Consultancy report **Ref: 6824 R03**

Initial Circulation:

National Physical Laboratory:
Dr R Barham, Project Manager
Mr T Sherwood

ISVR Consulting:
Dr M C Lower
Mr S J C Dyne, Manager

Prepared by:

BW Lawton
Senior Consultant

In accordance with our quality assurance procedures, this report is countersigned by:

Mr S J C Dyne
Manager, ISVR Consulting

Interim report

Variation of young normal hearing thresholds measured using various patterns of audiometric earphone: Implications for the acoustic coupler and the ear simulator (artificial ear)

September 2004

ISVR Consulting
Institute of Sound and Vibration Research
University of Southampton
Highfield
Southampton SO17 1BJ

Tel: 023 8059 2162 Fax: 023 8059 2728 Email: consultancy@isvr.co.uk

Summary

Calibration of audiometers requires the use of specific devices for the measurement of acoustic output from supra-aural audiometric earphones. The IEC 60318-3 acoustic coupler is specified for measurement of the sound output from only two earphones: the TDH39 and the Beyer DT48 (not known in British practice). The other calibration device is the IEC 60318-1 ear simulator (sometimes called an artificial ear) intended for all other patterns of audiometric earphone.

Is it worthwhile retaining the coupler (for calibration of the TDH39 only) while all other patterns of earphone, including circumaural and high-frequency earphones, require the use of an ear simulator? This study examines the *variation* of hearing threshold measurements using the TDH39, and using other patterns of earphone. If TDH39 measurements on samples of otologically normal young persons were to show *greater* spread than thresholds using other earphones, then the less reliable TDH39 might be retired from service, along with its required acoustic coupler.

A number of independent studies of young normal thresholds (and the associated standard deviations) were examined. The threshold samples from TDH39 earphones showed *smaller* variation than did samples using other patterns. Therefore, the TDH39 earphone and its acoustic coupler need not be retired on grounds of unreliability. Nevertheless, there are benefits to be gained from calibrating the TDH39 on the ear simulator (artificial ear), in contravention of the appropriate British and International Standards. To achieve these benefits, the Standards would have to be amended.

Table of Contents

1. Introduction

Present international practice for the calibration of audiometers requires the use of specific devices for the measurement of acoustic output from supra-aural audiometric earphones. The acoustic coupler is the older device, specified by International Electrotechnical Commission standard IEC 60318-3. It is intended for measurement of sound output from only two earphones: the TDH39 with specified cushions, and the Beyer DT48 (also used in the film and broadcast industries). The other device is the IEC 60318-1 ear simulator (sometimes called an artificial ear) intended for all other patterns of supra-aural earphone. The “audiometric zero” for supra-aural earphones is given in ISO 389-1. This Standard gives Reference Equivalent Threshold Sound Pressure Levels (RETSPLs) for individual test frequencies, from the TDH39 and from the Beyer, into the acoustic coupler. Separate RETSPLs are given for all other patterns of earphone, with output measured in the ear simulator. This paper addresses the question: Is it worthwhile retaining the coupler (for calibration of only two earphones) while all other earphone patterns require the use of the ear simulator?

The standard IEC 60318-3 describes an acoustic coupler for loading certain supra-aural audiometric earphones with a specified acoustic impedance in the frequency range of 125 Hz to 8 kHz. The coupler is a simple cavity, of specified shape and volume containing a calibrated microphone to measure the acoustic pressure developed within the cavity. The coupler may be constructed with one critical dimension chosen from values dependent upon the equivalent volume of the measurement microphone. Alternatively, the critical dimension may assume a specified fixed value, in which case the coupler is deemed equivalent to the US National Bureau of Standards 9A coupler (ANSI S3.7-1995).

American Standard ANSI S3.6-1996 illustrates a major shortcoming of the use of an arbitrary calibration cavity, not linked to the acoustic properties of the human ear. Each pattern of audiometric earphone must undergo subjective testing to establish an RETSPL in the NBS 9A coupler, for each audiometric frequency. The ANSI standard gives separate RETSPL values for the TDH39 earphone and for TDH49 and 50 earphones; the values are close but not identical for the 39 and 49/50 patterns. In addition, ANSI S3.6 also quotes competing RETSPL values for any “TDH type” earphone calibrated on the IEC 60318-1 ear simulator; these competing values for audiometric zero open the Standard to interpretation and choice.

In principle, RETSPL values would be rendered independent of earphone pattern if they were specified in an artificial ear having acoustical properties exactly simulating those of the average human ear. A device designed with this aim in view was first standardised in 1970 by the IEC; the present manifestation of this artificial ear is found in IEC 60318-1:1998. This Standard specifies the general characteristics of an ear simulator with overall acoustic impedance approximating that of the average human ear over the range 20 Hz to 10 kHz. Within the simulator, the plane of the microphone diaphragm is understood to represent the entrance of the mean human ear canal. These simulator features support the intention that any earphone might be calibrated to a universal audiometric zero (independent of earphone pattern).

It would simplify compliance with ISO and IEC standards if only one calibration device were specified for use with all patterns of supra-aural earphone. Such a state of affairs would necessitate retiring one calibration device: considering its arbitrary shape and volume, the acoustic coupler is the obvious choice.

This study presents an analysis of existing data samples of thresholds from young otologically normal ears, as measured using different earphone patterns, calibrated on the acoustic coupler or the artificial ear. The aim is to demonstrate differences attributable to the combination of earphone and calibration device; unfavourable descriptive statistics would provide a valid reason to retire one earphone-and-device in favour of another combination.

2. Method

Previous researchers have addressed the coupler -v- simulator question by electroacoustic investigations of the calibration devices themselves. This study takes a different approach, concentrating upon measurements from people rather than measurements on devices. Hearing threshold is the underlying physiological phenomenon of interest; audiometric equipment and calibration devices are simply tools to aid in the numerical description of that phenomenon. This paper will consider thresholds from samples of otologically normal young persons, from whom Reference Equivalent Threshold Sound Pressure Level values are derived. The accuracy of standardised RETSPL values will **not** be considered here. There are known inaccuracies in the RETSPLs for the TDH39 as calibrated on the acoustic coupler; these inaccuracies are manifest as non-zero threshold values for groups of otologically normal young persons. The greatest error is found at 6 kHz, where the coupler is said to interact with irregularities in the TDH39 frequency response. This error shows itself as a false hearing loss, a 6 kHz notch even for otologically normal young persons; the notch may be put right by use of a simple correction factor.

For present purposes, the **variation** of threshold measurements will indicate the more stable combination of earphone and calibration device. A number of independent samples of young normal thresholds (and the associated standard deviations) will be examined. These samples may be divided into four classes, depending on earphone and calibration device used, as may be seen in the following matrix:

		earphone type	
		TDH39	all others
calibration device employed	60318-3 (coupler)	(A) standard	(B)
	60318-1 (simulator)	(C)	(D) standard

Cell A represents standard practice for the TDH39 (and Beyer DT48). This cell will contain a large number of studies, each giving the standard deviations for threshold samples observed at each frequency. Cell D represents standard practice for all other patterns of earphone. Cell B represents calibration practice in a few countries where ISO standards do not guide national practice. Cell C is inconsistent with international audiometric practice.

If sufficient data were available to fill each cell of the matrix, then it might be possible to offer conclusions on the differences between patterns of earphone, and likewise between calibration devices. Critical analysis of the international literature shows that cells A, B and D together include virtually all studies; cell C contains only a few studies (with a correspondingly small number of ears). Such an imbalance between cells militates against simultaneous consideration of differences between earphones and calibration devices. A simpler approach will be made, comparing the sample variation of TDH39 threshold studies with those of investigations using other patterns of earphone. If the spread of young normal thresholds is greater for the TDH39, then that earphone might be taken out of use, leaving little function for the acoustic coupler in international audiometric practice.

The spread of observations from a sample may be specified as variance, or as standard deviation (the square root of the sample variance). The standard deviation of thresholds from a single sample of young otologically normal ears is one estimate of the standard deviation of thresholds from the population of **all** young otologically normal ears. If estimates were available from many samples, then the distribution of those standard deviations could be used as a rough-and-ready test to judge if other independent samples were drawn from the same underlying population. It is known that the variances of samples (size n), drawn from a population (size N, with $n < N$), will follow a Chi-squared distribution such as may be seen in Figure 1. The horizontal axis is sample variance, which is always positive. The vertical axis is probability density, giving the likelihood of observing any given value of variance for a sample drawn from a larger population. The area under the probability density function is equal to unity. Another view of the range of variances from samples drawn from a larger population may be seen in Figure 2. Here the horizontal axis is again variance, always positive; the vertical axis is cumulative probability, from zero to one. For any particular value of sample variance, the cumulative probability curve gives the likelihood of observing a variance of that value or less. The cumulative probability curve climbs asymptotically toward unity, representing the variances of samples drawn from a single population. Observations from a different population will follow a different cumulative probability curve.

