British Journal of Industrial Medicine, 1973, **30**, 271-275

Subject variability in short-term audiometric recording

B. P. R. HARTLEY[1], R. W. HOWELL[2], A. SINCLAIR,[3] D. A. D. SLATTERY[4]
British Steel Corporation

Hartley, B. P. R., Howell, R. W., Sinclair, A., and Slattery, D. A. D. (1973). *British Journal of Industrial Medicine*, **30**, 271-275. **Subject variability in short-term audiometric recording.** The reliability of a single audiogram at one examination has not been established under industrial conditions. It has previously been suggested that when audiograms are taken they should be performed at least three times, preferably not at one sitting, and that the mean level at each frequency should be taken as the definitive value of hearing level.

This study seeks to compare the reliability of three audiograms taken at a single session, but with a break between tests, with three audiograms taken at roughly weekly intervals. One hundred and thirty-two apprentices (average age 16 years) without occupational noise exposure were examined with a Peters audiometer, using one operator only at each of the two works involved. At a third works 45 men (average age 36 years), mostly with occupational noise exposure, had three audiograms taken within an hour using a self-recording audiometer. Not only did the mean of three audiograms from a single session show no practical difference when compared with the mean of three readings taken on separate occasions roughly a week apart, but the second audiogram of the first three was found to be generally representative of the mean of these three. In only 4 of 132 subjects did the second audiogram vary by more than 3 dB from the mean of the first three readings. It is suggested that single audiogram examination should be replaced by two audiograms routinely carried out at a single session, and that in the absence of any large difference (say 5 dB) between the two readings the second should be adopted.

In this series, variability between operators (at 3 and 4 kHz) exceeded mean subject variability. There appeared to be no reduction in subject variability when a self-recording machine was used.

Almost all current audiometry in industry in this country is based on a single audiogram at each subject's attendance. The reliability of a single audiogram has not been established under industrial conditions. Indeed, Burns and Robinson (1970) strongly advocate that when audiograms are taken they should be performed at least three times, preferably not at one sitting, and that the mean at each frequency be taken as the definitive value of hearing level. Howell and Hartley (1972) have suggested that the chances of getting a man released on three separate occasions may not be realistic, that repeat audiograms may have to be done at the initial attendance, and that further studies to ascertain the reliability and necessity of serial audiometry should be carried out in industrial settings. This paper presents the findings of one such study.

[1]Stocksbridge Works, Stocksbridge, near Sheffield
[2]Head Office, Grosvenor Place, London SW1
[3]Ebbw Vale Works, Ebbw Vale
[4]Rotherham Works, Rotherham

Methods and results

In this paper all measured hearing levels in excess of the International Organization for Standardization's (ISO) zero (i.e., elevations of threshold) are shown as positive figures, while hearing levels less than the ISO zero (lowering of threshold) are shown as negative figures. The term 'reliability' is used in the conventional sense of sound and consistent quality. Reliability can vary between absolute and zero. In audiometric recording and interpretation this assessment will often be a subjective evaluation which will vary with the individual reader's judgement and beliefs. For example, the value of the single audiometric examination commonly used in industry may well be regarded as reliable by some users, as doubtful by others, and as unreliable by a third group.

'Variability' as applied to audiometric measurements implies that readings may be neither exactly reproducible nor constant. Factors which affect reproducibility or variability include operator error, machine error, patient error, and biological variation as well as more appreciated factors such as pathological or noise-induced damage, presbycusis and wax in the ears.

Apprentices from two works, A (mean age 16·1 years) and B (mean age 16·8 years), were examined. They were free from occupational noise exposure and hence of the difficulties from temporary threshold shift which could arise in such a study. The manual audiometers used were the diagnostic audiometers manufactured by Peters, which offer air and bone conduction measurements and masking facilities. Three pure-tone, air conduction audiograms were carried out on each subject within an hour. A further audiogram was done on two subsequent occasions at roughly weekly intervals. Audiograms were carried out by the one operator at each works; the results of previous audiograms in the series were not available to her until the end of the study.

At a third works (C) 45 men of various ages (mean 36 years), mostly with occupational noise exposure, had three audiograms taken within an hour using a Medical Measurement Incorporated (MMI) self-recording audiometer. The MMI audiometer automatically records discrete frequency and air conduction thresholds at six frequencies for each ear. Each ear is tested for 30 seconds at each frequency and, in addition, the audiometer incorporates several other features including the choice of pulsed or continuous tones and manual or automatic validity checks. Recording is subject controlled. Instructions to the men were given by the one operator throughout the series.

All audiograms were carried out in audiometry booths in quiet rooms in medical departments, and all machines were calibrated to ISO standards.

Results

The comparison of mean audiometric readings by time interval is shown in Table 1.

The means of the readings at 3 and 4 kHz for the two works using a Peters audiometer, and the means of 1 and 2 kHz for all three works are given.

Means are based on individual hearing levels for both ears, aggregated and divided by 4. For example, a left ear reading of 5 dB at 3 kHz and 5 dB at 4 kHz plus a right ear reading of 0 dB at 3 kHz and 10 dB at 4 kHz would aggregate to 20 dB and give a mean of 5 dB for that individual.

