CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Texas A&M University

PILOT STUDY VALIDATION OF hEAR MOBILE HEARING SCREENING APPLICATION IN THE GENERAL POPULATION

A Thesis

by

LAKSHMI VAISHNAVI DAKURI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN PUBLIC HEALTH

Chair of Committee,	Adam W. Pickens
Committee Members,	Ranjana Mehta
	Hongwei Zhao
Head of Department,	Virender Sharma

May 2015

Major Subject: Occupational Health

Copyright 2015 Lakshmi Vaishnavi Dakuri

ABSTRACT

The purpose of this pilot study was to test the quality of data collected by a mobile hearing screening application (hEAR) against the gold standard of pure tone audiometry administered by a certified audiologist. hEAR used 7 pre-set frequencies (125 Hz, 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz and 8000 Hz), which were the independent variables, and recorded measurements as sound pressure levels in decibels (dB) during three trials.

In total, 30 subjects were recruited from the general population at Texas A&M University. Subjects were randomly assigned and counterbalanced in their assignment to a "quiet" room and a "noisy" room. Subjects used the hEAR mobile hearing screening application to self-administer hearing screening tests. Subjects also had hearing screening examinations performed by a certified audiologist at the identified pre-set frequencies.

Data were analyzed using a mixed effect model and testing for repeated measures at 95% confidence intervals, results were separated by room. It was found that the hEAR trials differed from the audiologist trial at almost all frequencies in a noisy environment, but only at 2000 Hz and 8000 Hz for the quiet environment. It was also found that the app trials were very similar to one another (trials 1&2, trials 1&3 and trials 2&3 similar to each other) in the noisy environment; while they statistically differed from one another at almost all frequencies except 125 Hz in a quiet environment.

Further research is needed so as to develop hEAR as an effective alternative to an audiologist-administered pure tone hearing test, which can consequently be used for better compliance with OSHA's hearing screening requirements.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Adam Pickens, and my committee members, Dr. Ranjana Mehta and Dr. Hongwei Zhao, for their guidance and support throughout the course of this research.

Thanks also go to Dr. Camille Peres and her graduate students for sharing their lab with me. I must also thank the faculty, staff and students at the School of Public Health and Texas A&M University, for making data collection possible. I would also like to thank Dr. Mark Benden, Dr. Thomas McDonald and Ms. Annamarie Bokelmann for encouraging me to pursue this path.

Finally, thanks to my parents for their encouragement and to my husband, Bobby Robertson, for putting up with me through the writing process.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

NOMENCLATURE

OSHA	Occupational Safety and Health Administration
HHIE	Hearing Handicap Inventory for the Elderly
HHIE-S	Hearing Handicap Inventory for the Elderly-Screening
WHO	World Health Organization
ASHA	American Speech-Language-Hearing Association
ONIHL	Occupational Noise Induced Hearing Loss
TS	Telessaúde
TAMU	Texas A&M University
IRB	Institutional Review Board
SPL	Sound Pressure Level
ANSI	American National Standards Institute
DJ	Disc Jockey

CHAPTER I
INTRODUCTION AND LITERATURE REVIEW

Introduction

Hearing impairment is one of the most common debilitating illnesses. There is evidence through prior research indicating that hearing loss or impairment may lead to social isolation, and consequently depression and withdrawal from daily activities (Mulrow & Lichtenstein, 1991). This leads us to endeavor to include hearing assessments in the normal health assessments for the older population. The overall objective would be to pinpoint individuals who would require further screening, assessments and consequently diagnostic corrections. Assessments would help narrow down, differentiate and correctly diagnose hearing impairments and hearing handicap. According to Schow (1991), hearing impairment is the "deficit in structure and/or function", while hearing handicap/disability is the "effect of such a deficit". Mulrow and Lichtenstein (1991), estimate that almost a quarter of the population above the age of 65 reports some form of decrease in auditory function. Usually this decrease is due to natural age-related deterioration or presbycusis, however, there is also increasing probability that this deterioration may be due to other factors such as occupational factors and non-work related factors.

Literature Review

According to the Bureau of Labor Statistics, the industry with the highest number of occupational hearing loss related OSHA recordables was the manufacturing and utilities sector. Within the manufacturing and utilities sector, metal manufacturing had the highest number of complaints, 33.8 cases/1000 full-time workers (Martinez, 2012).

The type of screening selected for audiometric tests depends on the testing criteria. However, pure tone audiometry is generally recognized as an industry gold standard (Dalton, et al., 2003), (Yueh, Shapiro, MacLean, & Shekelle, 2003). However, it is not an easily accessible option, and at times it may not be reimbursed by the

employer (Gates, Murphy, Rees, & Fraher, 2003). Therefore, it is highly probable that many practices are dependent upon self-administered tests. A benchmark for such tests is the Hearing Handicap Inventory for the Elderly (HHIE) developed by Ventry & Weinstein in 1982 (Gates et al., 2003). It is a 25-item/question survey that was devised for the assessment of "self-perceived psychosocial handicap of "hearing impairment" in the elderly population, and was meant to function as a supplement to pure tone audiometry to assess the effectiveness of hearing aids. It is one of the most validated and widely accepted screening methods; however, it was primarily developed as a method "to assess the effectiveness of amplification" (Gates et al., 2003).

Subsequent research led to the development of a shorter 10-question version of the HHIE called as the Hearing Handicap Inventory for the Elderly-Screening (HHIE-S), which was designed to work as the actual "screening tool" (Gates et al., 2003). Even with HHIE and HHIE-S, there is still a probability of prediction of false positives, and false negatives, that is to say, their sensitivity (true positive) and specificity (true negative) is not 100%. Several researchers have endeavored to formulate less time consuming tests with better positive predictive values than the currently existing standards. According to Gates et al. (2003), the global measure, "Do you have a hearing problem now" is one of the better screening methods, especially for the geriatric population.

Yueh et al. (2003) conducted an extensive review of various screening tests currently in practice, such as the whispered voice test (Mulrow, 1991) involves whispering words from behind the patient at varying distances. Hearing loss is determined by the farthest distance from which the patients could still satisfactorily reiterate what was whispered. The test is relatively easy to administer, however, the lack of any kind of standardization is one of its disadvantages, the other being low test-retest reliability.

The tuning fork test (Mulrow, 1991) is another widely applied screening test. It is similar to the whispered voice test in its execution except that a tuning fork is used to

test a patient, instead of whispering. Such a test also relies on the same principle of distance as the whispering test, and is more or less subjected to similar biases and problems. (Burkey, Lippy, Schuring, & Rizer, 1998).

The self-administered HHIE-S test as developed by Ventry and Weinstein, is another screening test. It is based on a point based system, wherein a "Yes is 4 points, a Maybe is 2 points, and a No is 0 points". The points range from 0-40 with hearing loss increasing in an ascending order. A score between10-24 depicts a 50% probability of hearing loss, while an increasing score consequently means an increasing probability of hearing loss. This test is most preferred due to its ease of administration, good test-retest reliability and inter-subject reliability.

However, compared to an audioscope, the HHIE-S has much lower sensitivity (Yueh et al., 2003). The audioscope is "a handheld combination otoscope and audiometer that delivers 25-40 dB pure tone at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz" (Yueh et al., 2003). The device costs $400-$600, and is held directly to the ear canal, and tones are repeated at every frequency mentioned. If the patients hear certain predetermined set of tones that indicative of hearing loss, they are referred to a specialist for an audiogram. Though Yueh et al. (2003) compared the audioscope and the HHIE-S to each other; they acknowledged that the two screening tests could be used to test for different spectrums of hearing loss. As explained by Yueh et al. (2003), the audioscope is focused on detection of physiologic loss, so it may identify patients with existing hearing loss but may not be able to identify those 'who are motivated to seek treatment"; while on the other hand HHIE-S may not be able to detect early disease. Therefore the researchers recommend using a combination of the two tests. However, most of the tests discussed here are meant for the elderly population.

Recommendations by Organizations

According to OSHA, any and all 'self-recording audiometers' should comply with appendix C of 29 CFR 1910.95 (OSHA, 29 CFR 1910.95 (h)). According to Appendix C, the requirements that the all such equipment should fulfill are as follows:

- The slewing rate for the audiometer attenuator shall not be more than 6 dB/sec except that an initial slewing rate greater than 6 dB/sec is permitted at the beginning of each new test frequency, but only until the second subject response.
- The audiometer shall remain at each required test frequency for 30 seconds (+ or - 3 seconds). The audiogram shall be clearly marked at each change of frequency and the actual frequency change of the audiometer shall not deviate from the frequency boundaries marked on the audiogram by more than + or - 3 seconds.
- It must be possible at each test frequency to place a horizontal line segment parallel to the time axis on the audiogram, such that the audiometric tracing crosses the line segment at least six times at that test frequency. At each test frequency the threshold shall be the average of the midpoints of the tracing excursions.

While these are available in a variety of settings, trained personnel still are needed to administer the tests. This is due to the fact that while the audiometers are self-recording in unto themselves, they are not self-administered by the individuals being tested. The World Health Organization (WHO) however, has requirements for both self-administered and self-recorded audiometric tests (Franks, 1995). According to the requirements, self-administered audiometry employs the use of an attenuator that can either increase or decrease the signal intensity at a fixed rate, and the listener has control over the attenuator. By pressing the 'response switch' the listener has the ability to decease the signal intensity, and upon release of the switch, the intensity increases. The listener's threshold is usually between the point of pressing and the point of releasing the switch. These recommendations are based on the best practices of self-administered tests identified by Békésy audiology test patterns, which is a type of hearing test in which the subject controls the intensity of a stimulus by pressing a button while listening to a pure-tone whose frequency moves through the entire audible range; and these test patterns have long been held as the self-administered test protocol standard (Franks, 1995).

Professional organizations like American Speech-Language-Hearing Association (ASHA) discourage the use of frequency at 6000 Hz during testing due to higher probability of prediction of false positives, especially at a lower screening decibel level (Meinke & Dice, 2007). Audiometric screening thresholds are not standardized for school testing, which could have several negative consequences, such as liability suits against the school system, or delay in timely treatment. Therefore, it is highly pertinent that audiometric testing screening requirements be standardized.

