

ICS 33.160.99

M72

备案号:

WH

中华人民共和国文化行业标准

WH/T 82—2019

演出用专业音响设备音质主观评价方法

The method of the subjective assessment of sound quality for the professional
sound equipment used for performing arts

[点击此处添加与国际标准一致性程度的标识](#)

(报批稿)

2019 - 01 - 02 发布

2019 - 06 - 01 实施

中华人民共和国文化和旅游部

发布

目 次

前言.....	III
1 范围.....	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 音质评价词.....	3
5 评价节目源.....	3
5.1 一般规定.....	3
5.2 节目制作规范.....	3
5.3 节目范围.....	3
5.4 典型节目.....	4
6 听音环境.....	4
6.1 一般规定.....	4
6.2 房间尺寸.....	4
6.3 混响时间.....	4
6.4 本底噪声.....	6
6.5 环境条件.....	6
6.6 扬声器系统的位置.....	6
6.7 听音员的位置.....	6
6.8 听音声级.....	7
7 听音员.....	7
7.1 听音员的条件.....	7
7.2 听音小组的组成.....	7
7.3 听音小组的人数.....	7
7.4 听音员的工作时间.....	7
7.5 听音员的预备工作.....	7
8 评分方法与评价尺度.....	8
8.1 评分方法的选择.....	8
8.2 等级评定法.....	8
8.3 成对比较法.....	10
8.4 可靠性分析.....	11
8.5 有效性分析.....	11
9 评价试验中的辅助设备.....	11
9.1 一般规定.....	11
9.2 传声器评价时的辅助设备.....	12

9.3 扬声器系统评价时的辅助设备.....	12
9.4 功率放大器评价时的辅助设备.....	12
9.5 其它设备评价时的辅助设备.....	12
10 评价结果的报告.....	12
附录 A（资料性附录） 等级评定方法的尺度构成方法.....	13
附录 B（资料性附录） 成对比较法的尺度构成方法.....	18

前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本标准由中华人民共和国文化和旅游部提出。

本标准由全国剧场标准化技术委员会（SAC/TC388）归口。

本标准负责起草单位：中国演艺设备技术协会演出场馆设备专业委员会。

本标准参加起草单位：中国传媒大学、北京第七九七音响股份有限公司、广州锐丰音响科技股份有限公司、广州飞达音响专业器材有限公司、广州励丰文化科技股份有限公司、宁波音王电声股份有限公司、东莞三基音响科技有限公司、深圳宝业恒实业发展有限公司、深圳迪斯声学股份有限公司、东莞爱高音响有限公司、恩平超达音响器材设备有限公司、中国电子科技集团公司第三研究所、北京联合大学、西安工程大学、浙江传媒学院。

本标准主要起草人：孟子厚、崔广中、李建伟、宋鸣、胡秉奇、龚惠哲、何伟峰、李志雄、刘永耀、黄永谦、夏德成、郑少凡、熊孟生、徐新国、周其麟、徐文学、吴帆、潘杨、石蓓。

演出用专业音响设备音质主观评价方法

1 范围

本标准规定了音质评价词、评价节目源、听音环境、听音员、评价方法与评价尺度，提出了评价试验中的辅助设备及评价结果的报告。

本标准适用于组成各类专业演出场所用扩声系统中的各类专业音响设备在规定环境下的音质主观评价。本标准也适用于同一制造商设计的同一类设备的不同样品的音质主观评价和不同制造商设计的同一类设备的音质等级评定。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅所注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB 3095 环境空气质量标准

GB 4854.1-2004/ISO 389-1:1998 声学 校准测听设备的基准零级 第1部分：压耳式耳机纯音基准等效阈声压级

GB 10240-1988 电声产品声音质量主观评价用节目源编辑制作规范

GB/T 18883 室内空气质量标准

GB/T 51091 试听室工程技术规范

GSB 16-3451 声音质量主观评价用节目源标准样品

3 术语和定义

3.1

音响 sound

音是声音，响是发出声音。声音的产生、传播、处理、达到听闻效果的过程中涉及到电声学、建筑声学、音乐声学、心理声学、生理声学等方面的内容。

[WH/T 57-2013, 定义 2.1.1]

注：现在通常把“音响”作为扩声设备、扩声系统或扩声效果等的简称。

3.2

音响设备 sound equipment

对声音信号拾取、记录或还原、处理加工、放大等所用的设备。

[WH/T 57-2013, 定义 2.1.4]

3.3

音质 sound quality

声音质量。声音通过人的听觉，由心理所确定的心理量。

[WH/T 57-2013, 定义 2.2.4]

3.4

音质主观评价 subjective assessment of sound quality

在规定条件下，通过听音者的主观感受来评价扩声系统和设备的声音质量水平。

[WH/T 57-2013, 定义 6.1.1]

3.5

听音试验 listening test

由一定数量的、经过训练的听音员，在规定声学特性的房间（或称之为试听室、听音室、审听室等）内，按照共同规定的听音试验方法，对音响设备、节目源、乐音或乐器音等的音质进行主观感觉的评定，最后用数理统计或其他方法对数据进行计算，评定出结果的试验。

[WH/T 57-2013, 定义 6.1.2]

3.6

典型节目 typical program

能代表某种类型节目特征的，同时对电声系统和产品声音质量主观评价有较大贡献的节目源。

[GB/T 9002-1996, 定义 12.2]

