Engineering Research Report

Evaluation of Smartphone Sound Measurement Applications

Chucri A. Kardous, MS, PE and Peter B. Shaw, Ph.D.

Division of Applied Research and Technology
Engineering and Physical Hazards Branch

EPHB Report No. 349-12a

December 2013

DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Institute for Occupational Safety and Health

Site Surveyed: N/A

NAICS Code: N/A

Survey Dates: N/A

Surveys Conducted By: Chucri A. Kardous

Contractor Representatives: N/A

Disclaimer

Table of Contents

Abstract

This report describes a pilot study to assess the functionality and accuracy of smartphone sound measurement applications (apps) and examines whether such mobile applications can be appropriately employed for occupational noise measurements. Testing was conducted in a reverberant acoustic chamber using pink noise at levels of 65-95 dB. A representative sample of smartphones and tablets on various platforms were acquired, more than 130 iOS applications were evaluated but only 10 apps met our selection criteria for functionality, measurement metrics, and calibration capability. Only 4 out of 62 Android apps partially met our selection criteria and were tested. None of the Windows-based mobile applications met our selection criteria. The results showed two iOS apps with mean differences of 0.07 dB (unweighted) and -0.52 dB (A-weighted) from the reference values. Two other iOS apps had mean differences within ± 2dB. The Android-based apps lacked the features and functionalities found in iOS apps and showed a wide variance between the same app measurements on different devices. Overall, the study suggests that certain apps may be appropriate for use in occupational noise measurements.

Background

Engineering and Physical Hazards Studies

The National Institute for Occupational Safety and Health (NIOSH) is the primary Federal agency engaged in occupational safety and health research. NIOSH is one of the Centers for Disease Control and Prevention (CDC) and is located in the Department of Health and Human Services. NIOSH was established by the Occupational Safety and Health Act of 1970. This legislation mandated NIOSH to conduct a number of research and education programs separate from the standard setting and enforcement functions carried out by the Occupational Safety and Health Administration (OSHA) in the Department of Labor. An important area of NIOSH research deals with methods for controlling occupational exposure to potential chemical and physical hazards. The Engineering and Physical Hazards Branch (EPHB) of the Division of Applied Research and Technology has been given the lead within NIOSH to study the engineering aspects of health hazard prevention and control.

Current Study

The ubiquity of smartphones and the sophistication of current sound measurement applications present a great opportunity to revolutionize current data collection and surveillance practices. Through the use of crowdsourcing techniques, workers around the world may be able to collect and share workplace (or task-based) noise exposure data using their smartphones. Scientists and occupational safety and health professionals could rely on such shared data to build job exposure databases and promote better hearing health and prevention efforts. In addition, the ability to acquire and display real-time noise exposure data raises people’s awareness about their work (and off-work) environment and allows them to make informed decisions about hazards to their hearing – addressing two recent major NIOSH initiatives (worker empowerment and total worker health).

NIOSH has received several requests from stakeholders, safety professionals, and the public to address the accuracy of the many sound measurement applications (apps) available for smartphones and whether they can be relied on to provide an accurate assessment of the ambient environment.

Occupational Exposure Limits and Health Effects

As a guide to the evaluation of the hazards posed by workplace exposures, NIOSH investigators use mandatory and recommended occupational exposure limits (OELs) when evaluating chemical, physical, and biological agents in the workplace. Generally, OELs suggest levels of exposure to which most workers may be exposed up to 10 hours per day, 40 hours per week for a working lifetime without experiencing adverse health effects. It is, however, important to note that not all workers will be protected from adverse health effects even though their exposures are maintained below these levels. A small percentage may experience adverse health effects because of individual susceptibility, a pre-existing medical condition, and/or hypersensitivity. In addition, some hazardous substances may act in combination with other workplace exposures, the general environment, or with medications or personal habits of the worker to produce adverse health effects even if the occupational exposures are controlled at the level set by the exposure limit. Combined effects are often not considered in the OEL. Also, some substances are absorbed by direct contact with the skin and mucous membranes, and thus can increase the overall exposure. Finally, OELs may change over the years as new information on the toxic effects of an agent becomes available.

Most OELs are expressed as a time-weighted average (TWA) exposure. A TWA exposure refers to the average airborne concentration of a substance during a normal 8- to 10-hour workday. Some substances have recommended short term exposure limit (STEL) or ceiling values which are intended to supplement the TWA where there are recognized toxic effects from higher exposures over the short-term.

In the U.S., OELs have been established by Federal agencies, professional organizations, state and local governments, and other entities. OSHA established permissible exposure limits (PEL) that are legally enforceable in workplaces regulated by the Occupational Safety and Health Act [29 CFR 1910]. NIOSH developed recommended exposure limits (REL) that are based on critical reviews of the scientific and technical information available on the prevalence of health effects, the existence of safety and health risks, and the adequacy of methods to identify and control hazards [NIOSH, 1992]. Other OELs that are commonly used and cited in the U.S. include the threshold limit values (TLVs) and biological exposure indices (BEIs) recommended by the American Conference of Industrial Hygienists (ACGIH) [ACGIH, 2012]. ACGIH TLVs and BEIs are considered voluntary guidelines for use by industrial hygienists and others trained in this discipline "to assist in the control of health hazards." The American Industrial Hygiene Association (AIHA) developed guidelines for chemical and physical agents called the workplace environmental exposure levels (WEELs) "when no other legal or authoritative limits exist." [AIHA, 2011].

OSHA requires an employer to furnish employees a place of employment that is free from recognized hazards that are causing or are likely to cause death or serious physical harm [Occupational Safety and Health Act of 1970, Public Law 91–596, sec. 5(a)(1)]. Thus, employers are required to comply with OSHA PELs. Some hazardous agents do not have PELs, however, and for others, the PELs do not reflect the most current health-based information. Thus, NIOSH encourages employers to consider the other OELs when making risk assessment and risk management decisions to best protect the health and safety of their employees. NIOSH also encourages the use of the traditional hierarchy of controls approach to eliminating or minimizing identified workplace hazards. This includes, in preferential order, the use of: (1) substitution or elimination of the hazardous agent, (2) engineering controls (e.g., local exhaust ventilation, process enclosure, dilution ventilation) (3) administrative controls (e.g., limiting time of exposure, employee training, work practice changes, medical surveillance), and (4) personal protective equipment (e.g., respiratory protection, gloves, eye protection, hearing protection).

Occupational Exposure Limits for Noise

OSHA's standard for occupational noise exposure (29 CFR 1910.95) specifies a maximum PEL of 90 decibels, A-weighted (dBA), averaged over an 8-hour time period. Noise generated from weapons is classified as impulse noise. The regulation uses a 5-dB exchange rate. This means that when the noise level is increased by 5 dBA, the amount of time a person can be exposed is cut in half. For example, a person who is exposed to noise levels of 95 dBA can be exposed to only 4 hours in order to be within the daily OSHA PEL (permissible exposure limit). The OSHA standard has an action level of 85 dBA, which stipulates that an employer shall administer a continuing, effective hearing conservation program when the 8-hour TWA equals or exceeds the action level. The program must include exposure monitoring, employee notification, observation, an audiometric testing program, hearing protection, training programs, and maintenance of records. The standard also states that when workers are exposed to noise levels in excess of the OSHA PEL of 90 dBA (8-hour TWA), feasible engineering or administrative controls shall be implemented to reduce workers' exposure levels. The OSHA standard states that exposure to impulse noise should not exceed 140 decibels (dB) sound pressure level (SPL).

