1994

Intersubject variability of real-ear sound pressure level: conventional and insert earphones

Michael Valente
Washington University School of Medicine in St. Louis

Lisa G. Potts
Washington University School of Medicine in St. Louis

L. Maureen Valente
Saint Louis University

William Vass
Saint Louis University

Joel Goebel
Washington University School of Medicine in St. Louis

Follow this and additional works at: https://digitalcommons.wustl.edu/audio_hapubs

Recommended Citation
https://digitalcommons.wustl.edu/audio_hapubs/18

This Article is brought to you for free and open access by the Division of Adult Audiology at Digital Commons@Becker. It has been accepted for inclusion in Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.
Intersubject Variability of Real-Ear Sound Pressure Level: Conventional and Insert Earphones

Michael Valente*
Lisa G. Potts*
Maureen Valente†
William Vass†
Joel Goebel*

Abstract
Measures of the sound pressure level (SPL) near the eardrum were determined at discrete frequencies between 500 and 4000 Hz on 50 ears using TOH-39P and ER-3A earphones with the attenuator of an audiometer fixed at 90 dB HL. Results revealed significant differences in the measured SPL between the two earphones at all test frequencies. Results also revealed large intersubject differences in the SPL measured near the eardrum for both earphones. The results of this study highlight the large intersubject variability associated with measuring the SPL at the eardrum and point out the difficulty in accurately predicting individual performance from averaged group data.

Key Words: Intersubject variability, loudness discomfort level (LDL), real-ear aided response (REAR), real-ear insertion response (REIR)

Real-ear measures (REM) have become increasingly popular over the past several years. Up to this point the primary use of this technology has been to determine if the measured real-ear insertion gain (REIG) “matched” a prescribed REIG. Recently, increased attention has been placed upon using REM to directly measure the sound pressure level (SPL) near the eardrum corresponding to the individual dynamic range between threshold and suprathreshold levels. This dynamic range, measured in dB SPL near the eardrum, could then serve as a “target” to determine if the real-ear aided response (REAR) for frequency-specific or composite speech signals was placed within the individual dynamic range using either single or multiple input levels (Hawkins, 1987; Seewald et al, 1987; Kawell et al, 1988; Feigin et al, 1989; Hawkins et al, 1989; Cox and Alexander, 1990; Hawkins et al, 1990; Gagné et al, 1991a, b; MacPherson et al, 1991; Stelmachowicz and Seewald, 1991; Stuart et al, 1991; Zelisko et al, 1992a, b; Valente et al, 1993a; Skinner et al, 1993, 1994).

Instead of direct measures of the SPL near the eardrum, several studies have suggested that the real-ear SPL near the eardrum can be predicted from audiometric thresholds measured in hearing level (dB HL) using either conventional or insert earphones (Etymotic ER-3A or E-A-R Tone® tubephones) by applying a set of average transformation values (Leijon et al, 1983; Walker et al, 1984; Libby, 1985; Cox, 1986, 1988; Hawkins et al, 1987; Kawell et al, 1988; Skinner, 1988; Bentler and Pavlovic, 1989; Hawkins et al, 1990; Gagné et al, 1991a, b; Seewald et al, 1991; Stuart et al, 1991; Seewald, 1992; Zelisko et al, 1992a). To illustrate this point, a probe microphone system (Audioscan, 1992) was recently introduced that contains software (i.e., Speechmap™) using threshold values (in dB HL) to calculate and display the predicted loudness discomfort level (LDL) in dB HL. In addition, the user can also obtain the
predicted SPL measured near the eardrum for both threshold and LDL.

One shortcoming of utilizing average transformation values to predict the individual real-ear SPL is the possibility of large intersubject variability of the SPL measured near the eardrum for either TDH or insert earphones. This may make it very difficult to predict accurately the individual thresholds (measured in either HL or SPL) from algorithms based upon average group data (Kamm et al., 1978; Dillon et al., 1984; Cox, 1985; Hawkins et al., 1987; Kavell et al., 1988; Ross and Seewald, 1988; MacPherson et al., 1991).

