Comparing two methods to measure preferred listening levels of personal listening devices

Darrin A. Worthington[a)]
Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois 60208

Jonathan H. Siegel and Laura Ann Wilber
Roxelyn and Richard Pepper Department of Communication Sciences and Disorders and The Hugh Knowles Center, Northwestern University, Evanston, Illinois 60208

Benjamin M. Faber
Faber Acoustical, LLC., Santaquin, Utah 84655

Kathleen T. Dunckley
Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois 60208

Dean C. Garstecki and Sumitrajit Dhar[b)]
Roxelyn and Richard Pepper Department of Communication Sciences and Disorders and The Hugh Knowles Center, Northwestern University, Evanston, Illinois 60208

(Received 9 September 2008; revised 7 April 2009; accepted 8 April 2009)

The potential risk to hearing that mass-storage personal listening devices (PLDs) pose remains unclear. Previous research in this area has either focused on maximum outputs of these devices or on ear-canal measurements of listening levels that could not be compared to standards of occupational noise exposure. The purpose of this study was to compare two standard measurement protocols [ISO 11904-1 (2002), Switzerland; ISO 11904-2 (2004), Switzerland] for the measurement of preferred listening levels of PLD. Noise measurements, behavioral thresholds, and oral interviews were obtained from 30 (18–30 years) PLD users. Preferred listening levels for self-selected music were determined in quiet and background noise using a probe microphone, as well as in the DB-100 ear simulator mounted in KEMAR. The ear-canal measurements were compensated for diffuse-field. Only one of the subjects was found to be listening at hazardous levels once their reported daily usage was accounted for using industrial workplace standards. The variance across subjects was the smallest in the ear-canal measurements that were compensated for diffuse-field equivalence [ISO 11904-1 (2002), Switzerland]. Seven subjects were found to be listening at levels above 85 dBA based on measurements obtained in the KEMAR and then compensated for diffuse-field equivalence.
 [DOI: 10.1121/1.3125798]

PACS number(s): 43.50.Hg, 43.50.Qp, 43.38.Lc, 43.50.Yw [BLM] Pages: 3733–3741

I. INTRODUCTION

The potential risk to hearing from mass-storage personal listening devices (PLDs) has been the subject of recent interest in the popular as well as in the scientific press (Nature Neuroscience, 2007; Kenna, 2008). However, the data on which this concern is based are not straightforward to interpret as listening levels are often measured in the ear canal and then have to be compared with standards created for the risk of occupational sound exposures measured in the free field. Here we report preferred listening levels from PLD of 30 college-age individuals and compare measurements made using two methods based on two ISO standards.

The exact number of at-risk PLD users notwithstanding, the scale of the issue with over 100×10^6 PLD being sold annually world wide is undeniable. The perceived importance of this potential problem has led the Council of Science and Public Health of the American Medical Association to issue a specific report and resolution on PLDs and Noise-Induced Hearing Loss (CSAPH, 2008). Arguments in support of this predicted risk include the potential for users to listen more frequently and for longer durations due to the longer battery life, mass-storage capabilities, and perhaps most importantly, the high maximum output levels (MOLs) achievable in many of these devices (Reuters Press, 2005). Many popular PLDs are capable of generating diffuse-field equivalent MOL of 91–121 dBA (Fligor and Cox, 2004) or estimated listener sound levels between 79 and 125 dBA (Keith *et al.*, 2008). Recent measures of listening levels, sampled randomly from individuals passing by on a public street using MP3 players, reveal a wide range of listening

[a)]Present address: North Chicago VA Medical Center, North Chicago, IL.
[b)]Author to whom correspondence should be addressed. Electronic mail: s-dhar@northwestern.edu

levels from 73.7 to 110.2 dBA (mean=86.1; SD=7.9) (Williams, 2005). More recent work by Hodgetts *et al.* (2007) has documented preferred listening levels in quiet and in background noise measured near the tympanic membrane for different types of headphones to be in the same approximate range.

The threat of these high outputs becomes obvious when compared against the *damage risk criteria* of 85 and 90 dBA advocated by the National Institute of Occupational Safety and Health (NIOSH) and the Occupational Safety and Health Administration (OSHA), respectively, for the length of a typical 8 h working day. The OSHA specifies 90 dBA time-weighted average (TWA) as the "permissible exposure level" but 85 dBA TWA as the "action level." All individuals exposed to levels greater than 90 dBA TWA during a workday must use hearing protection devices. On the other hand, hearing protection devices are required to be made available to individuals exposed to levels between 85 and 90 dBA TWA. These individuals must also receive annual audiograms and counseling as necessary. These individuals must also begin to use hearing protection devices if significant threshold shifts are detected after adjusting for age-related changes in hearing. However, the appropriateness of applying *occupational* risk criteria with *recreational* noise exposure remains problematic for three reasons. First, these standards were developed specifically for spectrally dense industrial noise with limited dynamic range and their application to music can be questionable. Second, the standards were developed based on free-field measurements whereas measurements of listening levels are obtained using a probe microphone placed near the tympanic membrane. Finally, the standards were developed as an exposure dose for an 8 h workday. Two international standards describe methods of measuring sound levels generated by sources close to the ear by either using a probe microphone [microphone in real ear (MIRE); ISO, 2002] or a manikin (ISO, 2004). Our goal is to compare measurements made using these two recommended procedures and then derive an index of personal music-related exposure based on reported daily usage.

