UNIVERSITY OF
Southampton
Institute of Sound and
Vibration Research

ISVR Technical Report

Variability of the threshold of hearing: It's importance in cases of noise induced hearing loss

B.W. Lawton

ISVR Technical Report No 336

April 2015

SCIENTIFIC PUBLICATIONS BY THE ISVR

Technical Reports are published to promote timely dissemination of research results by ISVR personnel. This medium permits more detailed presentation than is usually acceptable for scientific journals. Responsibility for both the content and any opinions expressed rests entirely with the author(s).

Technical Memoranda are produced to enable the early or preliminary release of information by ISVR personnel where such release is deemed to the appropriate. Information contained in these memoranda may be incomplete, or form part of a continuing programme; this should be borne in mind when using or quoting from these documents.

Contract Reports are produced to record the results of scientific work carried out for sponsors, under contract. The ISVR treats these reports as confidential to sponsors and does not make them available for general circulation. Individual sponsors may, however, authorize subsequent release of the material.

COPYRIGHT NOTICE

(c) ISVR University of Southampton All rights reserved.

ISVR authorises you to view and download the Materials at this Web site ("Site") only for your personal, non-commercial use. This authorization is not a transfer of title in the Materials and copies of the Materials and is subject to the following restrictions: 1) you must retain, on all copies of the Materials downloaded, all copyright and other proprietary notices contained in the Materials; 2) you may not modify the Materials in any way or reproduce or publicly display, perform, or distribute or otherwise use them for any public or commercial purpose; and 3) you must not transfer the Materials to any other person unless you give them notice of, and they agree to accept, the obligations arising under these terms and conditions of use. You agree to abide by all additional restrictions displayed on the Site as it may be updated from time to time. This Site, including all Materials, is protected by worldwide copyright laws and treaty provisions. You agree to comply with all copyright laws worldwide in your use of this Site and to prevent any unauthorised copying of the Materials.

UNIVERSITY OF SOUTHAMPTON
INSTITUTE OF SOUND AND VIBRATION RESEARCH
ISVR CONSULTING

Variability of the threshold of hearing:
Its importance in cases of Noise-induced Hearing Loss

by

B.W. Lawton
Associate Principal Consultant

ISVR Technical Report No: 336

April 2015

Authorised for issue by
Dr M.G. Smith
Manager, ISVR Consulting

© Institute of Sound & Vibration Research

Acknowledgement:

This review was prepared with financial support from BC Legal, through the Southend-On-Sea office. Its contents, including any opinions and/or conclusions expressed herein, are those of the author alone.

Variability of the threshold of hearing:
Its importance in cases of Noise-induced Hearing Loss

B. W. Lawton

Associate Principal Consultant

ISVR Consulting

University of Southampton

Southampton SO17 1BJ

ABSTRACT

This document reviews a number of research studies dealing with how Hearing Threshold Levels might change on repeat testing. The best estimate of the true threshold will be the lowest (most acute) value achieved during repeat testing at each particular audiometric frequency. The standard deviation of all threshold estimates at an audiometric frequency will give information regarding the dispersion of those estimates. The best estimates of the true thresholds are important in confirming the existence and degree of Noise-induced Hearing Loss.

KEY WORDS

repeated thresholds of hearing, standard deviation of estimates of threshold, Noise-induced Hearing Loss, noise notch

ABBREVIATIONS

BS	British Standard
dB	decibel
dB HL	decibels Hearing Level (from an audiometer)
dB HTL	decibels Hearing Threshold Level (from a person)
EN	European Standard
HTL	Hearing Threshold Level
ISO	International Organization for Standardization
kHz	kiloHertz
L	left
OAS	Otological Abnormality Score
PTA	pure tone audiometry
R	right
RETSPL	reference equivalent threshold sound pressure level
TDH	Telephonics Dynamic Headphone

intentionally blank

TABLE OF CONTENTS

intentionally blank

1. INTRODUCTION

1.1 Variability of Threshold

All measurements are subject to error. Measurement errors may be systematic or random, or a combination of the two. Systematic error is a repeatable displacement of the measured value from the true (and unknowable) value. In contrast, random error becomes apparent only when repeat measurements are performed; small variations appear as positive and negative differences about the estimated true value. In the absence of systematic error, repeated measurements of a variable subject to random variation usually follow a Gaussian or normal distribution, with observed values clustered about the true (but unknowable) value. These observed values will be displaced by positive and negative intervals about the true (but unknowable) value of the variable in question, as modified by any systematic error.

Hearing Threshold Levels (HTLs) are measurements, and thus subject to error, possibly systematic, and certainly random. In fact, determination of HTLs is a rather complicated process with a number of opportunities for error to enter the chain. A brief review will point out how or where the measurement process can (and does) go wrong, giving both systematic and random errors in estimating hearing sensitivity.

Audiometric signals are produced and controlled within the audiometer; an instrument which if uncalibrated or mis-calibrated will introduce systematic errors into the hearing threshold measurements. Audiometric earphones may be placed on the wrong ears (systematic) or improperly seated on the ears, making a less-than-efficient acoustic coupling (random). The audiometric method used to determine the HTL will influence the measured value; self-recording and manually determined thresholds are known to differ in a systematic manner. The state of mind of the listener or subject will, of course, play a primary role in the quality of the measurements. The subject may be newly learning the audiometric task, in which case the measurements will grow less (randomly) variable over a short span of time, perhaps several minutes. On the other hand, the subject may grow tired or impatient with the task, or lose concentration, with the result that the measurement becomes more (randomly) variable with time. Finally, the subject may simply be uncooperative, or may be malingering for some reason.

Even the audiometrician may introduce error into the measurement chain; just as the subject will have a criterion to accept a sound as heard, so may audiometricians have a criterion for accepting the patient's signal for 'sound heard'. An audiometrician working on a liberal or loose interpretation of the patient's signals will determine lower (systematically more acute) thresholds of a more variable nature. In contrast, the audiometrician who insists upon the patient's signals being in strict synchrony with the signal onset and offset will record relatively higher (less acute) thresholds which are comparatively stable on repeat.

1.2 Air-conduction Variability

What follows in this report are reviews of investigations into individual aspects of the variability of air-conduction threshold measurement. It will be apparent from the introduction above, and from the reviews following, that certain aspects of threshold variability are susceptible to study, while others are not. For example, the placement of earphones is a source of random error which is easily examined; threshold measurements are simply repeated with and without replacing the transducer between measurements. In contrast, the subject's listening strategy cannot be strictly controlled or changed; it is arguable that it cannot even be known, only observed.

In the end, there are two questions to be answered regarding any collection of hearing thresholds determined for an individual. How accurately do the measured HTLs represent this person's deviations from audiometric zero? When the thresholds are measured again, how big must the change be in order to be sure that any apparent worsening of hearing threshold is, in fact, real?

1.3 Thresholds for an individual: one-way error

The threshold of hearing is the minimum level of a sound which is **just audible** in given conditions on a specified fraction of presentations (conventionally 50%). The term is often understood to imply quiet listening conditions, that is, it represents the irreducible, absolute threshold.

Repeated measures of the threshold of hearing form a positively skewed distribution, with a long tail of observed values worse (less acute) than the true (but unknowable) absolute threshold. Valid responses cannot occur at presentation levels *below* the listener's true or absolute threshold. Responses will be given for presentations *above* the listener's true threshold, with the threshold values depending upon the listener's understanding of the instructions given, and upon the audiometrician's interpretation of those responses.

2. THE REVIEWS

Summaries are given here of research studies reporting descriptive statistics for groups of selected subjects: mean changes (positive or negative) of repeated HTLs from a base threshold, and the standard deviations of such changes, across the range of audiometric frequencies surveyed. Researchers do not report the raw data for individual subjects, but only as processed data for their screened subject groups.

BURNS and HINCHCLIFFE (1957)

The conservation of hearing against the effects of occupational noise is an important aspect of industrial medicine. This study set out to compare two audiometric methods thought suitable for large-scale screening, with subsequent examinations at intervals. The Békésy method used will not be considered here. [Continuous swept-frequency tones were presented to the test subjects; this method has been overtaken by fixed-frequency, pulsed tone presentations that are more easily recognised by the person under test.] Conventional manual pure-tone audiometry, with the signal presentations controlled by the audiometrician, was well established at the time.

The authors set out to answer several questions concerning the two test methods. Of present interest in respect of manual audiometry is the question: What threshold differences and their standard deviations would be observed for examinations separated by one week?

Twenty subjects were chosen: medical students and university teachers, aged 20 to 58 years. None had undergone audiometric testing before this study. All were screened to exclude severe hearing loss, as well as perforations or marked scarring of the tympanic membrane.