The present study measured the SPL near the eardrum for six discrete frequencies between 500 and 4000 Hz using conventional (TDH-39P) and insert earphones (ER-3A) with the attenuator of the audiometer fixed at 90 dB HL. As a result of these measurements, the magnitude of the intersubject variability was determined.

**METHOD**

**Subjects**

The experimental group included 50 ears from 25 adult subjects. The test ear of each subject demonstrated normal middle ear function (i.e., middle ear pressure within ± 50 daPa; static compliance between 0.6 and 1.8 mL) using a Y226 probe tone from a calibrated GS 1733 middle ear analyzer. Hearing thresholds were not an important factor for the purposes of this study and, therefore, are not reported.

**Procedures**

For each subject, measurements were obtained with TDH-39P (MX41/AR cushion) and ER-3A (50 ohm) earphones connected to a calibrated Maico MA 39 portable audiometer (ANSI, 1989) with the attenuator fixed at 90 dB HL. The SPL near the eardrum was measured for each earphone condition with continuous pure tones of 500, 1000, 1500, 2000, 3000, and 4000 Hz using a probe tube coupled to a probe microphone from a Frye 6500 real-ear analyzer. Measures were obtained only once because two previous studies reported excellent test–retest reliability for the equipment and procedures used in this study. One study reported mean intrasubject test–retest differences of less than 1 dB for the real-ear unaided response (Valente et al., 1990), while the second study reported the same results for the real-ear insertion response (Valente et al., 1991). In addition, numerous studies have reported on the test–retest reliability of the ER-3A. For example, Clark and Roeser (1988) reported that mean intrasubject test–retest reliability was less than 2 dB for the ER-3A and that the reliability of the ER-3A was equivalent to the TDH-50P. Larson et al. (1988) revealed that the standard error of estimate was 0.9 to 1.5 dB for the ER-3A and TDH-50P. Wilber et al. (1988) reported on the results of five studies and indicated that the standard deviation for threshold measures was equivalent for the TDH-39 and ER-3A earphones. Borton et al. (1989) reported that mean test–retest differences were less than 5 dB for the ER-3A. Finally, Lindgren (1990) and Frank and Vavrek (1992) reported that intratester test–retest reliability was within 3 dB at 500 and 4000 Hz for the ER-3A.

The probe tube was marked 30 mm from the tip, and this mark was placed on the intratragal notch. In the average adult ear, this would place the tip of the probe tube approximately 4 mm from the eardrum, which is necessary for accurate measures of SPL (Zemplenyi et al., 1985; Gilman and Dirks, 1986; Dirks and Kincaid, 1987). The probe tube was then taped into place to prevent movement. Great care was used to assure that the 30-mm mark remained in the same position as the diaphragm of the TDH-39P was placed over the orifice of the ear canal or when the immittance probe cuff from the ER-3A was placed into the ear canal.

The ER-3A was coupled to the ear canal using an appropriately sized Grason Stadler immittance probe cuff. For this study, an immittance cuff was placed on a plastic adapter (ER3-06) connected to the sound outlet tube and coupled to the ER-3A and then to the ear canal. Immittance cuffs were used for several reasons. First, the diameter of the ear canal of several subjects was either too large or small to successfully use the standard foam plug. In a recent article, Frank and Vavrek (1992) reported that 17 percent of their subjects had ear canals that would not allow the standard foam plug to be used successfully. On the other hand, the immittance cuffs used in this study have outside diameters of 2 to 22 mm. Insertion of the cuff so that the outside edge was flush with the bowl of the concha ensured a consistent insertion depth of 16 mm past the opening of the ear canal for all subjects. This depth is precisely the 15- to 16-mm inser-
tion depth recommended by the manufacturer for a "deep" insertion. Finally, Borton et al (1989) reported no significant differences in threshold when ER-3A earphones were connected to either foam plugs or immittance cuffs. To measure the SPL near the eardrum, the reference microphone was "dis-enabled," and the measured SPL was read directly from the video monitor when activating the "Calibrate Probe" software of the Frye 6500. Finally, the probe microphone was calibrated daily using the procedures suggested by the manufacturer and all treatment levels of earphone (TDH-39P and ER-3A) and frequency (500, 1000, 1500, 2000, 3000, and 4000 Hz) were counterbalanced.