Early research in this area focused on the maximum outputs from various types of headphones (Wood and Lipscomb, 1972; Katz *et al.*, 1982). Findings from these studies reported maximum output levels exceeding 120 dBA but were flawed because of the following: (i) diffuse-field equivalents were not calculated, an important conversion needed to compare findings to occupational risk criteria (Fligor and Cox, 2004); and (ii) the relationship between MOLs and preferred listening levels in quiet and in background noise remained uncertain. Multiple studies were conducted to examine if users of personal cassette players (PCPs) were operating these devices at harmful levels. Objective measurements of listening levels, incorporating the appropriate conversion factor and user accounts of listening habits, have shown the use-practices of the vast majority of users to be *safe*. Rice *et al.* (1987a) found only 5% of the users of PCPs to be at risk for noise-induced hearing loss. These results were supported by Wong *et al.* (1990) and again by Turumen-Rise *et al.* (1991a).

To make the OSHA and NIOSH standards more applicable, some researchers have carefully selected music with characteristics approximating that of industrial noise. For example, Hodgetts *et al.* (2007) made their measurements using the same song with all subjects. The song was chosen for its limited dynamic range thereby approximating (relatively) steady state industrial noise. Similarly, Farina (2007) used the signal specified by the IEC 60268-1 standard as well as a signal resembling the average spectrum of 30 GB of music popular with Italian teenagers to make measurements of listening levels. In the choice of music lies a paradox; choosing a piece of music for the listener may help in relating the results to established damage risk criteria but may not prompt the listener to set the volumes to habitual levels. Alternately, allowing the listener to choose music used during the measurements should lead to a more realistic volume setting. However, comparisons with industrial standards have to be made with caution.

Finally, individuals listen to music for different durations (hours/day) and with different frequencies (e.g., days/week). Standard computations for industrial noise exposure, however, are based on an 8 h exposure period. Appropriate modifications can be made to such calculations for exposures of other durations. One approach to accommodate this aspect has been to provide guidelines of maximum exposure duration based on listening level (Fligor and Cox, 2004). An alternate would be to combine measured listening levels and reported listening duration to then compute the best-estimated exposure dose for a given individual.

With annual worldwide sales already greater than 100 $\times 10^6$ units and projected growth rates of as much as 40% through 2010 (In-Stat, 2007), a sizable proportion of the global population will soon be using these devices. Accurate measurement techniques that provide realistic estimates of probable hearing damage need to be established. The goal of this research was to establish a benchmark protocol to measure preferred listening levels of mass-storage PLD users along with a set of preliminary data to demonstrate its validity. Our approach was to make empirical measurements of *preferred* listening levels using multiple techniques from a group of college students and use their reported daily listening durations to compute individual exposures comparable to OSHA and/or NIOSH standards.

II. METHODS

A. Subjects

Thirty subjects, 12 male and 18 female, between 18 and 30 years of age (mean=22, SD=3.44) were recruited using fliers posted around the Chicago and Evanston campuses of Northwestern University. Advertisement fliers stated that subjects would be answering questions about the typical use of their PLD and that measurements of preferred listening level would be taken. Subjects were required to have normal hearing sensitivity (<20 dB HL re: ANSI, 2004) between 250 and 8000 Hz, normal middle ear function as measured through tympanometry, and negative history for otologic problems ascertained through an interview. Subjects who failed the hearing screening were excluded to avoid any potential confound of listening levels being affected by the presence of a hearing loss. The Institutional Review Board for Protection of Human Subjects at Northwestern University approved this study. All subjects signed informed consent agreements and were compensated ($10) for their time.

23 out of 30 subjects (77%) reported a history of infrequent noise exposure (i.e., music concerts and clubs). The last reported exposure was no sooner than 14 h before the time of testing. The remaining seven subjects did not report any exposure to loud noises. Measurements were made from alternating left and right ears. Every subject screened qualified to participate in the study.

B. General methods

Measurements of hearing function as well as listening level were conducted in a double-walled IAC booth at the Northwestern University Hearing Clinic on the Evanston Campus. Admittance data were recorded using a Maico MI 26 Tympanometer/Audiometer combination unit (Maico Diagnostics, Inc.). Behavioral hearing thresholds were estimated using custom tracking software following Levitt (1971) using 2 dB steps.

Subjects were orally interviewed regarding mass-storage PLD use as well as previous PLD use (e.g., compact disc player, cassette player, etc.). All answers were recorded on a standard data sheet (Appendix A). During the initial interview no questions of hearing loss from PLD use were asked to try to reduce subject bias. All subjects were read identical instructions for each portion of the study including directions for setting the volume level for their PLD (Appendix B). Preferred sound pressure levels were measured in subjects' ear canals (MIRE technique; ISO, 2002) as well as in a Knowles Electronics DB100 ear simulator mounted on a KEMAR (ISO, 2004).

Subjects were asked to choose a preferred musical selection and adjust their ideal listening level in a quiet environment on their own device and headphones. Musical selection and genre were recorded on the data sheet. A 3 min sample was recorded in the subject's ear canal using a Knowles Electronics ER-7C probe microphone with the probe inserted within 2 mm of the tympanic membrane, confirmed using otoscopic inspection. A second 3 min recording of the same sample was made in the DB100 on a KEMAR. These recordings (in quiet) were then repeated in the presence of background noise (described below). Subjects were asked to adjust their preferred listening level to the previously identified musical selection in the presence of a digitally recorded environmental noise stimulus delivered via sound field speakers, with the subjects seated at 45° azimuth to the speakers. The level of the noise delivered to the subjects fluctuated from 78 to 81 dB sound pressure level (SPL) (81 dB SPL L_{eq}, SD=2.12) during the 3 min recording. These recordings were obtained in the ear canal and verified in KEMAR in the noise condition.