The manual audiometry was performed using the 'method of limits'. At each frequency, the threshold value was taken as the mean of 'tone-heard' ascending and descending presentation levels, using steps of 2 dB. The frequencies were 1 kHz, 2, 3, 4, 6 and 0.5 kHz in this order.

The manual test results are given below:

Comparison of first and second tests by manual pure-tone audiometry; tests separated by one week.

	frequency, kHz					
	0.5	1	2	3	4	6
mean difference, dB	1.0	2.2	1.5	2.0	1.4	-1.7
standard deviation of difference, dB	2.2	2.0	2.2	2.2	2.2	2.8

difference = (first test) minus (second test)

The authors opined that a second audiogram, performed one week after the first, gave a threshold lowering (improvement) of 1 to 2 dB only; such improvements were thought to be of marginal significance.

ATHERLEY and DINGWALL-FORDYCE (1963)

This paper considers the precision (repeatability) which may be expected in short-term serial measurements of audiometric thresholds.

Serial audiometric measurements are generally recommended as a means of monitoring the hazardous effect of noise. Their use for such a purpose requires the differentiation between systematic sources of variance, and random fluctuations due to the inherent variability of HTL measurements. The investigators took care to control systematic variation due to drift as the audiometer warmed up, audiometer calibration, drift of the earphone acoustic output, and extraneous noise in the test chamber. Careful attention to systematic variation allowed the investigators to concentrate on the chance fluctuations of the audiometric process.

The object of the investigation was, first, to confirm the extent of this inherent variability among a select group of subjects tested under carefully controlled clinical conditions, and then to consider, in the light of these findings, the ultimate provision of general tables of reference for the assessment of serial audiometric measurements.

The experimental subjects were 12 otologically-normal young male medical students in whom no ear, nose or throat pathology could be demonstrated, and who gave no history of noise exposure. The subjects' ages were not stated. Each subject performed four audiometric sequences at 1, 2, 3, 4, 6, 8, 0.5 and 1 kHz in each ear; the earphone type was not specified. For each individual, the time intervals between the *first* and *second* was 24 hours, followed by an interval of one week before the *third* test which was separated from the *fourth* by another 24 hours.

The final HTLs were taken to be the mean of the ascending and descending thresholds determined at each audiometric frequency on each ear, always starting with the left. Each response to each test tone, which was presented for about 1 second, was recorded, and the threshold was calculated at the end of each session. The test tones were presented four times at each HL explored; a two-out-of-four response criterion was used for both the descending and ascending thresholds. The audiometric test on each individual lasted between 15 and 20 minutes.

Returning to the authors' aims, the following table gives the "inherent variability" for the thresholds given by their selected group of subjects, tested under carefully controlled clinical conditions.

Within-ears standard deviations of repeated thresholds; 4 determinations, 24 ears

frequency, kHz						
0.5	1	2	3	4	6	8
standard deviations, dB						
2.9	2.7	2.6	2.4	4.1	4.4	4.8

The second aim was to provide a general table of reference for the assessment of serial audiometric measurements. Their analysis provided probability ranges about a single threshold determination; the true-but-unknowable threshold would fall **within** the given bands.

Estimation of true HTL from a single determination

probability that true threshold is outside band	frequency, kHz						
	0.5	1	2	3	4	6	8
	true threshold estimated to lie within the range of the single determination ± the following dB values						
0.05	6	5	5	5	8	9	9
0.01	7	7	7	6	10	11	12

Note that the bands are widest for the frequencies 4 and 6 kHz that quantify the depth of any ‘noise notch’ presumed due to injurious noise exposure, and at 8 kHz where recovery (lower threshold) establishes the existence of the ‘noise notch’.

ROBINSON and WHITTLE (1973)

The process of measuring the HTLs of an individual ear entails a chain of independent factors of which one only is the essential “unknown” in the process: the intrinsic acuity of the ear at the time of test.

This paper deals with the method of test, including:
1) the instructions as imparted to, or understood by, the subject: the concept of minimum audibility admits of a range of interpretations;
2) the type of audiometer used (manual or self-recording) and the particular character of the test tones and their time sequence; and
3) the manner in which the threshold level is deduced from the subject’s responses.

There is a particular complication: the measurement of HTLs can be affected to a significant extent by a person’s improvement at audiometry on successive occasions. This is of practical importance, particularly in industrial monitoring audiometry where testing time has to be kept to a minimum. Such improvements may obscure real changes of hearing acuity between one occasion and the next.

The experiments described were designed with two purposes: to attempt a systematic determination of the inherent difference between HTLs measured by self-recording audiometry and by the manual method, and to explore the influence of familiarization with the task.

Sixty-four test subjects were paid to participate in the experiment; they were mainly housewives and retired persons, aged 25 to 73 years. None had previously undergone audiometric testing. A rough matching into four groups of 16 was made in respect of sex and age. A replication of the experiment was performed approximately 15 months later. However, it was possible to recall only 48 of the subjects, who were divided into 12 individuals from each original group.

Manual audiometry and self-recording audiometry were administered using a Rudmose ARJ-5 audiometer equipped with TDH-39 earphones and MX-41-AR cushions. Half of the test subjects gave HTLs first by the manual method, followed by self-recording; the other half gave thresholds in the opposite test order. Half the subjects in each group were tested in the order left ear followed by right ear, and *vice versa* for the other half. A complete audiogram for both ears was obtained by one method before commencing the other method, after a pause of approximately 10 minutes. Tones were always presented in the order 0.25, 0.5, 1, 2, 4, 6 and 8 kHz.

For manuals thresholds, the first tone was presented at a plainly audible level, to elicit a press of the response button to acknowledge “tone heard”. At or near threshold, a descending technique was used to find the lowest level giving 2 positive responses out of 4 presentations. A similar technique was used to find the lowest level giving 2 acknowledgements out of 4 presentations when ascending from apparent silence (below threshold). The subject’s HTL for the test frequency was the average of the descending and ascending 2-out-of-4 presentations.

The instructions for the self-recording tests were:

> “In order to measure your hearing we ask you to listen to musical sounds in your head-phone and to signal by pressing a button all the time you hear the sounds. The object of the test is to find out the very faintest sound that you can hear.
>
> The sounds will be broken into a succession of ‘pips’ and they will be easy to hear at some times, but very faint indeed at others. The tones may be low-pitched, medium- or high-pitched, and you will hear them first in the one ear. Later you will hear them in the other ear. All you have to remember is that as long as you hear the ‘pips’ you keep the button pressed and when you hear nothing you let the button go.”

For the self-recording method, ascending and descending levels were plotted by the audiometer; reversals of the trace occurred in response to the subject’s response-button press or release. The HTL was determined as the mid-point of the ascending and descending excursions.

The analysis concentrated on the differences (self-recording threshold minus manual threshold) for each tone frequency between 0.25 and 8 kHz. Contrasts were made between right ear first and left ear first for each audiometric method, and between the ‘competing’ methods for the right and for the left.

The experiments described enabled three distinct components of variability in audiometry to be identified. The conclusions may depend in part on the audiometric procedures followed, in particular the use of an ascending order of test frequencies.

(i) After elimination of learning effects, the difference between HTLs obtained by self-recording and by manual audiometry was taken as zero at the first frequency tested and 3 dB at all other frequencies, the lower threshold being associated with the self-recording method.

(ii) Successive trials by either method resulted in lower thresholds. When the gross initial improvements exhibited by some subjects were disregarded, the effect of practice appeared to be a continuing process.

(iii) Amongst the audiometrically-naive subjects of this experiment, a certain proportion gave large initial errors, particularly at 0.25 and 8 kHz by self-recording audiometry and at 6 kHz by manual audiometry. These initial errors, which can be distinguished from the practice effect just summarised, were probably due to the subjects feeling their way towards the performance of a task which they found more difficult than may have been expected from the instructions given. In an industrial situation, the magnitude and prevalence of these errors may well be greater than in the experimental test group; the same conclusion might apply to audiometry in pursuit of a compensation claim. These points indicate the desirability of repeated audiometric tests, or some kind of initial training at the task.

ROBINSON, SHIPTON and HINCHCLIFFE (1981)

This study was undertaken to clarify a number of aspects of International standards regarding the audiometric zero used in the calibration of audiometers. The authors felt it advisable to return to first principles, making threshold measurements on a fairly large scale, and under carefully controlled conditions. To keep the project within manageable proportions, it was decided to concentrate on the important central part of the audiometric frequency range from 0.5 to 6 kHz.

Fixed-frequency automatic-recording audiometry (on the Békésy principle) was chosen to eliminate as much as possible the random error and subjective uncertainties of audiometry. This left the requirement for test subjects defined with deceptive simplicity in ISO 389: 1975. An otologically-normal subject is "*a person (within the inclusive age range 18 to 30 years) in a normal state of health who is free from all signs or symptoms of ear disease and from wax in the ear canal, and has no history of undue exposure to noise*". The authors' approach was to adopt a systematic method of classifying test subjects with respect to otological normality, and to explore the implications of criteria ranging from ultraconservative to clearly less conservative with respect to the definition above.