RESULTS AND DISCUSSION

Intersubject Variability

The range of intersubject variability in the SPL measured near the eardrum for the TDH-39P ranged from 9 dB at 1000 Hz to 36 dB at 4000 Hz (row 3; Table 1 and Fig. 1). In comparison, the range of intersubject variability for the ER-3A ranged from 12 dB at 1000 Hz to 29 dB at 2000 Hz (row 6; Table 1 and Fig. 2). Even if the highest and lowest data points from Figures 1 and 2 were removed, the intersubject variability would still remain rather large. By removing these extremes, the range of the intersubject variability for the TDH-39P was reduced to 20, 8, 14, 12, 21, and 23 dB at 500 to 4000 Hz, respectively. For the ER-3A, the range was reduced to 18, 8, 16, 17, 15, and 17 dB at 500 to 4000 Hz, respectively.

Several possibilities may account for the large intersubject variability of the SPL measured near the eardrum for the two earphones. First, the subjects included in this study had eardrum compliance that was within the normal range of 0.6 to 1.8 mL. Preves and Orton (1978) reported that small differences in eardrum compliance, even for those that are within the normal range (0.31 to 1.20 cc), can result in as much as a 6.5-dB difference between insertion and functional gain. These authors did not report if this variable was frequency dependent. Dirks and Kincaid (1987) report on the SPL measured for 3000 Hz at the eardrum and at

<table>
<thead>
<tr>
<th>Table 1 Mean, Standard Deviation (SD), and Range of Measured Real-Ear SPL for the TDH-39P and ER-3A Earphones at Six Test Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency (Hz)</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>TDH-39P</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>ER-3A</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Mean Difference 10.4 6.1 2.4 3.6 8.5 6.5

N = 50 ears.
*Also provided is the mean difference in the measured SPL between earphones.

Figure 1: Mean Interaural SPL Difference: Frontal Bone (FB) vs. Right Mastoid (RM) Placement

Figure 1 Individual SPL measured near the eardrum at 500 to 4000 Hz for the TDH-39P earphone (N = 50). The "O" represents the right ear and "X" represents the left ear. Also included are ± 1 and 2 standard deviations (SD).
varying probe positions in the ear canal for eardrums with average, low-normal, and high-normal impedance. They report that eardrums having high-normal impedance will result in SPLs that are lower than measured at the eardrum with average impedance. For eardrums with low-normal impedance, they report that the measured SPL will be higher than measured in an eardrum with average impedance. For both conditions, the difference increases as the distance from the probe to the eardrum increases.

Second, the procedure used in this study for probe placement assured that the distance from the orifice of the ear canal to the tip of the probe tube was equal across subjects. However, the distance from the end of the immittance cuff of the ER-3A and diaphragm of the TDH-39P to the eardrum probably varied quite widely across the 50 ears due to intersubject differences in actual canal length. Gilman and Dirks (1986) and Chan and Geisler (1990) report that the measured SPL at probe positions as far as 12 mm from the eardrum may be as much as 4 dB less at higher frequencies relative to probe positions closer to the eardrum. The likely presence of intersubject variability of canal length led Bruell et al (1976) to call for developing transfer functions based upon individual equivalent volumes of the residual ear canal and eardrum in order to more accurately predict the SPL at the eardrum. As noted by Bentler (1989), "wide intersubject variability of resonance amplitude may be related, in part, to the small, although significant, differences in probe-to-eardrum distance differences among subjects" (p. 286).

A third likely cause of the resulting intersubject variability may be related to slit leak, which affects the reliability of measures at 500 Hz and below (Bruell et al 1976; Borton et al, 1989).