Requirements of a Screening Protocol

To validate a screening protocol, it is important to understand why screening is needed in the first place. Mulrow and Lichenstein (1991) devised certain criteria to justify the need of screening. According to them, there are primarily two questions that a researcher should ask and be able to affirmatively answer, and these are listed below:

- "Does the burden of suffering warrant screening?" In case of testing of hearing loss and hearing handicap, the answer would be a resounding yes. Decrease in hearing ability leads to a decrease in the quality of life of the patients.
- "Are there any good screening tests?" The tests should be reliable, valid and acceptable. As mentioned earlier, there are certain 'validated' tests such as the whisper test, the tuning fork test, hand-held audioscope, the HHIE-S questionnaire and of course, pure tone audiometry.

Following these steps, there are certain criteria that are more applicable to follow up interventions after successful screening. These are:

- Persons with positive screening results should comply with the suggested interventions
- The effectiveness of the interventions should have been demonstrated in a randomized trial
- The interventions should have broad public health based implications

- The present health care system should be able to comply with the suggested interventions

The overarching objective of any screening protocol is to identify the section of the test population who may need further assessments. Therefore the objectives can be surmised as "identifying medical impairments needing referral, and finding potential hearing handicaps that need referral" (Schow R. , 1991). According to Ibrahim (1985), any screening protocol should include epidemiological principles that support the testing which are as follows:

- The condition must represent an important health problem
- The condition should have a preclinical or asymptomatic period that is identifiable by a test or a maneuver and should be amenable to intervention at this phase of its course.
- The intervention at this point should lengthen or improve the quality of life with respect to intervention when the condition becomes symptomatic.
- Detection of a risk factor or an early stage of a disease in an otherwise normal individual may also have consequences as a result of 'labeling' the individual as being at risk or having an early stage of a disease, and untoward effects of labeling should be weighed
- Screening tests when applied to large masses of population should be simple, safe, acceptable and cost effective.
- A screening test has several properties that must be understood and evaluated before a policy decision is made as to its inclusion/exclusion from a screening program.
- An intervention procedure must be available, accessible, and acceptable to the population for which it applies.

Validation of a screening test depends on the selection of the gold standard, which should be valid and reliable. As mentioned previously, in terms of hearing tests, the gold standard is pure tone audiometry administered by a certified audiologist. In such cases, the experimenter should decide on what frequency/frequencies are to be used, and this depends on the reason behind the testing. Most audiometric tests are conducted at 125 Hz, 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz, and 8000 Hz though there may be several other frequencies that are used. As far as the sound pressure levels are concerned, research has shown that for the elderly population a higher decibel screening point such as 40 dB should be used (Schow, Smedley, & Longhurst, 1990), but it may not be of much use for a considerably younger population, as it may result in several false positives and consequently a serious oversight in terms of the health of the working population.

As far as high risk groups such as an occupational work group identified in a noisy environment, or youth are concerned, then routine checkups would be more useful, so as to allow for timely detection of any anomalies. Schow et al. (1990) made recommendations on the utilization of various screening thresholds (or pure tone fences or sound pressure levels) for pure tone audiometry, namely low, medium and high. Low fences of 15-20 dB would benefit the youth and would allow for higher accuracy. Similarly, a mid level fence between 25-35 dB would help in successfully pinpointing any hearing related handicap in adults, as such a handicap "emerges with adults when thresholds at 1000 Hz and 2000 Hz exceed 25 dB." Furthermore, these findings may be validated at higher pure tone thresholds of 40 dB, thereby being self-validated.

Use of Available Software for Screening

Any viable screening method should be easily accessible first and foremost. In case of many occupational workforces there is a high probability that spatially these samples may be scattered, or may be located in rural areas, or in areas where there is limited outside penetration. In such cases, it is highly important that the testing platform be the most widely available and at the same time affordable. Pure tone audiometry

administered by an audiologist is at times very hard to access, especially if the study area is rural in nature. This is due to the centralization of most audiologists in high population density locations. Also, audiologist-administered pure tone audiometry is generally regarded as the only option as far as screening is concerned (Gates, Murphy, Rees, & Fraher, 2003). In this context, it is judicious to use/develop screening method that mirrors certain characteristics of pure tone audiometry, but is highly accessible at the same time. Ferrari, Lopez, Lopes, Aiello, & Jokura (2013) have analyzed options comparative to pure tone audiometry. In their study, they employed the use of Telessaúde (TS) audiometer which could be used with ordinary plug-in USB headphones. This makes the TS audiometer a cheaper option than audiologist-administered audiometry. The TS audiometer has been proven to have a high degree of sensitivity and specificity (Ferrari, Lopez, Lopes, Aiello, & Jokura, 2013). The TS audiometer was developed as software for computers, which included a calibration user interface, through which calibration parameters for certain headset models could be determined and stored. These parameters were used when audiometric screening was performed. The device (computer and headphones) were calibrated by an engineer who had experience with audiometric calibration. Since the TS audiometer was specifically developed as a robust, low-cost and more importantly mobile alternative to pure tone audiometry, it is an evidence of the need for more such alternatives to be readily available.

Hand-held smartphones subscriptions are steadily growing reaching 67 per 100 inhabitants globally, and is estimated to be almost double that number in the United States (WHO, 2011). The use of the iPhone and other mobile devices and platforms for such purposes has been well documented. With respect to similar applications available that focus on data collection, medical applications that monitor one's health, medications, doctors' appointments, hospital/physician visits are the most common. Several hospitals are now advocating the use of iPhones and iPads for their staff, so as to reduce any manual clerical errors that may at times be fatal. The use of such devices has been positively received by the target users as it 'less error prone', 'more hygienic', 'less

cumbersome', and usually have certain fail switch technology embedded in the program to allow for double checking or sometimes triple checking the entered data (Hamou, et al., 2010). With the success of such devices and their applications, it is prudent for us to use similar technological platforms to screen for hearing loss and hearing handicap.

Need for Interventions

Current literature on occupational noise induced hearing loss indicates that there is a need for an easy-to-administer test for pure tone audiometry. However, there is considerable conflict between methods and researchers as to what parameters should be used while administering these tests. Also, the inherent presence of inter and intra-rater variability makes the tests very subjective and at times unreliable (Leensen, de Laat, & Dreschler, 2011). There is a lack of standardization between different tests due to discrepancies between methods, which may include different ambient conditions (different subjects, difference in protocols such as length of administration, level of subject activity, difference in hearing gears etc.).Validity of a new application or procedure to test Occupational Noise Induced Hearing Loss (ONIHL) is hence a monumental task.

Ongoing research such as that by Ferrari et.al. (2013), and lack of an effective stand-alone alternative to audiologist-administered audiometry is an evidence of the need for more alternatives to audiologist administered pure tone audiometry, which are reliable and viable.

CHAPTER II

METHODS AND EQUIPMENT

Methods

Our application, hEAR, has been developed as a combination of best-practices for self-administered tests in accordance with the testing requirements as indicated by OSHA and the WHO. As is best practice with these recommendations, test tones initiate at inaudible levels and subjects respond to the attenuator control once they hear the tone. This is in contrast to the best practices of an audiologist-administered test where the tone is initiated at an audible level and lowered until it cannot be heard by the subject.

Objectives

The objectives for this study were the following:

- Validate the hEAR pure-tone mobile hearing screening application against Audiologist-administered pure-tone hearing screening data.
- Assess intra-subject variability for the use of the hEAR application to determine any learning effect of the end user.
- Determine whether or not the room in which the test was administered, has any effect on the results

This population prospective cohort study was the first of its kind.

Sample Size

Our assumptions were that *r* for the repeated measures ANOVA was 0.4, and there were three groups at alpha level 0.05 (three trials per person), also, effect size is 0.5. According to the power table (Li & Barker Bausell, 2006), for these values the sample size was 11. However, for the purposes of this project, the original sample size was chosen to be 20, which was later increased to 30 to better analyze our hypotheses.

Subject Characteristics

The study recruited 30 subjects from Texas A&M University (including students, faculty/staff). The subjects were in good physical health (self-reported). They were notified by email about the study through a TAMU-IRB-approved email script and those who were interested and who met the requirements scheduled a date and time to perform testing. Subjects received a $20 Target gift card for their participation, and the research team paid all charges incurred for the subject screening by a certified audiologist.

Out of the 30 subjects, 21 were male and 9 were female. Their ages ranged from 21 to 67, however, most (24) of the subjects were in the 21-28 age range. Most (22) of the subjects were undergraduate and graduate students from Texas A&M University. The subjects also completed a survey questionnaire prior to the collection of the audiology data. Along with demographic questions, the survey included questions about hearing-related medical and family history and estimates about the amount of time the subjects spent that exposed them to the hazardous sources. Each subject was assigned a participant ID which was a 7 digit random number generated by the uniform distribution random number generator for data collection/analysis purposes. Communication between the researchers and the audiologist used this identification number to maintain subject protection standards. The subjects were also sent to a local certified audiologist so as to undergo pure tone audiometry, which served as the gold standard. The scheduling procedure took place after the laboratory data collection for half of the subjects (15), whereas the other half of the subjects (15) underwent the audiologist test before they tried out the app. To reduce the possibility of either the treatment or any other factors affecting the results, the groups of subjects were counterbalanced among each other. To ensure scheduling efficiency, subjects were assisted with the scheduling.

Table 2.1: Socio-demographic information

DEMOGRAPHIC QUESTIONS	
What is your age?	
18-24 years	15
25-31 years	9
32-40 years	2
41-55 years	3
>55 years	1
Standard Deviation	10.597 (min=21, max=67)
What is your employment status?	
Employed full time	9
Employed part time	16
Self-employed	0
Not employed	5
Standard Deviation	6.75

Survey responses indicated that almost all of the subjects (28) had no previous health conditions, and one participant reported that he was 'clinically' deaf in one ear, due to which he was excluded from the analysis. In total, twenty-nine (29) subjects reported listening to music via headphones/earphones on a portable device. According to the questionnaire, 43% of the subjects (13/30) reported that they listened to their portable devices at 50-75% of the volume level, 30% (9/30) reported that they listened at a volume level of 25-50%, 13.3% of them (4/30) reported listening at 75-100% of the volume level, while 6.7% (2/30) and 3.3% (1/30) reported listening to their devices at 100% and <25% of the volume respectively. Of all the subjects, 63.3% reported that they 'sometimes' had trouble hearing normal conversation in noisy places, while 40% of the subjects reported that they 'sometimes' had trouble hearing conversation in normal (or less noisy) settings; 53.3% of the population reported that they 'sometimes' had trouble hearing when a speaker was talking softly, and 50% of the population said that they

could ‘sometimes’ hear but not understand what was being said. Of all the subjects, 56.67% of the population ‘sometimes’ had problems understanding someone if the speaker was not facing them and 50% of the population ‘sometimes’ had problems hearing on the cellphone.