3.7

听音员 listener

在进行听音试验时，对音质做出评价或按要求对声音刺激做出可重复性判断反应的人员，也称为受试者。

3.8

听音小组 listening group

听音试验中参与评分的全体听音员。

3.9

个体的可靠性 intra-individual reliability

某一听音员对某一评价对象重复评定的一致性。

3.10

个体间的可靠性 inter-individual reliability

不同听音员对同一个评价对象评定的一致性。

3.11

心理尺度 psychological scale

在声音质量的主观评价或听音试验中，通过统计处理得到的反映某个音质属性的心理量的数值。心理尺度一般是一个相对尺度，反映被评价对象之间在心理感知上的相对关系。

4 音质评价词

音质评价词的选择应根据被评价的设备以及它所可能具有的音质属性来确定。音质评价词应通俗易懂，避免生僻的词汇。

对一般性的设备音质评价，宜使用以下的音质评价词：

- a) 丰满 fullness
中、低音充分，高音适度，响度合适，听感厚实、温暖、舒适、有弹性。
- b) 明亮 brilliance
高、中音充分，但不过分，听感明朗、活跃、通透，谐音丰富。
- c) 清晰 articulation
语言声音节清晰、易懂度好；音乐声层次分明，乐器和人声音色清晰可辨，声音细节清楚。
- d) 柔和 mellow
听感上中低频能量充足，高音适度不刺耳。声音温和、松弛，没有紧张感。听感舒服、悦耳，不尖、不噪、不破。
- e) 力度 dynamics
声音坚实有力、通畅，构成整体音质的各声音元素之间协调一致，能充分表现声源的动态范围。
- f) 平衡 balance
高、中、低音搭配得当，均匀协和。能还原节目源各声部之间协调的比例。
- g) 真实 real
尽可能保持原有声音的时域和频域的音质特点。

5 评价节目源

5.1 一般规定

节目源的选取应根据所评价设备的音质属性来确定，听音节目一般由多种不同的素材组成，并能充分反映被评价设备的音质属性。素材的选择要结构合理，避免冗余和漏缺。

5.2 节目制作

节目制作应按 GB 10240 规定的要求执行。

5.3 节目基本类型

5.3.1 一般要求

对综合性的评价任务，节目源应至少包括以下几种类型。

5.3.2 语言声

用汉语普通话标准语速朗读的男、女语言声。语料的选取应基本符合日常语言交流中各种声、韵、调出现的概率分布。

5.3.3 声乐

无伴奏男低音和女高音各一段；无伴奏混声合唱一段；有传统乐器伴奏的代表性戏曲唱腔一段。

5.3.4 器乐

按 GB 10240-1988 中 4.2.3 的规定执行。

5.3.5 景观声

自然声或人类活动所造成的声音。可选用钟声、列车通过的声音、汽笛声、警报声、流水声、雷声、枪炮声、玻璃破碎声、虫鸟鸣叫声等。

5.4 典型节目

宜按 GSB 16-3451 规定的节目作为演出用专业音响设备音质主观评价的典型节目源。

6 听音环境

6.1 一般规定

音质主观评价在规定声学特性的室内环境下进行。试听室的房间声学特性应满足演出用专业音响设备音质主观评价的要求，即传输频率特性平滑，无可觉察的声染色；不应出现回声、颤动回声、声聚焦和共振等声学缺陷。

6.2 房间尺寸

应根据受评设备的类别、性质及听音员人数来确定试听室的合理建筑面积。

试听室的具体尺寸比例应按 GB/T 51091 规定的要求设计。

6.3 混响时间

试听室在 500Hz 的混响时间宜处于图 1 和图 2 中规定的范围，但最大不应超出图中的上限值 0.2s。

混响时间的频率特性，相对于 500Hz 的比值宜符合表 1 的规定。

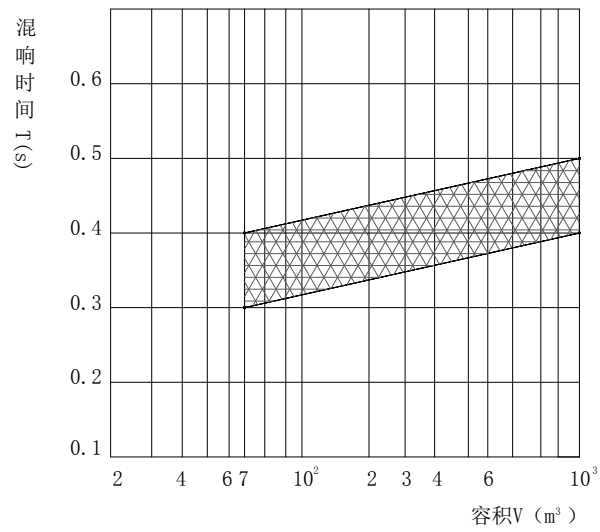


图 1 小型和中型试听室 500Hz 时的混响时间 T

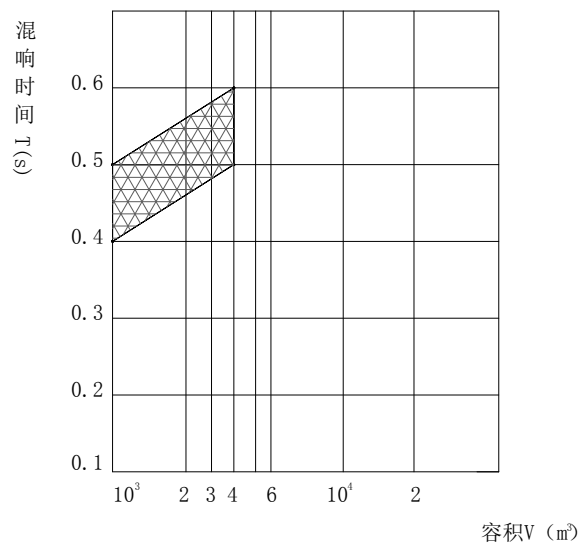


图 2 大型试听室 500Hz 时的混响时间 T

表 1 各频率混响时间相对于 500Hz 的比值

试听室	中心频率 Hz							
	63	125	250	500	1000	2000	4000	8000
小型/中型	1.0~ 1.2	1.0~ 1.1	1.0~ 1.1	1.0	1.0	1.0	1.0	—
大型	1.0~ 1.23	1.0~ 1.2	1.0~ 1.1	1.0	1.0	1.0	0.8~ 1.0	0.8

