Brit. J. industr. Med., 1963, **20**, 231.

THE RELIABILITY OF REPEATED AUDITORY THRESHOLD DETERMINATION

BY

G. R. C. ATHERLEY and I. DINGWALL-FORDYCE

From the Department of Occupational Health, University of Manchester

(RECEIVED FOR PUBLICATION JANUARY 21, 1963)

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

Twelve otologically normal young men were tested on four separate occasions at 1, 2, 3, 4, 6, 8, 0·5, and 1 kc/s. The tests were carried out in a mobile test room installed in a specially constructed vehicle chassis.

The acoustic output of the audiometer and ear-phones was measured at intervals throughout the investigation. Output stability with variations of mains supply voltage and drift during the warming-up period of the instrument were also measured.

It was concluded that the instrument variation had been extremely slight throughout the investigation.

The estimates of variance of repeated threshold determinations on a single ear were found to be 8·5 $(dB)^2$ at 0·5 kc/s, 6 $(dB)^2$ at 3 kc/s, and 23 $(dB)^2$ at 8 kc/s. Differences between consecutive determinations extended to 25 dB.

These results were obtained under conditions which practically precluded all sources of variation other than that due to the inherent uncertainty of audiometric measurements. It appears to follow, therefore, that if an apparent drop in auditory threshold in one ear is to be considered as significant evidence ($P = 1\%$) of a real change, the difference would have to be at least 17·5 dB at the higher frequencies. This level could possibly be reduced to 10 dB if the change occurred simultaneously at both 4 and 6 kc/s.

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 chance fluctuations due to the inherent variability of auditory threshold measurements. Many investigations into the reliability of audiometric measurements have been reported but the implications of the findings have seldom been extended to assist the practical evaluation of serial results. The object of the present 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 subjects were 12 male medical students in whom no ear, nose or throat pathology could be demonstrated and who gave no history of noise exposure. Attention was paid to the possibility of the subjects developing rhinitis, either coryzal or allergic in origin, during the course of the experiment. As far as is known no test was carried out on a subject suffering from either of these conditions. The possibility of a temporary threshold shift resulting from the recent use of noisy transport was also excluded by arranging to perform the tests between lectures, rather than on the individual's first arrival in the morning. Enquiry was also made concerning previous treatment by drugs with VIIIth nerve toxicity.

It was considered that this group comprised otologically normal subjects as defined in British Standards 2497 (1954) and this assumption is supported by the actual audiometric results, which showed no gross departure from the normal threshold of hearing.

For the purpose of this experiment it was necessary to demonstrate that the acoustic output of the audiometer did not vary during the course of the experiment, and an attempt was made to show that the performance of the audiometer and ear-phones was consistent.

It was not possible, within the practical limits of the equipment involved, to devise a wholly satisfactory way of checking the performance but a compromise technique was used. The precise calibration of an audiometer and ear-phones to exclude all doubts about the consistency of performance at normal threshold sound pressure is only within the scope of a highly specialized physical laboratory, and the technique evolved probably represents the most usual type of calibration method employed in work of this nature.

Acoustic output was measured using a coupler which adapted closely the contours of the MX/41 ear-caps to the condenser microphone fitted to a Bruel and Kjaer type 2203 sound level meter. This coupler was similar in design to the American N.B.S. type 9A. The procedure adopted was to record the acoustic output with the main audiometer attenuator set to 80 dB hearing loss, for both ear-phones. These readings were taken at irregular intervals throughout the experiment and the standard deviations of results obtained are set out in Table 1. Preceding each set of measurements the sound level meter was standardized with a Bruel and Kjaer "Pistonphone" standard sound source and it was assumed that the acoustic characteristics of the coupler did not alter. The characteristics of the coupler and the conformity of the audiometer output to normal threshold of hearing were determined by comparative measurements using an ear-phone of the same pattern which has been fully calibrated by the National Physical Laboratory. The acoustic output at an attenuator setting of 80 dB showed a good degree of consistency but it must be acknowledged that the conformity at threshold level can only be inferred. The electrical output of the audiometer was measured with a high impedance valve voltmeter, and the sensitivity of the ear-phones at a standard voltage input was determined to enable a comparison with the reference pattern to be made.