Table 2.2: Survey Responses on self-reported noise exposure

Questions	Responses			
	Very Often	Sometimes	Rarely	Never
How often do you partake in activities such as going to concerts, clubs, and music festivals etc.?	7 (23.3%)	9 (30%)	13 (43.3%)	1 (3.3%)
How often do you work in a noisy environment?	5 (16.7%)	9 (30%)	15 (50%)	1 (3.3%)
How often do you listen to music on a portable device?	15 (50%)	9 (30%)	4 (13.3%)	1 (3.3%)
When listening to music on your portable device, how often do you use earbuds?	17 (56.7%)	8 (26.7%)	3 (10%)	1 (3.3%)
When listening to music on your portable device, how often do you use headphones?	3 (10%)	5 (16.7%)	13 (43.3%)	8 (26.7%)
Do you have difficulty in hearing normal conversation in crowded places?	4 (13.3%)	19 (63.3%)	5 (16.7%)	2 (6.7%)
Do you have difficulty in hearing normal conversation in less noisy settings?	1 (3.3%)	12 (40%)	9 (30%)	8 (26.7%)
Do you have difficulty in hearing when people talk softly?	4 (13.3%)	16 (53.3%)	7 (23.3%)	3 (10%)
How often can you hear but not understand what is being said?	3 (10%)	15 (50%)	11 (36.7%)	1 (3.3%)
Do you trouble hearing if someone is not facing you?	4 (13.3%)	17 (56.7%)	6 (20%)	3 (10%)
Do you have trouble hearing on the cellphone?	2 (6.7%)	15 (50%)	10 (33.3%)	3 (10%)

Equipment

All test requirements and testing procedures were approved by the Texas A&M Institutional Review Board (IRB) for the Protection of Human Subjects. Two testing locations (SPH Lab 116 and Lab 113) were used. An Extech® HD600 Sound Data Logger was used to test the room's ambient sound pressure level. Each room was tested in the beginning of the first testing period, before the subjects were allowed in. Each room had five testing periods, and hence, each testing room's ambient noise level was logged five (5) times at five different locations in the room(s). After testing the ambient noise level for each testing environment, the quietest possible region of each room was subsequently chosen as the testing area.

Table 2.3: Measurement of sound pressure level in the testing room(s) using the Extech® 600

Room	Room Sound Pressure Level Measurement (dB)				
116	45*	52	48	55	58
113	13*	21	30	33	34

The Samsung Galaxy Tab™ 3.0 (Figure 2.1), an Android device was chosen to test the hEAR mobile application, because of its adequate and comfortable 8-inch screen and brilliant display. Along with the selected Android device, Bose® AE2 headphones (Figure 2.2) were used.

Figure 2.1: Samsung Galaxy Tab

Figure 2.2: Bose AE2 Headphones

Audiologist

A local audiologist was chosen based on the OSHA requirements that an audiologist is required to satisfy in accordance with 29 CFR 1910.95 are as follows:

> "Audiometric tests shall be performed by a licensed or certified audiologist, otolaryngologist, or other physician, or by a technician who is certified by the Council of Accreditation in Occupational Hearing Conservation, or who has satisfactorily demonstrated competence in administering audiometric examinations, obtaining valid audiograms, and properly using, maintaining and checking calibration and proper functioning of the audiometers being used. A technician who operates microprocessor audiometers does not need to be certified. A technician who performs audiometric tests must be responsible to an audiologist, otolaryngologist or physician." (OSHA, 29 CFR 1910.95(g)(3))

Procedures

The hEAR application works in accordance to Appendix C of 29 CFR 1910.95 (Monitoring of Occupational noise exposure). The sounds utilized in the testing are

calibrated in 1dB increments. Each test involved touching the device screen to begin the testing procedure, after which the application produced high and low frequency sounds, which were repeated randomly 4 times. In accordance with the WHO best-practices guidelines, tests for each frequency started at 45dB and decreased at the appropriate slewing rate. Subjects were instructed to maintain contact with the screen until they were unable to hear the sound. The subjects tested the app on the Samsung Galaxy device with the selected headphones so that the use of the instruments was standardized among the subjects. Each subject underwent at least 28 'mini-trials', the frequencies range from 125 Hz to 8000 Hz and the frequencies were fairly evenly distributed between the 28 mini-trials; the entirety of one trial ran for 15-20 minutes. The mini trials were administered randomly to the subject by the device. Each mini-trial lasted for 27-33 seconds in accordance with requirement D of Appendix C of 29 CFR 1910.95. There were measures in place to account for the possibility of a missed trial. A missed trial was defined as the result of accidentally letting go or tapping the screen before the end of a previous or on-going trial. The application accounted for a false positive/false negative scenario by adding an extra mini-trial for every 'missed' trial. In total, there were 3 complete trials (Trial 1, Trial 2 and Trial 3) and one audiologist administered pure-tone audiometric test (Trial 4) per subject. Testing procedures were carried out in the laboratories to meet the requirements of Appendix D of 29 CFR 1910.95, i.e. "the background sound pressure levels exceeding the values given in the table D1 (OSHA, 29CFR 1910.95, Appendix D)". Figure 2.3 represents the testing procedures.

Table 2.4: Max. allowable octave-band sound pressure levels for audiometric test rooms

Octave-band center					
Frequency (Hz)	500	1000	2000	4000	8000
Sound pressure level (dB)	40	40	47	57	62

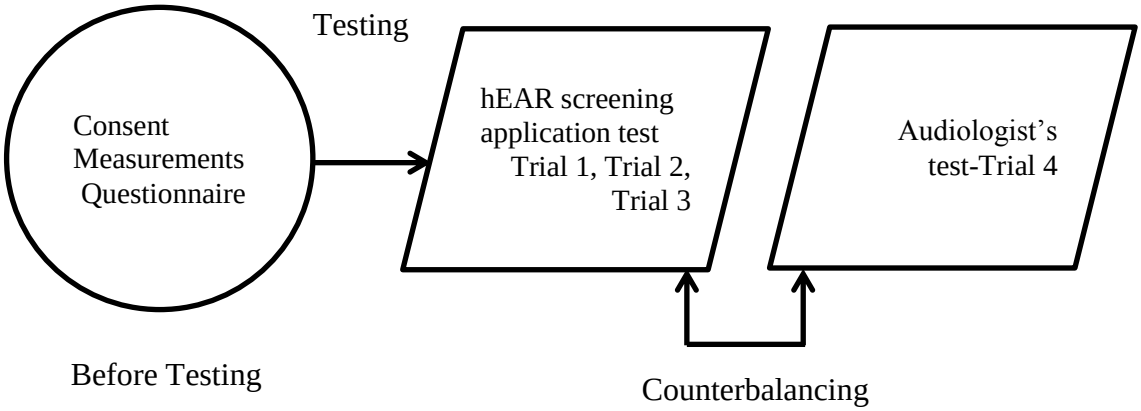

Figure 2.3: Procedure for testing hEAR application

Hypotheses

The analyses tested the following null hypotheses:

- H_0-1: There is no statistically significant difference between the results of audiologist-collected data, and the data collected with the hEAR app.
- H_0-2: There is no statistically significant learning curve between the first and the second trials, and the first and the third trials, as measured by the hEAR app.
- H_0-3: There is no statistically significant difference(s) between the results collected in a "noisy" environment or a "quiet" testing environment when compared to audiologist pure-tone data.

Statistical Analysis

In our model, Y (dependent variable) was the sound pressure level observations of the subjects (spldb), which were tested at 4 levels (trials) the levels being trial 1, 2, 3 and 4; and were a function of 7 repeated measures (frequencies) (independent variables). The effect of ear side and the individual subjects were assumed to be random effects. Therefore, we arrived at the following expression for our model:

*Y(spldb) = f (trial2 trial3 trial4 frequency1 frequency2 frequency3 frequency4 frequency5 frequency6)***

[**=Frequency 1=125 Hz, frequency 2= 250 Hz, frequency 3= 500 Hz, frequency 4= 1000 Hz, frequency 5= 2000 Hz, frequency 6= 4000 Hz, and frequency 7= 8000 Hz]

The model was then modified based on our hypotheses. To test our first null hypothesis, H_0-1 (there is no statistically significant difference between the results of audiologist-collected data, and the data collected with the hearing app); we assumed trial 1 to be the reference trial which was compared to the other trials at the reference frequency of 8000 Hz. The reference trial at reference frequency of 8000 Hz was then further compared to trial 4 at all other frequencies (interactions). The same process was repeated to compare trial 2 and trial 3 to trial 4. Therefore, we used the following model:

*Y(spldb) = f {(trial2 trial3 trial4 frequency1 frequency2 frequency3 frequency4 frequency5 frequency6) (trial 4*frequency1 trial 4*frequency2 trial 4*frequency3 trial 4*frequency4 trial 4*frequency5 trial 4*frequency6)}****

[trial 4*frequency1= audiologist's trial at frequency 125 Hz and so on]

To test our second null hypothesis, H_0-2 (there is no statistically significant learning curve between the first and the second trials, and the first and the third trials, as measured by the hearing app), the comparison was done between hEAR trials. For example, for the comparison between trials 1 and 2, our model was the following:

*Y(spldb) = f {(trial2 trial3 trial4 frequency1 frequency2 frequency3 frequency4 frequency5 frequency6) (trial 2*frequency1 trial 2*frequency2 trial 2*frequency3 trial 2*frequency4 trial 2*frequency5 trial 2*frequency6)}*****

[trial2*frequency1= trial 2 at frequency 125 Hz and so on]

Similar models were used for the comparison between trials 1 and 3 and trials 2 and 3.

To test our third null hypothesis, H_0-3 (there is no statistically significant difference(s) between the results collected in a "noisy" environment or a "quiet" testing environment when compared to audiologist pure-tone data), the effect of room was

assumed to be another fixed effect, and therefore, the models were sorted by their respective rooms,i.e. room 113 and room 116. The chosen α level was 0.05, and the 'Mixed' command was chosen to be run on SAS® statistical software, which performs mixed model analysis and repeated measures analyses 'by way of structured covariance models', where the default fitting method 'maximizes the restricted likelihood of the data under the assumption that the data are normally distributed and any missing data are missing at random'.

CHAPTER III

RESULTS

Sound Pressure Level Response Results

With respect to analysis of individual frequencies (125 Hz, 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz & 8000 Hz), the overall means per frequency were calculated for all subjects, as well as means per frequency per room. They are listed in Tables 3.1 and 3.2 below. It should be noted that results from the audiologist's test had missing values for the 125 Hz tests.