The NIOSH REL for noise (8-hour TWA) is 85 dBA using a 3-dB exchange rate (see OSHA regulations in previous section for an explanation of exchange rates). NIOSH also recommends that no exposure be allowed above 140 dBA [NIOSH, 1998].

Introduction

As of June 2013, smartphone penetration in the U.S. market has reached more than 60% of all mobile subscribers with more than 140 million devices. Apple iOS and Google Android platforms account for 93% of those devices [Nielsen, 2013]. Worldwide adoption is expected to hit 2 billion devices by 2015. Smartphones have evolved into powerful computing machines with exceptional capabilities: most now have built-in sensors such as microphones, cameras, global positioning system (GPS) receiver, accelerometers, gyroscopes, and proximity and light sensors. Smartphone developers now offer many sound measurement applications (apps) using the devices' built-in microphone (or through an external microphone for more sophisticated applications). Interest in such sound measurement applications is growing among audio enthusiasts, educators, acoustic and environmental researchers, and the general public.

Several government and research organizations have commissioned participatory noise pollution monitoring studies using mobile phones [Maisonneuve et al., 2009, 2010; European Environment Agency, 2013; Kanhere, 2013]. The success of these studies relies on the public to report data using their phones' audio and GPS capabilities. However, none of these studies documented the accuracy or the limitations of the sound measurement apps used and whether they can adequately perform measurements similar to current sound measurement instruments in the field.

Currently, occupational noise exposure assessments require the availability and use of specialized and expensive instrumentation such as noise dosimeters or sound level meters, industrial hygiene and data collection expertise, and hundreds of person-hours to assemble and analyze such data. In addition to the issues of instrumentation and expertise, workplace noise surveillance efforts have required extensive funding and large scale government support because of the need for human expertise, accessibility to workplaces, and the use of expensive sound measurement equipment [Sieber, 1991]. The ubiquity of smartphones and the adoption of smartphone sound measurement apps can have a tremendous and far-reaching impact in this area as every smartphone can be potentially turned into dosimeter/sound level meter [Maisonneuve, 2010]. However, in order for smartphone apps to gain acceptance in the occupational environment, the apps must meet certain minimal criteria for functionality, accuracy, and relevancy to the users in general and the worker in particular.

The possibilities associated with collecting real-time occupational and environmental noise data can have a great impact on hearing health, environmental noise pollution, noise source identification, and may also impact decisions related to public health in a manner that could not be envisioned just several years ago. There is a uniqueness about these types of apps that is not available (or applicable) to any other occupational or environmental hazard. Challenges remain with using smartphones to collect and document noise exposure data, mainly issues relating to privacy, motivation to participate in such studies, accuracy of applications, dealing with bad or corrupted data, and mechanisms for storing and accessing such data. Most of these issues are being carefully studied and addressed [Garcia et al., 2012; Drosatos et al., 2012; Huang et al. 2010].

Occupational and general purpose sound level measurements are conducted using type 1 (accuracy ± 1 dBA) or type 2 (accuracy ± 2dBA) sound measurement instruments that must meet the requirements of American National Standards Institute (ANSI) S1.4-1983 (R2007), Specifications for Sound Level Meters [ANSI, 1983 (R2007)]. ANSI S1.4 states the following: "the expected total allowable error for a sound level meter measuring steady broadband noise in a reverberant sound field is approximately ± 1.5 dB for a type 1 instrument and ± 2.3 dB for a type 2 instrument." For compliance with occupational and environmental noise requirements, standards and regulations in the United States require that instruments meet ANSI type 2 specifications. The Occupational Safety and Health Administration (OSHA) noise standard [29 CFR 1910.95] considers type 2 instruments to have an accuracy of ± 2 dBA.

NIOSH received several inquiries and requests from occupational safety and health professionals, stakeholders, and members of the working public to address the issues of smartphone sound measurement apps accuracy and whether such apps are appropriate for use in the occupational environment. This report describes a pilot study to assess the functionality and accuracy of smartphone sound measurement apps, examines the variability of device hardware on the accuracy of the measurements, and aims to determine whether these apps can be relied on to conduct participatory noise monitoring studies in the workplace.

Experimental Setup

We selected and acquired a representative sample of the popular smartphones and tablets on the market as of January 2013 (iPhone 3Gs, iPhone 4s, iPhone 5, iPad 4th generation, Samsung Galaxy S3, Samsung Note, Samsung Focus, HTC One X, and Motorola DROID RAZR). Smartphone apps were selected based on occupational relevancy criteria: (1) ability to report unweighted (C/Z/flat) or A-weighted sound levels, (2) 3-dB or 5-dB exchange rate, (3) slow and fast response, and (4) equivalent level average (Leq) or time-weighted average (TWA). Also, considerations were given to apps that allow calibration adjustment of the built-in microphone through manual input or digital upload files, as well as those with reporting and sharing features. For the purpose of this experiment, the apps were not calibrated to a reference sound level, they were tested with their original calibration settings to simulate a typical user experience who may not have access to a calibrated sound source. Ten iOS apps out of more than 130 apps were examined and downloaded from the iTunes store. The list of the 10 iOS apps tested and examined in this paper is shown in Table 1.

App	Developer	Features
Adv Decibel Meter	Amanda Gates	A/C weighting, Int/Ext mic, Calibration
Decibel Meter Pro	Performance Audio	A/C/Z weighting, Calibration
iSPL Pro	Colours Lab	A/C/SPL weighting, Calibration
Noise Hunter	Inter.net2day	A/C/SPL weighting, Int/Ext mic, TWA, Calibration
NoiSee	IMS Merilni Sistemi	A/C/Z weighting, ISO/OSHA, Dose, Calibration
Sound Level Meter	Mint Muse	A/C/SPL weighting, Calibration
SoundMeter	Faber Acoustical	A/C/SPL weighting, Leq, Int/Ext mic, Calibration
(Real) SPL Meter	BahnTech	A/C/SPL weighting, Calibration
SPL Pro	Andrew Smith	A/C weighting, Leq, Int/Ext mic, Calibration
SPLnFFT	Fabien Lefebvre	A/C/SPL weighting, Leq, Int/Ext mic, Calibration

Table 1. List of iOS smartphone sound measurement apps.

A total of 62 Android apps were examined and downloaded from the Google Play store but only 4 apps partially met our selection criteria (not all criteria elements highlighted above were available on all the apps). The Android apps are shown in Table 2. There were only two non-commercial apps available on both the iOS and Android platforms: Noise Exposure/Buller published by the Swedish Work Environment Authority, and NoiseWatch published by the European Environment Agency.

App	Developer	Features
SPL Meter	AudioControl	A/C Weighting, Int/Ext Mic, Calibration
decibel Pro	BSB Mobile Solutions	A Weighting, Calibration
dB Sound Meter	Darren Gates	Int/Ext Mic, Calibration
Noise Meter	JINASYS	A/SPL weighting, Calibration

Table 2. List of Android smartphone sound measurement apps.

Only a few apps were available on the Windows platform but none met our selection criteria. As a result, no testing was conducted on Windows-based devices or apps.