6.4 本底噪声

6.4.1 试听室无人占用时，听音员位置处噪声级的限值不宜超过表 2 的噪声评价曲线 NR 值的规定。各 NR 值的倍频带声压级如表 3 所示。

表 2 试听室内噪声限值

试听室环境	噪声限值
通风、空调设备关闭时	NR-15
通风、空调设备开启时	NR-25

表 3 噪声评价曲线 NR 值对应的各倍频带声压级

NR 值	1/1 倍频程中心频率 Hz	31.5	63	125	250	500	1000	2000	4000	8000
NR15	声压级 dB	65	47	35	25	19	15	11	9	8
NR20		69	51	39	30	24	20	16	14	13
NR25		72	55	43	35	29	25	21	19	18

6.4.2 感觉到的本底噪声不宜是冲击性的、周期性的或有音调感觉的。

6.5 环境条件

听音试验时的环境条件应尽可能接近设备（或设备所组成的系统）正常使用时的环境。试听室应满足的环境条件如下：

- a) 环境温度：18℃~27℃；
- b) 相对湿度：25%~75%；
- c) 大气压强：86kPa~106kPa；
- d) 空气质量：室内空气应按 GB/T 18883 规定的标准执行；环境空气污染物的浓度限值应不高于 GB 3095 规定的二级限值标准；
- e) 室内照明：在不能采用自然光的条件下，室内照明应中性、柔和，不应给听音员造成视觉上的刺激或不舒适感。

6.6 扬声器系统的位置

当所评价的设备是扬声器系统时，通常应按照制造商在用户手册中推荐的方法摆放。当缺乏这方面的信息以及扬声器系统是用来评价其它设备的音质时，则可按照以下方法摆放：

- a) 扬声器系统的高音单元应指向听音员处于听音状态时耳朵的平均高度，一般情况下为 1.2m；
- b) 扬声器系统应指向听音区的中央，除非扬声器系统的设计另有规定；
- c) 扬声器系统与墙面的距离应不小于 1m；
- d) 应避免把扬声器系统布置在试听室的对称主轴上。

6.7 听音员的位置

听音员的位置布置应使得所有的听音员尽可能感受到同样的音质，听音区距扬声器系统的距离应视具体的评价设备和放音条件而定。宜满足下列要求：

- a) 听音员距墙面的距离 $\geq 1\text{m}$;
- b) 听音员距扬声器的距离 $\geq 2\text{m}$;
- c) 所有听音员宜坐在同样的高度,当多排布置时,座位高度应逐排上升 12cm;
- d) 听音员之间的排距 $\geq 0.8\text{m}$,左右间距 $\geq 0.6\text{m}$;
- e) 听音员的座椅靠背不应高于肩;
- f) 有多个听音员时,应采取适当的措施避免或减少相互间的视觉交流;
- g) 在被评价的设备和听音员之间宜设置透声遮光的帘幕,避免设备的视觉信息对听音员的心理暗示和干扰。

6.8 听音声级

扬声器系统重放声级的大小,应为听音员的优选平均听音声级。

对一些声音输出能力不够的设备,听音试验宜在低于优选声级、具有更低本底噪声的环境下进行。

7 听音员

7.1 听音员的条件

听音员应同时满足以下条件:

- a) 应具备正常的听力条件,并且两耳听力应基本一致;
- b) 左右任意一耳在 125Hz~8000Hz 的频率范围内,听力级不应高于 GB/T 4854.1-2004 规定的 20dBHL 要求;
- c) 应能区别声压级为 80dB 时在 1000Hz 处,1dB 的声压级变化和 3.3Hz 的音调变化;
- d) 应具有临场听音的经验,包括有扩声和无扩声的现场演出;
- e) 应对音乐基础知识有一定的了解,且不应在音乐风格上有明显偏好;
- f) 经常参加听音试验的听音员应定期进行听力检查,宜经常聆听现场自然声音乐排练或演出。

7.2 听音小组的组成

听音小组应由听音员组成,可以从音响师、录音师、音乐工作者、声频工程师、声学工作者和相关专业的学生等中选取。成员组成宜包含 1/5~1/10 的非专业听众,男女比例和年龄宜均衡,且应至少包括 30 岁以下的男、女听音员各一名。

7.3 听音小组的人数

经过专业训练,有相同经验且成员个体间可靠性高的听音小组人数不应少于 7 人,否则不宜少于 20 人。

7.4 听音员的工作时间

在听音评价时,听音员的连续工作时间不宜超过 30min,每次休息时间不宜短于 20min。包括工作和休息时间在内的一次听音试验不宜超过 3h。每个听音员每天的累计听音时间不宜超过 3h。

7.5 听音员的预备工作

7.5.1 每次听音试验开始之前,每个听音员应熟悉听音过程,并了解以下事项:

- a) 被评价的设备类别;
- b) 采用的听音评价方法;