Table 3.1: Summary statistics for Sound Pressure Level (SPL) for Room 113 (*denotes missing values)

Frequency (Hz)	Trial 1 Mean SPL(db)	Trial 2 Mean SPL(db)	Trial 3 Mean SPL(db)	Audiologist's trial Mean SPL (dB)
125	18.9	18.05	17.4333	XX*
250	15.3833	13.9	11.9	9.25
500	10.95	10.0833	10	10.333
1000	10.6	9.46667	8.63333	10.916
2000	14.8	13.0667	12.3	10.75
4000	18.7	19.2667	17.7167	8.883
8000	17.9833	17.0833	17.0833	8.333

Table 3.2: Summary statistics for Sound Pressure Level (SPL) for Room 116 (*denotes missing values)

Frequency (Hz)	Trial 1 Mean SPL(db)	Trial 2 Mean SPL(db)	Trial 3 Mean SPL(db)	Audiologist's trial Mean SPL (dB)
125	28.2333	28.1833	25.6667	XX*
250	25.2667	25.2667	23.4	9.25
500	22.6	22.6	21.1333	10.333
1000	20.5333	20.5333	19.3167	10.916
2000	19.3833	19.3833	17.9833	10.75
4000	20.9167	20.9167	20.9	8.883
8000	21.1	19.4667	19.5	8.333

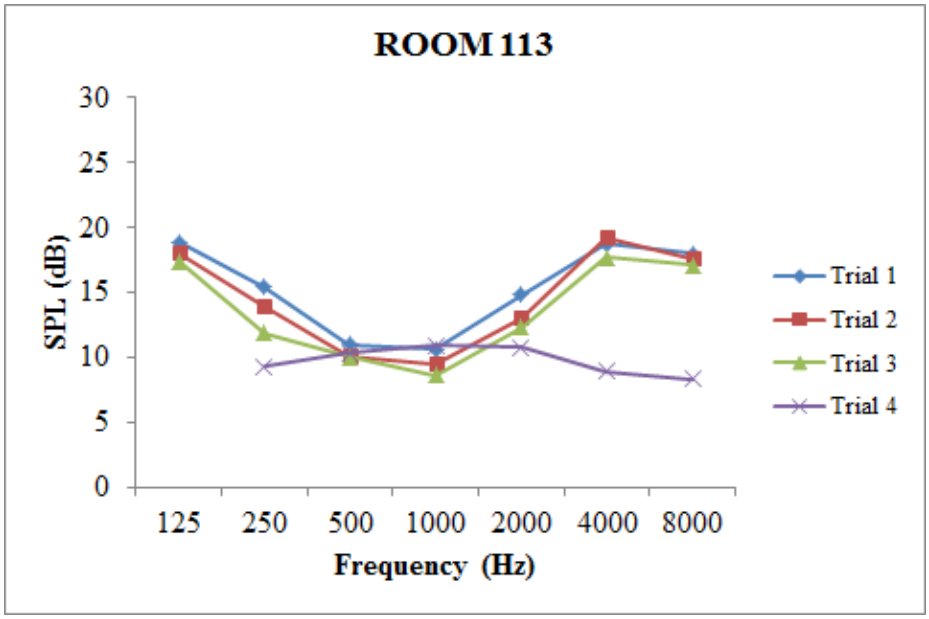

Figure 3.1: Frequency means for the sound pressure level observations for three trials for room 113 and the audiologist trial (125 Hz for audiologist missing data)

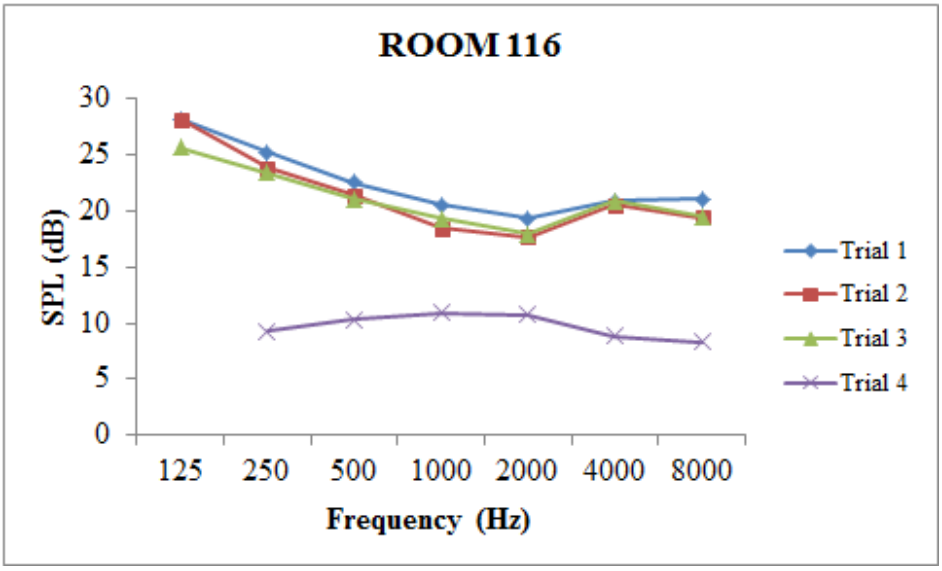

Figure 3.2: Frequency means for the sound pressure level observations for three trials for room 116 and the audiologist trial (125 Hz for audiologist missing data)

Figures 3.1 and 3.2 correspond to the overall means of the trials at the different test frequencies (125 Hz, 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz and 8000 Hz), as differentiated by rooms. There are clear differences between our trials and that of the audiologist, especially for room 116. Though our response SPL measurements correspond more or less to a normal hearing range, they are still on the higher side, especially in room 116, as compared to corresponding measurements from the audiologist's trial.

Repeated Measures ANOVA Results

After conducting a Repeated Measures ANOVA at 95% confidence intervals ($\alpha = 0.05$), we obtained the following results. It was observed that trials 2 and 3 were parallel to trial 1 throughout the frequencies (no interaction). The mixed model that was used to analyze the results compared trial 1 (reference trial) at a reference frequency of 8000 Hz (descending order of frequency). The results were separated by room, as it was found that the room where testing was conducted had a significant impact on the measurements.

Table 3.3: ANOVA test results between trial 1 and audiologist trial (trial 4), for room 116 (Freq1: 125 Hz, Freq2: 250 Hz, Freq3: 500 Hz, Freq4: 1000 Hz, Freq5: 2000 Hz, Freq6: 4000 Hz)

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	18.4286	2.1956	14	8.39	<.0001
Trial2	-0.8357	0.3828	766	-2.18	0.0293
Trial3	-1.7500	0.6530	766	-2.68	0.0075
Trial4	-10.0952	2.9880	766	-3.38	0.0008
Freq1	0.5611	2.0939	766	0.27	0.7888
Freq2	-3.8389	1.6047	766	-2.39	0.0170
Freq3	-7.2222	1.7511	766	-4.12	<.0001
Freq4	-8.0000	1.5656	766	-5.11	<.0001
Freq5	-4.1778	1.6112	766	-2.59	0.0097
Freq6	0.9944	0.6147	766	1.62	0.1061
Trial4*Freq1	0	.	.	.	.
Trial4*Freq2	5.1722	2.5856	766	2.00	0.0458
Trial4*Freq3	8.8889	2.8974	766	3.07	0.0022
Trial4*Freq4	10.0000	2.5117	766	3.98	<.0001
Trial4*Freq5	6.3444	2.1956	766	2.89	0.0040
Trial4*Freq6	0.6722	3.0524	766	0.22	0.8258

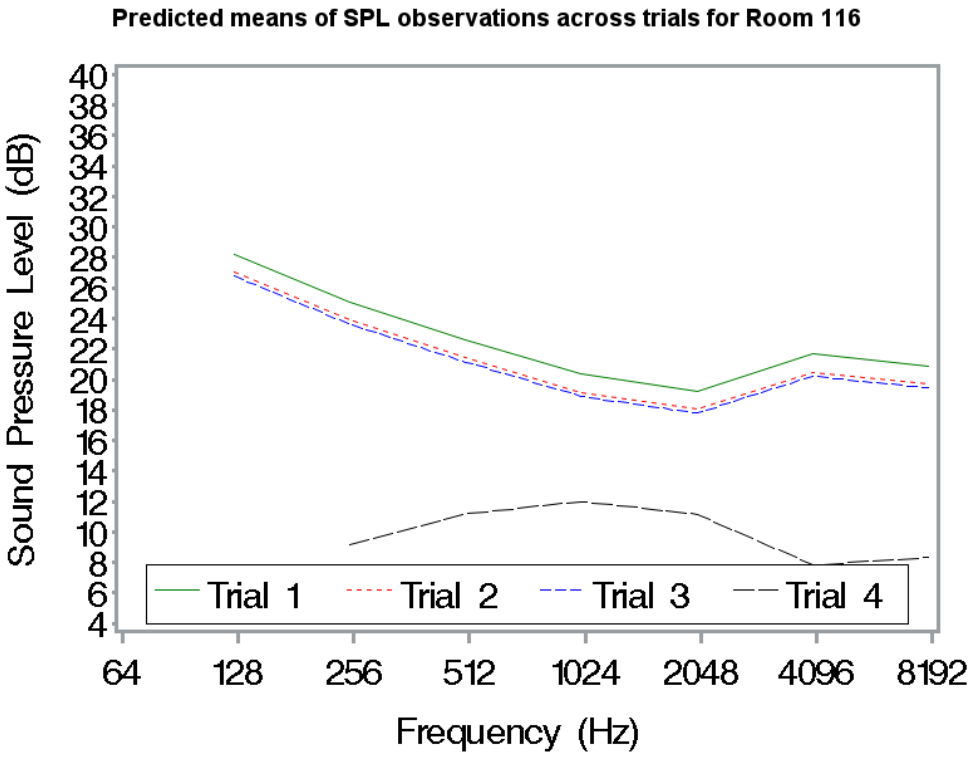

Figure 3.3: Predicted means ANOVA results between trial 1 and audiologist's trial (trial 4) for room 116

For room 116, there were statistically significant differences between trial 1 and trial 4 at all frequencies except 4000 Hz (p=0.4129)(Figure 3.3); and similar results were obtained for the comparison between trials 2 & 4 and 3 & 4, i.e. statistically significant differences at all frequencies except 4000 Hz.