Bearing in mind these considerations of the variances of samples, a set of hypotheses may be formulated for application to the threshold variation data (at individual audiometric frequencies) from samples of otologically normal young ears. A one-sided test offers a radical conclusion.

> null hypothesis: threshold variation with TDH39 earphones is not greater than threshold variation with other patterns
>
> alternative hypothesis: threshold variation with TDH39 earphones is greater than variation with other earphones

Rejection of the one-sided null hypothesis, on finding a larger variation for the TDH39 earphone, would indicate that the TDH39 couples erratically to human external ears, thus giving more wide-ranging thresholds than other patterns. Such a result would support withdrawal of the TDH39 (and its acoustic coupler) from audiometric practice, to be replaced by other patterns of earphone (calibrated on the ear simulator) associated with narrower threshold variation.

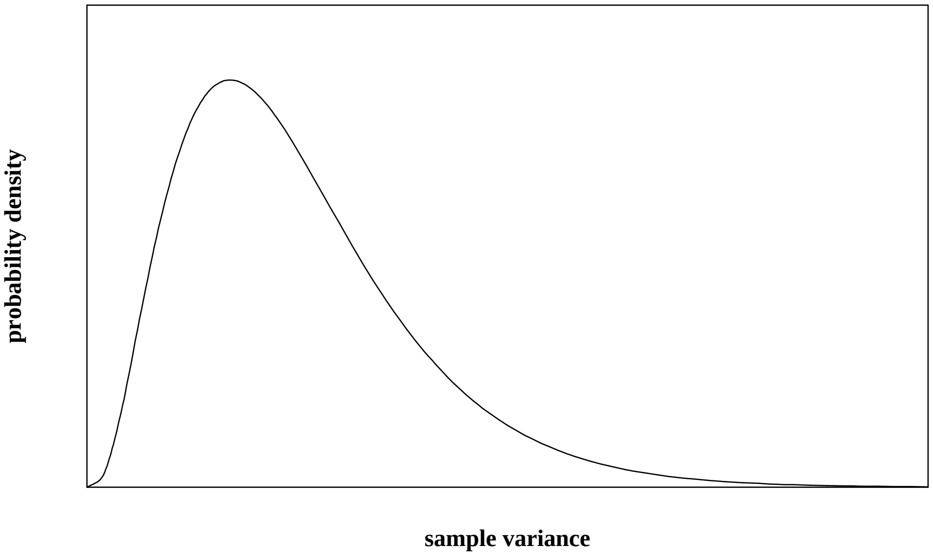

Figure 1. If all possible samples, size n, are drawn from a larger population, size N, then the variance of any sample may range from zero upwards; the distribution of all possible sample variances will follow the Chi-squared distribution. The probability density curve illustrates a general example of this distribution.

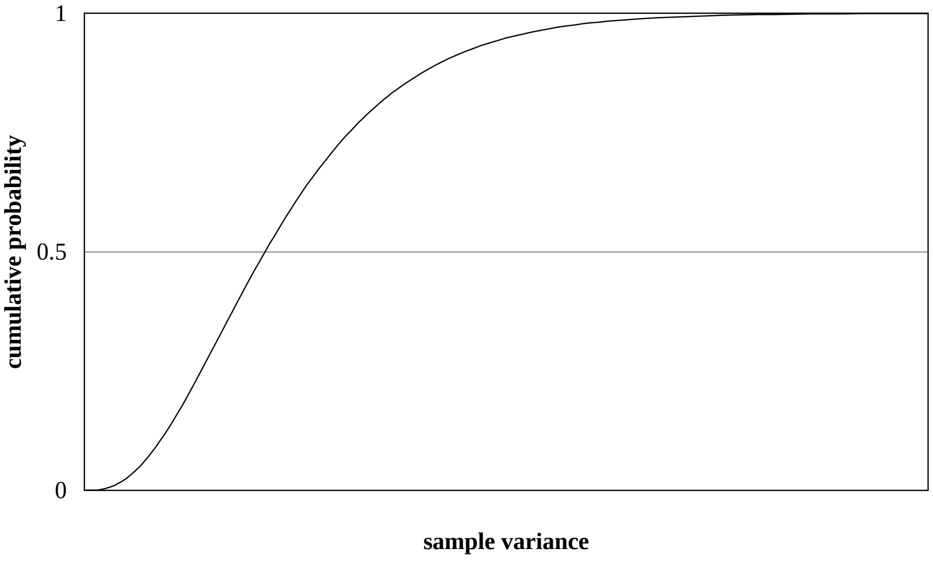

Figure 2. The Chi-squared distribution may also be represented in terms of cumulative probability. For any particular value of sample variance, the cumulative probability curve gives the likelihood of observing a variance of that value or less. The cumulative probability curve climbs asymptotically toward unity.

3. The Data

Audiological research reports, in English and covering a date range from 1959 to 2001, were searched for hearing threshold data from samples of young persons screened to match the conditions necessary and sufficient for otological normality. Twenty-three suitable data reports were found, although the search may not have been exhaustive. In all, threshold statistics were available for more than 5 100 ears, spread over the range of audiometric frequency and earphone pattern.

A specimen datafield is presented in Table 1. Arlinger (1991) gives the descriptive statistics representing Hearing Threshold Levels (in decibels Hearing Level, dB HL) over frequency, from 30 otologically normal subjects (18 males, 12 females) aged 18-26 years; the data given represent all 60 ears. The mean thresholds at each audiometric frequency may be seen in Figure 3, a group audiogram determined using TDH39 earphones calibrated on the acoustic coupler. For the audiometric frequencies 125 Hz to 4 kHz, the mean threshold is within ± 2 dB of the 0 dB HL line, representing the audiometric zero for young, otologically normal persons. Such small variation has little practical effect. At 6 kHz, however, the young people show a “phantom” hearing loss of 9 dB from the audiometric zero. For young normal ears, this threshold would be recorded as 10 dB HL during manual, pure-tone audiometry. For older ears, possibly with hearing pathology, the threshold recorded using a TDH39 would be inflated by 5-10 dB from the same threshold measured using a different pattern of earphone.

This 6 kHz notch in young ears (or irregularity in older ears) is typical of all samples tested with the TDH39 earphone calibrated using the acoustic coupler. It is the result of an incorrect (optimistically low) RETSPL in the calibration standard. The variation of threshold samples is, however, unaffected by the RETSPL value: the threshold variation resides in the subject sample and the audiometric test methods applied to those subjects, not the calibration standard.

Table 2 gives the standard deviations (in dB) of Hearing Threshold Levels determined with the TDH39 earphone, as reported in fifteen independent studies using screened young persons ranging in age from 10 to 30 years. The left-hand column lists the references; the adjacent column identifies the particular data within each cited reference. The remaining columns give the reported standard deviations of threshold samples, by audiometric frequency. Each entry is an *estimate of the true variation* associated with testing involving the TDH39 earphone, if the population of **all** otologically young persons were to give thresholds at the particular frequency. The bottom row of the Table gives the total number of ears associated with the data for each frequency; these totals range from 1596 for 125 Hz, to 2766 for the mid-frequencies 1, 2 and 4 kHz.

In contrast, Table 3 gives the standard deviations (in dB) of Hearing Threshold Levels determined with earphones other than the TDH39. Eight of the nine independent studies used TDH49 or TDH50 earphones to determine thresholds in screened samples of children and young people ranging in age from 10 to 31 years. As before, the main body of the Table gives the reported standard deviations of threshold samples, by audiometric frequency (as estimates of the underlying population variation); the total number of ears ranged from 342 at 125 Hz to 2378 for the mid-frequencies 1 and 2 kHz.

Table 1. An example of the TDH39 data available: Arlinger (1991) gives mean Hearing Threshold Levels (dB HL) and standard deviations (dB) from 30 otologically normal subjects (18 males, 12 females), aged 18-26 years.

source	comment		frequency, Hz								
			125	250	500	1k	2k	3k	4k	6k	8k
Arlinger (1991)	60 ears	mean	1.1	-1.2	-1.7	-1.2	0.9	1.4	0.9	8.9	4.1
		std. dev.	5.7	5.2	5.0	4.9	4.6	5.2	5.1	6.7	6.1

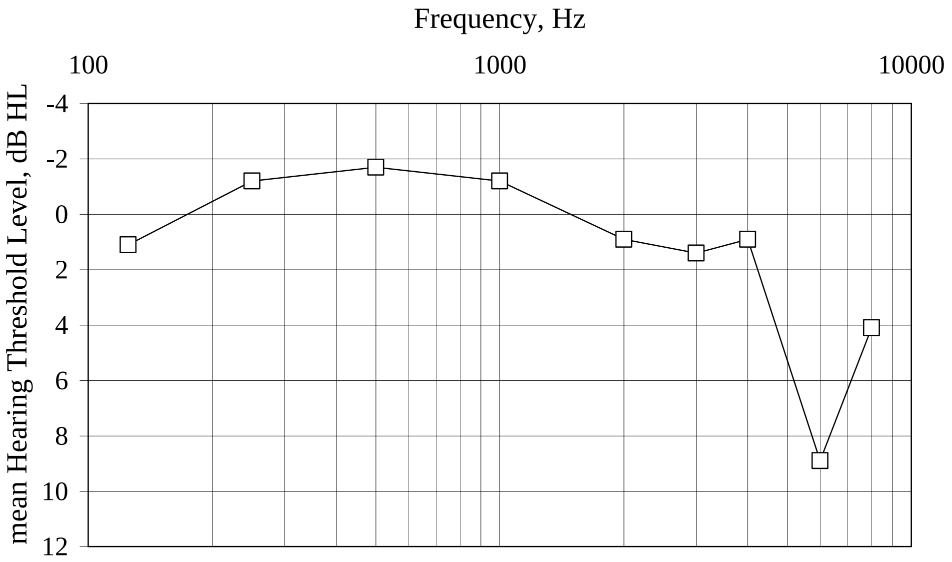

Figure 3. Mean thresholds over 60 ears for the Arlinger (1991) sample of 30 otologically normal subjects, aged 18-26 years.