- c) 采用的评分方法;
- d) 重点评价的音质属性和音质评价词。

7.5.2 听音试验时应避免听音员了解被评价的设备的品牌、生产厂商等详细情况。

8 评分方法与评价尺度

8.1 评分方法的选择

音质主观评价的目的是获得被评价设备的某个音质属性在听觉心理上的尺度。音质主观评价的基本要求是设计成可控的统计试验。评价尺度由评价数据的统计处理得到。

根据具体的评价任务设计不同的评分方法。在本标准的适用范围内,宜采用的常用评分方法为:等级评定法和成对比较法。

8.2 等级评定法

8.2.1 适用任务

等级评定法适用于对多个同类设备的音质进行综合评价的场合。这是一种具有分级意义的评价方法,需要对被评价设备的音质水平做出具有一致性的等级划分。在被评价的设备较多时,宜采用这种效率较高的方法。

8.2.2 听音前的练习

等级评定法中的各个节目是分别播放的,听音员在下一个节目播放前完成对当前设备的评价,听音员依靠对放音效果的记忆来做比较。因此,听音员宜事前进行一定的练习,熟悉评价方法,增强听觉记忆。

8.2.3 评分表格

等级评定法中推荐使用的评分记录表格如表4所示,其中等级的级数和音质属性的描述要根据具体的评价任务制定。

表4 等级评定法评分记录表格

演出用专业音响设备音质主观评价(等级评定法)用记分表							
评价地点:		设备类型:		听音编号:			
听音人姓名:		性别:		年龄:		职业:	
节目名称:							
听音评价时间:	年	月	日	时	分	-----	时 分
评价项目		评价等级					备注
		优	良	中	差	劣	
音质属性	丰满						
	清晰						
	明亮						

	柔和						
	力度						
	平衡						
	真实						
	… …						
总体效果							
评述							

8.2.4 等级设定

等级的设定应采用奇数个等级，通常设立五个等级，如：优、良、中、差、劣。

等级的设立和命名应便于构造和理解主观评价的心理尺度，这种尺度应是两极尺度，应避免多极尺度和模糊不清的尺度构造。

8.2.5 评述

以专业术语准确地描述听音评价中对被评价设备的具体感觉，包括正面的和负面的印象。对突出问题可以做比较详细的描述和分析。

8.2.6 备注

在此项评价属性上所遇到的问题，包括对等级设定的无所适从等。备注也可以是在此属性上被评价对象的某种得分（例如百分制得分）。备注信息在后续的数据分析中可作为可信度和有效性分析的依据。

8.2.7 评价尺度构成

在评价过程中，只要求听音员就某一音质属性在对应的等级上进行选择勾画，并不给出具体的分数。对所有听音员和所有被评测设备的结果运用统计学原理进行处理，最终给出的评价尺度值并不具有绝对的数值意义，但是给出的结果包括：

- a) 在某个音质属性的心理尺度上各个等级的边界和范围；
- b) 被评价设备在此音质属性的心理尺度上的位置。

附录 A 给出了一种构成等级评定法的评价尺度的统计处理方法。

8.3 成对比较法

8.3.1 适用任务

成对比较法适用于对音质差异比较小的设备进行相对比较的场合。成对比较法所给出的结果是各个被评价设备音质水平的相对排序和相对关系，并不具有绝对的意义。当被比较的设备超过 5 个时，成对比较法所需要的时间会大幅增加，比等级评定法需要的时间可能要更多。

成对比较法适合于对设备的音质水平进行更加细致的相对比较和评价。

8.3.2 听音前的练习

相对于等级评定法，成对比较法更适合于普通的听音员或未受过专业训练的听音员。在听音试验前，应进行基本的说明和练习。

8.3.3 评分表格

成对比较法中宜使用的评分记录表格示如表 5 所示。听音员只需就被评价的音质指标，在成对播放比较的两个声音（A、B）中根据自己的感觉做出偏爱选择。

表 5 成对比较法评分记录表格

演出用专业音响设备音质主观评价（成对比较法）用记分表											
评价地点:		设备类型:			听音编号:						
听音人姓名:		性别:	年龄:		职业:						
节目名称:											
听音评价时间:		年	月	日	时	分	-----	时	分		
评价项目		配对序号									
		1	2	3	4	5	6	7	8	
音质属性	丰满	A	A	A	A	A	A	A	A	A
		B	B	B	B	B	B	B	B	B
	清晰	A	A	A	A	A	A	A	A	A
		B	B	B	B	B	B	B	B	B
	明亮	A	A	A	A	A	A	A	A	A
		B	B	B	B	B	B	B	B	B
	柔和	A	A	A	A	A	A	A	A	A
		B	B	B	B	B	B	B	B	B
	力度	A	A	A	A	A	A	A	A	A
		B	B	B	B	B	B	B	B	B
	平衡	A	A	A	A	A	A	A	A	A
		B	B	B	B	B	B	B	B	B
	真实	A	A	A	A	A	A	A	A	A
		B	B	B	B	B	B	B	B	B
	A	A	A	A	A	A	A	A	A
		B	B	B	B	B	B	B	B	B
	总体效果		A	A	A	A	A	A	A	A
			B	B	B	B	B	B	B	B
评述											
注：如果有N个被评价的设备，则有 $N \times (N-1)$ 个配对要进行比较，若不考虑放音的先后顺序对评价的影响，则有 $N \times (N-1) / 2$ 个配对需要进行比较。											

8.3.4 评价尺度构成

在评价过程中，只要求听音员就某一音质属性在顺序播放的两个声音样本上（A、B）选择更偏爱的那个，并不给出具体的分数和等级。

对所有听音员和所有被评测设备的结果运用统计学原理进行处理，最终给出的评价尺度值并不具有绝对的数值意义，给出的结果是被评价设备在此音质属性的心理尺度上的相对位置；根据以上信息就可