Table 3.4: ANOVA test results between trial 1and audiologist trial (trial 4) for room 113 (Freq1: 125 Hz, Freq2: 250 Hz, Freq3: 500 Hz, Freq4: 1000 Hz, Freq5: 2000 Hz, Freq6: 4000 Hz)

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	17.9014	1.6449	13	10.88	<.0001
Trial2	-1.3316	0.5940	714	-2.24	0.0253
Trial3	-1.6224	0.7912	714	-2.05	0.0407
Trial4	-11.8299	2.1278	714	-5.56	<.0001
Freq1	9.7024	1.7284	714	5.61	<.0001
Freq2	6.2976	1.7739	714	3.55	0.0004
Freq3	3.6845	1.2439	714	2.96	0.0032
Freq4	0.5060	1.1741	714	0.43	0.6667
Freq5	-0.5833	0.9671	714	-0.60	0.5466
Freq6	0.8333	0.5584	714	1.49	0.1360
Trial4*Freq1	0	.	.	.	.
Trial4*Freq2	-3.4405	2.3415	714	-1.47	0.1422
Trial4*Freq3	0.9583	2.0267	714	0.47	0.6365
Trial4*Freq4	4.3155	1.9897	714	2.17	0.0304
Trial4*Freq5	3.7976	2.4349	714	1.56	0.1193
Trial4*Freq6	-1.0119	2.3684	714	-0.43	0.6693

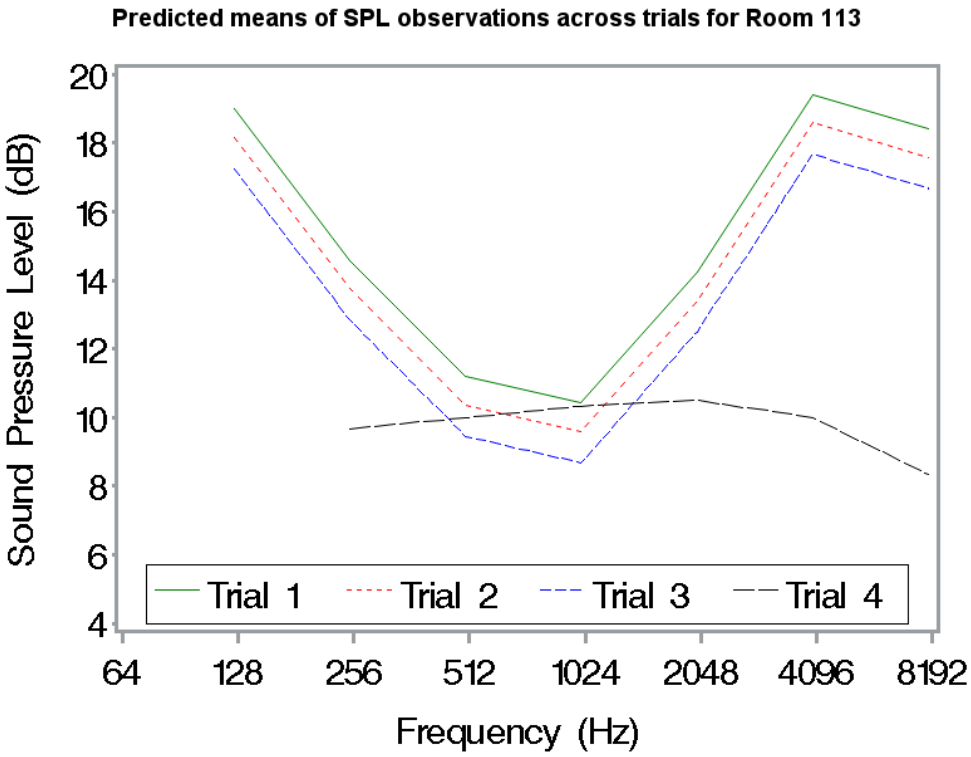

Figure.3.4: Predicted means ANOVA results between trial 1& 4 (audiologist trial) for room 113

For room 113(Figure 3.4), when trial 1 was compared to the audiologist's trial, there were statistically significant differences only for frequencies 1000 Hz and 8000 Hz; similar results were obtained for the comparison between trials 2 & 4 and trials 3 & 4, i.e. statistically significant differences at 2000 Hz (p= 0.0151) and 8000 Hz (p=<0.00001). The analysis returned significantly different results for 125 Hz; however, the missing data from the audiologist trials renders these results null.

Additionally, to examine if there was a learning curve effect between the trials, results from each trial was compared against each other. Results indicated statistically significant differences between trials, as can be interpreted from Figures 3.3 and 3.4, i.e. there was a statistically significant learning curve between trial 1 and trial 2, between trials 1 and 3, and between trials 2 and 3. For room 116, there were statistically

significant differences between trial 1 and trial 2 at frequencies of 125 Hz (0.045) and 1000 Hz (p=0.015) (Appendix E table 1); while for the same room, there were minimal statistically significant differences between trial 1 and trial 3 at 125 Hz (p=0.019) (Appendix E Table 2). However, there was only a statistically significant difference at one specific frequency, 125 Hz (p=0.019) between trials 2 and 3 (Appendix E Table 3) for room 116. For room 113, there were statistically significant differences between trial 1 and trial 2 (Appendix E Table 4) at all frequencies except 250 Hz (p=0.20) and 4000 Hz (p=0.166); while there were statistically significant differences between trial 1 and trial 3, for room 113 (Appendix E Table 5), at all frequencies except 4000 Hz (p=0.44) and 8000 Hz (p=0.45), and statistically significant differences at all frequencies except 4000 Hz (p=0.45) between trial 2 and trial 3 (Appendix E Table 6) for room 113.

The results from the analysis depict that the effect of the room is statistically significant on the observations. There is a difference between lab 113 and lab 116 (p=0.0012). Hence, we reject our null hypothesis (H_0-3) that there is no difference between the rooms.

CHAPTER IV

DISCUSSION, LIMITATIONS AND CONCLUSION

Discussion

Overall, our trials are statistically significant from that of the audiologist based on overall mean analysis. It is interesting to note, that the differences between the two trials, i.e. trial 1 and trial 4 are very clearly portrayed at specific frequencies, especially both at the lower and upper end of the frequency spectrum. Since our trials were observed by the statistical model to be parallel to each other, the pattern of the ANOVA results for trial 1 and 4 for both rooms was followed by trial 2 & 4, and trial 3 & 4, for both rooms. Due to this, we rejected our first null hypothesis (H_0-1). Our model also measured the difference between our trials, i.e. we analyzed if trial 1, 2 and 3 were statistically different among each other. Our results indicate that the trials differ amongst each other at specific frequencies. There is presence of a minimal learning curve effect between trial 1 and 2, and between trial 1and 3, and trial 2 and 3 for room 116. Because of the presence of a learning curve effect, we rejected our null hypothesis (H_0-2). The pattern of the learning curve between the two rooms is very interesting to note. It could be argued that over the time of a full and complete individual test (with three trials) in the relative absence of any ambient noise, the subjects were more likely to have a more 'focused' account for their responses during the trials and hence the statistically different responses between the three trials at almost all frequencies, whereas in the comparatively noisy room, the overall ambient noise could have probably negated any such effect.

For Lab 116, our trial measurements statistically differ from those of the audiologist trial, at almost all frequencies except for 4000 Hz, while for Lab 113, the measurements differ from those of the audiologist trial only at 2000 Hz and 8000 Hz. There could be several reasons for this. Firstly, for the audiologist trials, all the sound pressure level measurements at 125 Hz for all subjects were missing, indicating a potential issue with this data from the audiologist. These results effectively make that frequency null for the statistical tests, hence the statistically significance at 125 Hz for

both rooms. Also, the ambient noise in the testing rooms, especially in Lab 116, acted like a probable confounder, and may be responsible for the statistically significant differences at all other frequencies. Lab 113, which had lower ambient noise levels, had individual test frequency measurements that were not statistically different than the audiologist test. Based on these results, it is the conclusion based on hypothesis 3 (H_0-3), that testing environment has a significant effect on overall quality of data collection. We can see that the results for frequency 8000 Hz are inconsistent between the two rooms. This could be due to the headphones that were used, the confounding effect due to the ambient noise in the testing rooms, or a combination of both. The headphones were not noise cancelling, and were not optimized for reproducing sounds at higher frequency spectrum. To test the rooms' inherent combined SPL levels, no octave band spectral analysis was done, and hence, it is difficult to pinpoint the major cause of the disparity between results at higher frequencies.

Our results clearly indicated that the room that the tests were conducted in had a significant impact on the measurements. The conclusion is seen in the results comparison between rooms. Indications are that the higher the room's combined SPL levels, the significantly different the results are from the control data (audiologist-collected pure tone). This effect is well documented (OSHA 29 CFR 1910.95) and is commonly encountered while conducting mobile hearing tests, such as those conducted in mobile hearing booths etc. Due to this, both OSHA and ANSI have conducted extensive research and published guidelines to account for a room's combined SPL levels. These guidelines are tabulated in Tables 2.4 and 4.1. Another thing to note in our output would be the absence of a noise notch at 4000 Hz for the audiologist data. One explanation of this could be that noise notches are highly likely to be unilateral as proven by several studies (Wilson & McArdle, 2013). Also, the population expected hearing response for the audiologist-administered test shows a constant trend. Most of the hearing related research usually does not analyze an overall mean analysis or a "biaural population average" (Prince, Stayner, Smith, & Gilbert, 1997). However, NIOSH had used the approach (Figure 4.1) for conducting the Occupation Noise and Hearing Survey

(1968-1972), and subsequently the results were utilized by ANSI to formulate the standards for Occupational Noise Exposure. This study was also a "binaural population average" where the mean age range of the population was 28.9, and the population hearing response coincided with the second response curve in Figure 4.1.