Table 2. Standard deviations (dB) of hearing thresholds for samples of otologically normal young people, tested with TDH39 earphones.

source	comment	frequency, Hz								
		125	250	500	1k	2k	3k	4k	6k	8k
Hinchcliffe (1959)		7.0	6.2	5.4	5.9	5.8	6.4	6.2	9.2	13.3
Taylor et al. (1965)	control subjects	3.81	3.35	3.27	3.81	3.96	3.88	4.72	6.32	7.77
Rice, Coles (1966)	Rice			5.5	4.7	5.2	7.2	6.5	9.6	
	Coles		5.8	5.4	4.9	5.8	6.2	5.4	8.1	8.0
Taylor, Pearson, Mair (1967)	age 18-24	4.6	4.12	3.86	3.46	3.25	3.61	4.07	4.72	4.59
Kell, Pearson,	males	7.5	7.2	5.9	5.6	6.7	5.8	6.3	8.0	8.8
Taylor (1970)	females	6.3	6.4	6.5	6.1	6.2	5.5	6.4	6.9	7.6
Robinson, Shipton,	right			3.9	4.2	4.6	3.1	6.2	7.0	
Hinchcliffe (1979, 1981)	left-rpt			5.9	5.2	4.3	4.6	3.6	6.4	
Shipton, Robinson (1981)			4.0	4.5	4.4	5.0	5.2	5.8		
Arlinger (1982)				7.8	5.2	7.4	4.8	6.5	7.4	7.6
Lawton, Robinson (1986)			5.7	5.2	4.9	5.6	5.8	6.1	7.6	9.5
Arlinger, Kinnefors (1989)		3.6	3.8	3.5	3.7	3.5	3.7	4.2	5.2	5.2
Arlinger (1991)		5.7	5.2	5.0	4.9	4.6	5.2	5.1	6.7	6.1
Lutman, Davis (1994)			4.2	4.4	4.2	4.6	6.6	6.9	7.8	7.9
Qasem (1996)	TDH39		5.77	4.74	4.93	5.55	6.74	6.2	5.29	6.82
	TDH39P		5.89	4.27	4.87	4.97	6.05	5.78	5.82	6.76
Han, Poulsen (1998)		4.5	4.7	4.4	3.5	4.9	5.3	6.6	6.3	7.1
Rahko-Laitila et al. (2001)	right	5.47	5.21	4.8	4.91	5.51		6.12	7.74	8.38
	left	5.02	4.87	4.57	4.54	5.5		6.23	7.83	8.72
total number of ears		1596	2556	2766	2766	2766	1698	2766	2692	2502

Table 3. Standard deviations (dB) of thresholds for samples of otologically normal young people, using earphones other than the TDH39.

source	earphone	comment	frequency, Hz 125	250	500	1k	2k	3k	4k	6k	8k
Corso (1963)	PDR8	males		4.6	5.8	4.4	6.2	7.2	8.6	13.1	14.0
		females		4.3	5.0	4.5	4.6	5.8	5.4	6.3	8.4
Carter et al. (1978)	TDH49				5.7	5.7	6.0	6.2	7.2	10.2	
Carter et al. (1984)	TDH49	males right			6.8	6.3	4.5	6.6	12.0	16.6	
		males left			8.2	6.6	5.9	7.1	8.2	11.0	
		females right			4.8	4.3	4.9	5.4	6.2	7.8	
		females left			5.6	4.2	4.7	5.4	8.5	7.4	
Carter, Murray (1985)	THD49	10-12 yr children			5.4	5.2	5.4	5.3	6.9	9.6	
		18 yr			4.9	4.8	4.94	6.6	9.05	10.93	
		apprentice 1			9.8	11.4	11.4	11.4	12.3	13.2	
		apprentice 3			5.1	5.0	6.6	6.8	8.8	8.6	
		clerical			6.1	6.2	6.4	5.8		10.2	
Cox, McDaniel (1986)	TDH49			6.5	4.9	5.8	5.0		4.2		
Larson et al. (1988)	TDH50	lab 1	8.22	7.12	5.48	7.67	7.12	6.57	7.12	6.57	7.67
		lab 2	7.12	6.02	6.02	5.48	6.02	6.02	6.02	8.22	7.67
Frank, Vavrek (1992)	TDH49		8.1	7.6	6.7	5.8	6.1	6.8	7.0	7.0	7.2
Qasem (1996)	TDH49			6.63	5.18	4.82	5.17	7.58	5.79	5.29	6.77
	TDH49P			5.89	5.3	5.12	4.83	6.82	5.44	5.3	7.32
Smith et al. (1999)	TDH50	right	5.3	5.5	4.4	5.2	7.1	6.0	6.9	7.7	7.5
		left	5.6	5.9	5.0	5.0	5.7	6.8	8.3	8.8	7.6
total number of ears			342	480	2378	2378	2378	2358	2088	2358	460

4. Data Analysis

Each study considered here gives an estimate of the standard deviation (s.d.) of hearing thresholds, at each audiometric frequency, for the population of young otologically normal ears. Separating the studies by earphone pattern makes it possible to assess differences in the s.d. for each pattern, and thus judge if the samples for TDH39 and other earphones were drawn from the same population. Such an assessment may be made by considering how the sample s.d. values are distributed over the cumulative number of ears tested. Refer to Table 4, which presents 125 Hz data from the cited studies using the TDH39 and other earphone patterns. In the left-hand portion of the Table, the first and second columns give reported sample HTL s.d. values in ascending order, and the number of ears associated with each entry. The third column gives the total number of ears associated with an s.d. equal to or less than the stated value; at the bottom of this column is the total number of ears from all the cited studies. The fourth column gives the cumulative proportion of all ears with an s.d. equal to or less than the stated value (this cumulative relative frequency may be thought of as a practical manifestation of the cumulative probability curve seen in Figure 2). The right-hand side of Table 4 has another four columns, giving a similar analysis for studies using non-39 patterns of earphone.

The data of Table 4 are represented graphically by the cumulative distribution lines of Figure 4. The horizontal axis is standard deviation (of Hearing Threshold Levels at 125 Hz, for each of the investigations cited here) in dB; the vertical axis is the proportion of total number of ears (from all investigations reporting tests at 125 Hz). Studies using TDH39 earphones are represented as square symbols (with the solid line), and other earphone types as circles (with the dashed line). The lines of Figure 4 suggest that the variation associated with TDH39 thresholds is slightly less than that for non-39 patterns: TDH39 median ≈ 5.0 dB, non-39 median ≈ 5.6 dB. This observation suggests (by inspection) that the null hypothesis (variation for the TDH39 not greater than for other patterns) may not be rejected.

A number of following Tables and Figures present similar analyses for threshold variation at other audiometric frequencies, as indicated below.

freq., Hz	Table	Figure
250	5	5
500	6	6
1000	7	7
2000	8	8
3000	9	9
4000	10	10
6000	11	11
8000	12	12

Figures 4 to 12 may be interpreted (by simple inspection) to yield a central tendency of standard deviation for each earphone type. For all frequencies up to and including 6 kHz,

Table 4. Calculation of the cumulative proportion (cum. prop., by number n of otologically normal ears) for standard deviations (s.d.) of hearing threshold samples at 125 Hz, determined using TDH39 earphones or other patterns (not TDH39).