In-ear measurements were made using an Etymotic Research ER-7C probe-tube microphone with 0 dB gain at the pre-amplifier. Measurements in KEMAR were made by coupling the subjects' headphones to standard adult size artificial external pinnae. The DB-100 ear simulators of KEMAR were fitted with 0.5 in. occluding microphones (Etymotic Research ER-11, 0 dB gain). The diffuse-field inverse (DFI) filter on the ER11 pre-amplifier, enabled during these measurements, is an inverse filter of KEMAR's ear canal to diffuse-field response. Applying this inverse filter results in the diffuse-field equivalent of the recorded signal (Killion, 1979). The microphones (ER-7C or ER-11) were connected through their respective preamplifiers to an M-Audio Fast Track Pro, four-channel, 24 bit, firewire input/output device with a sampling rate of 48 000 Hz. Customized software (Electroacoustics Toolbox, Faber Acoustical Services and Technologies) controlled the digital recording of the signal and calculated noise dose, equivalent sound level (L_{Aeq}), measurements for the 3 min recording, with an exchange rate of 3 dB and using the *A*-weighting scale.

All recordings through the ER-7C and the ER-11 were stored and recordings from 23 of the 30 subjects could be processed offline to establish diffuse-field equivalence. A spectral average over the entire duration of each recorded sound file was measured using SysRes (Neely and Stevenson, 2002). With a sample rate of 44.1 kHz and an averaging buffer size of 2048 points, 3800 spectral averages were required to sample each 3 min recording. The power in each one-third octave band was weighted by the appropriate *A*-weighting factor and the total *A*-weighted sum of spectral energy was calculated. The DFI filter transfer function of the ER11 preamplifier was measured using SysRes and chirp excitation delivered through an ear-bud style earphone (SR =44.1 kHz, 100 averages into a 2048 point buffer), first with the DFI filter switched in, then with it switched out. This transfer function was used either to add or remove DFI filtering from the spectral averages of the ER-7C recordings prior to calculating *A*-weighted levels. This yielded measurements comparable to exposure standards developed for diffuse-field conditions. The recordings from the remaining seven subjects had been lost during data transfer between computers and hence could not be processed for this analysis.

The environmental noise was recorded using a Roland EDIROL R-1 portable digital recorder (24 bit, linear WAV file). The acoustic sensitivity of the recorder was calibrated using a 1000 Hz pure tone delivered via sound field speakers. The sound pressure level near the recorder's microphones was measured using a Quest Technologies 1800 Integrating sound level meter set to flat frequency weighting. Recordings and sound level measurements were made during rush hour on the Blue Line of the Chicago Transit Authorities (CTA) train system while going through the underground section, seated in the center of the train car, a common environment where PLDs are used in Chicago. The spectrum of the recorded noise is displayed in Fig. 1. The noise was played back at the level (81 dB SPL) measured in the train car at the time of the recording.

Following the recordings, subjects were interviewed regarding potential risk to hearing from using their mass-storage PLD. They were also asked whether they experience tinnitus following PLD use. Descriptive statistics were performed on the data set in the statistical computing environ-

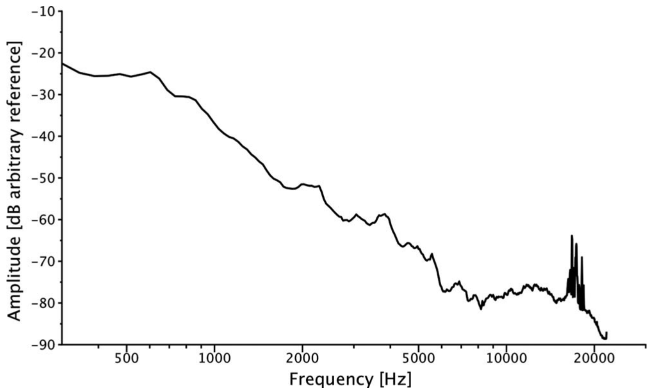

FIG. 1. Spectrum of noise recorded on a CTA train and played back at 80 dB SPL as the background noise.

ment R (R Development Core Team, 2006). A multiway-analysis of variance was also performed with L_{Aeq} measured in quiet and in noise as the dependent variables and gender, device, headphone type, and genre of music as the independent variables.

III. RESULTS

A. Listening Habits

23 subjects (77%) reported infrequent noise exposure (music concerts less than 2 times a year), and 7 subjects (23%) reported no exposure to noise in their everyday lives. Of the 23 subjects who reported infrequent noise exposure, only 4 (17% of the exposed subgroup; 13% of entire subject pool) reported routinely wearing some form of hearing protection while in noise, while an additional 4 (17% of exposed subgroup; 13% of entire subject pool) reporting infrequent use of hearing protection. 15 of the 23 subjects (65% of exposed subgroup; 50% of entire subject pool) reported not using any form of hearing protection while in noisy environments.

