	Listener 2				
No	1	2	3	4	5
1		0	0	1	1
2	1		1	1	0
3	1	0		1	0
4	0	0	0		0
5	0	1	1	1	
a_j	2	1	2	4	1
a_j^2	4	1	4	16	1

	Listener 3				
No	1	2	3	4	5
1		0	1	1	0
2	1		1	0	1
3	0	0		0	1
4	0	1	1		0
5	1	0	0	1	
a_j	2	1	3	2	2
a_j^2	4	1	9	4	4

	Listener 4				
No	1	2	3	4	5
1		0	1	1	0
2	1		0	1	1
3	0	1		1	0
4	0	0	0		0
5	1	0	1	1	
a_j	2	1	2	4	1
a_j^2	4	1	4	16	1

	Listener 5				
No	1	2	3	4	5
1		1	1	1	0
2	0		1	1	1
3	0	0		1	0
4	0	0	0		0
5	1	0	1	1	
a_j	1	1	3	4	1
a_j^2	1	1	9	16	1

The consistency coefficient K from Kendall

$$K = 1 - \frac{d}{d_{\text{max}}} \tag{A.2}$$

can be used to establish the judgement reliability, which compares the number d of circular triads with the maximum possible number d_{max}:

$$d_{\text{max}} = n(n^2 - 4)\frac{1}{24} \quad n\text{even} \tag{A.2a}$$

$$d_{\text{max}} = n(n^2 - 1)\frac{1}{24} \quad n\text{uneven} \tag{A.2b}$$

According to equation A.2b d_{max} = 5 for n = 5. According to equations A.1 and A.2, the 5 listeners have the following consistency coefficients:

Listener	$\sum a_j^2$	d	K
1	30	0	1,0
2	26	2	0,6
3	22	4	0,2
4	26	2	0,6
5	28	1	0,8

Listener 3 should be excluded from further evaluation due to low judgement reliability (4 circular triads out of 5 possible). Only listeners with consistency coefficients ≥0,6 may be taken into consideration in a compiled judgement.

Compilation of listeners 1, 2, 4 and 5 results in the following combined dominance matrix in which the judgements for corresponding loudspeaker sets are added:

Loudspeaker	**1**	**2**	**3**	**4**	**5**
1		1	3	4	1
2	3		3	4	3
3	1	1		4	0
4	0	0	0		0
5	3	1	4	4	
S_{uj}	7	3	10	16	4

The column totals (S_{uj}) for the combined dominance matrix indicate how frequently one loudspeaker was preferred to another in relation to one criterion by listeners with reliable judgements.

A.2 Calculation of scale values

First, the elements of the combined dominance matrix are standardized to the interval 0…1:

Loudspeaker	1	2	3	4	5
1		0,25	0,75	1	0,25
2	0,75		0,75	1	0,75
3	0,25	0,25		1	0
4	0	0	0		0
5	0,75	0,25	1	1	

Each element of this standardised dominance matrix should now be converted (z-conversion) according to 4.1 with the inverse distribution function of the standard distribution. However, since the values 0 and 1 would convert to ∞ the matrix is unusable in this special case. If a similar case occurs, the remedy is to use a function for conversion which converts the values 0 and 1 to finite z-values. One such suitable function is:

$$z = \arcsin\,(2p - 1)$$

This conversion results in the following z-matrix:

Loudspeaker	1	2	3	4	5
1		–0,52	0,52	1,57	–0,52
2	0,52		0,52	1,57	0,52
3	–0,52	–0,52		1,57	–1,57
4	–1,57	–1,57	–1,57		–1,57
5	0,52	–0,52	1,57	1,57	

The scale values D_i for the loudspeakers are then determined so that the double sum approaches a minimum:

$$\sum_{i=1}^{n}\sum_{j=1}^{n}(D_i - D_j - z_{ij})^2 \rightarrow \min$$

The solution for this problem leads to the well-known problem of compensation. If the matrix is completely filled the scale values D_j are equal to the average value of the coefficients of the corresponding column:

$$D_j = \frac{1}{n}\sum_{i=1}^{n} z_{ij}$$

In this special case, the following scale values result for the 5 loudspeakers:

Loudspeaker	1	2	3	4	5
Ranking order	3	4	2	1	4
D_j	–0,21	–0,63	0,21	1,26	–0,63

Bibliography

[1] ISO 389: *Acoustics – Reference zero for the calibration of audiometric equipment*

[2] GABRIELSON, A. *et al.*: *Judgement and dimension analyses of perceived sound quality of sound-reproducing systems* JASA 55 (1974)9 854 - 861

[3] BORTZ, J.: *Ein Verfahren zur Tauglichkeit von Ratingskalen*. *Psychologie und Praxis*, *Band XVI* (1972), Seite 49

[4] SIXTL, F.: *Meßmethoden der Psychologie.* Verlag Julius Beetz 1967

[5] GUILFORD, J. P.: *Psychometric Methods*. McGraw Hill 1954

[6] LIENERT, J. A.: *Verteilungsfreie Methoden der Biostatistik, Band II.* Verlag Anton Heim, Meisenheim 1979

IEC

Standards Survey

The IEC would like to offer you the best quality standards possible. To make sure that we continue to meet your needs, your feedback is essential. Would you please take a minute to answer the questions overleaf and fax them to us at +41 22 919 03 00 or mail them to the address below. Thank you!

Customer Service Centre (CSC)

International Electrotechnical Commission
3, rue de Varembé
1211 Genève 20
Switzerland

or

Fax to: **IEC**/CSC at +41 22 919 03 00

Thank you for your contribution to the standards-making process.

Nicht frankieren
Ne pas affranchir

Non affrancare
No stamp required

A Prioritaire

RÉPONSE PAYÉE

SUISSE

Customer Service Centre (CSC)
International Electrotechnical Commission
3, rue de Varembé
1211 GENEVA 20
Switzerland

Q1 Please report on **ONE STANDARD** and **ONE STANDARD ONLY**. Enter the exact number of the standard: *(e.g. 60601-1-1)*

....

Q2 Please tell us in what capacity(ies) you bought the standard *(tick all that apply).* I am the/a:

- purchasing agent ❑
- librarian ❑
- researcher ❑
- design engineer ❑
- safety engineer ❑
- testing engineer ❑
- marketing specialist ❑
- other...

Q3 I work for/in/as a: *(tick all that apply)*

- manufacturing ❑
- consultant ❑
- government ❑
- test/certification facility ❑
- public utility ❑
- education ❑
- military ❑
- other...

Q4 This standard will be used for: *(tick all that apply)*

- general reference ❑
- product research ❑
- product design/development ❑
- specifications ❑
- tenders ❑
- quality assessment ❑
- certification ❑
- technical documentation ❑
- thesis ❑
- manufacturing ❑
- other...

Q5 This standard meets my needs: *(tick one)*

- not at all ❑
- nearly ❑
- fairly well ❑
- exactly ❑

Q6 If you ticked NOT AT ALL in Question 5 the reason is: *(tick all that apply)*

- standard is out of date ❑
- standard is incomplete ❑
- standard is too academic ❑
- standard is too superficial ❑
- title is misleading ❑
- I made the wrong choice ❑
- other ....

Q7 Please assess the standard in the following categories, using the numbers:
(1) unacceptable,
(2) below average,
(3) average,
(4) above average,
(5) exceptional,
(6) not applicable

- timeliness ....
- quality of writing....................................
- technical contents.................................
- logic of arrangement of contents ....
- tables, charts, graphs, figures...............
- other ....

Q8 I read/use the: *(tick one)*

- French text only ❑
- English text only ❑
- both English and French texts ❑

Q9 Please share any comment on any aspect of the IEC that you would like us to know:

....
....
....
....
....
....
....
....
....
....
....
....

ISBN 2-8318-7976-0

9 782831 879765

ICS 33.160.50; 17.140.99

Typeset and printed by the IEC Central Office
GENEVA, SWITZERLAND