The basic population was a subset of the staff of at the National Physical Laboratory aged between 18 years 0 months and 30 years 11 months. Before any testing was done, all potential subjects completed a 19-item questionnaire on their relevant medical, otological and environmental histories. There were otoscopic examinations of the eardrum and external ear canal, as well as tuning fork tests.

An otological abnormality scoring system was devised to quantify the normality (or otherwise) of potential test subjects. This scoring system is given on the following page.

After elimination of any potential subjects with undue noise exposure, ear/mastoid surgery, upper respiratory tract infection, or with an accumulation of wax in one or both ears canals, 62 were enrolled. Also given on the next page is a table of the number of subjects listed by otological abnormality score.

Point system for otological abnormality score (OAS).

evidence		OAS
history (noise)	occupational	1 or 2*
	exposure to gunfire	2 or 3*
	exposure to explosions	3
	discotheques, other noisy recreations	0 or 1*
history (non-noise)	ear discharge	1
	persistent tinnitus	2
	episode(s) of unconsciousness	2
	family hearing disorder	1
	use of ototoxic drugs	2
	other 'ear trouble'	1
ear examination	scarring of drum(s)	2 or 4*
	retraction of drum(s)	6
	tuning fork: Rinne test unfavourable	5
	tuning fork: Bing test unfavourable	1
	binaural diplacusis	4
maximum		38

* depending on degree

Number of subjects listed by OAS.

OAS	number of subjects
0	11
1	11
2	10
3	7
4	5
5	3
6	4
7	3
8	2
9	1
10	2
11	1
12	2
total	62

In respect of international standardisation, the authors opined that OAS = 0 would be an impossibly strict condition for an *otologically normal* population; an OAS range 0 to 3 was felt to result in a more manageable rejection rate.

All 62 testees gave HTLs in both ears for all frequencies; a full-range practice test was performed in the L ear first, and then followed by HTLs in the R with an L-repeat series. The HTLs improved with repeat testing. The authors suggested that this improvement was due to learning, indicating increasing confidence with the audiometric task.

Mean improvement (in dB) of HTLs within test session (left minus left-repeat)

number of ears	frequency, kHz 0.5	1	2	3	4	6
	mean improvement with repeat testing, dB					
62	1.4	1.2	1.6	1.4	0.2	0.6

The authors reported the HTLs given as an average of the R and L-repeat values. As they suspected, the thresholds for subjects with OAS equal to 3 or less made a useful sample, with a tolerable rejection rate.

Mean thresholds of subjects grouped in OAS bands.

OAS bands	number of subjects	HTLs, dB HL average of R and L-repeat					
		frequency, kHz					
		0.5	1	2	3	4	6
0	11	-1.3	-4.3	-3.8	-4.1	-1.0	0.4
1	11	0.6	-1.1	-2.3	-3.2	-0.4	-0.1
2 - 3	17	0.1	-4.0	-3.8	-0.7	1.3	5.5
4 - 6	12	0.8	-1.8	-2.2	0.3	0.3	5.1
7 - 12	11	3.8	2.5	3.6	4.7	8.9	14.7
>12		—	—	—	—	—	—

Of significant interest here is the effect of increasing OAS upon the standard deviations of HTLs: the standard deviations increased for each audiometric test frequency. The deviations were greatest for the frequencies 4 and 6 kHz (which figure in NIHL).

Standard deviations of HTLs in selected cumulative ranges of OAS,
for the 62 test subjects (124 ears).

OAS range	number of ears	standard deviation, dB					
		frequency, kHz					
		0.5	1	2	3	4	6
0	22	4.9	4.6	4.3	4.0	5.1	7.1
0 - 3	78	4.9	5.1	5.4	6.4	5.8	8.5
0 - 6	102	5.7	5.4	5.9	6.9	6.7	9.7
0 - 9	114	5.5	5.3	6.5	7.0	8.9	11.3
0 - 12	124	8.4	8.2	8.1	9.5	11.3	12.7

These data lead to the observation that the greater values of standard deviations in the wider ranges of OAS (for instance 0 - 9, and 0 - 12) result from the inclusion of higher (worse) thresholds seen above in the table ‘Mean thresholds of subjects grouped in OAS bands’.

CHERMAK, DENGERINK and DENGERINK (1983)

The object of this investigation was to examine the test-retest reliability of HTLs and of measures of temporary threshold shift. Only the HTL facet of this study will be considered here.

Twenty young adults (10 males, 10 females) served as subjects. Otoscopic examinations were unremarkable; there were no reports of recent ear disease. Pure-tone HTLs were no poorer than 15 dB HL at all frequencies between 0.25 and 8 kHz. After meeting the subject selection criteria, pure-tone thresholds were determined at 4 kHz and 8 kHz three times for one ear of each individual (right or left ears were randomly chosen, but remained constant for all tests by each individual).

HTLs were determined using Békésy automatic audiometry for two tests within a single session, with further thresholds given one week later. TDH-49 earphones (with MX 41/AR cushions) were used for all tests; the authors did not report whether the earphones were re-positioned between tests 1 and 2. The HTL results are given in the table below.

Mean thresholds and standard deviations for the 20 test subjects;
for tests 1 and 2 within the same session, and for the single session one week later.

frequency, kHz	test	mean HTL, dB HL	standard deviation, dB
4	1	5.4	5.3
	2	6.2	3.8
	week later	2.3	5.3
8	1	5.1	5.2
	2	6.7	6.9
	week later	4.6	5.1

NOTE: the means involve all 20 test subjects; the
standard deviations are NOT within subjects.

Note that the mean thresholds (for all 20 subjects) increases slightly from test 1 to test 2 within the same session. This increase could be due to subject fatigue or boredom. The thresholds determined one week later are somewhat lower (more acute).

BS 6655-1986

This standard laid down requirements and procedures for conducting pure-tone air-conduction audiometry, appropriate to monitor the hearing of subjects exposed to noise at work. Methods were presented for conducting audiometric tests with manual audiometers, and with automatic-recording fixed-frequency audiometers. For manual audiometry, a bracketing method, and an ascending method were specified. Also specified were test methods using computer-controlled or other automated equipment.

Apropos the purpose of the present document, this standard gave measures of the reliability of audiometric measurements.

Typical values of the standard deviation of repeated threshold determinations according to this standard, given as a function of frequency.

frequency, kHz						
0.5	1	2	3	4	6	8
3.5	3.3	2.7	2.8	3.7	4.9	6.0

It is worth noting that threshold variability for repeat thresholds is relatively low for the middle frequencies, but somewhat higher for 6 and 8 kHz which would influence the appearance/aspect of a 'noise notch' expected in actual cases of NIHL. Bearing in mind the potential variability of thresholds at higher audiometric frequencies, the prudent audiometrician should produce repeat thresholds after removal and replacement of the earphones. If the initial and repeat values are not the same, a second repeat should be performed. The lowest value(s) should be reported as the 'true' threshold, that is, the lowest presentation acknowledged as heard.

JERIVALL and ARLINGER (1986)

The study was intended to make a comparison of the step sizes used in manual pure-tone audiometry, with regard to test-retest reliability. Audiograms were performed using the ascending method specified in ISO 6189 (see BS 6655: 1986), with step sizes of 2 dB and 5 dB.

Two subject groups participated: 10 individuals with normal hearing, and 10 with moderate high-frequency hearing losses. The 'normal' group comprised inexperienced listeners: 5 male and 5 female, aged 16 to 27 years, with HTLs up to a maximum of 20 dB HL. The hearing loss group was made up of 6 males and 4 female, aged 37 to 50 years; the mean hearing loss showed smoothly-increasing HTLs with increasing frequency, to values of 65 to 70 dB HL at 4 and 6 kHz, with slight recovery at 8 kHz.

The subjects gave thresholds on two different test occasions with a mean interval of 19 days in the normal-hearing group, and 8 days in the hearing-impaired group. On each test occasion, threshold determination were performed with the two step sizes. Half of the subjects in each test group started with the smaller step size and the other half with the larger step size. The right ear was always examined first. A complete audiogram with the selected step size was determined before changing the step size. The headset was not removed between the two step sizes at each test occasion. The test tones were presented in the sequence 1 kHz, 2, 3, 4, 6, 8, 0.5, 0.25, and 0.125 kHz, with a re-test at 1 kHz.

The audiometric test frequency 0.125 kHz is not usually used in cases of NHIL. The audiometric step sizes 'down by 4 dB, up by 2 dB' are not common in manual audiometry. The results for these two conditions are not reported below.