The results of this study seem to question the validity and clinical accuracy as advocated by some to predict the individual real-ear SPL for threshold and suprathreshold measures from averaged group data. This is clearly illustrated in Figures 3 and 4.

As mentioned earlier, a probe microphone system (Audioscan) recently introduced a new software package called Speechmap®. This software calculates the predicted LDL (dB HL) (left side of Figs. 3 and 4) as well as the predicted threshold and LDL in dB SPL measured near the eardrum (right side of Figs. 3 and 4) from audiometric threshold entered in dB HL.

The dashed upper line in the left side of Figure 3 is the threshold (dB HL) measured for one subject using the TDH-39P earphone. The lower solid line is the predicted LDL (dB HL). The lower dashed line represents the measured LDL (dB HL). As can be seen, the agreement between measured and predicted LDL is quite good. On the right side of Figure 3 is the predicted threshold (lower solid curve) and LDL (upper solid curve) measured in dB SPL near the eardrum. The lower dashed line represents the measured SPL for threshold, while the upper dashed line represents the measured SPL for LDL. Again, the agreement between measured and predicted SPL for threshold and LDL is quite remarkable.

Figure 4 reports the same measures for a second subject. The upper dashed line on the
left side of Figure 4 is the measured threshold (dB HL). The lower solid line is the *predicted* LDL (dB HL). The dashed line represents the *measured* LDL (dB HL). As can be seen, the measured LDL is considerably below the predicted LDL. If the predicted LDL were used to determine and calculate the appropriate

SSPL90, it is possible that the output would exceed the measured LDL. On the right side of Figure 4 is the *predicted* threshold (lower solid curve) and LDL (upper solid curve) in dB SPL measured near the eardrum. The lower dashed line represents the *measured* SPL for threshold, while the upper dashed line represents the

---

**Figure 3**  
A, LDL predicted from threshold using Speechmap™ software for Audioscan and the measured LDL for Subject 1 in dB HL. B, Threshold and LDL predicted from threshold (dB HL) using the Speechmap™ software for Audioscan and measured threshold and LDL for Subject 1 in dB SPL.

**Figure 4**  
A, LDL predicted from threshold using Speechmap™ software for Audioscan and the measured LDL for Subject 2 in dB HL. B, Threshold and LDL predicted from threshold (dB HL) using the Speechmap™ software for Audioscan and measured threshold and LDL for Subject 2 in dB SPL.
measured SPL for LDL. In this case, the measured SPL for threshold is in fairly good agreement with the predicted SPL above 1000 Hz. However, the measured SPL for LDL is significantly below the predicted SPL for LDL. Again, if the predicted LDL was used to verify that the REAR was below LDL, then it is quite possible that the output would exceed the measured LDL.

Although the above example is used to illustrate the potential error with one commercially available system, the same problem is likely to arise with the other commercially available software packages such as the Desired Sensation Level (Seewald et al., 1991) or by applying a set of average transformation values (Leijon et al., 1983; Walker et al., 1984; Libby, 1985; Cox, 1986, 1988; Hawkins et al., 1987; Kawell et al., 1988; Skinner, 1988; Bentler and Pavlovic, 1989; Hawkins et al., 1990; Gagné et al., 1991a, b; Stuart et al., 1991; Seewald, 1992; Zelisko et al., 1992b).

The only situation for which prediction of individual performance from average group data would be appropriate is in the case of evaluating children or the difficult-to-test population, where measurement of suprathreshold levels is not always possible or may be too time consuming.