The measurements were conducted in a diffuse sound field at a reverberant noise chamber at the NIOSH acoustic testing laboratory that meets the requirements of ANSI S12.6-2008. The diffuse sound field ensured that the location and size of the smartphones did not influence the results of the study. For our experimental setup, we generated pink noise with a 20Hz – 20kHz frequency range, at levels from 65 dB to 95 dB in 5-dB increments (7 different noise levels). The measurement range was chosen to reflect the majority of typical occupational noise exposures encountered in the workplace today. Noise generation and acquisition were performed using the Trident software (ViaAcoustics, Austin, TX). Noise was generated through three JBL XRX715 two-way loudspeakers oriented to provide maximum sound diffusivity inside the chamber. Reference sound level measurements were obtained using a ½-inch Larson-Davis (DePew, NY) model 2559 random incidence microphone. Additionally, a Larson-Davis Model 831 type 1 sound level meter was used to verify sound pressure levels. The microphone and sound level meter were calibrated before and after each measurement using G.R.A.S. (Holte, Denmark) model 42AP pistonphone. All the reference measurement instrumentation used in this study underwent annual calibration at a NIST accredited laboratory. Smartphones were set up on a stand in the middle of the chamber at a height of 4 feet and approximately 6 inches from the reference microphone as shown in Figures 1 and 2.

Figure 1. View of the setup inside the reverberant chamber showing the speakers and the smartphones stand with the reference microphone.

A close-up of the stand with the smartphones and reference microphone is shown in Figure 2.

Figure 2. The SoundMeter app on the iPhone 5 (left) and iPhone 4S (right) compared to ½" Larson-Davis 2559 random incidence type 1 microphone (center).

Statistical Design

The purpose of the statistical design of this study was to assess the extent to which the measurement of noise levels by the smartphone apps agreed with the actual noise levels as measured by the reference microphone. The approach was to use the difference between actual noise levels, as measured by the reference microphone, and the app measurement as the outcome variable, and then determine the effects of *app, device,* and *noise level* on this outcome. A difference equal to zero would indicate perfect agreement between the app measurement and the actual value. The larger the difference, the poorer would be the agreement between the app and the reference microphone. ANOVA (Analysis of variance) was used to analyze the data from this split-split-plot experiment; data from the study were balanced and complete. Both SAS's PROC ANOVA (Cary, NC) and Stata Software (College Station, TX) were utilized in the analysis.

The experimental design and analysis model was a split-split-plot with noise level as the whole plot experimental unit, device type as the split-plot experimental unit, and app as the split-split-plot experimental unit.

$$y_{hijk} = \mu + \phi_h + \alpha_i + (\phi\alpha)_{hi} + \beta_j + (\alpha\beta)_{ij} + (\phi\alpha\beta)_{hij} + \gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \varepsilon_{hijk}$$

Where y_{hijk} = observed difference between noise measured by lab meter and app (for block = *h*, noise level = *i*, device = *j*, and app = *k*),
μ = intercept parameter,
ϕ_h = effect due to block (*h* = 1,..., 6),
α_i = effect due to noise level (i=1,...,7),
$(\phi\alpha)_{hi}$ = interaction between block and noise level,
β_j = effect due to device (*j* = 1,...,4),
$(\alpha\beta)_{ij}$ = interaction between noise level and device,
$(\phi\alpha\beta)_{hij}$ = interaction between block, noise level, and device,
γ_k = effect due to app (*k* = 1,...,10),
$(\alpha\gamma)_{ik}$ = interaction between noise level and app,
$(\beta\gamma)_{jk}$ = interaction between device and app,
$(\alpha\beta\gamma)_{ijk}$ = interaction between noise level, device, and app, and
ε_{hijk} = error term.

Each block contained all possible noise levels, devices, and app combinations for a total of 280 samples. The order of noise level was randomized within each block. The order of device was randomized within each noise level. The order of app was randomized within each device. The experiment utilized six total replications (blocks). The STATA code used to generate the randomization tables can be found in Appendix I.

We examined both the unweighted (dB) and A-weighted (dBA) sound pressure levels on the smartphone apps and compared them to the unweighted and A-weighted sound levels measured by the reference microphone. We also used a Larson-Davis type 1 sound level meter and type 2 noise dosimeter to verify and compare the results of the reference and app data to commercially-available sound measurement instruments.

Results

a. A-weighted Sound Level Measurements (dBA) – iOS devices

The results of the ANOVA on differences in A-weighted sound levels between measurements by the apps and by the reference microphone are shown in Table 2. In carrying out the test for *noise*, the error term was that for the *block*noise* interaction, and in the tests for *device* and *noise*device* interaction, the error term was *block*noise*device* interaction in SAS). For all of the effects involving *app* the test was with the sub-sub-plot error term ("Error" in the table).

Source	DF	ANOVA SS	Mean Square	F Value	Pr > F
block	5	170.46133	34.09227		
noise	6	999.94133	166.65689	139.40	<.0001
block*noise	30	35.86617	1.19554		
device	3	5821.92605	1940.64202	1905.43	<.0001
noise*device	18	886.48795	49.24933	48.36	<.0001
block*device (noise)	105	106.94050	1.01848		
app	9	49431.38676	5492.37631	4459.98	<.0001
noise*app	54	637.37057	11.80316	9.58	<.0001
device*app	27	5409.17967	200.33999	162.68	<.0001
noise*device*app	162	1515.12967	9.35265	7.59	<.0001
Model	419	65014.69000	155.16632	126.00	<.0001
Error	1260	1551.66533	1.23148		
Corrected Total	1679	66566.35533			

Table 3. ANOVA on differences between measurements of noise levels from apps and from the lab meter using the dBA outcome (R^2 = 0.977).

The effect of primary interest, app, was highly significant ($p < 0.0001$). In order to see which apps provided measurements closest to the actual reference A-weighted sound levels, we compared the means of the differences using multiple pairwise Tukey comparisons, as shown below in Table 4. Use of the Tukey approach ensures an overall significance level of 0.05. Note that the means with the same letter in Table 4 are not significantly different.

App	Mean	Standard Error	N	Tukey Grouping	
Sound Level Meter	3.6083	0.27926	168		A
SPL Pro	2.4863	0.11935	168		B
SoundMeter	-0.5185	0.12852	168		C
NoiSee	-1.1280	0.25253	168		D
Noise Hunter	-1.9280	0.27227	168		E
SPLnFFT	-2.2744	0.25715	168	F	E
iSPL Pro	-2.5792	0.25884	168	F	
Adv Decibel Meter	-5.0464	0.27668	168		G
Real SPL Meter	-13.1327	0.27929	168		H
Decibel Meter pro	-13.1708	0.27644	168		H

Table 4. Multiple pairwise comparisons (Tukey) of the mean differences in dBA for the ten apps.

Thus we see that the *SoundMeter app* provides measurements closest to the actual values, and that its mean is significantly different from that of any of the other apps. The marginal predicted means of the differences in dBA with 95% confidence intervals are shown in Figure 3. The confidence intervals, which are very narrow, are on top of the dots.