"Rock" was the most popular genre amongst our subjects followed by "Pop" and "Alternative Rock." Eighteen subjects (60%) used various models of an Apple iPod with several other devices being represented in approximately equal proportion with one subject (3%) using a multipurpose personal digital assistant as their device of choice. Subjects reported using the device on which measurements were made for an average of 13.2 months (SD=8.37; range =11 days to 36 months). All subjects reported using another type of PLD with headphones (cassette player, compact disc player, and/or another MP3 player) prior to acquiring the current device, with 19 subjects (63%) reported using two or more of these devices frequently in the past. Headphone style, make, and model were similarly varied with a large proportion of the subjects using "earbud" style headphones (73%) and none of the subjects using noise-isolating headphones. Reported use per week ranged from 1 day to 7 days with an average of 4.06 days (SD=1.99) (Fig. 2, bottom panel). Daily use ranged from 15 min to 4 h per day with an average of 1.68 h per day (SD=1.04) (Fig. 2, top panel). Thus, the group of subjects used their devices an average of 6.8 h per week (SD=5.89).

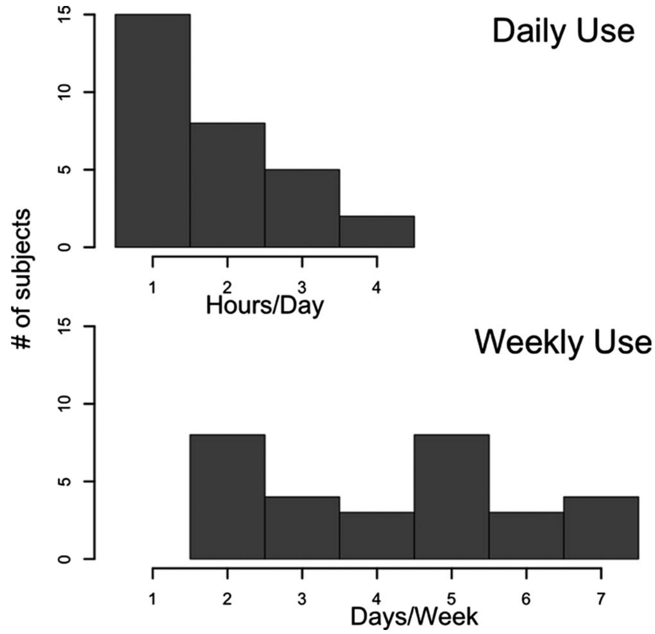

FIG. 2. Histograms of PLD use in hours/day (upper panel) and days/week (lower panel) for the 30 subjects.

B. Objective measurements

The *A*-weighted equivalent continuous noise level (L_{Aeq}) measured in the subjects' ear canal and in KEMAR in quiet is displayed in Fig. 3 along with a measure of exposure adjusted for the reported daily listening duration for each subject. L_{Aeq} is calculated as follows:

$$L_{eq} = 10 \log_{10}\left\{\left[(1/T)\int_{t1}^{t2} p_A^2(t)dt\right]/p_o^2\right\},$$

where $p_A^2(t)$ is the square of the instantaneous *A*-weighted sound pressure, in Pascals, as a function of time t for an averaging interval starting and ending at times t_1 and t_2, respectively, and p_o^2 represents the reference sound pressure of 20 μPa. If the exposure measurements are to be compared with NIOSH or OSHA standards, a transformation is necessary. In our application, we used the reported daily use for each subject to compute an L_{dpme} (daily personal music exposure) as follows:

$$L_{dpme} = L_{Aeq}(\text{in dBA}) + 10 \log[T/8],$$

where T is the reported use per day (in hours). Since we measured L_{Aeq} and then essentially used a 3 dB exchange rate to compute L_{dpme}, the values reported here are most applicable to the NIOSH risk criteria. In its derivation, L_{dpme} is identical to the computation for $L_{Aeq,d}$, where d is a duration other than 8 h. We choose to use the music-specific term L_{dpme} to emphasize the limitation of our measurement and computation to exposure related solely to the use of PLDs.

The box and whiskers in Fig. 3 represent the interquartile range with the whiskers extending to the greater of the range or 1.5 times the interquartile range. The symbols represent measurements from individual subjects with the open and closed symbols representing measurements below and above 85 dBA, respectively. The circles represent data from all 30 subjects while the squares represent the data from the

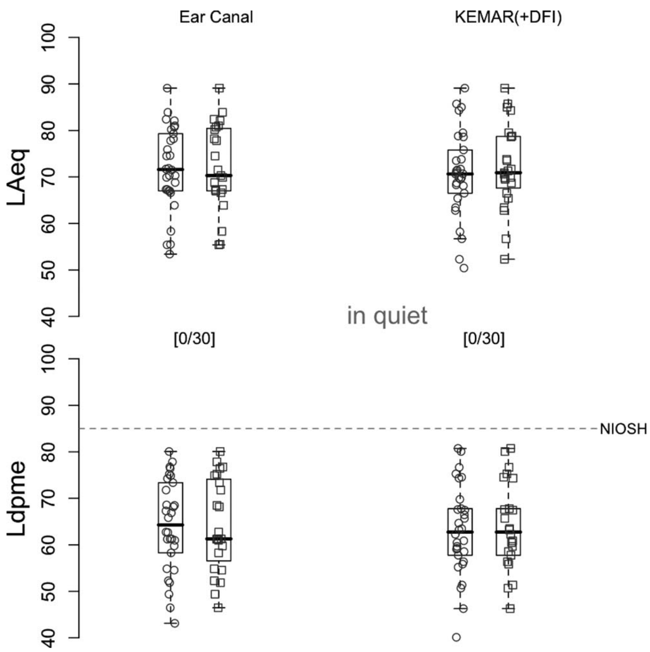

FIG. 3. Box-and-whisker plots comparing L_{Aeq} (top) and L_{dpme} (bottom) *in quiet* measured in the ear canal of individual subjects (left), and in KEMAR (right). The boxes span the interquartile range of the data with the whiskers extending to the greater of the range or 1.5 times the interquartile range. The solid line inside each box is the median. The dashed horizontal line in the bottom panel marks the limit of 85 dBA. See text for details on computation of L_{dpme}. The numbers in square brackets beside each data set show the number of subjects whose preferred listening level was equal to or greater than 85 dBA using that particular method of measurement (filled symbols). All open symbols represent data from subjects whose preferred listening level was lower than 85 dBA. The open and closed squares for each measure and measurement condition represent data from 23 of the 30 subjects, which were available for compensation for diffuse-field equivalent values. The circles, on the other hand, represent data from all 30 subjects.