Mean test-minus-retest threshold differences by audiometric frequency; for 10 subjects in each test group

test group	frequency, kHz							
	0.25	0.5	1	2	3	4	6	8
normal	2.3	-0.3	0.0	0.3	0.8	0.3	1.3	1.5
impaired	-2.8	0.8	0.3	0.0	1.3	0.3	-1.6	0.6

Over most frequencies, the test-minus-retest threshold differences are positive, that is, the first test HTL was greater than the retest HTL.

Standard deviations of threshold differences by audiometric frequency; for 10 subjects in each test group

test group	frequency, kHz							
	0.25	0.5	1	2	3	4	6	8
normal	5.0	4.1	4.3	3.0	4.0	3.4	7.2	5.6
impaired	5.5	3.7	3.0	3.2	3.9	4.4	4.7	5.6

No clear trend is seen in respect of which group, normal or high-frequency hearing-impaired, exhibited greater overall threshold variability.

CLARK and ROESER (1988)

The relevant aspect of this paper deals with a comparison of audiometric results obtained using the TDH-50P supra-aural earphone and the ER-3A insert earphone. Conventional supra-aural earphones provide little attenuation to lessen the masking effects of ambient noise in a hearing test environment; the Tubephone offers a greater degree of attenuation, due to its insertion into the ear canal (much like an earplug).

All participating subjects were determined to have normal hearing sensitivity by pulsed-tone Békésy audiometry; all subjects reported no recent otological disorders (within the six months before testing). Normal hearing sensitivity was defined as thresholds of 20 dB HL or better (lower) bilaterally for the octave frequencies at and between 250 and 8 000 Hz, and at the intermediate frequencies 3 000 and 6 000 Hz.

Subjects of interest here were 12 adults between the ages of 22 and 33 years (mean age 27.0 years), with an equal number of males and females. Other parts of the study concentrated on children aged between 7 and 12 years of age; these sub-studies will not be reported here.

The adult subjects performed test and re-test audiograms using the TDH-50P supra-aural earphone and the ER-3A insert phone. The authors do not state explicitly the time interval between the test and re-test audiograms. The mean differences between first and second audiometric test sessions are given by earphone type, in the two tables below.

Mean threshold differences for first and repeat thresholds
and standard deviations (of differences),
determined using the TDH-50P supra-aural earphone.
(negative difference indicates first threshold better/lower)

measure	frequency, kHz 0.25	0.5	1	2	3	4	6	8
mean difference	1.0	-0.5	0.5	-1.0	-1.0	-2.0	-0.5	0.0
standard deviation	3.0	1.0	3.0	1.0	1.9	2.4	4.5	2.9

Mean threshold differences for first and repeat thresholds
and standard deviations (of differences),
determined using the ER-3A insert earphone.
(negative difference indicates first threshold better/lower)

measure	frequency, kHz 0.25	0.5	1	2	3	4	6	8
mean difference	-5.8	1.2	1.5	-0.5	-1.5	-2.0	7.5	6.5
standard deviation	6.4	5.0	3.6	3.0	3.0	5.0	7.9	7.3

Over the lower-to-mid frequencies 250 Hz to 3 000 Hz, the two output transducers gave similar HTL values. However, at 4 000 Hz, the ER-3A insert Tubephone gave thresholds that were, on average, 6 dB lower (better) that the TDH-50P supra-aural earphone. At 6 000 Hz, the situation reversed: the TDH thresholds were, on average, 8 dB lower (better) than the insert phones.

LARSON *et al.* (1988)

These researchers undertook a relatively large study to determine Reference Equivalent Sound Pressure Levels (RETSPLs) for the calibration of the Etymotic ER-3A insert earphone used in audiometry.

Equal numbers of male and female subjects were selected with ages from 18 to 31 years; each gave a negative history of ear pathology and of significant noise exposure. Each of three laboratories collected thresholds from the right ears of 30 subjects; the total subject pool comprised 90 subjects. This work also included determining RETSPLs for standard supra-aural earphones, for comparison purposes. Two laboratories used the TDH-50; the third used the TDH-39.

For each earphone, the results were reported as the mean thresholds determined by each of the three laboratories, with standard errors of the means. The authors also gave the means and standard errors over all 90 subjects. The mean thresholds (SPL values), although useful to standardise the earphone RETSPLs over frequency, are of no interest here. The standard errors (SEs) of the thresholds will, however, yield standard deviations (SDs) by the simple calculation:

$$SE = SD \,/\, (\text{sample size})^{½}$$

$$SE \times (\text{sample size})^{½} = SD$$

Standard deviations of the mean threshold for all three laboratories,
total subject sample of 90 ears

earphone	frequency, kHz							
	0.25	0.5	1	2	3	4	6	8
TDH	6.6	5.7	6.6	6.6	6.6	6.6	7.6	8.5
ER-3A	5.7	5.7	5.7	6.6	4.7	5.7	7.6	6.6

The standard deviations for the ER-3A insert phones are equal to, or smaller than, those for the TDH supra-aural earphones. These data suggest that the insert phones would give more repeatable thresholds than would the supra-aural earphones.

ARLINGER and KINNEFORS (1989)

The authors felt that, as use of insert earphones was on the increase, it was worth determining a standardised set of reference equivalent threshold sound pressure levels, for calibration of such transducers. To fulfil this aim, HTLs were determined for both ears of 18 subjects, aged 18 to 28 years; there were equal numbers males and females. All subjects were determined to be otologically normal, with no history of significant exposure to loud noise.

The audiometric survey used two headphone types: the TDH-39 supra-aural earphone used in conventional audiometry, compared against the Etymotic Research ER-3A insert Tubephone. HTLs were determined according to the ascending method defined in the international standard current at the time-of-test.

The subjects' threshold values are of no interest for present purposes. The present document is concerned with threshold variability; the standard deviations of HTLs by both earphones are given below.

Standard deviations for the thresholds
by the two earphones, by frequencies

earphone	frequency, kHz							
	0.25	0.5	1	2	3	4	6	8
TDH-39	3.8	3.5	3.7	3.5	3.7	4.2	5.2	5.2
ER-3A	4.7	3.2	4.2	4.0	4.2	5.6	5.2	4.9

The audiometric test frequencies 1 to 6 kHz are important in determining the degree of hearing loss in cases of NIHL. For these frequencies, the TDH-39 threshold deviations were smaller than, or equal to, those of the insert phone.

LINDGREN (1990)

This study examined the accuracy and reliability of HTLs determined using an ER-3A insert earphone, and using a TDH-49P supra-aural earphone. Thirteen subjects, ages 26 to 61 years, gave no history of ear disease, and exhibited normal middle ear pressures. Each subject gave hearing thresholds for one ear only, by sweep-frequency self-recording audiometry; five sessions used the insert phone, five sessions used the supra-aural phone. The results are reproduced below.

Mean HTLs by frequencies, dB HL

audiometric earphone	frequency, Hz							
	250	500	1k	2k	3k	4k	6k	8k
TDH-49P	4.1	-1.7	-4.2	0.5	-0.4	-0.1	0.7	10.1
ER-3A	2.1	1.2	-1.4	1.8	1.3	0.5	14.1	18.6
difference	2.0	-2.9	-2.8	-1.3	-1.7	-0.6	-13.4	-8.5

The threshold difference at 6 and 8 kHz were found to be significantly different (p = 0.05). The ER-3A insert phone is not accurate for the frequencies 6 and 8 kHz; the manufacturer recommends corrections of -10 dB to any measured HTLs for these two frequencies. Without such corrections, the observed values at 6 and 8 kHz would distort the aspect of any audiogram, perhaps disguising an existing 'noise notch'.

In order to determine the reliability of the repeat measurements on each subject, pooled standard deviations were calculated for each frequency over all sessions using a specific earphone. The results are reproduced below.

Within-subject pooled standard deviations by frequencies, in dB

audiometric earphone	frequency, Hz							
	250	500	1k	2k	3k	4k	6k	8k
TDH-49P	5.7	2.6	2.4	2.0	2.6	3.7	3.6	4.4
ER-3A	4.8	3.6	3.5	3.4	3.4	3.3	3.3	3.8
difference	0.9	-1.0	-1.1	-1.4	-0.8	0.4	0.3	0.6

The insert and supra-aural phones give comparable reliability, with differences relatively close to zero across the frequency range.

STUART *et al.* (1991)

The limitations associated with the use of standard supra-aural audiometric earphones are well documented. These include little exclusion of background noise in the test environment, the possibility of cross-hearing (in the non-test ear) for high-level stimulus presentations to the ear under test, the possibility of ear canal collapse/closure, and patient discomfort with extended wear. Use of insert earphones circumvents these limitations. At the time of writing, however, the authors felt that measures of test-retest variability of hearing thresholds had not been explored for children using insert earphones; they set out to examine this topic.