以得到各个设备音质水平的相对关系。

附录 B 给出了一种构成成对比较法的评价尺度的统计处理方法。

8.4 可靠性分析

对缺乏足够经验和训练或者听音试验时状态欠佳的听音员应进行评价数据的可靠性检验。因此，听音试验可能需要重复进行。可通过某一听音员对某一评价对象重复评定的一致性来检验评价数据的可靠性，可靠的听音员在评价过程中进行重复选择的一致性宜在 75% 以上。

8.5 有效性分析

当有多个评价指标同时进行评价，并且对评价指标的含义可能出现歧义时，应对每个指标评价结果的有效性进行检验。

评价结果的有效性可以通过听音员个体间的可靠性来考查，不同听音员对同一个评价指标评定的一致性在一定程度上反映了评价结果的有效性。如果听音员两两之间的一致性不能达到一个合理的水平，评价结果的有效性难以保证。

当各个评价指标之间的含义明显不同，而且其表现也明显不受相互影响时，可通过评价指标之间的相关性分析来检验评价结果的有效性。对两个完全独立的评价指标，如果其评价数据表现出明显的相关性，则评价结果的有效性值得怀疑。

9 评价试验中的辅助设备

9.1 一般规定

音质主观评价需要放音，放音设备是一个系统的设备组合。除了被评价的设备作为放音系统的一个环节外，其它设备都是必要的辅助设备。为了保证音质主观评价的可靠性，辅助设备的音质应得到保证，不应对被评价设备的音质造成影响。

9.2 传声器评价时的辅助设备

9.2.1 录音环境

传声器的评价采取同步录放的方式进行，录音环境的声学条件不应影响对传声器音质的评价。如果不能在消声室录音，则录音环境的混响时间不应长于听音室的混响时间。

9.2.2 音源

录音用的音源应是可重复的。如果不具备可重复的自然音源，则可以使用扬声器系统作为音源，但此时扬声器系统的各项客观性能指标不应劣于听音用的扬声器系统的各项客观性能指标。

9.2.3 听音扬声器系统

听音扬声器系统应采用监听扬声器系统，其音质特点应能充分反映音源信号的特点。

9.3 扬声器系统评价时的辅助设备

9.3.1 音源设备

音源设备应采用专业级的设备，宜符合 5.3 规定的标准节目源和与其配套的音源设备。

9.3.2 功率放大器

对无源扬声器系统的评价，应采用专业级的功率放大器。功率放大器的各项电声性能指标应与被评价扬声器系统匹配，功率放大器的音质水平应达到优的等级。

9.4 功率放大器评价时的辅助设备

9.4.1 音源设备

应符合 9.3.1 的规定。

9.4.2 听音扬声器系统

评价功率放大器的音质水平时，放音扬声器应采用专业级的监听扬声器系统。扬声器系统的各项电声性能指标应与功率放大器匹配。扬声器系统本身的各项音质水平应达到优的等级。

9.5 其它设备评价时的辅助设备

当被评价的设备不属于 9.2~9.4 规定的范围时，其评价过程中的辅助设备的要求可以参考 9.2~9.4 中对辅助设备的规定。基本要求是辅助设备的电声性能应相互匹配，辅助设备的音质不应影响对被评测设备的判断。

10 评价结果的报告

对音质主观评价的结果应做出专门的报告，至少应包括以下内容：

- a) 被评价设备的信息和电声性能客观指标；
- b) 听音评价的时间；
- c) 听音室的声学特性；
- d) 听音室的环境条件；
- e) 听音小组的构成；
- f) 节目源信息；
- g) 设备的摆放和听音员的布置；
- h) 辅助设备的电声性能指标和音质等级；
- i) 听音声级；
- j) 听音评价的方法；
- k) 听音评价的步骤、程序和指导用语；
- l) 数据的统计处理方法；
- m) 评价结果的可信度和有效性分析；
- n) 被评价设备的得分（尺度值）和评语。

附 录 A
(资料性附录)
等级评定方法的尺度构成方法

设有 r 个对象 a_1, a_2, \dots, a_r ，以及按从小到大或从好到坏等顺序排列的 m 个等级（类别） C_1, C_2, \dots, C_m 。一组评价人将被评价的对象在这 m 个等级上进行分类排序。以这些分类数据为基础进行尺度化的方法为等级评定法。当只有一名评价人时，则反复进行多次测量。等级评定法中的等级数目，在两极尺度的情况下，多采用 5、7、9 等奇数。其中中间的等级多设计为中立项目。

以下通过一个具体的例子来说明等级评定法的基本原理。

50 个人就 6 个对象在 5 个等级上进行选择。等级分类统计如表 A.1 所示，表 A.2 是根据表 A.1 所得出的左侧等级的累积人数，表 A.3 是由表 A.2 得出的累积百分数表，这个表是等级评定法尺度构成的基础。

表 A.1 意见分类统计

意见	优 C_1	良 C_2	中 C_3	差 C_4	劣 C_5	合计 (人)
a_1	35	9	4	1	1	50
a_2	29	12	5	2	2	50
a_3	25	13	6	3	3	50
a_4	19	14	8	3	6	50
a_5	12	13	10	5	10	50
a_6	4	8	9	6	23	50