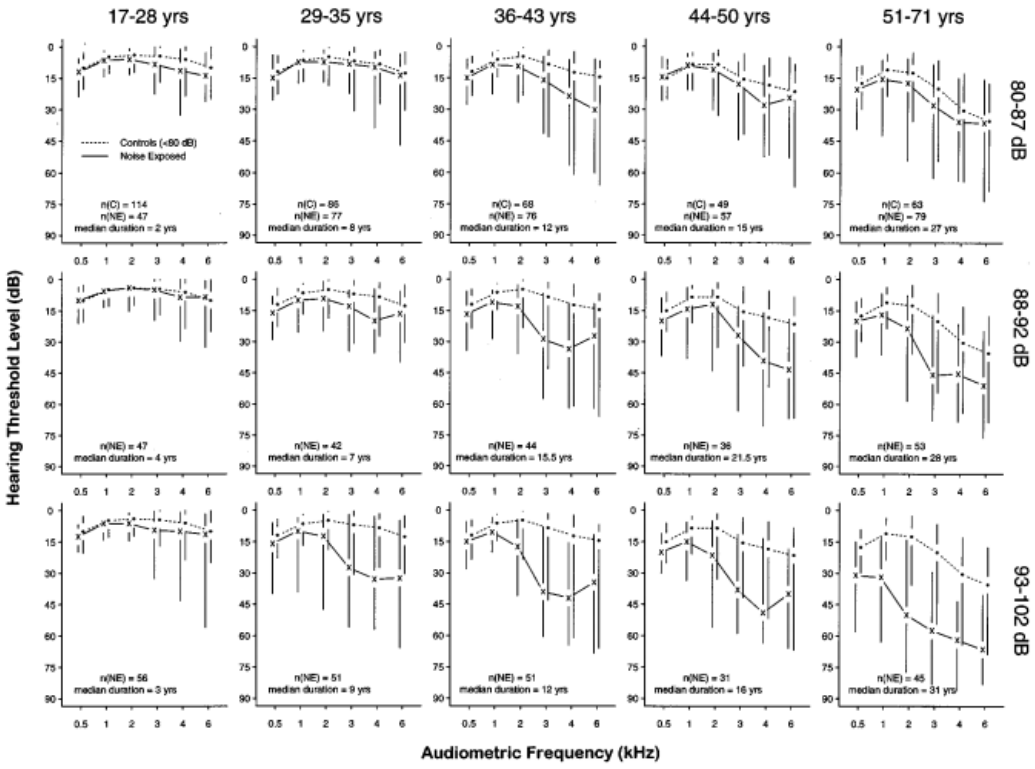

Figure 4.1: Biaural population mean response curve as recorded by NIOSH for the Occupational Noise and Hearing Survey of 1968-1972.

Limitations and Confounders

This study is the first of its kind, and therefore, had certain limitations and confounders that could not be accounted for. This study requires further research, and it is imperative that the study design be conducted after a thorough octave band spectral analysis of the testing areas to accurately account for background ambient noise, which was one of our biggest confounders. As an octave band spectral analysis was not

conducted, confounding due to background ambient noise or headphone acoustics could not be confidently determined. Researchers who have assessed self-recording hearing tests have noted that most such tests have confounding due to these factors. Masalski and Krecicki (2013) conducted a study where they validated a web-based pure tone self-test, they found that the test sound pressure level observations were "greatly exaggerated" with respect to pure tone audiometry, but were still under the normal ranges of hearing thresholds, and therefore they recommended using the web based test in conjunction with a pure tone test, and not by itself.

Ordinary, meant for day-to-day use headphones were chosen to conduct this study, as noise cancellation headphones may not have offered a realistic reproduction of a work/occupational environment. There have been documented issues with the use of 'ordinary' headphones for self-administered hearing tests. Ferrari et al. (2013) observed a certain degree of variability in the sensitivity and specificity of the TS audiometer when different headphones were used. Similarly, in a study conducted by Choi, Sohn, Ku, Kim and Lee (2013), they observed different results with different sets of headphones while testing their phoneme based testing application. To control for such effects, they conducted their tests in a sound-proof booth, and advised the usage of ear protection muffs in the absence of a sound-proof booth/room.

Another of our important limitations was our inability to use results for 125 Hz in the audiologist's test as all the data at that frequency was missing. Also, the audiologist used other uncommon frequencies during the pure tone test (namely 750 Hz, 1500 Hz, 3000 Hz, and 6000 Hz) which were not accounted for in our application. It would be useful to repeat the experiment with these frequencies, even though literature suggests against using 6000 Hz as a test frequency because of a higher probability of false positives at that particular frequency.

Future studies should focus also on limitation of subject headphone use or exposure to loud noises for a set time period prior to data collection. OSHA recommends no occupational noise exposure for a 14-hour period prior to data collection

(29 CFR 1910.95(g)(5)(iii). Future studies, to ensure minimization of temporary threshold shifts, should include a similar limitation.

Background ambient noise in testing rooms

Typically, if we assume that the testing rooms were soundproof, high SPL levels as assessed by an audiometric test of any kind, would indicate a certain degree of hearing loss in the respondents (-10 dB to 20 dB is normal hearing range, whereas the higher the SPL, the more the severity of hearing loss).

However, background noise in audiometric testing rooms is a concern to the testers. American National Standard (ANSI S3.1-1999; Table 4.1) has defined the acceptable ambient noise levels and the associated errors in the measurements they (ambient noise) create. The standard is based on several objective measurements and includes detailed options that allow for adjustment of the tabulated values. OSHA (Table 2.4) also has recommendations for the background noise levels.

Table 4.1: ANSI S3.1-1999 Maximum allowable octave band sound pressure levels for audiometric test rooms.

Center Frequency (Hz)	Octave-band levels (dB)		
	125 to 8000 Hz	250 to 8000 Hz	500 to 8000 Hz
125	29	35	44
250	21	21	30
500	16	16	16
1000	13	13	13
2000	14	14	14
4000	11	11	11
8000	14	14	14

As is evident, the OSHA values are 13-25 dB higher than the ANSI standard values. Most research done on this topic indicates that the frequency that gets masked due to high background noise levels is 500 Hz (Frank & Williams, 1994). However, we found that there was appropriate response (even after accounting for background noise) for that particular frequency. This was possibly due to good frequency reproduction in that frequency range of the headphones that were used to test subjects.

The ambient noise in the testing rooms was measured and the results were respectively between 13-34 dB for room 113, and 45-58 dB for room 116. As can be seen in Tables 3.1 and 3.2, some of the sound pressure level observations as assessed by the hearing app were on the higher side and these higher SPL levels corresponding to higher frequencies. This could possibly due to the high ambient noise levels in the rooms. The testing environment's inherent sound pressure levels are also potentially responsible for the difference in the test SPLs responses (as there is a difference between the test SPLs in room 113 and the test SPLs in room 116, the test SPLs in room 116 being higher), which in turn is the cause of rejecting our third null hypothesis. As is evident by the ANSI and OSHA tables, the noise levels in room 113 correspond to both the recommendations, while the noise levels in room 116 corresponded to OSHA recommendations, but not to the more stringent ANSI recommendations. One way of combatting this issue would be to either use a sound proof/insulated room to measure the subjects' responses (SPL); or use professional quality noise cancellation headphones to compensate for the ambient noise level of the testing rooms. Additionally, rooms could have been tested with an octave band analyzer to test different frequency spectrum SPLs. This data would help account for any specific frequency-related differences between the test scenario and the audiologist-collected data.

Another limitation that we encountered was not accounting and subsequently testing for other not-so commonly used frequencies, namely 750 Hz, 1500 Hz, 3000 Hz, and 6000 Hz. The pure tone audiometric test that was administered by the audiologist for this exercise did use those frequencies, in addition to the other commonly used ones (125 Hz, 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz, 8000 Hz). The recommendation for an

improved test via the hearing application would be to use these other frequencies, in addition to the ones it already accounted for. That may improve the robustness of the experimental design and could potentially result in more reliable results.

Data Outliers

The data that we collected was normally distributed, however, as with all data; our data also had outliers as seen in Tables 3.1 and 3.2. These were sound pressure level (spldb) observations that were higher than normal. Most of these outlying observations could be explained in correspondence to the rooms where they were observed (an overall higher sound pressure level in room 116); however, even then, there were some observations that stood out. As iterated earlier, most of our participant population had no prior history of any hearing related problems. However, two of our subjects were alerted of mild hearing loss (for one it was age related, and for the other participant it was 'occupation' related, as that participant was a part-time Disc Jockey or DJ). One of our participants was clinically deaf in one ear, whose observations were subsequently excluded from statistical analysis. All these subjects had participated in the Pure Tone Audiometry Test, and then subsequently, participated in the app trials. Their observations (spldb) on these occasions were correspondingly higher as compared to the similar ones by other subjects. The observations from these subjects corresponded to high sound pressure levels in response to the tested frequencies, and these were therefore the outliers.

Headphone Acoustics

We used Bose® AE2 over ear headphones for this study. These headphones are from the Quiet Comfort Line of headphones, where the main focus is comfort for the user and distortion free audio performance. As a result, these headphones offer no noise cancellation; however, these do afford the listener "a pleasant frequency response with a deeper, sculpted sound signature". Also, unlike a lot of the similar products in the same price range, these headphones' low frequency response is more powerful than their high frequency response. These headphones were built to emphasize upon bass and sub-bass

level frequencies. As a result of this characteristic, these headphones were found to not be able to effectively reproduce mid and high frequency sounds (2500 Hz and above) to offer strong reliability in those ranges, as found out by a study conducted by Inner Fidelity in 2012. The headphones were fairly competent at reproducing low end frequency sounds, studies have shown that they offer almost no attenuation at those frequencies; which could be responsible for significant differences in the low frequency test results in Lab 116, and for some in Lab 113. Overall, the headphones are robust and "good for daily use", however, since the headphones offer somewhat unreliable reproduction of higher frequencies, it would be prudent to perhaps try to use other headphones, for example, ones that offer noise cancellation, or headphones (such as those used by professional DJs) that offer a better high frequency spectrum reproduction, and repeat the experiment.

Conclusion

The data collected by the application is different from that collected by the audiologist. As much as we would have liked to have perfect congruence between the two, our data is a relative representation of 'real life situations' in an occupational cohort. For this study, the results for the 'quiet' environment are promising as they deviate least from the audiologist data. This would imply that under slightly different conditions, such as a more representative sample of an occupational cohort, better hardware in terms of headphones and/or better environmental analysis with respect to ambient noise; the app would be able to successfully identify/diagnose hearing loss to some extent.

This however, does not mean that the results for the 'noisy' environment are completely invalid. In fact, such an environment would offer a more representative view of workplace noise conditions. Therefore, with slightly different conditions, especially octave band spectral analysis of ambient noise, and its subsequent countermeasure in the test procedure (such as a mobile hearing test location like a van), it is possible to have more realistic results.

Further studies that are formulated with the appropriate countermeasures to the above mentioned limitations, and large sample size that is more representative of an occupational cohort, should result in hEAR being released as a suitable and competitive alternative to an audiologist-administered pure tone hearing test, especially for rural occupational cohorts, and consequently, in an easy-to-follow hearing screening occupational programs.

REFERENCES

Burkey, J. M., Lippy, W. H., Schuring, A. G., & Rizer, F. M. (1998). Clinical utility of the 512-Hz Rinne Tuning Fork Test. *American Journal of Otology, 19*(59), 59-62.