The mean maximum variations in the individual's readings for three and five audiograms are shown in Table 2. For example, a subject with a mean score at 3 and 4 kHz of 5 dB at the first reading, 2·5 dB at the second reading, and 10 dB at the third reading would be credited with a maximum variation of 7·5 dB (10 −2·5) over three readings.

The three readings suggested by Burns and Robinson (1970) as not at one sitting, are shown in this series as audiograms 1, 4, and 5 (Table 3). The means of the three audiograms taken within an hour (1, 2, and 3) are compared on the lines of Table 2 with the means of readings 1, 4, and 5. If the two series are found to be acceptably close in practical terms then the advantages of a single session over

TABLE 1

COMPARISON OF MEAN AUDIOMETRIC READINGS BY TIME INTERVAL AND FREQUENCY (kHz) AT THREE WORKS

	Works	*Audiometer*	*No. of men*	*Readings* 1	2 (*within 1 hr*)	3	4 (*after 1 wk*)	5 (*after 1 wk*)
Mean reading at 1 and 2 kHz (both ears)	A	Peters	69	2·8	2·0	2·0	2·1	2·1
	B	Peters	63	2·2	1·1	0·5	1·5	1·2
	C	MMI	45	9·4	8·0	8·2		
Difference between the means of A and B				0·6	0·9	1·5	0·6	0·9
Mean reading at 3 and 4 kHz (both ears)	A	Peters	69	2·5	1·5	0·8	1·7	1·6
	B	Peters	63	−0·2	−1·7	−2·6	−1·1	−1·8
Difference between the means of A and B				2·7	3·2	3·4	2·8	3·4

TABLE 2

Mean of Individual Subjects' Maximum Variation between Audiograms

	Works	*Audiometer*	*1-2 kHz (both ears)*	*3-4 kHz (both ears)*
3 readings within 1 hr	A	Peters	2·0 dB	2·9 dB
	B	Peters	3·5 dB	4·1 dB
	C	MMI	4·5 dB	
5 readings over 3 weeks (approx.)	A	Peters	3·4 dB	4·5 dB
	B	Peters	5·2 dB	5·8 dB

TABLE 3

Comparison of Means of Audiograms 1 + 2 + 3, 1 + 4 + 5, and 2 only

	Works	*Audiograms (dB)*		
		1, 2, 3	*1, 4, 5*	*2*
1-2 kHz (both ears)	A	2·2	2·3	2·0
	B	1·2	1·6	1·1
3-4 kHz (both ears)	A	1·6	1·9	1·5
	B	−1·5	−1·4	−1·7

multiple attendances must be very substantial in industry. Since Table 1 suggested that audiogram 2 was likely to be a fair indication of the mean of numbers 1, 2, and 3, this has also been shown in Table 3 for comparison.

Not only are the means of the three audiogram groupings shown in Table 3 very similar, especially when considered in the context of total audiometric reliability, but the mean maximum variations between the 1, 2, 3 and the 1, 4, 5 sets are also very similar, with slightly smaller average variation in the 1, 2, 3 group at both works at both 1-2 kHz and 3-4 kHz. Thus the 1, 4, 5 cluster of audiograms hardly differs from the three taken within the hour. Not only were the means similar, but the distribution of the individual subjects' variations was very similar (Table 4).

Discussion

The position of audiometry in hearing conservation programmes in Britain is far from established. The hearing conservation programme of one large concern requires that 'Any such applicant (for employment) should be finally accepted only after satisfactory audiometric examination and any offer of employment should be made conditional upon

TABLE 4

Distribution of Numbers of Men by Stated Variation between Individual Audiograms

	Individual subjects' maximum variation between 3 readings (dB)											*Total men*
	0	*1¼*	*2½*	*3¾*	*5*	*6¼*	*7½*	*8¾*	*10*	*11¼*	*12½*	
1-2 kHz (both ears)												
Works A												
Audiograms 1, 2, 3	9	30	15	13		1		1				69
1, 4, 5	5	19	24	14	6			1				69
Works B												
Audiograms 1, 2, 3	2	8	19	17	8	5	4					63
1, 4, 5	2	16	10	11	12	6	2	2	2			63
3-4 kHz (both ears)												
Works A												
Audiograms 1, 2, 3	8	17	18	10	6	8	1		1			69
1, 4, 5	4	18	15	14	5	5	5	1	1		1	69
Works B												
Audiograms 1, 2, 3	5	7	15	11	10	5	4	3	2	1		63
1, 4, 5	5	7	16	8	12	4	6	3	1	0	1	63

satisfactory completion of hearing tests'; and 'Audiograms should be updated at intervals of one year' (Pelmear, 1973). At the same time Atherley (1972) suggests that current audiometry has limitations as a basis for medical standards; that a clear distinction between normal and abnormal is not always possible with this technique; that audiometry performed outside the laboratory, as perhaps in industry, might not always provide reliable results; and that it can be shown numerically that even the most sensitive ears would have to be tested several times over a long period before any change could be accepted with confidence.