limited set of 23 subjects whose ear-canal recordings were available for transformation to diffuse-field-equivalent values.

Measured L_{Aeq} in quiet (top left) ranged from 53.4 dBA to 89.1 dBA with a mean of 71.9 dBA (SD=8.9). One subject had L_{Aeq} values greater than or equal to 85 dBA. L_{Aeq} measured in KEMAR with DFI compensation are reported in the top right quadrant of Fig. 3. The data are presented in the same format as described for the ear-canal measures. L_{Aeq} in quiet in KEMAR ranged from 50.4 to 89.1 dBA with a mean of 70.7 dBA (SD=9.2). Three subjects had L_{Aeq} values greater than or equal to 85 dBA. Neither of these sets of measures is comparable to regulatory standards as listening duration is not taken into account, a hurdle in reporting personal damage risk in previously published work (Hodgetts *et al.*, 2007; Ahmed *et al.*, 2007).

Measured L_{dpme} in quiet are presented in the bottom panels of Fig. 3 for both ear-canal and KEMAR measurements. The ranges of L_{dpme} were 36.3–76.0 dBA and 39.3–76.08 dBA with mean (and standard deviations) of 57.5 (8.9) and 56.3 (9.8) dBA in the ear canal and the KEMAR, respectively. Although the measurements made in the KEMAR were lower than those made in individual ear canals (not DFI compensated) in the vast majority of our subjects, these differences were not found to be significantly different. The average L_{dpme} values (accounting for reported listening durations) were lower than the L_{Aeq} values by approximately 14 dB. The ear-canal measurements would have been lower (~6–8 dB estimate from our measurements) had we compensated them to express their diffuse-field equivalents.

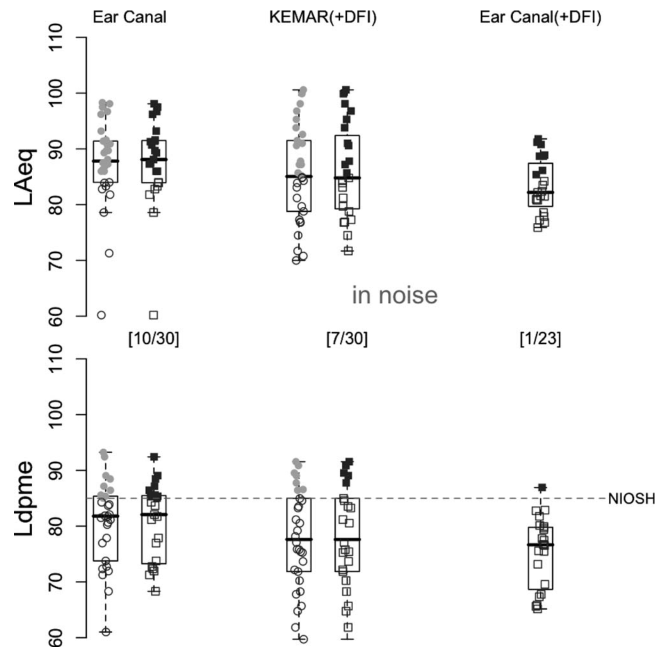

FIG. 4. Data similar to those in Fig. 3 but measured in noise. Measurements made in the ear canal and in KEMAR with DFI are presented with circles. Additionally, the ear-canal data from 23 subjects are presented in the right-most column after equivalent diffuse-field compensation. This limited data set are for the ear-canal and KEMAR measurements are also presented in squares. Closed symbols represent preferred listening levels equal to or greater than 85 dBA. Circles and squares present data from all 30 subjects and the 23 subjects whose data were available for transformation, respectively.

Results of objective measurements made *in noise* are presented in Fig. 4. The format of the figure is similar to that of Fig. 3. In addition to the ear-canal and KEMAR data, the ear-canal data compensated for diffuse-field equivalence are presented in the right column. For the L_{Aeq} measurements (top row), mean (and standard deviation) values were 87.5 (7.8), 85.4 (8.5), and 82.01 (8.4) dBA for the ear canal, KEMAR, and DFI-compensated ear-canal recordings, respectively. For the L_{dpme} measurements (bottom row), mean (and standard deviation) values were 79.9 (9.1), 77.75 (8.9), and 74.14 (9.5) dBA for the ear canal, KEMAR, and compensated ear-canal recordings, respectively. The numbers of subjects at risk for L_{dpme} measured using each technique (ear canal, KEMAR, or DFI-compensated ear canal) are presented in square brackets in the bottom panel of Fig. 4. The portion of our subjects found to be at risk while listening in background noise varied greatly depending on the measurement used. 33% of our subjects were found to be at risk based on the L_{dpme} measured in the ear canal. In contrast, one (4%) of the subjects were found to be at risk based on the L_{dpme} computed after transforming the ear-canal recording to diffuse-field equivalence.