Choi, J.M., Sohn, J., Ku, Y., Kim, D., & Lee, J. (2013). Phoneme-based self hearing assessment on a smartphone. *IEEE Journal of Biomedical and Health Informatics*, 17(3), 526-529.

Hamou, A., Guy, S., Lewden, B., Bilyea, A., Gwadry-Sridhar, F., & Bauer, M. (2010). Data collection with iPhone Web Apps. *Institute of Electrical and Electronics Engineers.*, 235-239.

Dalton, D. S., Cruikshank, K. J., Klein, B. E., Klein, R., Wiley, T. L., & Nondahl, D. M. (2003). The impact of hearing loss on quality of life in older adults. *The Gerontologist, 43*, 661-668.

Daniel, E. (2007). Noise and hearing loss: A review. *Journal of Social Health, 77*(5), 225-231.

Ferrari, D. V., Lopez, E. A., Lopes, A. C., Aiello, C. P., & Jokura, P. R. (2013). Results obtained with a low cost software-based audiometer for hearing screening. *International Archives of Otorhinolaryngology*, 257-264.

Frank, T., & Williams, D. L. (1994). Ambient noise levels in industrial audiometric test rooms. *American Industrial Hygiene Association Journal, 55*(5), 334-337.

Franks, J. R. (1995). Hearing Measurement. In W. H. (WHO), *Occupational exposure to noise: Evaluation, Prevention, and Control* (pp. 183-232). Geneva, Switzerland: WHO.

Gates, G. A., Murphy, M., Rees, T. S., & Fraher, A. (2003). Screening for handicapping hearing loss in the elderly. *The Journal of Family Practice, 52*(1), 56-62.

Leensen , M. C., de Laat, J., & Dreschler, W. A. (2011, June 3). Speech-in-noise screening tests by internet, Part 1: Test evaluation. *International Journal of Audiology, 50*, 823-834.

Li, Y.-F., & Barker Bausell, R. (2006). *Power Analysis for Experimental Research: A Practical Guide for the Biological, Medical and Social Sciences.* New York: Cambridge University Press.

Martinez, L. (2012). Can you hear me now? Occupational hearing loss, 2004-2010. *Monthly Labor Review*, 48-55.

Masalki, M., & Krecicki, T. (2013). Self-Test Web-Based Pure-Tone Audiometry: Validity Evaluation and Measurement Error Analysis. Journal of Medical Internet Research, 15(4).

Meinke, D. K., & Dice, N. (2007). Comparison of audiometric screening criteria for the identification of noise induced hearing loss in adolescents. *American Journal of Audiology, 16*, S190-S202.

Mulrow, C. (1991). Screening for Hearing Impairment in the Elderly: Rationale and strategy. *Journal of General Internal Medicine, 6*, 249-257.

Mulrow, C. D., & Lichtenstein, M. J. (1991). Screening for hearing impairment in the elderly. *Journal of General Internal Medicine*, 249-258.

OSHA. (2008). Audiometric Measuring Instruments. *29 CFR 1910.95: Occupational Noise Exposure*. Washington DC.

OSHA. (2008). Audiometric Test Rooms. *29 CFR 1910.95: Occupational Noise Exposure*. Washington DC.

Prince, M. M., Stayner, L. T., Smith, R. J., & Gilbert, S. J. (1997). A re-examination of risk estimates from the NIOSH Occupational Noise and Hearing Survey (ONHS). *The Journal of the Acoustical Society of America, 101*(2).

Schow, R. (1991). Considerations in selecting and validating an adult/elderly hearing screening protocol. *Behavioral Audiology., 12*(5), 337-348.

Schow, R. L., Smedley, T. C., & Longhurst, T. M. (1990). Self-assessment in adult/elderly hearing screening-recent data and new perspectives. *Ear and Hearing.*, 17S-27S.

WHO. (2011). *Mobile phone use: A growing problem of driver distraction.* Cincinnati: World Health Organization.

Wilson, R. H., & McArdle, R. (2013). Characteristics of the audiometric 4,000 Hz notch (744,553 veterans) and the 3,000, 4,000 and 6,000 Hz notches (539,932 veterans). *Journal of Rehabilitation Research & Development, 50*(1), 111-132.

Yueh, B., Shapiro, N., MacLean, C. H., & Shekelle, P. G. (2003). Screening and management of adult hearing loss in primary care. *Journal of American Medical Association, 289,* 1976-1985.

APPENDIX A: APPROVED CONSENT FORM

Consent Form for Participation in a Research Study
Texas A&M University

Evaluation of hEAR mobile application as an effective alternative to audiometry

Description of the research and your participation

You are invited to participate in a research study conducted by Dr. Adam W. Pickens, a researcher from Texas A&M University School of Public Health and Lakshmi V. Dakuri. The purpose of this research is to assess the validity of a mobile hearing screening application in collecting quality hearing screening data. The information in this form is provided to help you decide whether or not to take part. If you decide to take part in the study, you will be asked to sign this consent form. If you decide you do not want to participate, there will be no penalty to you, and you will not lose any benefits you normally would have.

Why is this study being done?
The purpose of this study is to analyze the effectiveness of a mobile hearing screening application at collecting hearing screening data.

How many people will be asked to be in this study?
The research team recruited via email and fliers locally at Texas A&M University. It is the intent of the research team to recruit a total of approximately 22 total study participants.

What are the alternatives to being in this study?
The alternative to being in the study is not to participate. If you choose not to participate, you will not receive any penalty or lose any benefits you normally would have.

What will I be asked to do in this study?
Your participation will involve testing the mobile application on an Android tablet. There will be three (3) trials per person. Each trial has 28 mini-trials, which run for 27-33 seconds each. Each trial, therefore, runs for 15-20 minutes each. Breaks of 10-15 minutes will be given after the completion of each trial, and as and when requested by the participant. Additionally an appointment with a certified audiologist in the Bryan/College Station area will be scheduled for you at your convenience by the investigators. Before the day of the appointment, participants will visit SPH to sign the consent form. The audiometry test at the audiologist will similarly be 1-2 hours. At the appointment you will be asked to participate in a pure tone audiometric test. The test will

be employ frequencies similar to the mobile application, to test participants' hearing. The total study duration will therefore be 3 days. All expenses regarding the test will be borne by the research team.

Testing Locations
TAMHSC School of Public Health, Department of Environmental and Occupational Health, Laboratory Building Room No. 116, Adriance Lab Road, Raymond Stotzer Parkway, College Station, TX 77843-1266.
Listen Hear Audiology Center, 3091 University Drive East, Suite 410, Bryan, TX 77802
.

Risks and discomforts

The potential harm of the testing procedure, if any, would be that the tones utilized in the trials may be a little aggravating, if at all.

Potential benefits

The benefit to the subjects would be that their hearing would be tested, and if there is a previously undiscovered hearing disability, then it can be addressed by a qualified physician at a later date.

Protection of confidentiality

The paper reports will be kept in a secure locked cabinet where only the principle investigators will have access to the lock. The paper reports will be assigned a 7-digit uniformly distributed random number that will be unique to each subject. All digital information will be encrypted and stored on the primary investigator's computer that will be password protected, the password only known to primary investigators. Participant's identity will not be revealed in any publication resulting from this study. No identifiers linking you to this study will be included in any sort of report that might be published. People who have access to your information include the Principle Investigator and research study personnel. Representatives of regulatory agencies such as the Office of Human Research Protections (OHRP) and entities such as the Texas A&M University Human Subjects Protection Program may access your records to make sure the study is being run correctly and that information is collected properly. Information about you and related to this study will be kept confidential to the extent permitted or required by law.

Voluntary participation
Your participation in this research study is voluntary. You may choose not to participate and you may withdraw your consent to participate at any time. You will not be penalized in any way should you decide not to participate or to withdraw from this study.

You will compensated with a $20 Target gift card for your time for being in this study.

Contact information

If you have any questions or concerns about this study or if any problems arise, please contact the study PI, Dr. Adam Pickens, at 979-845-0203 or by email at pickens@tamhsc.edu. You can also contact the Research Assistant, Lakshmi V. Dakuri at 979-587-8650 or by email at dakuri@sph.tamhsc.edu. If you have any questions or concerns about your rights as a research participant, please contact the Texas A&M University Institutional Review Board at 979.458.4117.

Consent

I agree to be in this study and know that I am not giving up any legal rights by signing this form. The procedures, risks, and benefits have been explained to me, and my questions have been answered. I know that new information about this research study will be provided to me as it becomes available and that the researcher will tell me if I must be removed from the study. I can ask more questions if I want. A copy of this entire form will be given to me if I so request.

Participant's signature_____ Date:_____

INVESTIGATOR'S AFFIDAVIT:

Either I have or my agent has carefully explained to the participant the nature of the above project. I hereby certify that to the best of my knowledge the person who signed this consent form was informed of the nature, demands, benefits and risks involved in his/her participation.

Signature of Presenter:_____ Date_____

APPENDIX B: APPROVED QUESTIONNAIRE

1. What is your age?

    - Under 18 years of age
    - 18-24
    - 25-31
    - 32-40
    - 41-55

2. What is the highest degree of education that you have received?

    - High school diploma
    - Associate's degree
    - Bachelor's degree
    - Master's degree
    - Post graduate degree

3. Employment status:

    - Employed Part-time
    - Employed Full-time
    - Self-employed
    - Student

4. How often do you work in a noisy environment?

    - Very often
    - Sometimes
    - Rarely
    - Never

5. How would you describe your state of physical health?

    - Excellent
    - Very good
    - Good
    - Fair
    - Poor

6. Are you currently taking any medications for the following (select that apply) :

- □ Heart disease
- □ Hypertension
- □ Diabetes
- □ Depression
- □ Insomnia

7. Have you ever been diagnosed with a hearing loss?

- □ Yes
- □ No

8. How often do you listen to music?

- □ Very Often
- □ Sometimes
- □ Rarely
- □ Never

9. What kind of music do you like to listen to (select all that apply)?

- □ Pop
- □ Jazz
- □ Classical Music
- □ Country
- □ Hip-Hop/R'n'B
- □ Rock/Heavy metal
- □ Disco/Dance music
- □ Electronic Dance Music

10. How often do you listen to music on a portable device?

- □ Very often
- □ Sometimes
- □ Rarely
- □ Never

11. While listening to music on your portable device, what approximate volume do you listen on?

- 100%
- 75-100%
- 50-75%
- 25-50%
- <25%

12. When listening to music on your portable device, how often do you use earbuds?
    - Very often
    - Sometimes
    - Rarely
    - Never
13. When listening to music on your portable device, how often do you use headphones?
    - Very often
    - Sometimes
    - Rarely
    - Never

14. How often do you partake in activities such attending concerts, going to clubs, music festivals etc.?

    - Very often
    - Sometimes
    - Rarely
    - Never

15. Do you have difficulty hearing normal conversations in crowded places?

    - Very often
    - Sometimes
    - Rarely
    - Never

16. Do you difficulty hearing normal conversations in less noisy settings?

    - Very often
    - Sometimes
    - Rarely
    - Never

17. Do you difficulty hearing when people speak softly?

- Very often
- Sometimes
- Rarely
- Never

18. How often can you hear but not understand what is being said?

    - Very often
    - Sometimes
    - Rarely
    - Never

19. Do you have trouble understanding someone if they are not facing you?

    - Very often
    - Sometimes
    - Rarely
    - Never

20. Do you have trouble hearing on the telephone?

    - Very often
    - Sometimes
    - Rarely
    - Never

APPENDIX C: RECRUITMENT EMAIL

From: Lakshmi V. Dakuri

Subject: Research Participation Invitation: Validation of Android Application to test mobile device hearing data collection

This email message is an approved request for participation in research that has been approved or by the Texas A&M University Institutional Review Board (IRB).