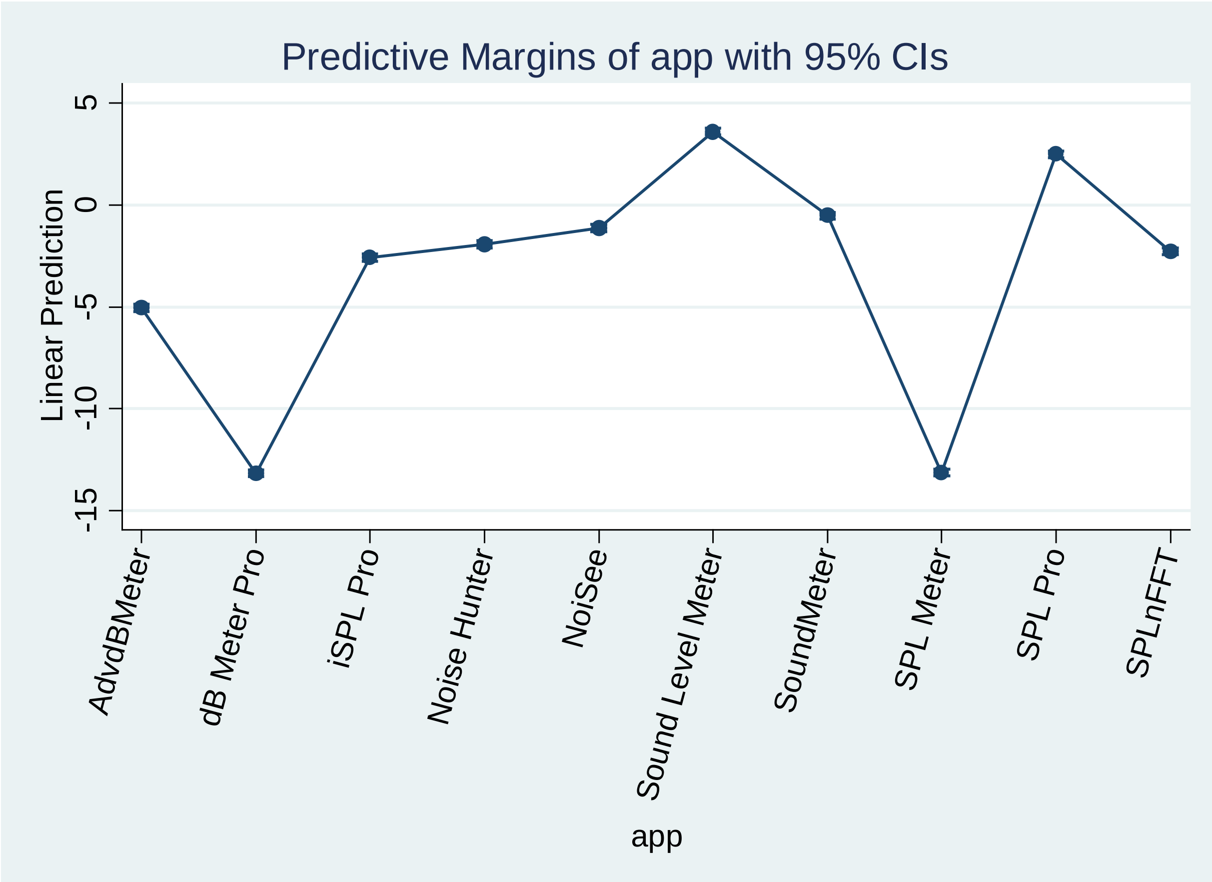

Figure 3. Marginal predicted means of differences in dBA for the ten apps.

Figure 4 shows box plots of the distribution of differences between the reference microphone and the app sound level measurements in dBA.

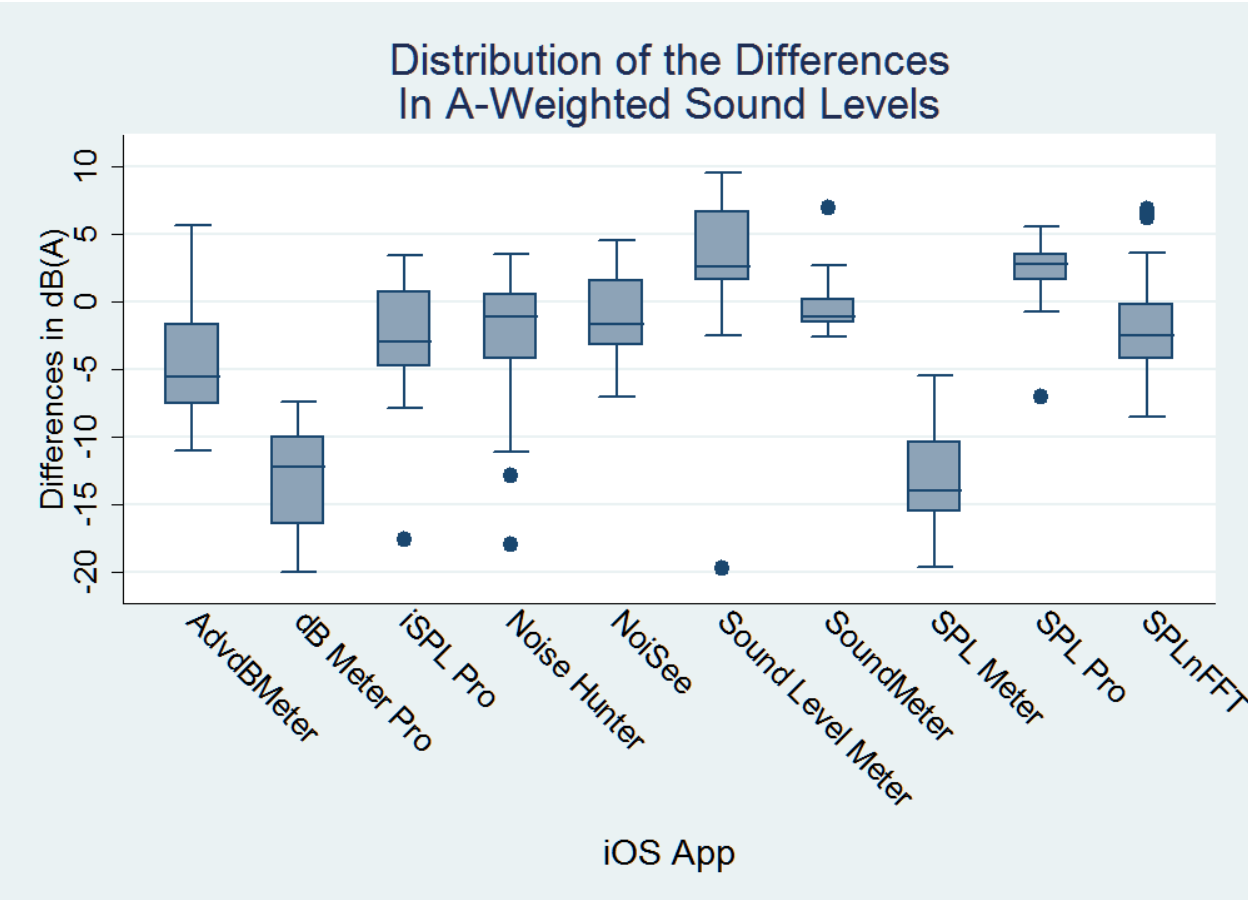

Figure 4. Box plots of differences in A-weighted sound levels (dBA) between reference microphone and app measurements by app. For each box plot, the upper edge of the box represents the upper quartile (75th percentile) and the lower edge the lower quartile (25th percentile). The horizontal line inside the box is the median. A whisker extends upwards from the upper quartile to the largest data point less than the upper quartile + 1.5*(upper quartile – lower quartile). The small horizontal line at the top of the upper whisker is the upper adjacent value. Any data points greater than the upper adjacent value are shown as dots directly above the upper adjacent value. The whisker below each box extends down to the smallest data point greater than the lower quartile – 1.5*(upper quartile – lower quartile). The lower adjacent value is represented by the small horizontal line at the tip of the lower whisker. Any data points less than the lower adjacent value are shown as dots directly below the lower adjacent value.

The effect of device is also quite substantial. Again, in order to see which devices provided measurements closest to the actual reference sound level, we compared the means of the differences using the Tukey multiple pairwise procedure. A total of 420 sample combinations of different apps and noise levels were used to calculate the means of the differences for each device as shown in Table 5.

Device	Tukey Grouping	Mean	N
iP3Gs	A	-0.70810	420
iP4s	B	-2.57071	420
iPhone 5	C	-4.80548	420
iPad 4thGen	D	-5.38905	420

Table 5. Multiple pairwise comparisons of means of differences in dBA for the four devices tested.