We have presented the data from the subset of 23 subjects for all conditions to demonstrate that the range and/or variance of the data set was not affected significantly due to the omission of the 7 subjects whose data were not available

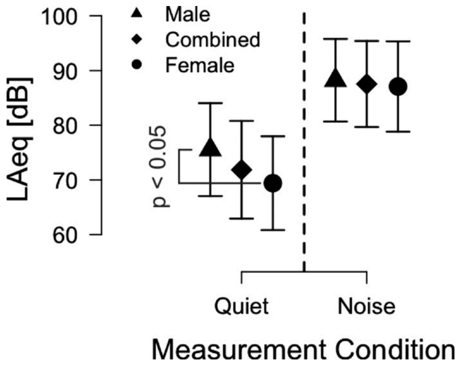

FIG. 5. Mean L_{Aeq} values in quiet (left) and in noise (right) for the entire subject pool and those for the male and female subjects. The error bars represent ±1 standard deviation.

for transformation to free-field equivalence. It is notable that the variance in the data is significantly reduced for the ear-canal data once compensated for diffuse-field equivalence.

The L_{Aeq} in the male participants was found to be 6.1 dB higher than the females in the quiet condition, and this difference was statistically significant ($p<0.05$) in a general linear model where gender, genre, device, and headphone types were all included as predictor variables without any interaction terms (Fig. 5). The main effect of gender was, however, not statistically significant ($p=0.06$) when the order of variables was changed on the right-hand-side of the model or when gender was used in isolation. A test of power revealed that a sample size of 32 in each gender-group would be needed to attain an effect size of 0.80 (confidence interval=95%) for gender if the difference between the two means (6.1 dB) and the standard deviation in the two samples (~8.5 dB) remained unchanged. The difference in L_{Aeq} between male and female subjects was smaller and not statistically significant in the noise condition. Similarly, no significant differences between male and female subjects were observed for L_{dpme}. The introduction of the background noise caused an average increase in L_{Aeq} by 15.7 dB.

IV. DISCUSSION

The possibility of increased incidence of hearing loss due to PLD use has been of recent interest in the popular media. Audiologists and other hearing health care professionals are often sought out as experts in this area. It is our responsibility to provide a realistic estimate of the potential problem and such estimates have to be based on accurate ecologically relevant measurements. Such empirical measurements then should be related to criteria for possible hearing damage for public consumption. Unfortunately, the only standards that allow any determination of damage risk were developed specifically for industrial noise and their application to music is not straightforward.

Previous work had time and again estimated that a vast majority of users of portable devices listen to them at safe levels (Rice *et al.*, 1987b; Wong *et al.*, 1990; Turumen-Rise *et al.*, 1991b). However, the alarm regarding the recent generation of PLDs was perhaps justified based on the capability of these devices to store more music and play the music back at hazardously high sound pressure levels (Fligor and Cox, 2004, Keith *et al.*, 2008). Objective measurements of preferred listening levels using modern-day PLDs have been missing from the debate. One peer-reviewed publication put 25% of the population "at risk" of listening at hazardous levels based on measurements made in KEMAR after randomly seeking participants from the street and not allowing the participants to adjust the volume control of their device prior to the measurement (Williams, 2005). Hodgetts *et al.* (2007) made similar measurements near the tympanic membrane of 38 individuals but could not relate their findings to NIOSH or OSHA standards as their measurements were not converted to their diffuse-field equivalent values. Further, one piece of music, limited in dynamic range to mimic industrial noise, was used for all subjects. While this approach makes the measures (when converted to diffuse-field equivalents) more comparable to NIOSH and OSHA standards, whether the volume setting is an accurate representation of an individual's habitual listening level is questionable. For example, a listener may "turn it up" for music they like but "turn it down" when the music is chosen by someone else and not liked as well. We deliberately chose to allow the individual to select the music thereby reducing the possibility of a bias towards lower listening levels. While this probably allowed greater accuracy in characterizing habitual listening levels, it makes comparison with OSHA and NIOSH standards less transparent.

Individual use over time (L_{dpme} rather than L_{Aeq}) and diffuse-field equivalency have been accounted for in this study to estimate the at-risk population. These estimates were based on standards-based measurement techniques and self-reported duration of use. Once these essential accommodations were made, one (4%) of the participants was found to be at risk *when listening in noise* using the stricter NIOSH benchmark for occupational noise exposure. In contrast, 7 of the 30 (23%) subjects would be categorized to be at risk if the measurements made in the KEMAR and then compensated for free-field equivalency. Our results are most directly comparable to, and in close agreement with those of Williams (2005). Williams (2005) reported an average L_{Aeq} of 86.1 dB measured in a manikin in the presence of background noise, comparable to the 85.4 dB reported here. Similarly, Williams (2005) reported an average value of 79.8 dB after accommodating for individual listening durations, comparable to the average L_{dpme} of 77.75 measured in the KEMAR in this data set. None of the participants would be considered at risk when listening in quiet, either with KEMAR or ear-canal recordings.