TDH39				not 39			
s.d.	n	cum. n	cum. prop.	s.d	n	cum. n	cum. prop.
3.6	36	36	0.0226	5.3	93	93	0.2719
3.81	58	94	0.0589	5.6	93	186	0.5439
4.5	62	156	0.0977	7.12	30	216	0.6316
4.6	92	248	0.1554	8.1	96	312	0.9123
5.02	534	782	0.4900	8.22	30	342	1.0000
5.47	534	1316	0.8246				
5.7	60	1376	0.8622				
6.3	72	1448	0.9073				
7.0	62	1510	0.9461				
7.5	86	1596	1.0000				

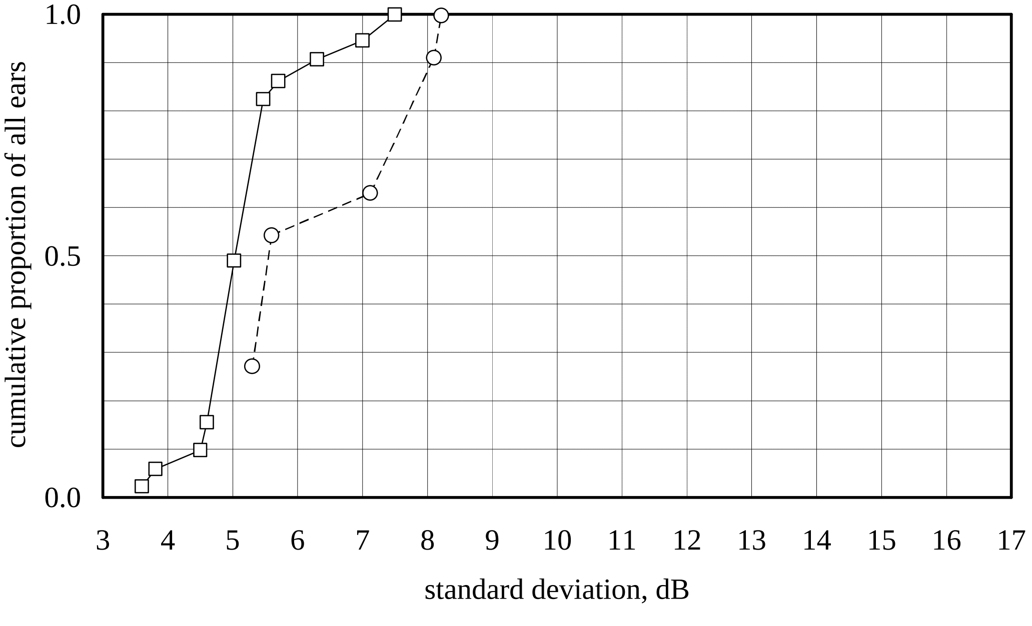

Figure 4. Cumulative distributions of the standard deviations of threshold levels at 125 Hz, reported for independent studies using the TDH39 earphone (squares with the solid line) and other patterns of earphone (circles with the dashed line).

Table 5. Calculation of the cumulative proportion for standard deviations of hearing threshold samples at 250 Hz, determined using TDH39 earphones or other patterns.

TDH39 s.d.	n	cum. n	cum. prop.	not 39 s.d	n	cum. n	cum. prop.
3.35	58	58	0.0227	4.3	42	42	0.0875
3.8	36	94	0.0368	4.6	36	78	0.1625
4	74	168	0.0657	5.5	93	171	0.3563
4.12	92	260	0.1017	5.89	20	191	0.3979
4.2	482	742	0.2903	5.9	93	284	0.5917
4.7	62	804	0.3146	6.02	30	314	0.6542
4.87	534	1338	0.5235	6.5	20	334	0.6958
5.2	60	1398	0.5469	6.63	20	354	0.7375
5.21	534	1932	0.7559	7.12	30	384	0.8000
5.7	162	2094	0.8192	7.6	96	480	1.0000
5.77	20	2114	0.8271				
5.8	202	2316	0.9061				
5.89	20	2336	0.9139				
6.2	62	2398	0.9382				
6.4	72	2470	0.9664				
7.2	86	2556	1.0000				

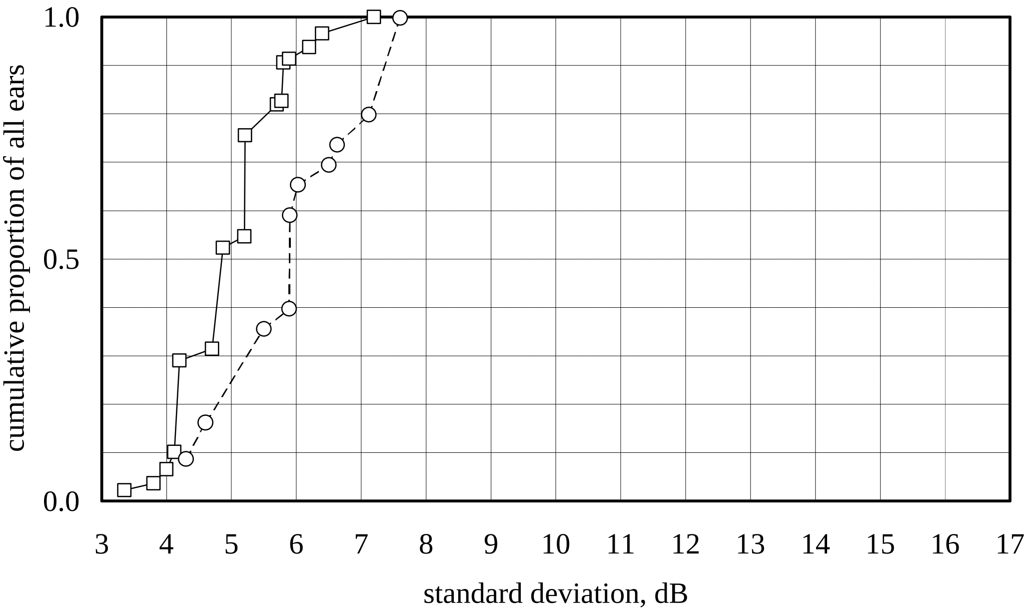

Figure 5. Cumulative distributions of the standard deviations of threshold levels at 250 Hz, reported for independent studies using the TDH39 earphone (squares) and other patterns of earphone (circles).

Table 6. Calculation of the cumulative proportion for standard deviations of hearing threshold samples at 500 Hz, determined using TDH39 earphones or other patterns.

TDH39 s.d.	n	cum. n	cum. prop.	not 39 s.d	n	cum. n	cum. prop.
3.27	58	58	0.0210	4.4	93	93	0.0391
3.5	36	94	0.0340	4.8	54	147	0.0618
3.86	92	186	0.0672	4.9	78	225	0.0946
3.9	11	197	0.0712	4.9	20	245	0.1030
4.27	20	217	0.0785	5.0	42	287	0.1207
4.4	482	699	0.2527	5.0	93	380	0.1598
4.4	62	761	0.2751	5.1	367	747	0.3141
4.5	74	835	0.3019	5.18	20	767	0.3225
4.57	534	1369	0.4949	5.3	20	787	0.3310
4.74	20	1389	0.5022	5.4	211	998	0.4197
4.8	534	1923	0.6952	5.48	30	1028	0.4323
5.0	60	1983	0.7169	5.6	55	1083	0.4554
5.2	162	2145	0.7755	5.7	353	1436	0.6039
5.4	62	2207	0.7979	5.8	36	1472	0.6190
5.4	202	2409	0.8709	6.02	30	1502	0.6316
5.5	168	2577	0.9317	6.1	290	1792	0.7536
5.9	86	2663	0.9628	6.7	96	1888	0.7939
5.9	11	2674	0.9667	6.8	49	1937	0.8146
6.5	72	2746	0.9928	8.2	50	1987	0.8356
7.8	20	2766	1.0000	9.8	391	2378	1.0000

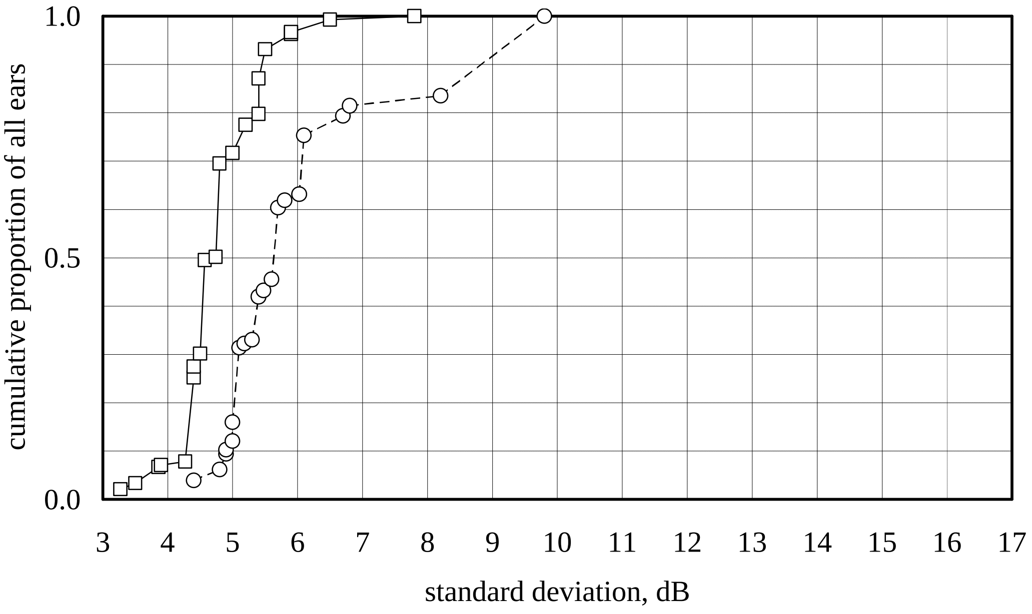

Figure 6. Cumulative distributions of the standard deviations of threshold levels at 500 Hz, for studies using the TDH39 earphone (squares) and other patterns of earphone (circles).