Figure 5 shows the overall pattern of the differences between the app measurements and the reference microphone in A-weighted sound levels for the four different iOS devices tested.

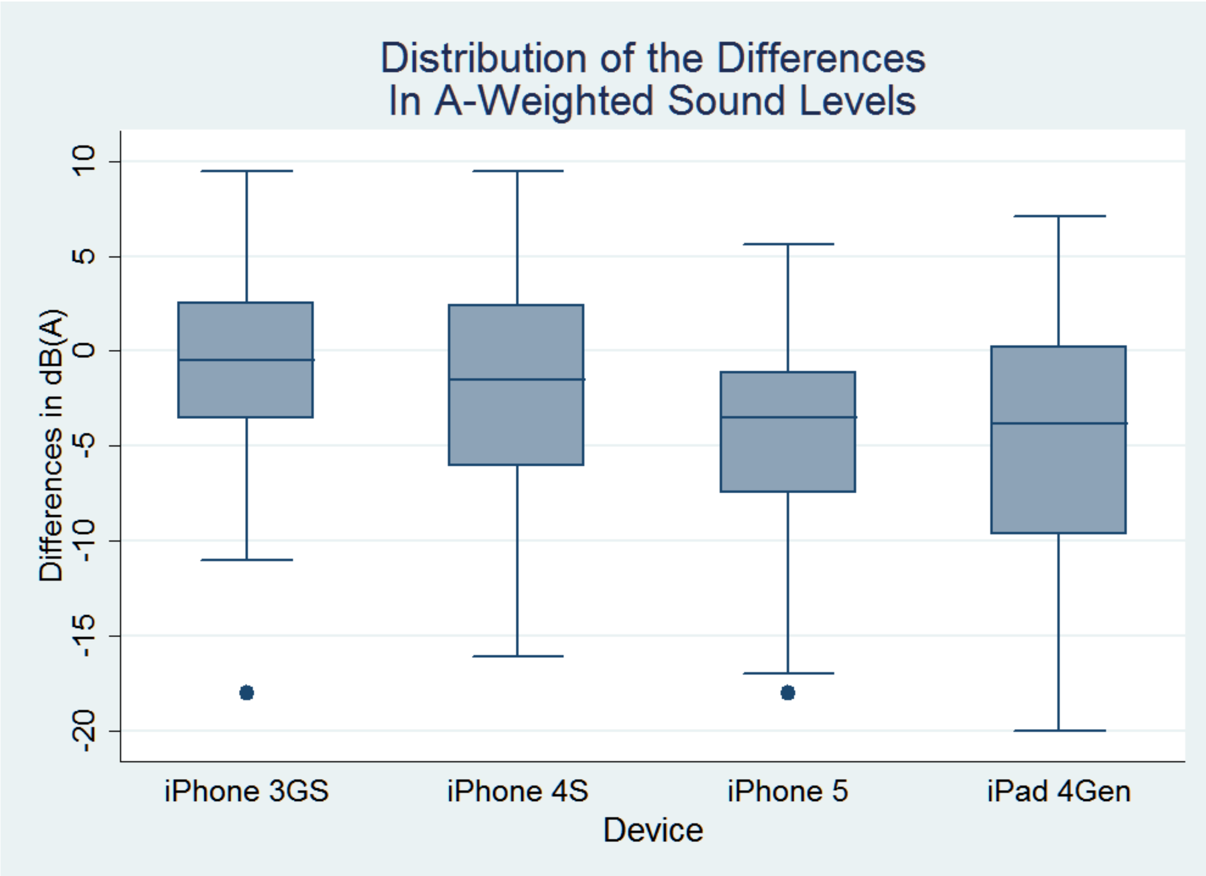

Figure 5. Box plots of differences in A-weighted sound levels (dBA) between reference microphone and app measurements by device.

b. Unweighted sound level measurements (dB) – iOS devices

The results of the ANOVA on the differences in dB between measurements from apps and from the reference microphone are shown in Table 6. Tests were conducted with the same error terms as used above for the dBA measurements.

Source	DF	ANOVA SS	Mean Square	F Value	Pr > F
block	5	306.78446	61.35689		
noise	6	5824.16270	970.69378	982.53	<.0001
block*noise	30	29.63858	0.98795		
device	3	4201.45235	1400.48412	1685.01	<.0001
noise*device	18	715.35511	39.74195	47.82	<.0001
block*device (noise)	105	87.26979	0.83114		
app	9	60016.15715	6668.46191	5965.78	<.0001
noise*app	54	1004.42039	18.60038	16.64	<.0001
device*app	27	9373.83616	347.17912	310.60	<.0001
noise*device*app	162	1116.00846	6.88894	6.16	<.0001
Model	419	82675.08516	197.31524	176.52	<.0001
Error	1260	1408.40883	1.11778		
Corrected Total	1679	84083.49399			

Table 6. ANOVA on differences between measurements of noise levels from apps and from the lab meter using the dB outcome (R^2 = 0.983).

As with the A-weighted sound level measurements, the effect of primary interest, app, was again highly significant ($p < 0.0001$). As shown below in Table 7, the Tukey procedure was used to compare means of differences for unweighted sound levels. The means with the same letter in Table 7 are not significantly different.

App	Mean	Standard Error	N	Tukey Grouping
Sound Level Meter	6.7649	0.29457	168	A
Adv Decibel Meter	3.7875	0.25718	168	B
SPL Pro	2.7851	0.23576	168	C
NoiSee	1.9702	0.29079	168	D
SoundMeter	1.7595	0.23338	168	D
SPLnFFT	0.0696	0.35569	168	E
Real SPL Meter	-5.5857	0.30416	168	F
iSPL Pro	-7.4274	0.27222	168	G
Decibel Meter Pro	-8.6500	0.32718	168	H
Noise Hunter	-12.2161	0.33186	168	I

Table 7. Multiple pairwise comparisons (Tukey) of the mean differences in unweighted sound levels (dB) for the ten apps.

For the differences in unweighted sound levels, we observe that the app *SPLnFFT* provides the closest agreement with actual noise levels. Furthermore, the mean for SPLnFFT is significantly different from all of the other means. The marginal predicted means for the differences in dB with 95% confidence intervals are shown in Figure 6.