**Mean Differences Between Earphones**

Table 1 reports the mean, standard deviation, and range of the real-ear SPL measured for the two earphones. The differences in the measured real-ear SPL between the two earphones ranged from 2.4 dB at 1500 Hz to 10.4 dB at 500 Hz. A two-factor repeated measures ANOVA (earphone by frequency) revealed a significant earphone by frequency interaction (F = 14.6; df = 5,490; p < .01), indicating that the mean differences in measured SPL between the earphone conditions were not constant across test frequencies. A one-factor repeated measure ANOVA was performed at each test frequency to determine if the mean differences in measured real-ear SPL between each earphone was significantly different. Results revealed that the mean real-ear SPL produced by the TDH-39P was significantly (p < .01) greater than the real-ear SPL produced by the ER-3A at each test frequency, with the exception of 1500 Hz (p < .05). No study could be found that reported on differences in the SPL measured near the eardrum between TDH and ER-3A earphones. However, Frank and Vavrek (1992) reported that the mean SPL corresponding to threshold for a TDH-49 earphone, as measured in an NBS-9A coupler, was 5.9, 4.0, 6.6, 5.1, and 4.1 dB greater at 500 to 4000 Hz (1500 Hz was not reported) than the mean SPL measured in an HA-2 coupler for an ER-3A earphone.

The significant differences in measured SPL between the two earphones points out the effect earphone type (insert versus supra-aural) may have upon the SPL measured near the eardrum. This finding also points out the potential problem associated with plugging an insert earphone into the output of an audiometer for which it may not have originally been calibrated. This common practice can create significant differences between the input impedance of the ER-3A and the specified load impedance at the output of the audiometer (Lilly and Purdy, 1993).

For example, readers should be aware that the ER-3A is available in several versions varying in impedance (10, 50, and 300 ohms). For the audiometer used in this study, either the 10- or 50-ohm version could be used. While the 10-ohm earphone is preferred for this audiometer so that thresholds (dB HL) between the ER-3A and TDH-39P can be directly compared, the 50-ohm earphone can still be used with this audiometer. However, the measured output for the 50-ohm earphone would be approximately 6 dB greater than that measured for the 10-ohm earphone. At this facility, the measured SPL in an HA-1 coupler for the 50-ohm earphone was 5.9, 6.1, 5.9, 6.1, 6.4, and 5.8 dB greater than the 10-ohm earphone for the test frequencies used in this study. In this case, a correction table would need to be generated so that 5 dB would be added to the measured thresholds (dB HL) relative to the 10-ohm ("standard") earphone. In addition, the 6-dB difference was revealed, on average, for five subjects when measurements of the SPL near the eardrum were completed for the 50-ohm earphone in comparison to the 10-ohm earphone. It is important to note that clinicians can use the 50-ohm earphone for audiometers designed to use the 10-ohm earphone (after making the necessary coupler calibrations and corrections), but clinicians should not use the 10-ohm earphone on audiometers designed to use the 50-ohm earphone. The latter error can lead to excessive distortion and damage to the audiometer (Etymotic Research, personal communication).

Because the primary focus of this study was intersubject variability, using a 10- or 50-ohm earphone was not thought to be critical because
the range of the intersubject variability revealed in Table 1 for the ER-3A should be the same regardless of the impedance of the earphone. To illustrate this point, the ranges of the intersubject variability, reported in Table 1, were similar for the ER-3A (50 ohm) and TDH-39P (10 ohm). However, if the 10-ohm ER-3A were used instead, the mean SPL measured near the eardrum would have been 6 dB less, and the mean differences between earphones appearing at the bottom of Table 1 would have been 6 dB greater, resulting in even larger mean differences between the TDH-39P and ER-3A earphones.

Table 2 further illustrates the problems associated with arbitrarily plugging an insert earphone into the output of an audiometer for which it was not calibrated. In our clinic, three portable audiometers are available for daily use (audiometers B–D). The same ER-3A earphone used in this study was coupled to an HA-12-cc coupler and the output measured with the ER-3A plugged into the earphone output of each of the three audiometers with the attenuator set at 70 dB HL. For comparison, coupler measures are provided for the ER-3A when it was plugged into the audiometer for which it was calibrated (audiometer A). The first row provides the interim standard for measuring the ER-3A in a HA-1-2-cc coupler (ANSI, 1989) with the attenuator fixed at 70 dB HL. As can be seen from Table 2, the coupler measures were within 3 dB at all test frequencies for audiometer A. However, coupler measures varied quite widely among the other three audiometers relative to the ANSI (1989) standard and among themselves. In fact, the differences among the four audiometers ranged from 0.8 dB at 3000 Hz to 12.6 dB at 4000 Hz. Thus, for the same ER-3A earphone it would have been possible to arrive at different SPLs measured in the ear canal when coupled to different audiometers with the attenuator fixed at some predetermined level.