Table 7. Calculation of the cumulative proportion for standard deviations of hearing threshold samples at 1 kHz, determined using TDH39 earphones or other patterns.

TDH39				not 39			
s.d.	n	cum. n	cum. prop.	s.d	n	cum. n	cum. prop.
3.46	92	92	0.0333	4.2	55	55	0.0231
3.5	62	154	0.0557	4.3	54	109	0.0458
3.7	36	190	0.0687	4.4	36	145	0.0610
3.81	58	248	0.0897	4.5	42	187	0.0786
4.2	11	259	0.0936	4.8	78	265	0.1114
4.2	482	741	0.2679	4.82	20	285	0.1198
4.4	74	815	0.2946	5.0	367	652	0.2742
4.54	534	1349	0.4877	5.0	93	745	0.3133
4.7	168	1517	0.5484	5.12	20	765	0.3217
4.87	20	1537	0.5557	5.2	211	976	0.4104
4.9	202	1739	0.6287	5.2	93	1069	0.4495
4.9	162	1901	0.6873	5.48	30	1099	0.4622
4.9	60	1961	0.7090	5.7	353	1452	0.6106
4.91	534	2495	0.9020	5.8	20	1472	0.6190
4.93	20	2515	0.9093	5.8	96	1568	0.6594
5.2	11	2526	0.9132	6.2	290	1858	0.7813
5.2	20	2546	0.9205	6.3	49	1907	0.8019
5.6	86	2632	0.9516	6.6	50	1957	0.8230
5.9	62	2694	0.9740	7.67	30	1987	0.8356
6.1	72	2766	1.0000	11.4	391	2378	1.0000

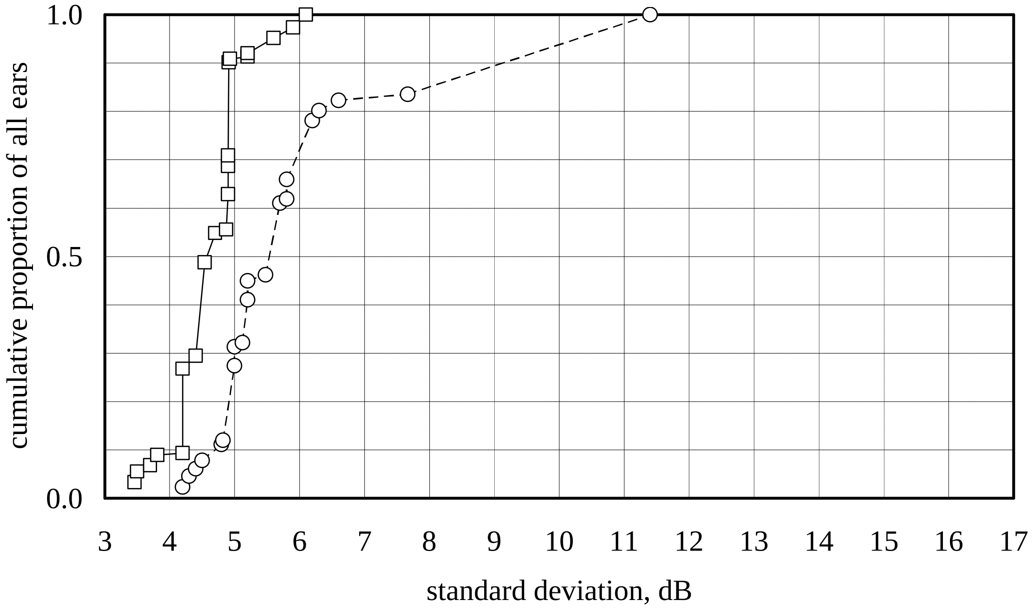

Figure 7. Cumulative distributions of the standard deviations of threshold levels at 1 kHz, for studies using the TDH39 earphone (squares) and other patterns of earphone (circles).

Table 8. Calculation of the cumulative proportion for standard deviations of hearing threshold samples at 2 kHz, determined using TDH39 earphones or other patterns.

TDH39 s.d.	n	cum. n	cum. prop.	not 39 s.d	n	cum. n	cum. prop.
3.25	92	92	0.0333	4.5	49	49	0.0206
3.5	36	128	0.0463	4.6	42	91	0.0383
3.96	58	186	0.0672	4.7	55	146	0.0614
4.3	11	197	0.0712	4.83	20	166	0.0698
4.6	11	208	0.0752	4.9	54	220	0.0925
4.6	60	268	0.0969	4.94	78	298	0.1253
4.6	482	750	0.2711	5.0	20	318	0.1337
4.9	62	812	0.2936	5.17	20	338	0.1421
4.97	20	832	0.3008	5.4	211	549	0.2309
5.0	74	906	0.3275	5.7	93	642	0.2700
5.2	168	1074	0.3883	5.9	50	692	0.2910
5.5	534	1608	0.5813	6.0	353	1045	0.4394
5.51	534	2142	0.7744	6.02	30	1075	0.4521
5.55	20	2162	0.7816	6.1	96	1171	0.4924
5.6	162	2324	0.8402	6.2	36	1207	0.5076
5.8	62	2386	0.8626	6.4	290	1497	0.6295
5.8	202	2588	0.9356	6.6	367	1864	0.7839
6.2	72	2660	0.9617	7.1	93	1957	0.8230
6.7	86	2746	0.9928	7.12	30	1987	0.8356
7.4	20	2766	1.0000	11.4	391	2378	1.0000

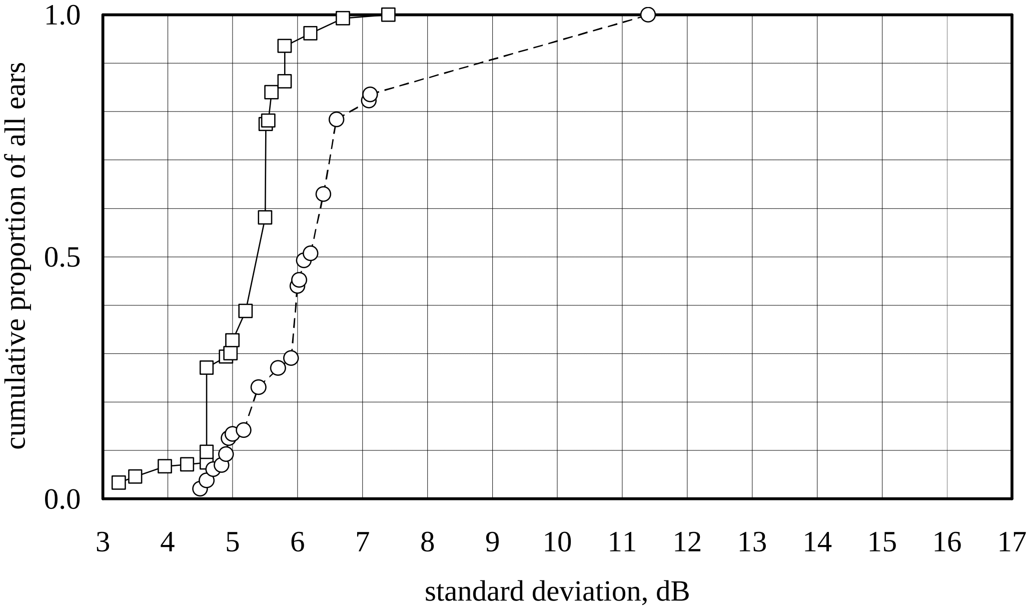

Figure 8. Cumulative distributions of the standard deviations of threshold levels at 2 kHz, for studies using the TDH39 earphone (squares) and other patterns of earphone (circles).

Table 9. Calculation of the cumulative proportion for standard deviations of hearing threshold samples at 3 kHz, determined using TDH39 earphones or other patterns.