表 A.2 表 A.1 中左侧等级的累计人数

意见	优 C_1	良 C_2	中 C_3	差 C_4	劣 C_5
a_1	35	44	48	49	50
a_2	29	41	46	48	50
a_3	25	38	44	47	50
a_4	19	33	41	44	50
a_5	12	25	35	40	50
a_6	4	12	21	27	50

表 A.3 累积百分数

意见	优 C_1	良 C_2	中 C_3	差 C_4	劣 C_5

α_1	0.70	0.88	0.96	0.98	1.00
α_2	0.58	0.82	0.92	0.96	1.00
α_3	0.50	0.76	0.88	0.94	1.00
α_4	0.38	0.66	0.82	0.88	1.00
α_5	0.24	0.50	0.70	0.80	1.00
α_6	0.08	0.24	0.42	0.54	1.00

表 A.4 表 A.3 的符号化表

意见	优 C ₁	良 C ₂	中 C ₃	差 C ₄	劣 C ₅
α_1	P ₁₁	P ₂₁	P ₃₁	P ₄₁	P ₅₁
α_2	P ₁₂	P ₂₂	P ₃₂	P ₄₂	P ₅₂
α_3	P ₁₃	P ₂₃	P ₃₃	P ₄₃	P ₅₃
α_4	P ₁₄	P ₂₄	P ₃₄	P ₄₄	P ₅₄
α_5	P ₁₅	P ₂₅	P ₃₅	P ₄₅	P ₅₅
α_6	P ₁₆	P ₂₆	P ₃₆	P ₄₆	P ₅₆

在心理测量模型中，对象 a_i 的偏爱度是偏爱性尺度上的随机变量 X_i ，假定这个变量是正态分布，其均值和散布为 (μ_i, σ_i) 。

这里设定第 g 号等级和第 $g+1$ 号等级的分界线是主观偏爱性尺度上的随机变量 T_g ，它服从正态分布，均值和方差为 (t_g, σ_g^2) 。(g=1, 2, 3, ..., m-1; 不考虑第一个等级的下限和最后一个等级的上限)。

由此，对被评价对象 a_j ，如果 $T_g - X_j \geq 0$ ，则将其归类为第 g 号以下的等级。如果 $T_g - X_j < 0$ ，则将其归类为第 $g+1$ 号以上的等级。

根据与成对比较法的统计数学原理相类似的方法，可以得到关系式：

$$t_g - f(a_j) = z_{gj} \dots\dots\dots (A. 1)$$

其中 z_{gj} 是对象 a_j 在第 g 号等级以下的累积百分率 p_{gj} 所对应的正态分布的横坐标。即：

$$p_{gj} = \phi(z_{gj}) \dots\dots\dots (A. 2)$$

将 (A. 1) 式用表 A. 5 表示如下：

表 A.5 等级判断模型

$a_i \backslash C_i$	C_1	C_2	...	C_{m-1}
a_1	$t_1 - f(a_1) = z_{11}$	$t_2 - f(a_1) = z_{21}$...	$t_{m-1} - f(a_1) = z_{m-11}$
a_2	$t_1 - f(a_2) = z_{12}$	$t_2 - f(a_2) = z_{22}$...	$t_{m-1} - f(a_2) = z_{m-12}$
...
a_r	$t_1 - f(a_r) = z_{1r}$	$t_2 - f(a_r) = z_{2r}$...	$t_{m-1} - f(a_r) = z_{m-1r}$
合计	$rt_1 - \sum f(a_j)$	$rt_2 - \sum f(a_j)$...	$rt_{m-1} - \sum f(a_j)$
平均	$t_1 - \frac{1}{r} \sum f(a_j)$	$t_2 - \frac{1}{r} \sum f(a_j)$...	$t_{m-1} - \frac{1}{r} \sum f(a_j)$

对第 g 列求和得：

$$rt_g - \sum_j f(a_j) = \sum_j z_{gj} \dots\dots\dots (A. 3)$$

两边求平均：

$$t_g - \frac{1}{r} \sum_j f(a_j) = \bar{z}_g \dots\dots\dots (A. 4)$$

由此得：

$$t_g - t_{g-1} = \bar{z}_g - \bar{z}_{g-1} \dots\dots\dots (A. 5)$$

特别地，当等级数是奇数时， $m=2k+1$ ，第 $k+1$ 号等级在中央，将 t_k 和 t_{k+1} 的中央定为 0，

$$c = \frac{\bar{z}_k + \bar{z}_{k+1}}{2} = \frac{t_k + t_{k+1}}{2} - \frac{1}{r} \sum_j f(a_j) \dots\dots\dots (A. 6)$$

从而得
$$t_g^* = \bar{z}_g - c \dots\dots\dots (A. 7)$$

(A. 5) 式和 (A. 7) 式的意义在于说明，等级的边界和宽度并不是一个固定的值，也是一个统计变量，与被评价的样本和全体参评人的整体统计特性有关，当然也受到实验程序和实验中其它因素的影响。

对前面所举的意见分类的例子，表 A. 6 是由表 A. 3 做成的 z_{gj} 表

表 A.6 表 A.3 的 z_{gj} 表

$a_i \backslash t_j$	t_1	t_2	t_3	t_4	合计	平均	$f^*(a_j)$
a_1	0.52	1.18	1.75	2.05	5.50	1.375	-1.472
a_2	0.20	0.92	1.41	1.75	4.28	1.070	-1.167
a_3	0.00	0.71	1.18	1.55	3.44	0.860	-0.957
a_4	-0.31	0.41	0.92	1.18	2.20	0.550	-0.647
a_5	-0.71	0.00	0.52	0.84	0.65	0.163	-0.260
a_6	-1.41	-0.71	-0.20	0.10	-2.22	-0.555	0.458
合计	-1.71	2.51	5.58	7.74	13.85		
平均	-0.285	0.418	0.930	1.245		0.577	
t_g^*	-0.959	-0.256	0.256	0.571		-0.097	