Thirty normal-hearing subjects were recruited to participate in this study. Two of the groups comprised children and adolescents; these subjects will not be considered here. The third group comprised 10 young adults, with mean age 24.1 years; their results are of interest here. The researchers defined normal hearing sensitivity as HTLs lower than or equal to 20 dB HL in at least one ear for the octave audiometric frequencies at and between 0.25 and 8 kHz. For each participant, one ear was selected randomly for this study; the HTLs were then determined by manual audiometry. After a short break, the HTLs were measured a second time, by a different audiometrician who was unaware of the results of the previous test.

The thresholds were determined for three earphone conditions: the conventional TDH-50 supra-aural earphone, and the Etymotic Research ER-3A insert phone coupled to the test ear by a probe cuff or a foam insert. The conventional supra-aural earphone is more likely to be used for HTLs in cases of NIHL. The mean differences between test and re-test threshold values are given below; the values are all positive, indicating that the re-test was the lower (better, more acute) threshold.

Mean differences between first and second thresholds, for 10 young adults.

earphone	test minus re-test, dB					
	frequency, kHz					
	0.25	0.5	1	2	4	8
TDH-50	0.0	1.0	2.0	1.5	1.0	4.0

The standard deviations are given below. No earphone gave consistently lower standard deviations across all frequencies.

Standard deviations of test-retest differences in HTLs, by 10 young adults

earphone	frequency, kHz					
	0.25	0.5	1	2	4	8
TDH-50	3.3	5.2	4.2	5.3	4.6	4.6
ER-3A (probe cuff)	4.7	4.1	3.4	5.3	4.1	7.1
ER-3A (foam tip)	4.4	3.2	6.2	5.7	2.6	5.9

FLOTTORP (1995)

The author based this study upon a belief that, considering industrial monitoring audiometry for people exposed to potentially injurious levels of occupational noise, the thresholds at 6 and 8 kHz are essential to determine if NIHL has developed since any previous test. Thresholds at these frequencies also figure in cases where calculations must be performed to quantify disablement benefit or compensation for hearing injury. Such calculations should be based upon the best (lowest, most acute) thresholds available.

As part of an industrial monitoring audiometry project spanning 10 years, the author tried a second placement of the supra-aural earphone in certain cases, to reduce the measured HTLs. This was done only when the audiograms showed "suspicious" losses at frequencies above 3 kHz, *e.g.* a normal threshold in one ear, but a contralateral hearing loss at 6 kHz and/or 8 kHz. Earphone replacements were also made on testees who, in previous years, had shown cause to reposition the supra-aural earphone(s).

From a total of 2 708 manual audiograms from 576 employees (54% male, 46% female) over 10 years, the author had cause to move the headphones (and repeat the high-frequency thresholds) in 870 cases (32%). For the cases of moving-and-retesting, the repeats gave threshold improvements between zero and 40 dB.

At 6 kHz, audiograms and yearly retests comprised 4 466 ears tested. The phones were repositioned for 790 of the ears tested; over these retests, a total improvement of 5 855 dB was determined, giving an average improvement of 7.4 dB. The improvement of 5 855 dB (for 790 ears) gives an average of 1.3 dB if all 4 466 ears are considered.

At 8 kHz, there were 4 691 ears tested, for which the phones were repositioned on 1 015 of the ears tested. A total improvement of 7 350 dB was found, giving an average improvement of 7.2 dB. For all 4 691 ears with HTLs at 8 kHz, the improvement gives an average of 1.6 dB.

The author recognised that "… a new threshold determination at those frequencies after moving the headphone may have a significant influence." Such improvements at 6 and/or 8 kHz could indeed influence the amount of any compensation payable in a personal injury claim for occupational NIHL.

POULSEN and HAN (2000)

The hearing threshold for a number of pure tones was determined under conditions of binaural listening in a free progressive wave, with each subject directly facing the sound source (frontal incidence), and with the sound pressure level measured in the free progressive wave at the centre position of the listener's head with the listener absent.

Thirty-one student-subjects participated in the tests, 14 females and 17 males, in the age range from 18 to 25 years. Each was selected as 'otological normal' by history and screening tests. As part of the screening process, HTLs were determined for the subject group by manual audiometry using the shortened version of the ascending method given in ISO 8253-1: 1989, that is, the threshold determined by two responses at the same level out of three ascending presentation series. The step size was 5 dB. The HTLs were determined using TDH-39 earphones with MX41-AR cushions. The dispersions of the manual thresholds are given below.

Standard deviations of HTLs by conventional manual audiometry, by test frequency.

frequency, kHz	0.25	0.5	1	2	3	4	6	8
standard deviation, dB	4.7	4.4	3.5	4.9	5.3	6.6	6.3	7.1

The main intention of the authors was to determine and report the subjects' thresholds by binaural free-field listening. **This test method would NOT be used to quantify a person's HTLs for any compensation claim involving NIHL.** The dispersion data, reported below, are included to confirm the relative change of standard deviations with audiometric test frequency found with conventional manual audiometry, used to screen the 31 student-subjects.

For the free-field tests, the signal duration was 1 second measured between the half-amplitude points of the rise and fall course. The rise and fall times lasted 50 milliseconds. The pause between presentations was randomly chosen between 1 s and 2.4 s. The observed thresholds are of no interest here; the standard deviations are of interest, and are reproduced below.

Standard deviations of free-field thresholds of hearing, by test frequency.

frequency, kHz	0.25	0.5	1	2	3	4	6	8
standard deviation, dB	4.2	3.9	3.8	5.0	6.5	5.9	6.2	5.6

BS EN ISO 8253-1: 2010

Annex A of this reference gives measurement uncertainty of hearing threshold levels.

"During routine audiometry, the hearing threshold level of a test subject at a certain frequency is usually determined just once for each ear. However, based upon empirical knowledge, the following standard uncertainties for repeated measurements under identical test conditions can be assumed … for air conduction audiometry: 2.5 dB at frequencies up to 4 kHz and 4 dB at frequencies above 4 kHz."

The HTL of a test subject, for a certain frequency, is said to follow a normal or Gaussian distribution. The term "standard uncertainty" is used here to allow combination with other standard uncertainties (not following a Gaussian distribution) for other test factors.

VINAY *et al.* (2015)

This study investigated test–retest variability and time efficiency for a new automated hearing assessment method; the automated test results were contrasted with thresholds and variability from conventional manual Pure Tone Audiometry (PTA). Only these PTA results will be reported here.

Thirty-two participants were verified as having normal hearing sensitivity, with audiometric thresholds better than 20 dB HL for pure tone frequencies from 0.25 to 8 kHz in one or both ears. For the study itself, air conduction thresholds were measured once by two experienced audiometricians at 1, 3, 4 and 6 kHz using Sennheiser HAD 200 circumaural headphones. Two testers were involved in order to eliminate the effects of tester bias on the resulting HTLs. Subjects were allowed a brief break between the two test measurements, PTA1 and PTA2. The order of frequencies was randomised in both tests across all participants. Randomization across ears was done for subjects who were tested in both ears. For each participant, all testing was done within the same day.

The audiometric thresholds were measured using a modified Hughson-Westlake technique. The initial presentation of the signal was 40 dB HL; the level was varied both up and down to determine the thresholds. Subjects were asked to acknowledge hearing the test tone by pressing a response button only when they were sure that they heard the tone. A positive response from the test subject was followed by a decrease in level by 4 dB; a no-response was followed by an increase in the tone level by 2 dB. The hearing threshold was defined as the minimum (lowest) level at which the test subject responded to the presence of the tone at least two out of three times, either during the ascent or the descent method. All participants gave average thresholds less than or equal to 15 dB HL.

The results of interest here are given in the table below. The reader should note that positive differences indicate that PTA2 were lower (better) thresholds; this is probably a learning effect.

Average test-retest threshold differences with pure-tone manual audiometry (PTA) by 39 subjects.

PTA1 minus PTA2	frequency, kHz			
	1	3	4	6
Mean difference	2.6	2.7	3.1	3.0
Standard deviation of differences	2.5	2.3	2.9	2.5

3. CONCLUSIONS

The table below brings together the important results from the reviews, in respect of differences between repeated thresholds. At the bottom of the table are listings, by frequency, of the range of values, and the middle observation within this range (with interpolation in the case of an even number of observations).