TDH39				not 39			
s.d.	n	cum. n	cum. prop.	s.d	n	cum. n	cum. prop.
3.1	11	11	0.0065	5.3	211	211	0.0895
3.61	92	103	0.0607	5.4	54	265	0.1124
3.7	36	139	0.0819	5.4	55	320	0.1357
3.88	58	197	0.1160	5.8	42	362	0.1535
4.6	11	208	0.1225	5.8	290	652	0.2765
4.8	20	228	0.1343	6.0	93	745	0.3159
5.2	74	302	0.1779	6.02	30	775	0.3287
5.2	60	362	0.2132	6.2	353	1128	0.4784
5.3	62	424	0.2497	6.57	30	1158	0.4911
5.5	72	496	0.2921	6.6	49	1207	0.5119
5.8	86	582	0.3428	6.6	78	1285	0.5450
5.8	162	744	0.4382	6.8	367	1652	0.7006
6.05	20	764	0.4499	6.8	96	1748	0.7413
6.2	202	966	0.5689	6.8	93	1841	0.7807
6.4	62	1028	0.6054	6.82	20	1861	0.7892
6.6	482	1510	0.8893	7.1	50	1911	0.8104
6.74	20	1530	0.9011	7.2	36	1947	0.8257
7.2	168	1698	1.0000	7.58	20	1967	0.8342
				11.4	391	2358	1.0000

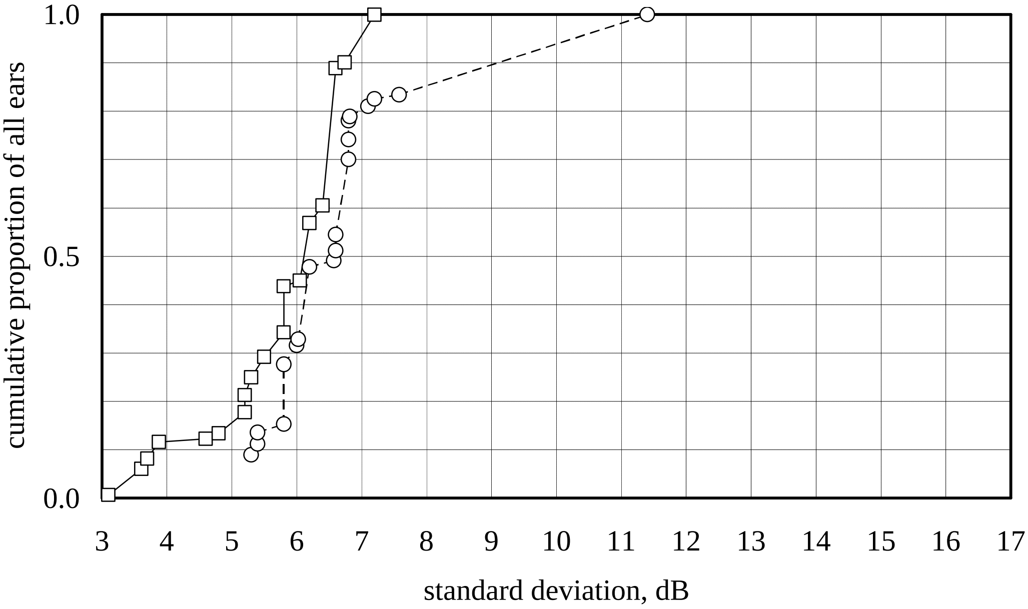

Figure 9. Cumulative distributions of the standard deviations of threshold levels at 3 kHz, for studies using the TDH39 earphone (squares) and other patterns of earphone (circles).

Table 10. Calculation of the cumulative proportion for standard deviations of hearing threshold samples at 4 kHz, determined using TDH39 earphones or other patterns.

TDH39				not 39			
s.d.	n	cum. n	cum. prop.	s.d	n	cum. n	cum. prop.
3.6	11	11	0.0040	4.2	20	20	0.0096
4.07	92	103	0.0372	5.4	42	62	0.0297
4.2	36	139	0.0503	5.44	20	82	0.0393
4.72	58	197	0.0712	5.79	20	102	0.0489
5.1	60	257	0.0929	6.02	30	132	0.0632
5.4	202	459	0.1659	6.2	54	186	0.0891
5.78	20	479	0.1732	6.9	211	397	0.1901
5.8	74	553	0.1999	6.9	93	490	0.2347
6.1	162	715	0.2585	7.0	96	586	0.2807
6.12	534	1249	0.4516	7.12	30	616	0.2950
6.2	62	1311	0.4740	7.2	353	969	0.4641
6.2	11	1322	0.4779	8.2	50	1019	0.4880
6.2	20	1342	0.4852	8.3	93	1112	0.5326
6.23	534	1876	0.6782	8.5	55	1167	0.5589
6.3	86	1962	0.7093	8.6	36	1203	0.5761
6.4	72	2034	0.7354	8.8	367	1570	0.7519
6.5	168	2202	0.7961	9.05	78	1648	0.7893
6.5	20	2222	0.8033	12.0	49	1697	0.8127
6.6	62	2284	0.8257	12.3	391	2088	1.0000
6.9	482	2766	1.0000				

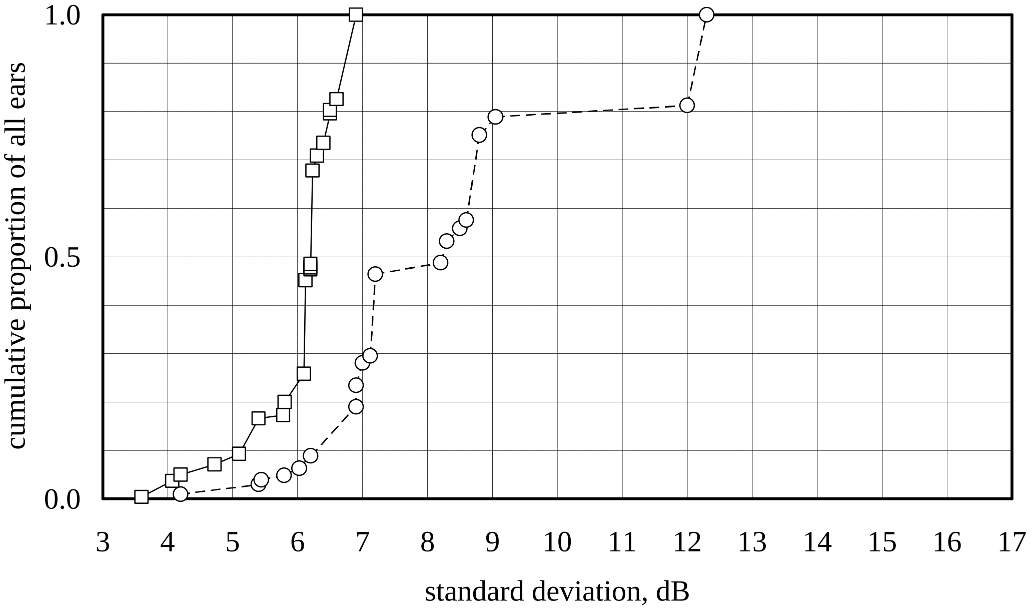

Figure 10. Cumulative distributions of the standard deviations of threshold levels at 4 kHz, for studies using the TDH39 earphone (squares) and other patterns of earphone (circles).

Table 11. Calculation of the cumulative proportion for standard deviations of hearing threshold samples at 6 kHz, determined using TDH39 earphones or other patterns.

TDH39 s.d.	n	cum. n	cum. prop.	not 39 s.d	n	cum. n	cum. prop.
4.72	92	92	0.0342	5.29	20	20	0.0085
5.2	36	128	0.0475	5.3	20	40	0.0170
5.29	20	148	0.0550	6.3	42	82	0.0348
5.82	20	168	0.0624	6.57	30	112	0.0475
6.3	62	230	0.0854	7.0	96	208	0.0882
6.32	58	288	0.1070	7.4	55	263	0.1115
6.4	11	299	0.1111	7.7	93	356	0.1510
6.7	60	359	0.1334	7.8	54	410	0.1739
6.9	72	431	0.1601	8.22	30	440	0.1866
7.0	11	442	0.1642	8.6	367	807	0.3422
7.4	20	462	0.1716	8.8	93	900	0.3817
7.6	162	624	0.2318	9.6	211	1111	0.4712
7.74	534	1158	0.4302	10.2	353	1464	0.6209
7.8	482	1640	0.6092	10.2	290	1754	0.7439
7.83	534	2174	0.8076	10.93	78	1832	0.7769
8.0	86	2260	0.8395	11.0	50	1882	0.7981
8.1	202	2462	0.9146	13.1	36	1918	0.8134
9.2	62	2524	0.9376	13.2	391	2309	0.9792
9.6	168	2692	1.0000	16.6	49	2358	1.0000

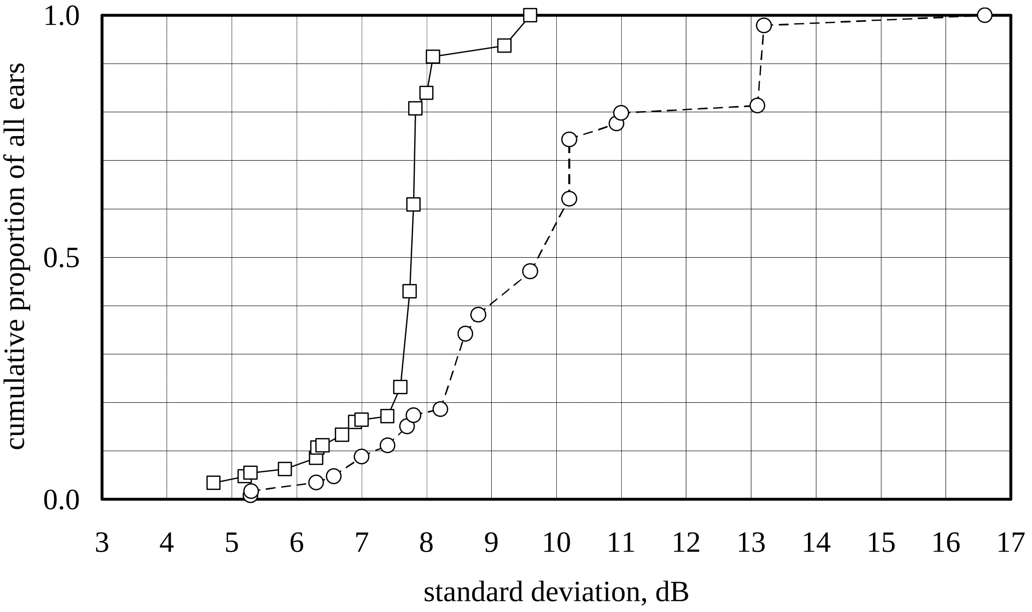

Figure 11. Cumulative distributions of the standard deviations of threshold levels at 6 kHz, for studies using the TDH39 earphone (squares) and other patterns of earphone (circles).