You are invited to participate in a thesis study that seeks to validate an Android application that tests the ability of a mobile device to collect acceptable hearing data. The study will require participants to devote a maximum of one hour (with breaks) on one day to use the application on an android device with headphones, and responding to the various prompts on the screen. This data collection will take place in a laboratory at the Texas A&M School of Public Health. In addition, the participants would be requested to go to a (pre-selected) registered audiologist in the Bryan/College Station area to undergo pure tone audiometry. The total duration spent by the participants will be 3 days. The investigation team will cover all costs for the hearing screening performed by the audiologist.

Participants must fit the following criteria:

- Must be over 18 years of age
- Must be willing to answer questions about earbud/headphone usage and cellphone usage
- Must be willing to answer questions about their hobbies (non-occupational time)
- Must be willing to undergo pure tone audiometric test by a registered audiologist
- Must not be involved in any similar studies

This project was approved by the Texas A&M University IRB. Pertinent questions or concerns about the research, research participants' rights, and/or research-related injuries to participants should be directed to the project chair, Dr. Adam W. Pickens (979-845-0203 - pickens@sph.tamhsc.edu).

If interested, please contact Lakshmi V. Dakuri (979-587-8650 – dakuri@sph.tamhsc.edu) directly for scheduling.

APPENDIX D: RECRUITMENT FLIER

Participants Needed

Study: Validation of a Mobile Application for Mobile Hearing Screening

You are invited to participate in a study evaluating the validity of data collected by a mobile hearing screening application conducted by Dr. Adam Pickens at Texas A&M School of Public Health (SPH).

The study involves one (1) approximately one-hour visit to SPH and one (1) approximately one-hour visit to a local certified audiologist, for a total of 3 study days.

If you are over 18, are willing to undergo audiometric testing, willing to answer brief questions about your activities related to noisy environments, and would like more information about participation, contact Lakshmi Dakuri at 979-587-8650 or dakuri@sph.tamhsc.edu.

APPENDIX E: ADDITIONAL DATA TABLES

Table 1: ANOVA test results between trial 1 and trial 2 for room 116

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	20.7210	3.2950	14	6.29	<.0001
Trial2	-1.2544	1.0322	765	-1.22	0.2246
Trial3	-1.4476	0.7575	765	-1.91	0.0564
Trial4	-11.7821	1.6973	765	-6.94	<.0001
Freq1	6.9528	2.6380	765	2.64	0.0086
Freq2	2.9667	2.4367	765	1.22	0.2238
Freq3	1.9889	2.1030	765	0.95	0.3446
Freq4	0.9722	1.4792	765	0.66	0.5112
Freq5	-0.1333	1.2537	765	-0.11	0.9153
Freq6	0.2389	0.9134	765	0.26	0.7937
Trial2*Freq1	1.7639	1.0409	765	1.69	0.0906
Trial2*Freq2	1.4833	0.9366	765	1.58	0.1137
Trial2*Freq3	-0.02222	0.8782	765	-0.03	0.9798
Trial2*Freq4	-1.9056	0.8855	765	-2.15	0.0317
Trial2*Freq5	-1.6000	1.0938	765	-1.46	0.1440
Trial2*Freq6	0.8278	1.3591	765	0.61	0.5427

Table 2: ANOVA test results between trial 1 and trial 3 for room 116

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	20.5571	3.2319	14	6.36	<.0001
Trial2	-1.1762	0.5744	765	-2.05	0.0409
Trial3	-1.0571	0.7079	765	-1.49	0.1358
Trial4	-11.5950	1.6957	765	-6.84	<.0001
Freq1	8.2394	2.6999	765	3.05	0.0024
Freq2	3.1500	2.3836	765	1.32	0.1867
Freq3	2.1000	2.0767	765	1.01	0.3122
Freq4	0.7222	1.4606	765	0.49	0.6211
Freq5	-0.2056	1.1102	765	-0.19	0.8532
Freq6	0.1278	1.0010	765	0.13	0.8985
Trial3*Freq1	-2.0727	0.9890	765	-2.10	0.0364
Trial3*Freq2	0.7500	0.9939	765	0.75	0.4507
Trial3*Freq3	-0.4667	0.9849	765	-0.47	0.6358
Trial3*Freq4	-0.9056	0.7101	765	-1.28	0.2026
Trial3*Freq5	-1.3111	1.0950	765	-1.20	0.2315
Trial3*Freq6	1.2722	1.4132	765	0.90	0.3683

Table 3: ANOVA test results between trial 2 and trial 3 for room 116

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	16.2717	1.4975	13	10.87	<.0001
Trial1	1.3316	0.5940	713	2.24	0.0253
Trial3	0.03189	0.7755	713	0.04	0.9672
Trial4	-9.6288	1.7011	713	-5.66	<.0001
Freq1	10.5982	1.5512	713	6.83	<.0001
Freq2	5.2321	1.3701	713	3.82	0.0001
Freq3	4.0119	1.0170	713	3.94	<.0001
Freq4	1.8214	1.0745	713	1.70	0.0905
Freq5	0.6310	0.8156	713	0.77	0.4394
Freq6	0.2321	1.0670	713	0.22	0.8278
Trial3*Freq1	-2.1161	1.0587	713	-2.00	0.0460
Trial3*Freq2	0.8214	1.0623	713	0.77	0.4396
Trial3*Freq3	-0.3512	1.0485	713	-0.33	0.7378
Trial3*Freq4	-0.9464	0.7596	713	-1.25	0.2132
Trial3*Freq5	-1.0595	1.1439	713	-0.93	0.3546
Trial3*Freq6	1.3929	1.5090	713	0.92	0.3563

Table 4: ANOVA test results between trial 1 and trial 2 for room 113

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	15.5777	1.9754	13	7.89	<.0001
Trial2	0.5651	0.9314	713	0.61	0.5442
Trial3	-1.9362	0.6725	713	-2.88	0.0041
Trial4	-4.9398	1.7247	713	-2.86	0.0043
Freq1	3.3011	2.2051	713	1.50	0.1348
Freq2	-1.2321	1.8409	713	-0.67	0.5035
Freq3	-3.0893	1.7381	713	-1.78	0.0759
Freq4	-3.6905	1.4653	713	-2.52	0.0120
Freq5	-0.9643	1.5047	713	-0.64	0.5218
Freq6	0.3274	0.8954	713	0.37	0.7147
Trial2*Freq1	-1.6761	1.0564	713	-1.59	0.1131
Trial2*Freq2	-1.3036	1.5492	713	-0.84	0.4004
Trial2*Freq3	-3.0357	1.1996	713	-2.53	0.0116
Trial2*Freq4	-3.2381	1.1288	713	-2.87	0.0042
Trial2*Freq5	-2.4643	1.2517	713	-1.97	0.0494
Trial2*Freq6	1.3869	1.4302	713	0.97	0.3325

Table 5: ANOVA test results between trial 1 and trial 3 for room 113

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	15.5159	1.9859	13	7.81	<.0001
Trial2	-0.9107	0.4027	713	-2.26	0.0240
Trial3	-0.1766	1.1643	713	-0.15	0.8795
Trial4	-4.9762	1.7419	713	-2.86	0.0044
Freq1	3.1448	2.1719	713	1.45	0.1481
Freq2	-0.7500	1.7140	713	-0.44	0.6618
Freq3	-3.3155	1.6972	713	-1.95	0.0511
Freq4	-3.6250	1.3513	713	-2.68	0.0075
Freq5	-1.0060	1.3539	713	-0.74	0.4577
Freq6	0.6369	0.8106	713	0.79	0.4323
Trial3*Freq1	-1.3056	1.2168	713	-1.07	0.2836
Trial3*Freq2	-3.2321	1.0698	713	-3.02	0.0026
Trial3*Freq3	-2.1310	1.1247	713	-1.89	0.0585
Trial3*Freq4	-3.5000	1.1298	713	-3.10	0.0020
Trial3*Freq5	-2.2976	1.2159	713	-1.89	0.0592
Trial3*Freq6	0.1488	0.9903	713	0.15	0.8806

Table 10: ANOVA test results between trial 2 and trial 3 for room 113

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	14.6052	2.0212	13	7.23	<.0001
Trial1	0.9107	0.4027	713	2.26	0.0240
Trial3	0.7341	1.0796	713	0.68	0.4967
Trial4	-4.0655	1.7222	713	-2.36	0.0185
Freq1	3.1448	2.1719	713	1.45	0.1481
Freq2	-0.7500	1.7140	713	-0.44	0.6618
Freq3	-3.3155	1.6972	713	-1.95	0.0511
Freq4	-3.6250	1.3513	713	-2.68	0.0075
Freq5	-1.0060	1.3539	713	-0.74	0.4577
Freq6	0.6369	0.8106	713	0.79	0.4323
Trial3*Freq1	-1.3056	1.2168	713	-1.07	0.2836
Trial3*Freq2	-3.2321	1.0698	713	-3.02	0.0026
Trial3*Freq3	-2.1310	1.1247	713	-1.89	0.0585
Trial3*Freq4	-3.5000	1.1298	713	-3.10	0.0020
Trial3*Freq5	-2.2976	1.2159	713	-1.89	0.0592
Trial3*Freq6	0.1488	0.9903	713	0.15	0.8806