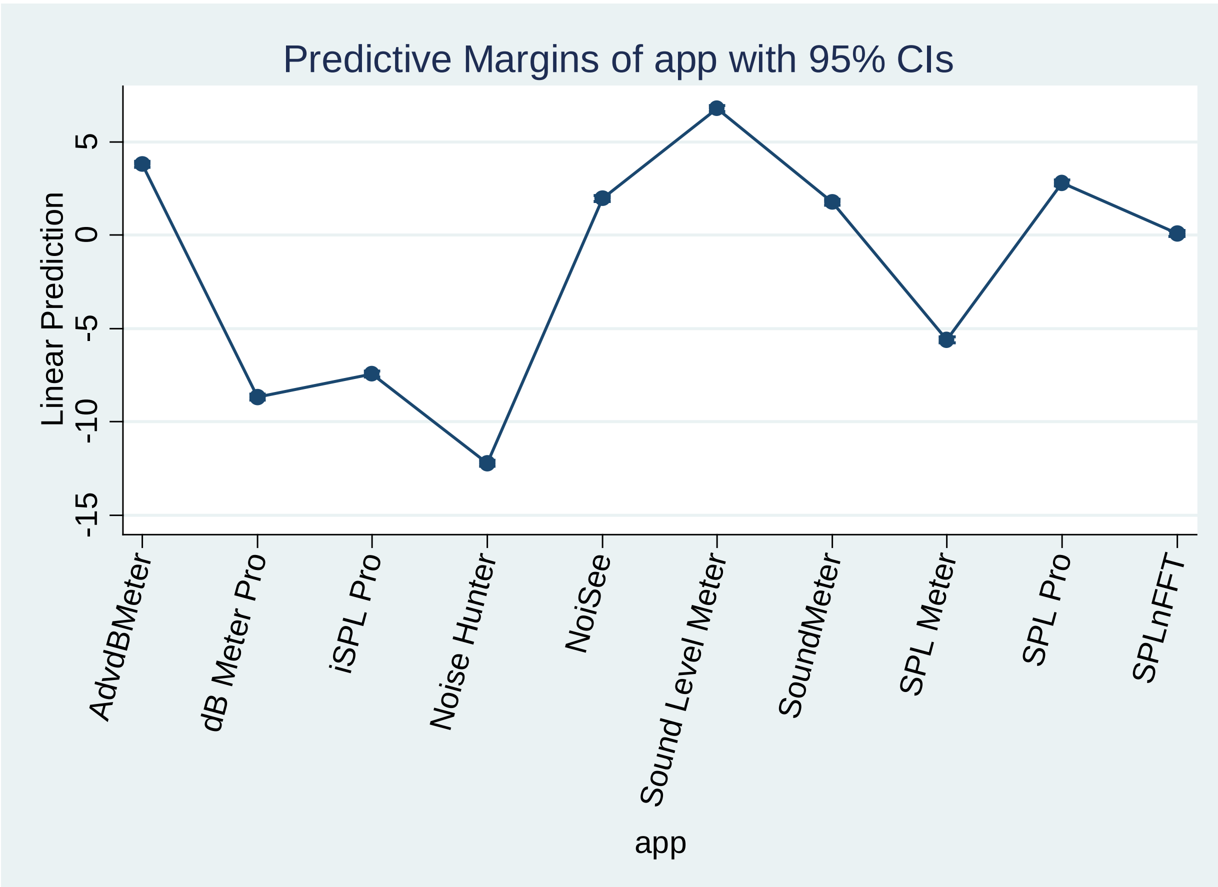

Figure 6. Marginal predicted means of differences in unweighted sound levels (dB) for the ten apps.

As with the A-weighted measurements, Figure 7 shows box plots of the distribution of differences between reference microphone and app sound level measurements in dB.

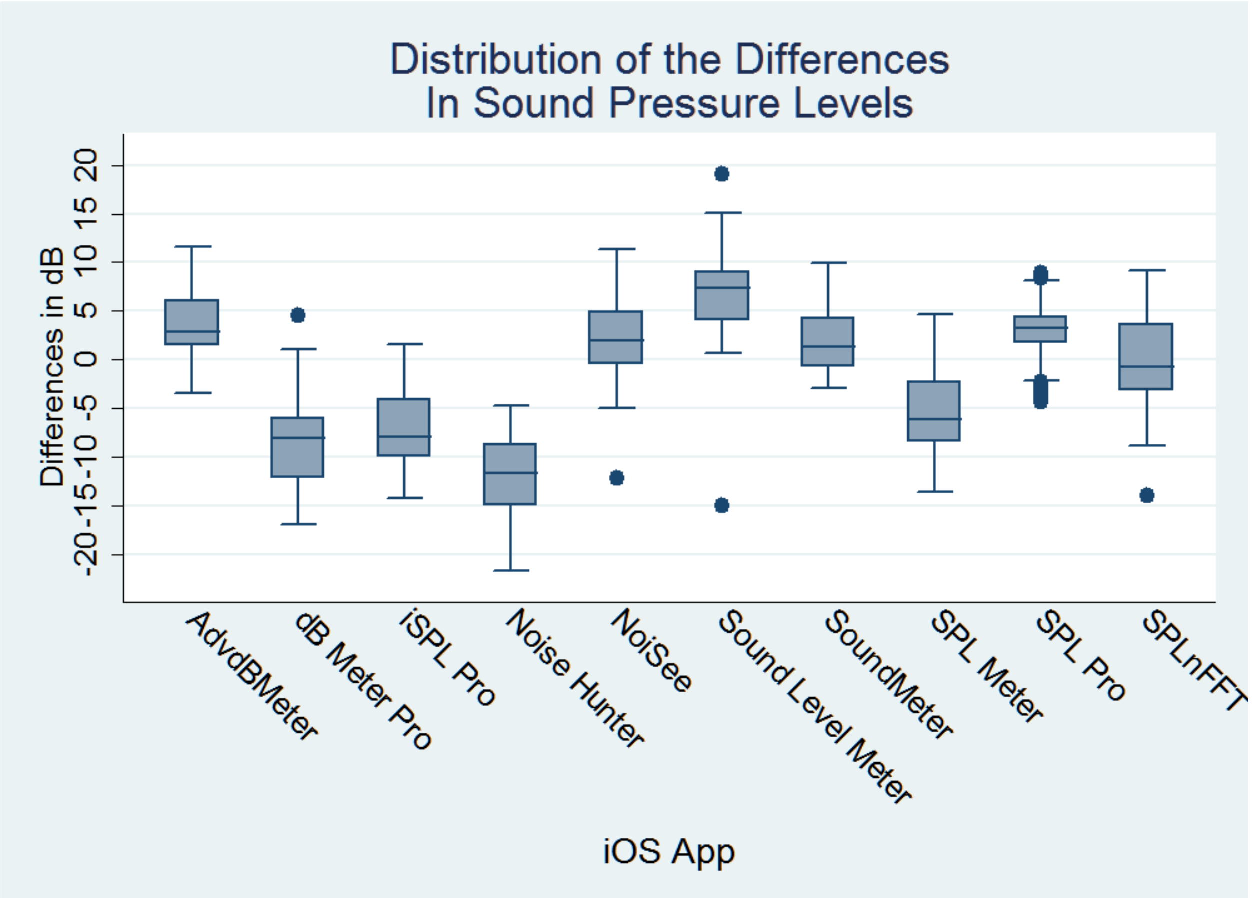

Figure 7. Box plots of differences in unweighted sound levels (dB) between reference microphone and app measurements by app.

As with the dBA differences, there was a highly significant effect due to device for the differences in unweighted sound levels. The means for the different devices are shown in Table 7. As indicated by the Tukey groupings, all of the means are significantly different from one another.

Device	Tukey Grouping	Mean	N
iPhone 3Gs	A	0.44024	420
iPhone 4s	B	-0.83190	420
iPhone 5	C	-3.62786	420
iPad 4thGen	D	-2.67738	420

Table 8. Multiple pairwise comparisons of means of differences in unweighted sound levels (dB) for the four devices tested.

Figure 8 shows the overall pattern of the differences between the app measurements and the reference microphone sound levels in dB for the four different iOS devices tested.