Table 12. Calculation of the cumulative proportion for standard deviations of hearing threshold samples at 8 kHz, determined using TDH39 earphones or other patterns.

TDH39 s.d.	n	cum. n	cum. prop.	not 39 s.d	n	cum. n	cum. prop.
4.59	92	92	0.0368	6.77	20	20	0.0435
5.2	36	128	0.0512	7.2	96	116	0.2522
6.1	60	188	0.0751	7.32	20	136	0.2957
6.76	20	208	0.0831	7.5	93	229	0.4978
6.82	20	228	0.0911	7.6	93	322	0.7000
7.1	62	290	0.1159	7.67	30	352	0.7652
7.6	72	362	0.1447	7.67	30	382	0.8304
7.6	20	382	0.1527	8.4	42	424	0.9217
7.77	58	440	0.1759	14.0	36	460	1.0000
7.9	482	922	0.3685				
8.0	202	1124	0.4492				
8.38	534	1658	0.6627				
8.72	534	2192	0.8761				
8.8	86	2278	0.9105				
9.5	162	2440	0.9752				
13.3	62	2502	1.0000				

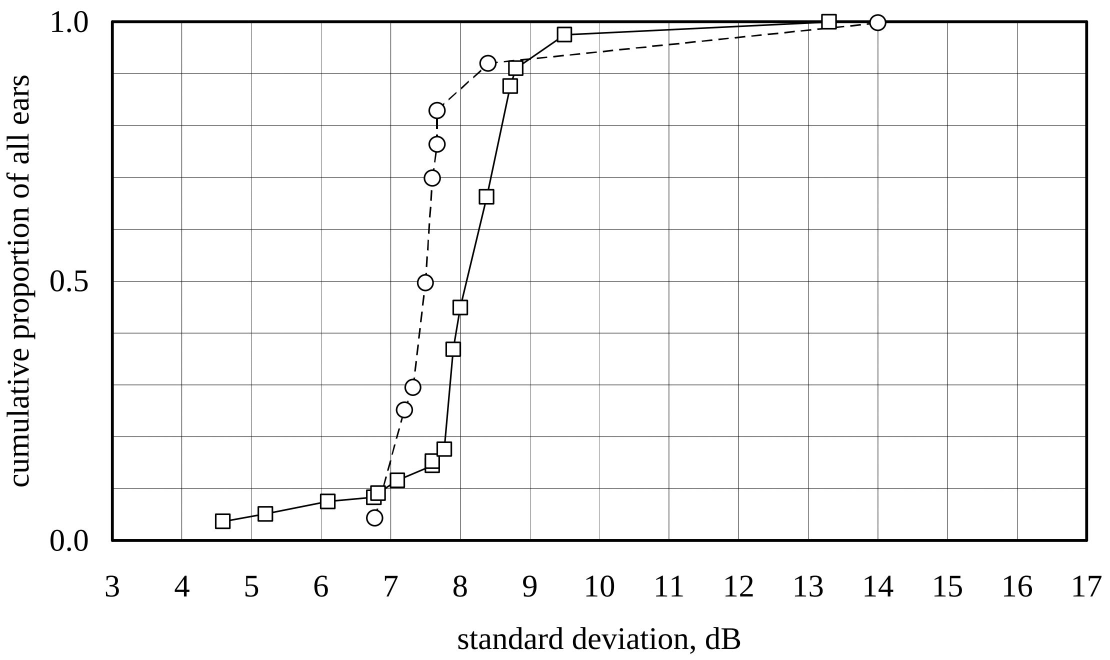

Figure 12. Cumulative distributions of the standard deviations of threshold levels at 8 kHz, reported for independent studies using the TDH39 earphone (squares) and other patterns of earphone (circles).

the outcome is the same: the variation of Hearing Threshold Levels measured with TDH39 earphones is less than the variation associated with other earphones. For 8 kHz, the relation is reversed.

Recall from the Method section that hypotheses were formulated for application to the standard deviations of threshold values (at individual audiometric frequencies) from samples of otologically normal young ears.

> null: threshold variation with TDH39 earphones is not greater than threshold variation with other patterns
>
> alternative: threshold variation with TDH39 earphones is greater than variation with other earphones

For audiometric frequencies up to and including 6 kHz, statistical tests for homogeneity of variance are unnecessary: the conclusions were obvious by simple visual inspection. For eight of the nine audiometric frequencies examined here, threshold variation of the TDH39 earphone may be judged as **less than** the variation of other earphone patterns. The standard deviations of thresholds determined by TDH39 and non-39 earphones provide no reason to reject the null hypothesis. Indeed, the data support the unorthodox (and statistically improper) action of rejecting the one-sided alternative hypothesis, deliberately chosen as the expected outcome of the investigation.

5. Discussion

In the Introduction to this paper, a question was asked: Is it worth retaining the acoustic coupler for calibration of the TDH39 only, while all other patterns of earphone require the use of an ear simulator? In attempt to answer this question, data were gathered to show the reliability of hearing thresholds obtained using TDH39 earphones, for comparison against those measured using other patterns. The standard deviation of threshold samples using the TDH39 was found to be generally **less** than the deviations associated with other earphones. There is no compelling reason to abandon the TDH39, and therefore no need to retire its specified calibration device, the acoustic coupler.

In an effort to reduce the number of calibration devices, there are several options:

- Continue to use the TDH39, and cease testing with other patterns (with higher threshold variability). This retrograde step would fossilise earphone technology; worldwide audiology could not seriously consider such a course of action.
- Calibrate TDH39 earphones on the ear simulator (artificial ear). The TDH39 does indeed fulfil all eight geometrical and function requirements imposed upon all other patterns, as listed in International Standard ISO 389-1.

This second option brings consequences, both desirable and potentially troublesome. On the positive side, linking the TDH39 to the ear simulator would minimise the effects of high-frequency irregularities in the TDH39 frequency response (as seen in acoustic coupler measurements). Calibrating the output of TDH39 earphones to the ear simulator RETSPLs eliminates the “phantom” 6 kHz audiometric notch (resulting from an optimistic RETSPL specified for the acoustic coupler). This benefit may be seen in the data of one data source from the present study; see Table 13 and Figure 13. These TDH39/simulator data indicate an improvement in audiometric accuracy for both clinical and medico-legal applications.

Allowing calibration of TDH39 earphones on the ear simulator would bring some benefits, but also a troublesome requirement. International Standard ISO 389-1 would need to be amended to eliminate the possible choice of coupler or simulator for use with the TDH39. With the present wording of ISO 389-1, competing RETSPLs would be applicable to the TDH39 earphone. Calibration target values must not be a matter of choice: standards must not be open to interpretation.

6. Final Considerations

This paper started with a question: Is it worth retaining the acoustic coupler solely for calibration of two patterns of earphone used in audiometry, viz. TDH39 and Beyer DT48, while all other patterns require the use of an ear simulator? Examination of threshold statistics gave no compelling reason to abandon the TDH39 (and its specified calibration device, the acoustic coupler). However, in countries where ISO and IEC standards influence national practice, the acoustic coupler must fall into disuse over time. Considering British and international practice, the coupler has one function only; when TDH39 earphones are no longer available, the coupler will cease to have a function. On the other hand, the ear simulator (artificial ear) will have increasing importance in audiometric practice. As new types of supra-aural earphones come into use, the simulator will be required for their calibration, as specified in ISO 389-1. As high-frequency audiometry becomes a useful clinical and research tool, the simulator (with flat plate adaptors, see IEC 60318-2) will be required for verification of the earphone output levels against the audiometric zero specified in ISO 389-5. As circumaural earphones come into audiometric use, the simulator (with flat-plate adaptor) will also be required for their calibration; see ISO 389-8 which gives RETSPLs for calibration of one such earphone. Wherever International Standards are in force, the ear simulator (artificial ear) has a future. The acoustic coupler is an evolutionary dead end.