TABLE 1

ACOUSTIC OUTPUT OF EAR-PHONES MEASURED IN dB, re. 0·0002 dyn/cm^2., ON 7 OCCASIONS. ATTENUATOR SET AT 80 dB HEARING LOSS

f(c/s)	Right Ear-phone Mean Output (dB)	Right Ear-phone S.D. (dB)	Left Ear-phone Mean Output (dB)	Left Ear-phone S.D. (dB)
250	109·3	0·44	110·0	0·22
500	95·5	0·33	96·6	0·23
1,000	88·6	0·39	89·5	0·16
2,000	87·2	0·44	88·1	0·30
3,000	88·1	0·45	88·0	0·64
4,000	87·7	0·36	88·3	0·60
6,000	90·1	0·46	90·5	0·63
8,000	88·7	0·49	88·3	0·70

It was also considered possible that fluctuations in mains voltage might affect the performance of the audiometer and this was accordingly investigated. From the ear-phone sensitivity curve it was apparent that 6 kc/s was the frequency most likely to be sensitive to frequency and output voltage change, and this was therefore used as the test frequency. The supply voltage was adjusted with a variable transformer used in conjunction with a voltmeter, and the acoustic output was recorded in the usual way. The audiometer is fitted with an output control which is intended to be used with an indicating meter which apparently monitors the oscillator side of the attenuator. A set of results (Table 2) are presented which indicate that if the manufacturer's instructions are followed in the use of this control the output of the instrument can be held constant over wide fluctuations of mains supply voltage.

Another possible source of error which was considered was that of drift occurring during the warm-up period of the instrument. Output was measured at 6 kc/s repeatedly over a 30-minute warm-up period and the variation was found to be less than 1 dB (Table 3).

Frequency accuracy was also investigated. The audiometer is a beat frequency oscillator with visual presentation of the beat frequency. The frequency selector is provided with a zero setting and an adjuster. The procedure recommended is to set the selector at zero and adjust the frequency indicator to null point. Using an oscilloscope and mains frequency as a standard source, the selector error and frequency drift were assessed. The selected frequencies were found to be within a few cycles of standard and it was concluded that if the manufacturer's instructions were followed frequency deviation was not likely to be a cause of error.

TABLE 2

OUTPUT STABILITY WITH VARIATION OF MAINS SUPPLY VOLTAGE

Output (dB)	Mains Supply Voltage 250	240	230	220	210	200	190	180	170
Control not used	91·4	91·4	91·4	91·4	91·4	91·2	90·6	90·0	89·0
Control used	91·4	91·4	91·4	91·4	91·4	91·4	91·4	91·4	91·4

TABLE 3
DRIFT MEASUREMENT

	Time (min.)											
	00	02	04	06	08	10	12	14	16	20	24	28
Output (dB)	92	92	92	91·8	91·8	91·9	91·8	91·2	91·2	91·1	91·0	91·0

Frequency = 6 kc/s (frequency adjuster not used).
At 28 min. frequency adjuster was reset to zero c/s and the output became 91·5 dB.

The foregoing evidence supports the conclusion that instrument variation did not produce the results obtained in this experiment. Instrument variation was more important for the purpose of this experiment than the absolute calibration level, although this was checked and, as far as could be ascertained, it conformed to the recommendation of British Standards 2497 (1954).

The audiometric tests were carried out in a mobile test room installed in a specially constructed vehicle chassis. The construction and acoustic properties have been fully described by Lee *et al.* (1963). They concluded that when used in textile mill yards the test environment conformed to the criteria laid down for the threshold audiometry at all frequencies above 125 c/s. For the purpose of the experiment described here the vehicle was placed in a quiet car park.

Four audiometric tests were performed with each individual, the time intervals between the first and second being 24 hours, followed by an interval of one week before the third test which was separated from the fourth by 24 hours.

The threshold measurements were carried out in the way recommended by Hinchcliffe and Littler (1958). An approximate descending threshold was determined. The final threshold was taken to be the mean of the ascending and descending thresholds, which were determined as exactly as possible. Test frequencies were 1, 2, 3, 4, 6, 8, 0·5 and 1 kc/s taken in this order on alternate ears, always commencing 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 intensity level explored. A two out of four response criterion was used for both the descending and ascending threshold. The audiometric test on each individual lasted between 15 and 20 minutes and was carried out at least 24 hours after removal of wax and clinical examination. Before each test the individual was asked about upper respiratory tract symptoms.