Some justification for this last statement derives from the laboratory study of Atherley and Dingwall-Fordyce (1963). Additionally, Howell and Hartley (1972) concluded from their industrial study that with current variability in audiometric recording it seems unlikely that small changes in recorded hearing levels will give confident early indication of deterioration in a susceptible ear. Thus the case for widespread early or frequent serial audiometry is not presently strong. The pre-employment audiogram with aural examination may be useful to detect those whom it would be unwise, for medicolegal reasons at least, to place in noisy areas because of existing hearing levels or defects. Follow-up audiograms may help to detect an abnormally high rate of deterioration, but the cost-benefit of frequent audiograms has yet to be demonstrated in industry in the United Kingdom. The arbitrary period between audiograms, and the question of whether or not a single audiogram at each examination is reliable, may also drastically affect the cost of these exercises. It would be difficult and costly for a man to be released from work on a number of occasions within a short period. One of the additional difficulties would be getting the man at times when he is free of temporary threshold shift, and this difficulty would be enhanced if he had to make several attendances within a short period. If the concept of taking two audiograms at a single session is accepted, the avoidance of temporary threshold shift probably remains the major problem. Industries with a large employee turnover may find frequent routine audiometry expensive.

Table 1 suggests that with the Peters audiometer there is a learning effect when three readings are taken within an hour. Some of the learning effect may well have lapsed by the fourth and fifth audiograms, so that threshold levels tend to revert to something of the order of the second test. This is true of the means both at 1 and 2 kHz and at 3 and 4 kHz. On the other hand, Delaney (1970), in small laboratory studies, found learning effect improvement even after 10 tests. He also found that the first audiometric test on each occasion is likely to give a mean hearing level about 1 dB higher than second or subsequent tests on that day. Table 1 confirms this difference between first and second audiograms at the same session. For technical reasons at the time of the survey, readings at 3 and 4 kHz were not available on the MMI machine, but readings at 1 and 2 kHz suggest perhaps that the learning effect had reached a maximum on average by the second recording. The figures provide some evidence that using the same operator, variation in the individual subject over a relatively short period of time is not as great as might have been feared but is still a factor operating against the early detection of the sensitive ear. The mean individual variation at works A and B was less than the mean difference (at 3 and 4 kHz) between their operators when using unexposed males of similar age and social class from the same geographical background.

Tables 3 and 4 suggest that in the context of total audiometric variation in industry (Howell and Hartley, 1972) at both works and at both frequency groupings there was no great difference between the mean of three readings taken at a single session, the mean of three readings taken at separate attendances over a period of three weeks, and the second of three readings taken at one attendance. If these findings are confirmed elsewhere in industry then they are of great significance as industrial audiometry could be reliably limited to a single session for each examination.

If it is generally accepted that small changes in threshold levels cannot be taken as a significant indication of either a susceptible or damaged ear, then on economic grounds it might be thought (Table 1) that a single audiometric reading might well be considered satisfactory. This would be especially so where there is continuity of operator. However, while a relatively small difference between these two readings is true of a series, Table 4 shows that the variation in individuals over three readings was occasionally as high as 12·5 dB. The maximum variation between first and second readings was 11·25 dB. It follows then that the single audiogram may well be misleading, particularly if it is a pre-employment examination which will be used as a reference level at subsequent examinations. In only four out of 132 subjects did the second audiogram vary by more than 3 dB from the average of the first three readings at 3 and 4 kHz. A reasonable and economic conclusion to be drawn from this study might well be that a subject attending for audiometry should have two tests at the one examination and that in the absence of any large difference (say 5 dB) between the two sets of readings, the second should be adopted. Any sizeable discrepancy would lead to further examination of the individual. Earlier results should not be available to the operator when tests are being carried out.

Continuity of operator and monitoring the operator's standards may well be an important factor in reducing variability (Howell and Hartley, 1972). The present study suggests that in any busy programme for noise-exposed men, the avoidance of temporary threshold shift may be more of a practical problem than the single-attendance examination.

There was no evidence in this study that the MMI self-recording audiometer gave less variability than the manually operated Peters machine, using a single operator at each works.

We are grateful to Dr. K. P. Duncan, Chief Medical Officer, for encouraging this study and to the three anonymous operators who made these conclusions possible.

References

Atherley, G. R. C. (1972). Personal communication.
——, and Dingwall-Fordyce, I. (1963). The reliability of repeated auditory threshold determination. *British Journal of Industrial Medicine*, **20**, 231-235.
Burns, W., and Robinson, D. W. (1970). *Hearing and Noise in Industry*. H.M.S.O., London.
Delaney, M. E. (1970). On the stability of auditory threshold. N.P.L. Aero Report Ac 44.
Howell, R. W., and Hartley, B. P. R. (1972). Variability in audiometric recording. *British Journal of Industrial Medicine*, **29**, 432-435.
Pelmear, P. L. (1973). Hearing conservation. *Journal of the Society of Occupational Medicine*, **23**, 22-26.

Received for publication July 27, 1972.
Accepted for publication December 14, 1972.

Requests for reprints to R. W. Howell.