7. Acknowledgement

This work was supported financially by the Department of Trade and Industry of the United Kingdom government. The opinions expressed are those of the author only.

Table 13. Threshold data determined using the TDH39 earphone calibrated on the ear simulator (artificial ear) to RETSPLs intended for all other earphone patterns; mean Hearing Threshold Levels (dB HL) and standard deviations (dB) for 81 otologically normal males, aged 16-27 years.

source	comment		freq, Hz							
			250	500	1k	2k	3k	4k	6k	8k
Lawton, Robinson (1986)	162 ears	mean	0.8	0.5	-0.8	-0.5	0.1	0.6	2.4	4.1
		std. dev.	5.7	5.2	4.9	5.6	5.8	6.1	7.6	9.5

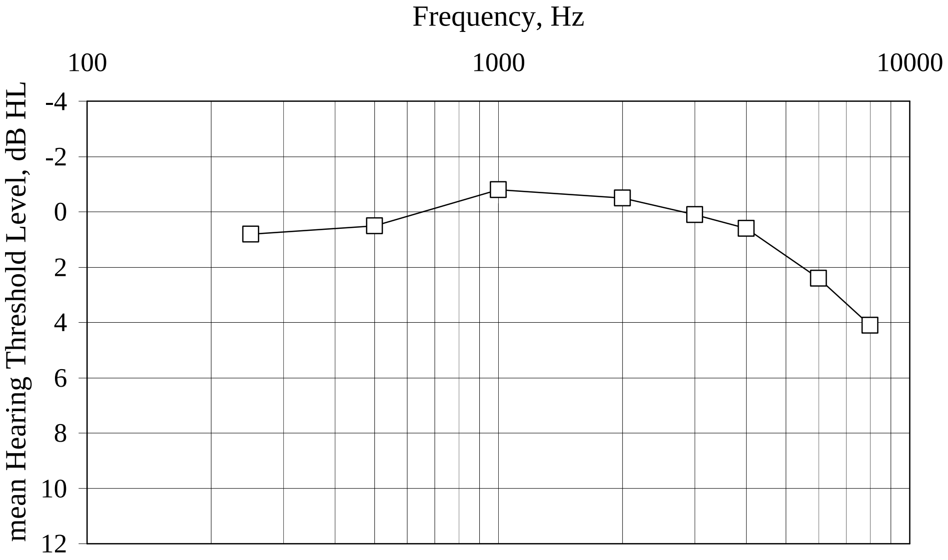

Figure 13. Mean thresholds for the Lawton and Robinson (1986) sample of 81 otologically normal males, aged 16-27 years.

8. References

American National Standards Institute 1995. ANSI S3.7-1995 *American National Standard method for coupler calibration of earphones*. New York: ANSI.

American National Standards Institute 1996. ANSI S3.6-1996 *American National Standard specification for audiometers*. New York: ANSI.

Arlinger, S. 1982. Normal threshold of hearing at preferred frequencies. *Scand Audiol*, 11, 285-286.

Arlinger, S.D. 1991. Normal hearing thresholds in the low-frequency range determined by an insert earphone. *J Acoust Soc Am*, 90, 2411-2414.

Arlinger, S. & Kinnefours, C. 1989. Reference equivalent threshold sound pressure levels for insert earphones. *Scand Audiol*, 18, 195-198.

Carter, N.L., Keen, K., Waugh, R.L. & Bulteau, V.G. 1978. *Hearing levels and oto-rhino-laryngological findings in students of the University of Sydney under age 21*. National Acoustic Laboratories Report No. 72. Sydney, Australia: NAL.

Carter, N., Murray, N., Khan, A. & Waugh, D. 1984. A longitudinal study of recreational noise and young people’s hearing. *Aust J Audiol*, 6, 45-53.

Carter, N.L. & Murray, N.L. 1985. *Reference equivalent threshold sound pressure levels for the TDH-49 earphone - Survey data and comment on Australian and International standards*. National Acoustics Laboratory Non-commissioned Report No. 105. Sydney, Australia: NAL.

Corso, J.F. 1963. Age and sex differences in pure-tone thresholds. *Arch Otolaryngol*, 77, 385-405.

Cox. R.M. & McDaniel, D.M. 1986. Reference equivalent threshold levels for pure tones and 1/3 oct noise bands: Insert earphone and TDH-49 earphone. *J Acoust Soc Am*, 79, 443-446.

Frank, T. & Vavrek, M.J. 1992. Reference threshold levels for an ER-3A insert earphone. *J Am Acad Audiol*, 3, 51-59.

Han, L.A. & Poulsen, T. 1998. Equivalent threshold sound pressure levels for Sennheiser HDA 200 earphone and Etymotic Research ER-2 insert earphone in the frequency range 125 Hz to 16 kHz. *Scand Audiol*, 27, 105-112.

Hinchcliffe, R. 1959. The threshold of hearing of a random sample rural population. *Acta Otolaryngol*, 50, 411-422.

International Electrotechnical Commission 1998. IEC 60318-1. *Electroacoustics — Simulators of human head and ear — Part 1: Ear simulator for the calibration of supra-aural earphones*. Geneva: IEC.

International Electrotechnical Commission 1998. IEC 60318-2:1998. *Electroacoustics — Simulators of human head and ear — Part 2: An interim acoustic coupler for the calibration of audiometric earphones in the extended high-frequency range*. Geneva: IEC. [Note: Despite its title, this standard does not describe an acoustic coupler: it specifies two adapters for use with specific earphones on a 60318-1 ear simulator].

International Electrotechnical Commission 1998. IEC 60318-3:1998. *Electroacoustics — Simulators of human head and ear — Part 3: Acoustic coupler for the calibration of supra-aural earphones used in audiometry*. Geneva: IEC.

International Organization for Standardization 1998. ISO 389-1:1998. *Acoustics – Reference zero for the calibration of audiometric equipment – Part 1: Reference equivalent threshold sound pressure levels for pure tones and supra-aural earphones*. Geneva: ISO.

International Organization for Standardization 2001. ISO 389-5:2001. *Acoustics — Reference zero for the calibration of audiometric equipment — Part 5: Reference equivalent threshold sound pressure levels for pure tones in the frequency range 8 kHz to 16 kHz*. Geneva: ISO.

International Organization for Standardization 2004. ISO 389-8:2004. *Acoustics – Reference zero for the calibration of audiometric equipment – Part 8: Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones*. Geneva: ISO.

Kell, R.L, Pearson, J.C.G. & Taylor, W. 1970. Hearing thresholds of an island population in North Scotland. *Int Audiol*, 9, 334-349.

Larson, V.D., Cooper, W.A., Talbott, R.E., Schwartz, D.M., Ahlstrom, C., et al. 1988. Reference threshold sound-pressure levels for the TDH-50 and ER-3A earphones. *J Acoust Soc Am*, 84, 46-51.

Lawton, B.W. & Robinson, D.W. 1986. *An investigation of tests of susceptibility to noise-induced hearing loss*. ISVR Technical Report 138. Southampton, UK: University of Southampton.

Lutman, M.E. & Davis, A.C. 1994. The distribution of Hearing Threshold Levels in the general population aged 18-30 years. *Audiology*, 33, 327-350.

Qasem, H.Y.N. 1996. *Comparison of hearing threshold levels using TDH-39 and TDH-49 earphones*. MSc Dissertation, ISVR. Southampton, UK: University of Southampton.

Rahko-Laitila, P., Karma, P., Laipalla, P., Salmelin, R., Sapilä, M., et al. 2001. The pure-tone hearing thresholds of otologically healthy 14-year-old children. *Audiology*, 40, 171-177.

Rice, C.G. & Coles, R.R.A. 1966. Normal threshold of hearing for pure tones by earphone listening with a self-recording audiometric technique. *J Acoust Soc Am*, 39, 1185-1187.

Robinson, D.W., Shipton, M.S. & Hinchcliffe, R. 1979. *Normal hearing threshold and its dependence on clinical rejection criteria*. National Physical Laboratory Acoustics Report Ac 89. Teddington, UK: NPL.

Robinson, D.W., Shipton, M.S. & Hinchcliffe, R. 1981. Audiometric zero for air-calibration: A verification and critique of international standards. *Audiology*, 20, 409-431.

Shipton, M.S. & Robinson, D.W. 1981. *Thresholds of hearing by air-conduction and by bone-conduction under different masking conditions*. National Physical Laboratory Acoustics Report Ac 100. Teddington, UK: NPL.

Smith, P., Davis, A., Ferguson, M. & Lutman, M. 1999. Hearing in young adults: Report to ISO/TC43/WG1. *Noise Health*, 4, 1-10.

Taylor, W., Pearson, J., Mair, A. & Burns, W. 1965. Study of noise and hearing in jute weaving. *J Acoust Soc Am*, 38, 113-120.

Taylor, W., Pearson, J. & Mair, A. 1967. Hearing thresholds of a non-noise-exposed population in Dundee. *Br J Ind Med*, 24, 114-122.