在此场合， C_3 是中央等级，因而：

$$c = \frac{0.418 + 0.930}{2} = 0.674$$

从而：

$$t_g^* = \bar{z}_g - 0.674$$

各个等级的界限为：

等级	下限	上限
C_1 :		-0.959
C_2 :	-0.959	-0.256
C_3 :	-0.256	0.256
C_4 :	0.256	0.571
C_5 :	0.571	

各个对象的尺度值可以从表 A.5 的横向求和得到。对表 A.5 中第 j 行的合计有

$$\sum_g t_g - (m-1)f(a_j) = \sum_g z_{gj} \dots\dots\dots (A.8)$$

记 $\bar{t} = (\sum_g t_g)/(m-1)$, $\bar{z}_j = (\sum_g z_{gj})/(m-1)$

则 $f(a_j) = \bar{t} - \bar{z}_j \dots\dots\dots (A. 9)$

当使用 t_g^* 时, $\bar{t}^* = (\sum_g t_g^*) / (m - 1)$ 尺度值计算如下:

$$f^*(a_j) = \bar{t}^* - \bar{z}_j \dots\dots\dots (A. 10)$$

对表 A. 6 的场合, $\bar{t}^* = -0.097$, 则

$$f^*(a_j) = -0.097 - \bar{z}_j$$

等级评定法不但可以给出各个被评价对象的相对评价尺度和排序, 同时也给出了各个等级边界的统计值。除了两端的等级外, 其它等级的宽度并不都是相等的。

附 录 B
(资料性附录)
成对比较法的尺度构成方法

成对比较法构造心理尺度的基本原理是认为心理量是符合正态分布的随机变量。

设有 r 个对象： a_1, a_2, \dots, a_r 将其中任意两个两两配对组合， (a_i, a_j) 组合数目为 $m=n(n-1)/2$ ，若考虑到前后顺序的不同，则组合数目为 $m=n(n-1)$ 。就某个评价指标来说，设 a_i 优于 a_j 的概率为 $P_r(a_i > a_j)$ 。

当实验者是一个人时，可从其 m 次 (a_i, a_j) 的重复比较实验中统计概率值 $P_r(a_i > a_j)$ 。当实验者是一群人时，可从 n 个人的 (a_i, a_j) 的比较实验中统计概率值 $P_r(a_i > a_j)$ 。以 $P_r(a_i > a_j)$ 为基础，可以构造优势和等级判断的心理尺度。

在对象 a_1, a_2, \dots, a_r 中选定一个基准，比如 a_1 ，成对比较法的目的是对每一个对象找出一个反映它优越性的相对心理尺度

$$f(a_i) = F [P_r(a_i > a_1)] \dots\dots\dots (B. 1)$$

这个心理量是某一对象相对选定对象的优越概率的函数。函数 F 满足如下的关系：

$$F[P_r(a_i > a_j)] + F[P_r(a_j > a_1)] = F[P_r(a_i > a_1)] \dots\dots\dots (B. 2)$$

$$F[P_r(a_i > a_j)] = F[P_r(a_i > a_1)] - F[P_r(a_j > a_1)] = f(a_i) - f(a_j) \dots\dots\dots (B. 3)$$

假设 F 的逆函数存在，

$$P_r(a_i > a_j) = F^{-1} [f(a_i) - f(a_j)] \dots\dots\dots (B. 4)$$

心理变量是服从正态分布的随机变量 $X(\mu, \sigma)$ ，而对象 a_i 的优势尺度是这个随机变量的均值。

$$f(a_i) = \mu_i \quad (i = 1, 2, \dots, r) \dots\dots\dots (B. 5)$$

根据以上的假设， $(a_i > a_j)$ 的判断可通过 $X_i - X_j$ 的正负判断来实现。而 $X_i - X_j$ 也是一个正态分布的随机变量，其均值和方差为

$$E(X_i - X_j) = \mu_i - \mu_j \dots\dots\dots (B. 6)$$

$$\sigma^2(X_i - X_j) = \sigma_i^2 + \sigma_j^2 - 2\sigma_{ij} \dots\dots\dots (B. 7)$$

将 $X_i - X_j$ 化成均值为 0 方差为 1 的随机变量

$$Z_{ij} = \frac{(X_i - X_j) - (\mu_i - \mu_j)}{\sqrt{\sigma_i^2 + \sigma_j^2 - 2\sigma_{ij}}} \dots\dots\dots (B. 8)$$

设 $X_i - X_j = 0$ 时的 Z_{ij} 的值为 $-z_{ij}$

$$-z_{ij} = -\frac{\mu_i - \mu_j}{\sqrt{\sigma_i^2 + \sigma_j^2 - 2\sigma_{ij}}} \dots\dots\dots (B. 9)$$

$(a_i > a_j)$ 的概率可由 $X_i - X_j > 0$ 的概率求得

$$P_r(a_i > a_j) = \frac{1}{\sqrt{2\pi}} \int_{-z_{ij}}^{\infty} e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_{ij}} e^{-t^2/2} dt = \phi(z_{ij}) \dots\dots (B. 10)$$

同时，由前面的假设可以得到

$$f(a_i) - f(a_j) = z_{ij} \sqrt{\sigma_i^2 + \sigma_j^2 - 2\sigma_{ij}} \dots\dots\dots (B. 11)$$