Differences (in dB) between repeated thresholds, from the reviews
(test minus retest: a positive value indicates lower/improved threshold on retest)

	frequency, kHz							
	0.25	0.5	1	2	3	4	6	8
BURNS and HINCHCLIFFE (1957)		1.0	2.2	1.5	2.0	1.4	-1.7	1.0
ROBINSON, SHIPTON and HINCHCLIFFE (1981)		1.4	1.2	1.6	1.4	0.2	0.6	
CHERMAK, DENGERINK and DENGERINK (1983)						3.1		0.5
JERIVALL and ARLINGER (1986)								
normal	2.3	-0.3	0.0	0.3	0.8	0.3	1.3	1.5
impaired	-2.8	0.8	0.3	0.0	1.3	0.3	-1.6	0.6
CLARK and ROESER (1988)								
TDH	1.0	-0.5	0.5	-1.0	-1.0	-2.0	-0.5	0.0
ER-3A	-5.8	1.2	1.5	-0.5	-1.5	-2.0	7.5	6.5
STUART *et al.* (1991)								
TDH	0.0	1.0	2.0	1.5		1.0		4.0
ER-3A	-0.5	1.0	0.0	1.0		2.0		2.5
FLOTTORP (1995)							1.3	1.6
VINAY *et al.* (2015)			2.6		2.3	3.1	3.0	
highest	2.3	1.4	2.6	1.6	2.7	3.1	3.0	6.5
middle value from the table	-0.25	1.0	1.2	0.65	1.35	0.65	0.6	1.05
lowest	-2.8	-0.5	0.0	-1.0	-1.0	-2.0	-1.7	0.0

Considering the highest table values, the threshold differences show modest improvements (on retest) across the frequency range, save for the larger threshold improvement at 8 kHz. The reader should be aware that the mean differences may hide large improvements. ROBINSON and WHITTLE (1973) disregarded “gross initial improvements exhibited by some subjects”. FLOTTORP (1995) investigated the need to remove and replace the audiometric earphones for retest thresholds; the repeats gave threshold improvements between zero and 40 dB.

Across the frequencies, the lowest differences show mixed results, either no improvement on retest, or higher (less acute) thresholds on retest.

The middle values show mostly small improvements on retest. Such small improvements are less than the decibel increments used to determine HTLs; such small threshold improvements might not show in the reported HTLs.

The table below brings together the important results from the reviews, in respect of the standard deviations for repeated thresholds within selected subject groups.

Standard deviations (in dB) from the reviews

	frequency, kHz							
	0.25	0.5	1	2	3	4	6	8
ATHERLEY and DINGWALL-FORDYCE (1963)		2.9	2.7	2.6	2.4	4.1	4.4	4.8
ROBINSON, SHIPTON and HINCHCLIFFE (1981) OAS 0-3		4.9	5.1	5.4	6.4	5.8	8.5	
CHERMAK, DENGERINK and DENGERINK (1983)						5.3		5.1
BS 6555-1986		3.5	3.3	2.7	2.8	3.7	4.9	6.0
LARSON *et al.* (1988)								
TDH	6.6	5.7	6.6	6.6	6.6	6.6	7.6	8.5
ER-3A	5.7	5.7	5.7	6.6	4.7	5.7	7.6	6.6
ARLINGER and KINNEFORS (1989)								
TDH	3.8	3.5	3.7	3.5	3.7	4.2	5.2	5.2
ER-3A	4.7	3.2	4.2	4.0	4.2	5.6	5.2	4.9
LINDGREN (1990)								
TDH	5.7	2.6	2.4	2.0	2.6	3.7	3.6	4.4
ER-3A	4.8	3.6	3.5	3.4	3.4	3.3	3.3	3.8
STUART *et al.* (1991)								
TDH	3.3	5.2	4.2	5.3		4.6		4.6
ER-3A	4.4	3.2	6.2	5.7		2.6		5.9
POULSEN and HAN (2000)								
TDH	4.7	4.4	3.5	4.9	5.3	6.6	6.3	7.1
field	4.2	3.9	3.8	5.0	6.5	5.9	6.2	5.6
BS EN ISO 8253-1: 2010	2.5	2.5	2.5	2.5	2.5	2.5	4.0	4.0

highest	6.6	5.7	6.6	6.6	6.6	6.6	8.5	8.5
middle value from the table	4.7	3.55	3.75	4.45	3.95	4.6	5.2	5.15
lowest	2.5	2.5	2.4	2	2.4	2.5	3.3	3.8

It is worth noting, from the summary section immediately above, that the standard deviations at 6 and 8 kHz are uniformly greatest for the highest, middle and lowest values. In respect of NIHL, the threshold at 6 kHz plays a part in quantifying the depth of any ‘noise notch’; the threshold at 8 kHz is useful in confirming the existence of a ‘noise notch’. As these two frequencies are subject to substantial threshold variation, repeat tests would prove informative.

4. RECOMMENDATIONS

For an individual test subject, a single audiogram is an unconfirmed determination of that individual's state-of-hearing in both ears. Put more starkly, **a single audiogram is a guess**.

For an individual test subject, a second audiogram may be expected to result in different thresholds at some frequencies. Such repeat audiometry should be done after a break/rest from the concentration required for the difficult audiometric task. Such a break would offer several advantages:

- avoidance or minimisation of fatigue in the testee;
- offer the opportunity for the audiometrician to remove and replace the earphones; and
- provision of an interval during which the audiometrician may offer encouragement or re-instruction to the subject.

In order to minimise the total concentration time of the listener, the repeat thresholds could be restricted to the frequency range 3 kHz to 8 kHz.

After the completion of two audiometric series (full frequency range in both ears, restricted range repeated in both ears), the prudent audiometrician should produce a **composite audiogram** giving the better (lower, more acute) thresholds from the first and repeat audiograms. The reported audiogram should make a clear statement of the composite feature.

A third audiogram series might produce even lower thresholds at some frequencies, but at the likely cost of subject fatigue, impatience and frustration with the difficult listening task. Such an attempt to obtain ever-lower thresholds would probably fail.

REFERENCES

Atherley GRC, Dingwall-Fordyce I (1963). The reliability of repeated auditory threshold determination. *Brit J indust Med*, **20**, 231-235.

British Society of Audiology (2011). *Recommended Procedure: Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking*. BSA, Reading.

BS EN ISO 8253-1:1989. *Acoustics – Audiometric Test Methods Part 1: Basic pure tone air and bone conduction audiometry*. British Standards Institution, London.

BS 6655:1986, EN 26189:1991, ISO 6189:1983. *Specification for pure tone air conduction audiometry for hearing conservation purposes*. British Standards Institution, London

BS EN ISO 8253-1:2010. *Acoustics – Audiometric Test Methods Part 1: Pure-tone air and bone conduction audiometry*. British Standards Institution, London.

Burns W, Hinchcliffe R (1957). Comparison of the auditory threshold as measured by individual pure-tone and Bekesy audiometry. *J Acoust Soc Am*, **29** (12), 1274-1277.

Chermak GD, Dengerink JE, Dengerink HA (1983). Test-rest reliability of auditory threshold and temporary threshold shift. *Scand Audiol*, **12**, 237-240.

Clark JL, Roeser RJ (1988). Three studies comparing performance of the ER-3A Tubephone with the TDH-50P earphone. *Ear and Hearing*, **9**(5); 268-274.

Lawton BW and Robinson DW (1999). *A Concise Vocabulary of Audiology and allied topics*. ISBN 0854326839, Institute of Sound and Vibration Research, University of Southampton. Online: http://www.isvr.co.uk/reprints/vocab.htm

Lindgren F (1990). A comparison of the variability in thresholds measured with insert and conventional supra-aural earphones. *Scand Audiol*, **19**, 19-23.

Flottorp B (1995). Improving audiometric thresholds by changing the headphone position at the ear. *Audiol*, **34**, 221-231.

ISO 389: 1975. *Acoustics – Standard reference zero for the calibration of pure-tone audiometers*. International Organization for Standardization, Geneva.

Larson VD, Cooper WA, Talbott RE, Schwartz DM, Ahlstrom C, De Chicchis AR (1988). Reference threshold sound-pressure levels for the TDH-50 and ER-3A earphones. *J Acoust Soc Am*, **84** (1), 46-51.

Robinson DW, Whittle LS (1973). A comparison of self-recording and manual audiometry: Some systematic effects shown by unpractised subject. *J Sound Vib*, **26** (1), 41 – 62.

Robinson DW, Shipton MS, Hinchcliffe R (1981). Audiometric zero for air conduction: A verification and critique of International Standards. *Audiol*, **20**, 409-431.

Vinay SN, Svenson UP, Kvaløy O, Berg T (2015). A comparison of test–retest variability and time efficiency of auditory thresholds measured with pure tone audiometry and new early warning test. *App Acoust*, **90**, 153–159

APPENDIX: methods for obtaining thresholds of hearing

This section contains excerpts from a British Standard and from the British Society of Audiology, giving summaries of recommended methods for obtaining thresholds of hearing.