It was remarkable that the variance was significantly reduced in our free-field equivalent measures from the ear canal (MIRE technique, ISO, 2002). This transformation was only possible in 23 of the 30 subjects due to loss of data between computers. However, it is clear that the reduced variance is *not* a sampling error as the range of this reduced set is equivalent to the complete data set as seen by comparing the circles and squares for any measurement in Figs. 3 and 4. This would suggest that the subjects adjusted the listening levels to compensate for individual ear-canal acoustics, earphone frequency response, and fit and program material. The preferred listening level was relatively consistent between subjects. It should also be noted that the transformation to diffuse-field equivalence does not scale all measures by a constant but rather is dependant on the spectrum of the music. The significant and important differences between the measurements made in KEMAR with the DFI filter activated (ISO, 2004) and the DFI-compensated ear-canal measurements (ISO, 2002) suggest that the latter is clearly the superior measure when assessing exposure from PLDs even with the use of an average transfer function, rather than individually measured ones. Our results seem to suggest that listeners choose a listening level from a relatively narrow range of levels which was by and large safe for our subjects, given their reported duration of use.

Given our small sample size and a biased population of a group of university students, our estimates of population at risk should not be extended to the general population. Caution has to be exercised in interpreting results from studies such as this one and projecting probable incidence of hearing loss from them. Even after the measurement is made using an appropriate diffuse-field compensation and an estimate of the daily listening duration is incorporated in computations, the comparison made is with standards developed specifically for industrial noise. Further, the use of specific types of background noise cannot be easily generalized to all listening situations. Perhaps most importantly, the bias introduced by participating in an experiment and selecting a preferred listening level knowing that an objective measurement is forthcoming cannot be accounted for. Finally, any additional exposure from other sources of noise is not accounted for in these computations, and hence may underestimate the true daily exposure of an individual.

Male participants were listening to the PLDs at a significantly higher level than female participants in quiet ($p < 0.05$). No gender-based difference was detected when the measurements were made in background noise. The difference in preferred listening level between male and female subjects was approximately 5 dB in our data set—in close alignment with other reports (Williams, 2005; Hodgetts *et al.*, 2007). No significant effects of gender were found on age and hearing thresholds, thus making the 5 dB difference between males and females an interesting yet unexplained phenomenon. The identification of physiological and psychological factors defining individuals who prefer to listen to music at levels higher than the general population may be an interesting and important area of research.

In the exit interview, close to three-quarters of our subjects were certain that they were not using their PLDs at levels harmful to their hearing and indeed they were not. However, 17% reported that they believed that they were using their devices at harmful levels. Encouragingly, an overwhelming majority of our subjects said that they would reduce their listening level or shorten their listening session if asked by a professional. These reported trends suggest that at least in a population equivalent in education and socioeconomic status to our subject pool, public awareness and education can play a very important role in preventing hearing loss due to over exposure from music delivered by PLDs.

V. CONCLUSIONS

Audiology and other professions related to hearing health care have received tremendous attention in the recent past with regard to PLD and the potential hazard they pose. Following a standardized and accepted protocol for accurate and consistent measurements that can be related to standards of damage risk is essential. We have demonstrated that the estimation of the population at risk varies significantly with the method used for measuring preferred listening levels. After comparing three methods of measurement, we recommend measurements made in the ear canal and then compensated for free-field equivalence, as described in ISO 11904-1 (ISO, 2002), as the method of choice. Further, we recommend incorporating daily listening durations, as we have done in our L_{dpme} computation, in risk assessment. We hope that such ecologically valid and accurate characterizations of listening levels will become the basis of risk projection and counseling.

ACKNOWLEDGMENTS

The research reported here was conducted as a Capstone project in the partial fulfillment of the Au.D. degree at Northwestern University (DW). The authors would like to thank Brian Fligor, Sc.D., for helpful discussions on topics related to his paper. They also thank the clinical staff at the Northwestern Hearing Clinic (Evanston) for the use of their testing suites, the members of the Auditory Research Laboratory at Northwestern University for their support. The Au.D. program at Northwestern University provided financial support to the project. The project was also partially supported by the NIH/NIDCD, Grant No. DC005692-04 to S.D.

APPENDIX A: DATA RECORDING SHEET

DATA SHEET
Subject ID#: ________ Date: ___________
Filename: ___________ Tester Initials: ___
Subject Demographics:
Age:______
Sex: ______
Hearing History:

(1) Do you have any concerns with your hearing?
(2) History of ear infections?
(3) Have you ever been exposed to loud noises for long periods of time?
(4) In what environment?
(5) Do you wear hearing protection?
(6) Are you on any ototoxic medications (Aspirin, etc.)?
(7) Any ringing in your ears?
(8) Any head or neck surgeries within the last year?

Transducers: Make: _____________
Model #: _____________
Type:
☐ Circumaural
☐ Supra Aural
☐ Earbuds
☐ Ear Inserts

☐ Noise-Isolating
☐ Other
Initial Oral Interview:

(1) How long have you been using this particular device?
(2) On average for how many hours a day do you use your device under headphones?
(3) On average how many days a week do you use your device under headphones?
(4) What genre(s) of music do you generally listen to a majority of the time under headphones?
(5) In what situation do you listen to the device under headphones the majority of the time (e.g., home, train, while exercising, etc.)
(6) Do you listen to music through speakers? (i.e., computer speakers and stereo system)
(7) Have you previously used any of the following PLDs?:
☐ Personal cassette player
☐ Portable compact disc player
☐ Personal AM/FM Radio that required the use of headphones
☐ Other

Exit Oral interview:

(1) Can loud sounds cause damage to your hearing?
(2) Have you ever been educated on the risks to hearing from loud sounds?
(a) If so, when?
(b) Were you ever taught in middle or high school health class?
(3) Do you wish that you could turn your device up louder at anytime?
(a) If yes, when?
(4) Have you ever experienced a ringing (tinnitus) in your ears after using your device?
(a) If yes, how long did it last?
(5) Do you feel that you listen to your device at harmful levels?
(6) Do you feel that you are harming your hearing by using this device?
(a) If yes, do you plan to continue using this device?
(7) Do you feel that manufacturers should put a warning label about the potential risk to hearing from the maximum outputs of these devices?
(8) If you were warned that you are listening to your device under headphones at harmful levels and advised to turn it down/limit your daily use would you?