The ear-phones headband was not changed throughout the experiment and each individual was allowed to adjust the head set for maximum comfort. Robinson (1960) studied the effect of variations in the force of application of the ear-phones and concluded that the principal effect was on frequencies below 1 kc/s although the effect of an "indifferent fit" was also found in the higher frequencies. In this context indifferent fit is taken to mean an inexact opposition of the ear-phone aperture and the external auditory meatus. In this investigation it was not possible to control closely the position of the ear-phone relative to the external auditory meatus but it is assumed that the position of maximum comfort does not vary very much from time to time in the same individual. In any event routine audiometry does not usually involve any more stringent precautions than were observed in our procedures.

Results

For the present purpose it was considered more satisfactory to assess the disagreement between measurements rather than the agreement. The use of analysis of variance and the standard deviation has therefore been preferred to intraclass correlation coefficients, discussed by High, Glorig, and Nixon (1961).

The major source of variation was the range of thresholds among individuals. The variations in measurement obtained from the same individual at different times are given in Table 4, which shows an analysis of the four audiograms performed on 24 ears. The "within ears" estimates of variance are of the same order as those found by other workers (Jackson, Fassett, Riley, and Sutton 1962; Robinson, 1960; and Brown, 1948).

TABLE 4
ANALYSIS OF VARIATIONS IN REPEATED THRESHOLD DETERMINATIONS: FOUR DETERMINATIONS ON 24 EARS

f (c/s)	Estimates of Variance ((dB)²)		
	Within Ears	Robinson (1960)	Brown (1948) Experiment III
500	8·5	9	9
1,000	7·5	9	17
2,000	7	5	10
3,000	6	—	—
4,000	17	5	14
6,000	19·5	22	—
8,000	23	31	25

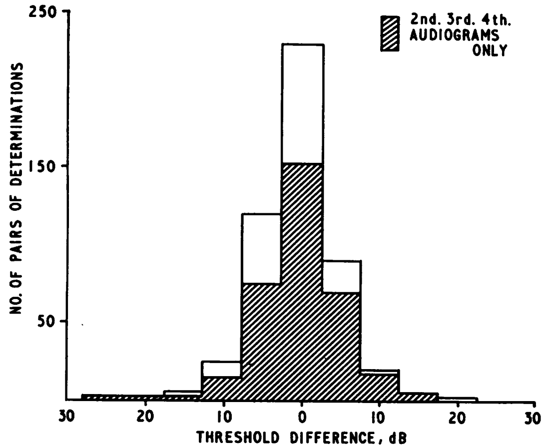

FIG. 1.—Differences in thresholds obtained from consecutive determinations at frequencies from 500 to 8,000 c/s.

The extent of the differences found between pairs of consecutive determinations is illustrated in Fig. 1. These differences ranged from −25 dB to + 20 dB.

The median thresholds of individuals ranged from −10 dB to 25 dB hearing level over the frequency range used. There was no evidence that the amount of variation was related to the threshold level.

The relation between changes of threshold measurements occurring at the same frequency in both ears was slight but was found to be significant at the 1% level, both at 500 c/s and when all frequencies were considered together. There was little evidence of a relation in adjacent frequencies in the same ear (Table 5).

There was some evidence of a facilitation effect but differences between the first and subsequent audiograms were not significant and made only negligible practical impact on the results as a whole. A study of this effect and other souofrces variation has been carried out and is the subject of another paper.

TABLE 5

CORRELATION BETWEEN THRESHOLD SHIFTS IN (1) LEFT AND RIGHT EARS AND (2) ADJACENT FREQUENCIES

f (c/s)	Degrees of Freedom	Correlation Coefficient	Significance
Left and right ears			
500	46	0·39	$P < 0{\cdot}01$
1,000	46	0·29	$P < 0{\cdot}05$
2,000	46	0·19	
3,000	46	0·17	
4,000	46	0·12	
6,000	46	−0·06	
8,000	46	0·24	
Total audiograms	334	0·16	$P < 0{\cdot}01$
4,000 and 6,000 c/s	94	0·07	
6,000 and 8,000 c/s	94	0·14	

Discussion

For a drop in auditory threshold to be considered possibly significant and not a chance result due to the inherent variability of audiometric measurements, the change would have to be at least 12·5 dB at 4 kc/s and 15 dB at 8 kc/s.