BS EN ISO 8253-1:2010
ISO 8253-1: 2010(E)

A.1 Instruction of test subjects

In order to achieve reliable test results, it is essential that relevant instruction in the test procedure be given unambiguously, and that these are fully understood by the test subject.

The instructions shall be phrased in language appropriate to the listener and shall normally indicate:

a) the response task;
b) the need to respond whenever the tone is heard in either ear, no matter how faint it may be;
c) the need to respond as soon as the tone is heard and to stop responding immediately once the tone is no longer heard;
d) the general pitch sequence of the tones; and
e) the ear to be tested first.

The response from the test subject indicating when the tone is heard and when it is no longer heard shall be clearly observable. Examples of commonly used responses are: pressing and releasing a signal switch; or raising and lowering of a finger or hand.

Test subjects shall also be instructed to avoid unnecessary movements so as to obviate extraneous noise. After the instructions have been given, the test subject shall be asked if he or she has understood. The test subject shall be informed that he or she may interrupt the test in the case of discomfort. If there is any doubt, the instructions should be repeated.

A.2 Placement of transducers

In advance of testing, the following actions should be undertaken: spectacles, head ornaments, and hearing aids shall be removed. Hair shall be moved from between the head and the sound transducers, i.e. earphones and bone vibrators, if possible. The transducers shall be fitted by the tester to ensure that they are properly positioned and subjects shall be instructed not to touch the transducers thereafter. The sound opening of an earphone shall face the ear canal entrance.

A.3 Air conduction hearing threshold level determinations using fixed-frequency audiometry

The audiometric test may be carried out using a manual audiometer or an automatic-recording audiometer.

The order of presentation of test tones when the audiometer settings are made manually shall be from 1 000 Hz upwards, followed by the lower frequency range, in descending order. A repeat test shall be carried out at 1 000 Hz on the ear tested first.

Preferably, automatic-recording audiometers should present test tones in the same sequence as in manual audiometry.

A.3.1 Manually controlled threshold determination: presentation and interruption of test tones

The test tone shall be continuous and presented for a duration of 1 second to 2 seconds. When a response is given, the interval between tone presentations shall be varied but shall not be shorter than the test tone duration.

Automatically pulsed tones are sometimes used as an alternative stimulus. The use of such stimuli should be noted on the audiogram.

A3.2 Initial familiarization

The test subject shall be made familiar with the task prior to threshold determination by presenting a signal of sufficient intensity to evoke a definite response. By using the familiarization step, the tester can be sure that the test subject understands and can perform the response task.

EXAMPLE The following method of familiarization can be used:

a) present a tone of 1 000 Hz at a hearing level which is clearly audible, *e.g.* 40 dB HL for a normal hearing test subject;
b) reduce the level of the tone in steps of 20 dB until no response occurs;
c) increase the level of the tone in steps of 10 dB until a response occurs;
d) present the tone again at the same level.

If the responses are consistent with the tone presentation, the familiarization is complete. If not, it should be repeated. After a second failure, the instructions should be repeated.

A4 Procedure for hearing threshold measurements without masking

Two audiometric test procedures with a manual audiometer are specified: a *bracketing* method and an *ascending* method. These methods differ only in the sequence of the levels of the test tones presented to the test subject.

In the ascending method, consecutive test tones having ascending levels are presented until a response occurs.

In the bracketing method, consecutive test tones having ascending levels are presented until a response occurs, after which test tones having levels in a descending sequence are presented.

When properly carried out, both methods result in substantially the same hearing threshold levels.

Measurements using the ascending method differ from those of the bracketing method only in step 2 of the measurements presented below.

If the hearing threshold level measurements result in a hearing level of 40 dB or more in either ear at any frequency, these results should be interpreted with caution due to the phenomenon of cross-hearing. Contralateral masking can then be necessary.

A4.1 Step 1

Present the first test tone at a level which is 10 dB below the lowest level of the response of the test subject during the familiarization session. After each failure to respond to a test tone, increase the level of the test tone in steps of 5 dB until a response occurs.

A4.2 Step 2

Ascending method

After the response, decrease the level in steps of 10 dB until no response occurs. Then begin another ascent with 5 dB steps. Continue until three responses occur at the same level out of a maximum of five ascents. This level is then defined as the hearing threshold level.

If less than three responses out of five ascents have been obtained at the same level, present a test tone at a level 10 dB higher than the level of the last response. Then repeat the general test procedure: 10 dB down after a response, 5 dB up until a response occurs.

A shortened version of the ascending method has been shown to yield nearly equivalent results and may be appropriate in some cases. In this shortened version, continue the testing until at least two responses occur at the same level out of three ascents.

Bracketing method

After the response, increase the level of the test tone by 5 dB and begin a descent in which the level of the tone is decreased in steps of 5 dB until no response occurs. Then decrease the level of the test tone another 5 dB and begin the next ascent at this level. This should be continued until three ascents and three descents have been completed.

Shortened versions of the bracketing method may be appropriate in some cases. Shortening consists of omitting the further descent of 5 dB after no response occurs or requiring only two ascents and two descents in series provided that the four minimal response levels differ by no more than 5 dB.

A4.3 Step 3

Proceed to the next test frequency at an estimated audible level, as indicated by the previous responses, and repeat step 2. Finish all test frequencies on one ear.

Finally, repeat the measurement at 1 000 Hz. If the results at 1 000 Hz of the repeat measurement for that ear agree to 5 dB or less with those of the first measurements for the same ear, proceed to the other ear. If 10 dB or more improvement or worsening in hearing threshold level is discernible, retest at further frequencies in the same order until agreement to 5 dB or less has been obtained.

A4.4 Step 4

Proceed until both ears have been tested.

A.5 Calculation of hearing threshold level

The hearing threshold levels for each frequency and ear shall be determined in accordance with the following procedures, dependent upon the measurement method used.

Ascending method
For each frequency and ear, determine the lowest level at which responses occur in more than half of the ascents. This level is defined as the hearing threshold level.

If the lowest response levels span more than 10 dB at a given frequency, the test should be considered of doubtful reliability and should be repeated. This should be noted on the audiogram.

Bracketing method
For each frequency and ear, average the lowest levels at which responses occur in the ascents. Again, for each frequency and ear, average the lowest levels at which responses occur in the descents. Determine the mean value of the two averages obtained in this way for each frequency and ear. This mean value, rounded to the nearest 5 dB step, is taken as the hearing threshold level for that frequency and ear.

If the lowest response levels in the ascents deviate by more than 10 dB among themselves or if the lowest response levels in the descents deviate by more than 10 dB among themselves, the test should be repeated.

British Society of Audiology (2011)

This document presents a Recommended Procedure by the British Society of Audiology (BSA). Its purpose is to describe standard procedure and recommendations for effective pure-tone audiometry carried out in most audiological contexts.

B1 Before testing:

The test-subject must be asked about any exposure to loud noise during the previous 24 hours, as this can cause a temporary hearing loss. If the results may have been affected by recent noise exposure, then it may be necessary to re-test the subject at a time when that person has had no recent exposure to noise.

Subjects shall be asked if they have tinnitus, as this may affect their ability to detect tones in one or both ears.

Subjects shall be asked if they have better hearing in one ear; if so testing should commence with that ear, otherwise testing may start in either ear.

Instructions must give clear information about the task. This could be as follows:

> "I am going to test your hearing by measuring the quietest sounds that you can hear. As soon as you hear a sound (tone), press the button. Keep it pressed for as long as you hear the sound (tone), no matter which ear you hear it in. Release the button as soon as

you no longer hear the sound (tone). Whatever the sound and no matter how faint the sound, press the button as soon as you think you hear it, and release it as soon as you think it stops."

The subject's response to the test tone should clearly indicate when the test tone is heard and when it is no longer heard. The subject should be asked if he or she understands the instructions.

B2 Initial familiarisation

To ensure the subject is familiar with the task, present a tone of 1 kHz that is clearly audible (*e.g.* at 40 dB HL for a normally-hearing subject, or approximately 30 dB above the estimated threshold for a subject with a hearing impairment. If there is no response, increase the level of the tone in 10 dB steps until a response occurs. If the responses are consistent with the tone presentation (*i.e.* onset and offset), the subject is familiarised with the task. If not, repeat. If after this repeat, the responses are unsatisfactory, re-instruct the subject.

B3 Test order

Start with the better-hearing ear (according to the subject's account) and at 1 kHz. Next, test 2 kHz, 4 kHz, 8 kHz, 500 Hz and 250 Hz in that order. Then, for the first ear only, retest at 1 kHz. If the retest value is no more than 5 dB different from the original value, take the more sensitive threshold as the final value. [If the retest value differs from the original value by more than 5 dB, the reason for the variation shall be investigated. The subject may need to be re-instructed and the full test repeated for that ear.]