APPENDIX B: INSTRUCTIONS TO SUBJECTS

1. TRACKING

This test you will hear some tones. Once you begin to hear them, press and hold the mouse button. When you can no longer hear them, release the mouse button. Once you begin hearing them again, press and hold the mouse button until you can no longer hear them. Follow this pattern until you hear a bell, we will then move on to another tone. Do you have any questions?

2. CONDITION 1

You will set your preferred listening volume to your preferred musical selection. Please let me know the song title, band, and genre of music you will be selecting. I will be making a 3 min recording while your music is playing. During that time it is important that you do not talk. After the recording has ended, pause your selection and do not change the volume. I will take your device and place it on the manikin and make another 3 min recording. Any questions?

3. CONDITION 2

We will do the same thing again, but this time I will be playing train noise through the speakers. Again, set your preferred listening volume to your previously used musical selection. The recording will last 3 min. After the recording has ended, pause your selection and do not change the volume. I will make another recording on the manikin. Any questions?

Ahmed, S., Fallah, S., Garrido, B., Gross, A., King, M., Morrish, T., Pereira, D., Sharma, S., Zaszewska, E., and Pichora-Fuller, K. (**2007**). "Use of portable audio devices by university students," Can. Acoust. **35**, 35–52.
American National Standards Institute (ANSI) (**2004**). "Standards for audiometers," Standsard S3.6–2004.
Farina, A. (**2007**). *A Study of Hearing Damage Caused by Personal MP3 Players* (Audio Engineering Society, New York).
Fligor, B. J., and Cox, L. C. (**2004**). "Output levels of commercially available portable compact disc players and the potential risk to hearing," Ear Hear. **25**, 513–27.
Hodgetts, W. E., Rieger, J. M., and Szarko, R. A. (**2007**). "The effects of listening environment and earphone style on preferred listening levels of normal hearing adults using an MP3 player," Ear Hear. **28**, 290–297.
In-Stat (**2007**). "MP3 player market to reach 286 million units by 2010," In-Stat Market Track Research Press Release, http://www.instat.com/press.asp?ID=1648&sku=IN0603155ID (Last viewed May 24, 2006).
ISO (**2002**). "Acoustics—Determination of sound immission from sound sources placed close to the ear Part 1: Technique using a microphone in a real ear (MIRE technique)," ISO 11904-1, Switzerland.
ISO (**2004**). "Acoustics—Determination of sound immission from sound sources placed close to the ear Part 2: Technique using a manikin," ISO 11904-2, Switzerland.
Katz, A. E., Grestman, H. L., Sanderson, R. G., and Buchanan, R. (**1982**). "Stereo earphones and hearing loss," N. Engl. J. Med. **307**, 1460–1461.
Keith, S. E., Michaud, D. S., and Chiu, V. (**2008**). "Evaluating the maximum playback sound levels from portable digital audio players," J. Acoust. Soc. Am. **123**, 4227–4237.
Kenna, M. A. (**2008**). "Music to your ears: Is it a good thing?," Acta Paediatr. **97**, 151–152.
Killion, M. C. (**1979**). "Equalization filter for eardrum-pressure recording using a KEMAR manikin," J. Audio Eng. Soc. **27**, 13–16.
Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467.
Nature Neuroscience (**2007**). "More noise than signal," Nat. Neurosci. **10**, 799.
Neely, S. T., and Stevenson, R. (**2002**). *SysRes* (Boys Town National Research Hospital, Omaha, NE), Technical Memo No. 19.
R Development Core Team (**2006**). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, http://www.R-project.org (Last viewed September 10, 2008).
Reuters Press (**2005**). "Limit use of iPod earbuds to protect your ears," MSNBC Online, http://www.msnbc.msn.com/id/10648715/ (Last viewed May 26, 2006).
Rice, C. G., Breslin, M., and Roper, R. G. (**1987a**). "Sound levels from personal cassette players," Br. J. Audiol. **21**, 273–278.
Rice, C. G., Rossi, G., and Olina, M. (**1987b**). "Damage risk from personal cassette players," Br. J. Audiol. **21**, 279–288.

Turumen-Rise, I., Flottorp, G., and Tvete, O. (**1991a**). "Personal cassette players ('walkman'). Do they cause noise-induced hearing loss?," Scand. Audiol. **20**, 239–244.

Turumen-Rise, I., Flottorp, G., and Tvete, O. (**1991b**). "A study of the possibility of acquiring noise-induced hearing loss by the use of personal cassette players (walkman)," Scand. Audiol. **34**, 133–144.

Williams, W. (**2005**). "Noise exposure levels from personal stereo use," J. Acoust. Soc. Am. **44**, 231–236.

Wong, T. W., Van Hasselt, C. A., Tang, L. S., and Yiu, P. C. (**1990**). "The use of personal cassette players among youths and its effects on hearing," Public Health **104**, 327–330.

Wood, W. S. and Lipscomb, D. M. (**1972**). "Maximum available sound pressure levels from stereo components," J. Acoust. Soc. Am. **52**, 484–487.