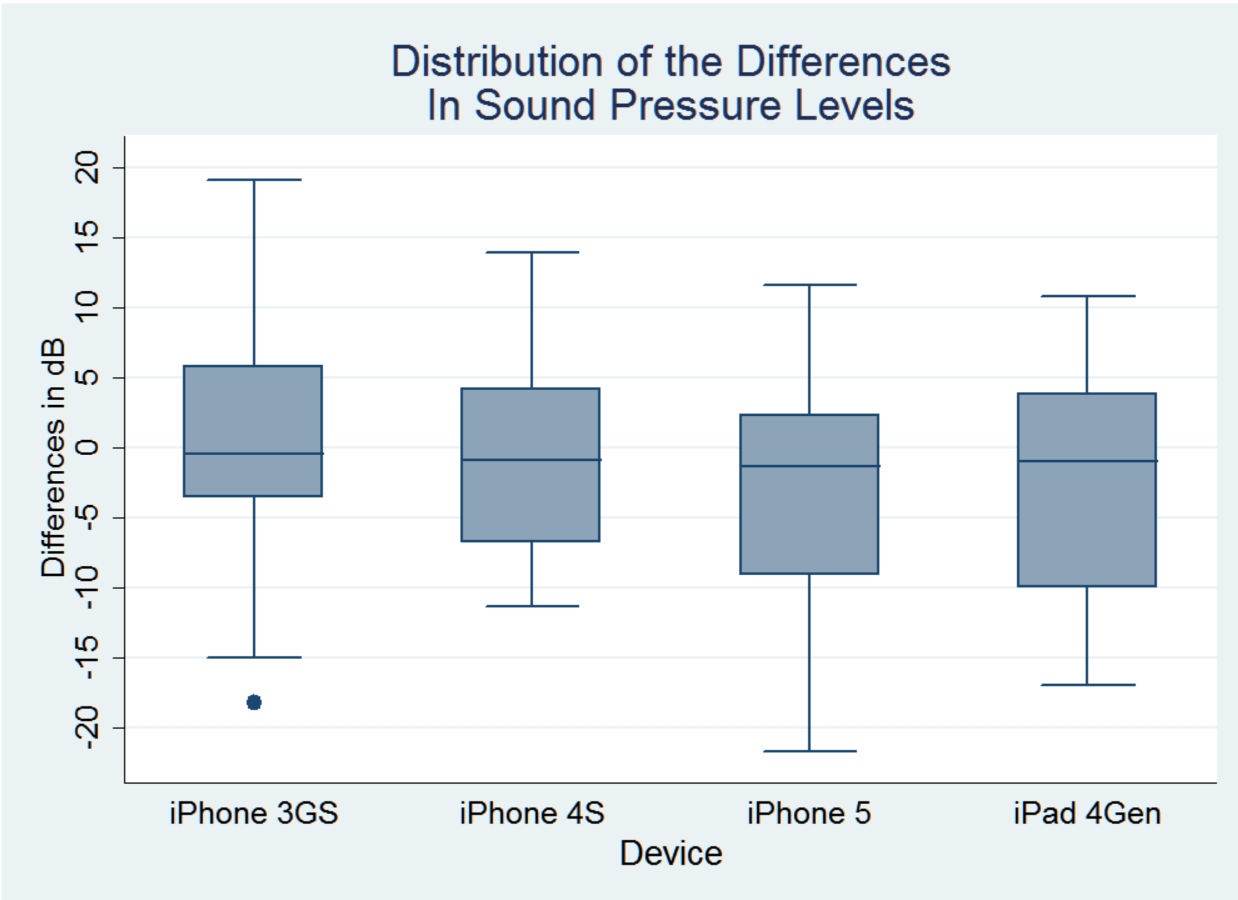

Figure 8. Box plots of differences in unweighted sound levels (dB) between reference microphone and app measurements by device.

c. Android devices

Four Android based apps, (out of a total of 62 that were examined and downloaded) partially met our criteria and were selected for additional testing. Only one app, SPL Meter by AudioControl met our criteria; the other apps did not offer all the features and functions that would be relevant to occupational sound level measurements. Some of the apps offered either unweighted or A-weighted measurements, but not both. As a result, a comprehensive experimental design and analysis similar to the iOS devices and apps study above was not possible. In addition to the low number of apps available with similar functionality, there was a high variance in measurements and a lack of conformity of features of the same apps between different devices. Table 9 shows the extent of the results from testing on Android devices and apps.

App	**Sound Level (dB)**	**Samsung S3**	**HTC One X**	**Motorola Droid**	**Samsung Note**
SPL Meter	70	62.5 ± 2.5	72.4 ± 2	73.6 ± 2.5	66.7 ± 1.2
	80	83.4 ± 1.8	76.6 ± 1.7	85 ± 3	75.6 ± 1.8
	90	91.2 ± 2.2	85.4 ± 1.5	93.6 ± 2.8	92 ± 1.6
deciBel Pro (dBA)	70	69.8 ± 1.5	71 ± 0.8	81 ± 0.6	68.5 ± 1.2
	80	76.1 ± 1.5	79 ± 1	84.9 ± 0.8	75.8 ± 1
	90	87.2± 1.5	85 ± 1.2	82 ± 0.6	86.5 ± 1.5
dB Sound Meter	70	71±1	80 ± 1.5	66 ± 0.5	69 ± 1
	80	78±1	91 ± 1.3	80 ± 0.7	77 ± 1
	90	87 ± 1	92 ± 1.2	93 ± 0.4	86 ± 1
Noise Meter	70	61 ± 0.8	63 ± 1.2	66 ± 0.9	60.6 ± 0.6
	80	68.5 ± 1.2	71 ± 1	75.6 ± 0.6	69 ± 1.1
	90	77.8 ± 1	80.2 ± 1.4	82.2 ± 1	78.6 ± 1.2

Table 9. Measurements of Android based apps and devices at selected unweighted sound levels (dB).

There were only two non-commercial apps available on both the iOS and Android platforms: Noise Exposure/Buller published by the Swedish Work Environment Authority, and NoiseWatch published by the European Environment Agency. The apps did not meet our criteria but testing of the same apps showed a variance of ± 6 dB between Android and iOS devices. Figure 9 shows the NoiseWatch app and the difference between the sound levels measured by a Samsung S3 Android device and the iPhone 5. Reference sound level was 70 dB.

Figure 9. NoiseWatch app on the Samsung S3 Android device (left) and the iPhone 5 (right).

Discussion

The results reported in Table 4 show that the SoundMeter app had the best agreement, in A-weighted sound levels, with a mean difference of -0.52 dBA from the reference values. The SPLnFFT app had the best agreement, in un-weighted sound pressure levels, with a mean difference of 0.07 dB from the actual reference values (Table 7). For A-weighted sound level measurements, Noise Hunter, NoiSee, and SoundMeter had mean differences within ± 2dBA of the reference measurements. For un-weighted sound level measurements, NoiSee, SoundMeter, and SPLnFFT had mean differences within the ± 2 dB of the reference measurement. The agreement with the reference sound level measurements shows that these apps may be considered adequate (over our testing range) for certain occupational noise assessments. The evidence suggests that for A-weighted data, SoundMeter is the app best suited for occupational and general purpose noise measurements. In addition to having the smallest mean difference for the A-weighted data, SoundMeter had one of the narrowest distributions of differences, as shown by the box plot (Figure 4). The apps with differences outside the ± 2 dB/2dBA are considered not to be in good agreement with un-weighted and A-weighted measurements.

The effect of the 4 different iOS devices used in this study on sound level measurements as demonstrated in Tables 5 and 8 and Figures 5 and 8 show that the older iPhone 3GS model produced the best overall agreement between app and reference sound level measurements, with mean differences of 0.4402 dB and -0.7081 dBA between the apps and reference microphone measurements. The variability in the results could be due to the different microphone elements in each device as Apple moved to a new supplier of microphones with the introduction of the iPhone 5 and iPAD 4th Generation devices. The differences could also be related to the introduction of a new operating system (iOS 6) that allowed developers to bypass speech filters and input gain control on older devices.