The evidence of an actual change in auditory acuity would be strengthened if the observed change occurred at the same frequency in both ears. In the present experiment there was an overall correlation ($r = 0{\cdot}16$) between shifts in left and right ears. Robinson found a slightly greater degree of association ($r = 0{\cdot}35$). An overall correlation of $r = 0{\cdot}2$ has been assumed in the compilation of the significance levels for simultaneous shifts in both ears. These indicate that each ear would have to show a deterioration of at least 10 dB at 4 kc/s for the effect to be judged significant.

If an observed change should occur at more than one frequency the evidence of an actual change would be considerably increased. Levels have been calculated only for the frequencies of most diagnostic significance; changes in the same direction of at least 10 dB at both 4 kc/s and 6 kc/s would have to occur for the effect to be considered significant, a correlation of $r = 0{\cdot}1$ between shifts at adjacent frequencies being assumed. Full details of such criteria are given in Table 6.

TABLE 6

ESTIMATED LEVELS OF SIGNIFICANCE FOR DROPS IN AUDITORY THRESHOLD
(Given to next highest 2·5 dB step)

f (c/s)	One Ear		Two Ears ($P<0{\cdot}01$)	One Ear Two Adjacent Frequencies ($P<0{\cdot}01$)
	$P<0{\cdot}05$	$P<0{\cdot}01$		
500	10	12·5	7·5	
1,000	10	10	7·5	
2,000	7·5	10	7·5	
3,000	7·5	10	7·5	
4,000	12·5	17·5	10	10
6,000	12·5	17·5	12·5	10
8,000	15	17·5	12·5	12·5

It is considered that this table provides a preliminary basis for the evaluation of audiometric results whilst acknowledging that their application in a wider context should be justified by further investigations among other populations. Factors such as otological pathology, noise exposure, and longer inter-audiogram time intervals might well increase the variation. (Small groups of five noise-exposed, and four otologically impure young men were tested in the same way. The results, shown in Table 7, tend to confirm that the variation found in the original group is probably the least that could be expected.)

TABLE 7

DROPS IN AUDITORY THRESHOLDS ON REPEATED DETERMINATIONS IN GROUPS OF FIVE NOISE-EXPOSED AND FOUR OTOLOGICALLY IMPURE YOUNG MEN

f (c/s)	Magnitude of Drop in dB*					
	Under 7·5	7·5	10	12·5	15	17·5
Noise-exposed						
500	29	1				
1,000	29	1				
2,000	30					
3,000	27	2			1	
4,000	30					
6,000	27	2			1	
8,000	27	1	1		1	
Otologically Impure						
500	16					
1,000	14	1	1			
2,000	16					
3,000	15			1		
4,000	16					
6,000	13	1	1	1		
8,000	13	1			1	1

(Stepped lines: 1% Significance levels of previous groups)

*There were no bilateral drops greater than 7·5 dB.

TABLE 8

ESTIMATION OF TRUE AUDITORY THRESHOLD FROM A SINGLE DETERMINATION

f (c/s)	True threshold estimated to lie within the range of the single determination ± the following, with probability values of:		
	0·05	0·01	0·001
	(dB)	(dB)	(dB)
500	6	7	10
1,000	5	7	9
2,000	5	7	9
3,000	5	6	8
4,000	8	10	13
6,000	9	11	15
8,000	9	12	16

The estimation of the true auditory threshold from a single determination is subject to considerable uncertainty as shown in Table 8.

We wish to acknowledge the help and advice we have received from Professor R. E. Lane, Nuffield Department of Occupational Health, University of Manchester, Professor V. F. Lambert, Department of Otolaryngology, University of Manchester, and Mr. D. L. Chadwick, Department of Otolaryngology, University of Manchester. We are also most grateful to the students, without whose co-operation the experiment could not have taken place.

REFERENCES

British Standards Institution (1954). B.S. 2497.
Brown, R. E. C. (1948). *J. Laryng.*, **42**, 487.
High, W. S., Glorig, A., and Nixon, J. (1961). *J. Auditory Res.*, **4**, 247.
Hinchcliffe, R., and Littler, T. S. (1958). *Ann. occup. Hyg.*, **1**, 114.
Jackson, J. E., Fassett, D. W., Riley, E. C., and Sutton, W. L. (1962). *J. acoust. Soc. Amer.*, **34**, 218.
Lee, W. R., John, J. E., and Fowweather, F. (1963). *Brit. J. industr. Med.*, **20**, 57.
Robinson, D. W. (1960). *Ann. occup. Hyg.*, **2**, 107.