此公式是成对比较法的基本心理尺度构成公式。但是只凭以上公式还不能计算心理尺度，所以引入以下的假设条件

$$\sigma_i^2 = \sigma^2 \quad \rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j} = \rho \quad \sigma \sqrt{2(1-\rho)} = 1 \dots\dots\dots (B. 12)$$

这个假设条件的意义是不同的被评价对象的统计特性是一样的，它们两两之间就心理尺度的离散关系也是恒定的。则

$$f(a_i) - f(a_j) = z_{ij} \dots\dots\dots (B. 13)$$

取 $j=1$ ，并且规定 $f(a_1) = 0$ ，则

$$f(a_i) = z_{i1} \dots\dots\dots (B. 14)$$

例如，若 $P_r(a_i > a_1) = 0.7 = \phi(0.524)$ ，则 $f(a_i) = 0.524$

成对比较法心理尺度构造的具体过程可以通过以下五个对象的实验数据示例说明。设有五个对象，两两比较得表 B. 1。

表 B. 1 概率 $P_r(a_i > a_j)$

$j \backslash i$	a_1	a_2	a_3	a_4	a_5
a_1	$P_r(a_1 > a_1)$	$P_r(a_2 > a_1)$	$P_r(a_3 > a_1)$	$P_r(a_4 > a_1)$	$P_r(a_5 > a_1)$
a_2	$P_r(a_1 > a_2)$	$P_r(a_2 > a_2)$	$P_r(a_3 > a_2)$	$P_r(a_4 > a_2)$	$P_r(a_5 > a_2)$
a_3	$P_r(a_1 > a_3)$	$P_r(a_2 > a_3)$	$P_r(a_3 > a_3)$	$P_r(a_4 > a_3)$	$P_r(a_5 > a_3)$
a_4	$P_r(a_1 > a_4)$	$P_r(a_2 > a_4)$	$P_r(a_3 > a_4)$	$P_r(a_4 > a_4)$	$P_r(a_5 > a_4)$
a_5	$P_r(a_1 > a_5)$	$P_r(a_2 > a_5)$	$P_r(a_3 > a_5)$	$P_r(a_4 > a_5)$	$P_r(a_5 > a_5)$
一般项	$P_r(a_1 > a_j)$	$P_r(a_2 > a_j)$	$P_r(a_3 > a_j)$	$P_r(a_4 > a_j)$	$P_r(a_5 > a_j)$

此表根据正态分布表，变换为表 B. 2。

表 B. 2 由 $P_r(a_i > a_j) = \phi(z_{ij})$ 所导出的正态分布变量 z_{ij}

$j \backslash i$	a_1	a_2	a_3	a_4	a_5
a_1	z_{11}	z_{21}	z_{31}	z_{41}	z_{51}
a_2	z_{12}	z_{22}	z_{32}	z_{42}	z_{52}
a_3	z_{13}	z_{23}	z_{33}	z_{43}	z_{53}
a_4	z_{14}	z_{24}	z_{34}	z_{44}	z_{54}
a_5	z_{15}	z_{25}	z_{35}	z_{45}	z_{55}
合计	$\sum z_{1j}$	$\sum z_{2j}$	$\sum z_{3j}$	$\sum z_{4j}$	$\sum z_{5j}$
平均	\bar{z}_1	\bar{z}_2	\bar{z}_3	\bar{z}_4	\bar{z}_5
$f(a_i)$	$\bar{z}_1 - \bar{z}_1$	$\bar{z}_2 - \bar{z}_1$	$\bar{z}_3 - \bar{z}_1$	$\bar{z}_4 - \bar{z}_1$	$\bar{z}_5 - \bar{z}_1$

上表可以用更直接的形式表示出来，如表 B. 3 所示。

表 B. 3 表 B. 2 的等价表示

$i \backslash j$	a_1	a_2	a_3	a_4	a_5
a_1	$f(a_1) - f(a_1)$	$f(a_2) - f(a_1)$	$f(a_3) - f(a_1)$	$f(a_4) - f(a_1)$	$f(a_5) - f(a_1)$
a_2	$f(a_1) - f(a_2)$	$f(a_2) - f(a_2)$	$f(a_3) - f(a_2)$	$f(a_4) - f(a_2)$	$f(a_5) - f(a_2)$
a_3	$f(a_1) - f(a_3)$	$f(a_2) - f(a_3)$	$f(a_3) - f(a_3)$	$f(a_4) - f(a_3)$	$f(a_5) - f(a_3)$
a_4	$f(a_1) - f(a_4)$	$f(a_2) - f(a_4)$	$f(a_3) - f(a_4)$	$f(a_4) - f(a_4)$	$f(a_5) - f(a_4)$
a_5	$f(a_1) - f(a_5)$	$f(a_2) - f(a_5)$	$f(a_3) - f(a_5)$	$f(a_4) - f(a_5)$	$f(a_5) - f(a_5)$
合计	$5f(a_1) - \sum f(a_i)$	$5f(a_2) - \sum f(a_i)$	$5f(a_3) - \sum f(a_i)$	$5f(a_4) - \sum f(a_i)$	$5f(a_5) - \sum f(a_i)$
平均	$f(a_1) - \frac{1}{5} \sum f(a_i)$	$f(a_2) - \frac{1}{5} \sum f(a_i)$	$f(a_3) - \frac{1}{5} \sum f(a_i)$	$f(a_4) - \frac{1}{5} \sum f(a_i)$	$f(a_5) - \frac{1}{5} \sum f(a_i)$

概率分布到正态随机变量的变换可查正态分布变换表。上表中的最后一行就是所要构造的被评价对象的心理尺度值。