Where needed and practicable, test also at intermediate frequencies 750 Hz, 1.5 Hz, 3 kHz and 6 kHz [3 kHz and 6 kHz may be required in cases of high-frequency hearing loss such as NIHL]. Test the opposite ear in the same order. The retest at 1 kHz is normally not required in the second ear unless tests in the first ear revealed significant variation.

B4 Timing of the test stimuli

The duration of the presented tone shall be varied between 1 and 3 seconds. The interval between the tones shall be varied between 1 sec and at least 3 sec. The tester must ensure that the timing of each tone is not predictable; random variations in durations (both on and off) are intended as a check against false-positive responses. It is important that the tester does not stop the signal as soon as the subject responds, signals must be of the full duration, and the subject must respond for the full duration of each presentation.

B5 Method for finding threshold

1. Following a satisfactory positive response (tone heard), reduce the level of the tone in 10 dB steps until no further response occurs.

2. Increase the level of the tone in 5 dB steps until a response occurs.

3. After the first response using an ascending approach, decrease the level by 10 dB and begin another ascending 5 dB series until the subject responds again.

4. Continue to decrease the level by 10 dB and increase by 5 dB until the subject responds at the same level on two out of two, three or four (i.e. 50 % or more) responses on the ascent. This is the hearing threshold level.

5. Proceed to the next frequency, starting at a clearly audible level (e.g. 30 dB above the adjacent threshold, and use the 10 dB down, 5 dB up sequence described in Step 4 until the threshold criterion is satisfied.

6. Complete the frequency order in this first ear, and proceed to the other ear.

Care should be taken not to fatigue the subject as this can affect the reliability of the test results. If the test time exceeds 20 minutes, subjects may benefit from a short break.

B6 Audiometric descriptors

The hearing threshold levels of an individual ear may be described in descriptive terms rather than using the actual HTL values at different frequencies on a pure-tone audiogram. Recommendations are made below to associate particular descriptors with bands of average hearing impairment.

Four audiometric descriptors are given. These are based on the average of the pure-tone hearing threshold levels at 250 Hz, 500, 1 kHz, 2 k and 4 k. Averages do not imply any particular configuration of hearing loss and do not exclude additional terms (*e.g.* profound high-frequency hearing loss) being used.

Descriptor	**Average hearing threshold levels (dB HL)**
Mild hearing loss	20-40
Moderate hearing loss	41-70
Severe hearing loss	71-95
Profound hearing loss	in excess of 95

Average hearing threshold levels of less than 20 dB HL do not necessarily imply normal hearing.

DEFINITIONS of audiological terms

These definitions are excerpts from Lawton and Robinson (1999).

audiometer

An electroacoustical instrument, equipped (for air conduction) with earphones and headband, which provides pure tones at specified frequencies and known sound pressure levels, used to determine hearing threshold levels, one ear at a time.

(a) Manual audiometer: one in which the signal presentations, the selection of frequency and hearing level, as well as the noting of the subject's responses, are performed manually by the audiometrician.

(b) Automatic-recording audiometer (also called self-recording audiometer): one in which the signal presentations and the changes of hearing level and frequency are implemented automatically at set rates; only the direction of hearing level change is under the subject's control. Recording of the subject's responses is also done automatically. An automatic-recording audiometer may have facilities for presenting fixed frequencies or a continuously-variable (sweep) frequency, or both; it may also provide continuous as well as pulsed tone outputs.

(c) Computer-controlled audiometer: one in which the test procedure is controlled by a computer or microprocessor. Often, the hearing threshold levels are calculated by a pre-set programme, for display or print-out.

Békésy audiometry

A form of automatic-recording pure-tone audiometry employing a continuous frequency sweep (glide tone). Use of the eponymous term should be confined to its original meaning to distinguish Békésy audiometry from fixed-frequency automatic-recording audiometry, sometimes called self-recording audiometry.

pure-tone audiometry

A technique for determining a person's hearing threshold levels for pure tones by behavioural means, usually understood to employ a manual technique as described under **audiometer (a)**. Sound may be applied monaurally by means of an earphone (termed air-conduction audiometry), or vibrations may be applied to the skull by a bone vibrator (termed bone-conduction audiometry).

decibel (dB)

The unit for measuring the relative magnitude of a quantity based on a logarithmic scale. Commonly used to specify sound pressure level; A-weighted sound pressure level; hearing level; and hearing threshold level.

hearing level (HL)

Of a pure tone generated by a specified type of transducer for a specified frequency and manner of application, the sound pressure level (or the vibratory force level) of the tone, produced by the transducer in a specified calibration device (artificial ear, acoustic coupler or mechanical coupler) minus the appropriate reference equivalent threshold sound pressure level (or reference equivalent threshold force level). Thus, dial settings in hearing level of a correctly-calibrated audiometer indicate decibels relative to audiometric zero.

hearing threshold level (HTL)

Of a given ear, for a pure tone of specified frequency, the threshold of hearing at that frequency, expressed as hearing level. Note that hearing threshold level is a property of the ear under test whereas hearing level refers only to the sound (or vibration) generated by an audiometer.

audiometric frequencies

The series of frequencies conventionally employed in pure-tone audiometry. The series comprises the preferred frequencies at one-octave intervals from 125 to 8 000 Hz, usually supplemented by 3 000 and 6 000 Hz.

otological normality

An ideal definition of otological normality would be the notional state of complete freedom from derangement, both of form and function, of the auditory system. Natural biological variability and ageing are not exclusions. In practice, normality is based upon a specified set of exclusion criteria, applied with greater or lesser rigour, to a total population.

In the principal audiological Standards, an *otologically normal person* is defined as:

> A person in a normal state of health, who is free from all signs or symptoms of ear disease and from obstructing wax in the ear canals, and who has no history of undue exposure to noise.

For the purpose of specifying calibration levels for audiometric instruments (audiometric zero), the otologically normal population is restricted to the age range 18-30 years. Recent extensions of the Standards (*e.g.* to higher frequencies and to speech audiometry) have introduced a narrower age range (18-25 years) and additional exclusion criteria, *viz.* history of familial hearing loss and known exposure to potentially ototoxic drugs.

threshold of hearing

The minimum level of a sound which is just audible in given conditions on a specified fraction of trials (conventionally 50%). The term is often understood to imply quiet listening conditions, that is, it represents the irreducible, absolute threshold.

acoustic trauma

Instantaneous injury to, or destruction of, a component or components of the auditory system resulting from exposure to a very high transient sound pressure, e.g. from explosion or weapons fire. The term is **not to be confused** with noise-induced hearing loss from chronic exposure.

audiometric zero

For pure-tone air-conduction audiometry: a set of sound pressure levels of pure tones at audiometric frequencies, intended to typify the threshold of hearing of young otologically normal persons. For each frequency, the value is expressed by the sound pressure level measured in an acoustic coupler or artificial ear when the earphone, driven by a specific electrical signal, is placed on it.

This value is known as the reference equivalent threshold sound pressure level (RETSPL) for the frequency in question. The specific electrical signal is such that the sound pressure level it generates under the earphone when placed on the average human ear corresponds to the modal value of the thresholds of hearing of a group of young otologically normal persons of both sexes within a specified age range.

The term audiometric zero is also taken to mean the 0 dB HL line on audiogram charts.

frequency

The rate of oscillation of an acoustic or vibratory signal, symbol f. The unit is the hertz (Hz), one complete oscillation per second. High frequencies can conveniently be expressed in kilohertz (kHz).

hearing loss

The amount in decibels by which an individual's hearing threshold level changes for the worse, commonly understood to refer to the combined loss from all causes. The term may also be applied to that part of the overall loss which is attributable to a known influence (for example, noise-induced hearing loss) or a combination of contributing causes (for example, age-associated hearing loss). The related term *threshold shift* implies before-and-after comparison, whereas hearing loss commonly assumes a notional starting point such as audiometric zero.

In a qualitative sense, hearing loss is used loosely to mean a symptom of hearing disorder, and is often modified by descriptors, e.g. conductive, sensorineural, etc.

hearing threshold level (HTL)

Of a given ear, for a pure tone of specified frequency, the threshold of hearing at that frequency, expressed as hearing level in dB HL. Note that hearing threshold level is a property of the ear under test whereas hearing level refers only to the sound (or vibration) generated by an audiometer.

notch, noise notch

Of a pure-tone audiogram, a colloquial term for a sharply elevated (worse) threshold level over a narrow frequency range, flanked by lower (better) thresholds. A notch is often associated with noise-induced hearing loss.

sound pressure level

The sound pressure level of a sound in air, in decibels (dB), is equal to 20 times the logarithm to the base 10 of the ratio of the root-mean-square sound pressure to the reference sound pressure (20 micropascals). The reference sound pressure is zero dB on the scale of sound pressure level.