There is one final point on the issue of measured differences between earphones. From a clinical standpoint, the differences in the SPL measured near the eardrum between earphones coupled to the same or different audiometers may not be as critical when the attenuator is varied to measure individual threshold and suprathreshold levels. This is because the SPL necessary to elicit a response corresponding to a loudness perception of “threshold” or “loudness discomfort” would be the same regardless of the transducer/audimeter combination. However, the audiometer dial reading necessary to obtain these threshold levels may be quite different, depending upon the earphone/audimeter variables discussed earlier. A search for the answer to this issue is the subject of another study (Valente et al, 1993b).

**CONCLUSION**

The results of this study suggest:

1. The presence of large intersubject differences in the SPL measured near the eardrum questions the validity of predicting individual performance based upon averaged group data. Intersubject differences were independent of the type of earphone used to make the measure.

2. The measure of the SPL near the eardrum is related to the type of earphone used to make the measure. In this study, the measured SPL was significantly higher for the TDH-39P at 500 to 1000 Hz and 2000 to 4000 Hz. This study also pointed out some of the problems associated with arbitrarily plugging in an insert earphone to an audiometer for which it was not originally calibrated.

---

**Table 2 Calibration of an ER-3A Earphone in an HA-1 Coupler at 500 to 4000 Hz for Right and Left Earphones for Four Audiometers**

<table>
<thead>
<tr>
<th>Audiometer/Earphone</th>
<th>Frequency (Hz)</th>
<th>500</th>
<th>1000</th>
<th>1500</th>
<th>2000</th>
<th>3000</th>
<th>4000</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANSI (1989)</td>
<td></td>
<td>78.5</td>
<td>73.5</td>
<td>-</td>
<td>76.5</td>
<td>75.5</td>
<td>71.5</td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>79.5</td>
<td>74.0</td>
<td>74.1</td>
<td>77.3</td>
<td>75.3</td>
<td>70.4</td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>79.1</td>
<td>74.3</td>
<td>74.5</td>
<td>77.5</td>
<td>73.2</td>
<td>72.6</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>81.8</td>
<td>77.9</td>
<td>77.9</td>
<td>80.8</td>
<td>74.9</td>
<td>72.3</td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>81.4</td>
<td>78.1</td>
<td>78.3</td>
<td>81.4</td>
<td>72.9</td>
<td>69.9</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>82.2</td>
<td>78.4</td>
<td>78.6</td>
<td>82.0</td>
<td>75.3</td>
<td>70.4</td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>82.3</td>
<td>78.6</td>
<td>78.1</td>
<td>81.8</td>
<td>74.1</td>
<td>69.4</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right</td>
<td>74.3</td>
<td>74.2</td>
<td>74.0</td>
<td>76.0</td>
<td>74.5</td>
<td>64.0</td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>74.0</td>
<td>74.5</td>
<td>74.3</td>
<td>76.8</td>
<td>72.3</td>
<td>60.1</td>
<td></td>
</tr>
</tbody>
</table>

Maximum Difference

| Right   | 7.9 | 4.4 | 4.6 | 6.0 | 0.8 | 8.3 |
| Left    | 8.3 | 4.3 | 4.0 | 5.0 | 2.2 | 12.6 |

*The attenuator is fixed at 70 dB HL. Also reported are the maximum and minimum differences in the measured SPL and the interim reference threshold levels for insert earphones measured in an HA-1 coupler (ANSI, 1989).
REFERENCES