Almost all smartphone manufacturers use microelectromechanical systems (MEMS) microphones in their devices. MEMS microphones typically have a sensitivity between 5 mV/Pa and 17.8 mV/Pa and can capture signals as low as 30 dB SPL and as high as 120-130 dB SPL (signal-to-noise ratio > 60 dB). MEMS microphones also have a flat frequency response similar to ceramic and condenser microphones used in type 2 noise dosimeters. With the introduction of the iOS 6 operating system in late 2012, Apple allowed developers to bypass the high-pass filter that degraded the quality of acoustical measurements on older iPhones. This development also allows users of Apple smartphones to connect external microphones through the headset input jack. External microphones such as the MicW i436 (Beijing, China) Omni-directional measurement microphone comply with IEC 61672 class 2 sound level meter standard. An extension of this study is planned to examine the effect of external microphones on the overall accuracy of sound measurements apps.

Overall, the Android-based apps lacked the features and functionalities found in iOS apps. This is likely due to the development ecosystem of the Android marketplace and users' expectations for free or low priced apps. A comprehensive testing procedure could not be carried out to show conclusive evidence of differences, since not all apps shared features and metrics that met our selection criteria. The limited testing showed a wide variance between the same app measurements on different devices. This can likely be attributed to the fact that Android devices are built by several different manufacturers and that there is a lack of conformity for using similar microphones and other audio components in their devices.

Challenges remain with using smartphones to collect and document noise exposure data. Some of the main issues encountered in recent studies relate to privacy and collection of personal data, sustained motivation to participate in such studies, the overall accuracy of the sound applications, bad or corrupted data, and mechanisms for storing and accessing such data. Most of these issues are being carefully studied and addressed (Maisonneuve N., et al., 2009) (Kanjo, 2010).

This study is not a comprehensive assessment of the mobile sound measurement apps marketplace. Apps are added and removed on a daily basis and features and updates occur regularly. This study had several limitations, mainly because of the small number of devices that were acquired and tested. Furthermore, this study examined these apps in a controlled noise environment. Field measurement results may vary greatly due to the effect of temperature, humidity, long-term use, object interference, and overall stability of the microphone and electronics in these devices. Finally, smartphone apps cannot be relied upon to conduct compliance assessments in the workplace until the devices and apps meet national and international standards for sound measurement instrumentation such as ANSI S1.4 and IEC 61672-1.

Conclusion

This study showed that certain sound measurement apps for Apple smartphones and tablets can be considered accurate and reliable to be used in the field in lieu of more expensive professional equipment. Currently, Android and Windows developers do not offer apps that meet the functionality needed for occupational noise assessments due to lack of features and functionality required for such measurements (A/C/Z weighting, average or TWA calculation, calibration capability). Recent developments in the use of crowdsourcing and participatory noise monitoring techniques of environmental noise suggest that these apps may also be appropriate for use in certain occupational environments to improve awareness of workplace noise and help advance the hearing health of workers.

References

ACGIH [2013]. 2013 TLVs® and BEIs®: threshold limit values for chemical substances and physical agents and biological exposure indices. Cincinnati, OH: American Conference of Governmental Industrial Hygienists.

AIHA [2011]. 2011 Emergency Response Planning Guidelines (ERPG) & Workplace Environmental Exposure Levels (WEEL) Handbook. Fairfax, VA: American Industrial Hygiene Association.

ANSI [1983]. Specification for Sound Level Meters, ANSI S1.4-1983. American National Standards Institute, New York. (1983 (R2007)).

ANSI [1996]. Determination of occupational noise exposure and estimation of noise-induced hearing impairment, ANSI S3.44-1996. American National Standards Institute, New York.

CFR. 29 CFR 1910. Code of Federal Regulations. Washington, DC: U.S. Government Printing Office, Office of the Federal Register.

Drosatos, G., Efraimidis, P. S., Athanasiadis, I. N., D'Hondt, E., & Stevens, M. [2012]. A privacy-preserving cloud computing system for creating participatory noise maps. In Computer Software and Applications Conference (COMPSAC), 2012 IEEE 36th Annual (pp. 581-586). IEEE.

European Environment Agency. (2013, June 23). eyeonearth.org. Retrieved June 23, 2013, from eyeonearth.org: http://eyeonearth.org/map/NoiseWatch/

Garcia Martí, I., Benedito, M., Trilles, S., Beltran Fonollosa, A., Díaz Sánchez, L., Huerta Guijarro, J., & Rodríguez, L. E. [2012]. Mobile Application for Noise Pollution Monitoring through Gamification Techniques.In Entertainmnet Computing-ICEC 2012 (pp.562-571). Springer Berlin Heidelberg.

Huang, K. L., Kanhere, S. S., & Hu, W. [2010]. Are you contributing trustworthy data? the case for a reputation system in participatory sensing. In Proceedings of the 13th ACM international conference on Modeling, analysis, and simulation of wireless and mobile systems (pp. 14-22). ACM.

IEC [2002]. Electroacoustics – Sound level meters 0 part 1: Specifications, IEC 61672-1, International Electrotechnical Commission, Geneva, Switzerland.

Kanhere, S. S. [2013]. Participatory sensing: Crowdsourcing data from mobile smartphones in urban spaces. Distributed Computing and Internet Technology. Springer Berlin Heidelberg, 19-26.

Kanjo, E. [2010]. NoiseSPY: A Real-Time Mobile Phone Platform for Urban Noise Monitoring and Mapping. Mobile Networks and Applications, 562-574.

Maisonneuve N., Matthias N. [2010]. Participatory noise pollution monitoring using mobile phones. Information Polity, 51-71.

Maisonneuve N; Mattias S.; Nissen M.E.; Hanappe P.; & Steels L. [2009]. Citizen Noise Polution Monitoring. In Proceedings of the 10th Annual International Conference on Digital Government Research: Social Networks: Making Connections between Citizens, Data and Government (pp. 96-103). Digital Government Society of North America.

Nielsen [2013]. Mobile Majority: U.S.smartphone ownership tops 60%. Retrieved June 23, 2013, from http://www.nielsen.com/us/en/newswire/2013/mobile-majority--u-s--smartphone-ownership-tops-60-.html

NIOSH [1998]. Criteria for a recommended standard—Occupational noise exposure (revised criteria 1998). Cincinnati, OH: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health (NIOSH) Pub. No. 98–126.

NIOSH [1992]. Recommendations for occupational safety and health: compendium of policy documents and statements. Cincinnati, OH: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health (NIOSH) Pub. No. 92-100.

OSHA [1992]. Occupational Noise Exposure. Code of Federal Regulations 29 CFR 1910.95. U.S. Government Printing Office, Office of the Federal Register. OSHA, Washington, D.C.

Sieber, W. K. [1991]. Development, use, and availability of a job exposure matrix based on national occupational hazard survey data. American Journal of Industrial Medicine, 20(2), 163-174.

Appendix I. STATA Code for randomization tables

```
clear
set seed 19
set obs 756

egen block = seq(), from(1) to(7) block(126)
gen rannum=runiform()
egen block_n = seq(), from(1) to(126)
egen dBA_grp = seq(), from(1) to(7) block(18)
sort block dBA_grp
egen tag_grp = tag(block dBA_grp)

* randomize the sound levels:

by block: egen rank_grp =rank(rannum) if tag_grp==1
by block dBA_grp: egen dBA_grp_rank=max(rank_grp)

* randomize by devices:

egen device = seq(), from(1) to(3) block(6)

sort block dBA_grp device
egen tag_device = tag(block dBA_grp device)
by block dBA_grp: egen rank_device = rank(rannum) if tag_device==1
by block dBA_grp device: egen device_rank=max(rank_device)

* randomize by app:
by block dBA_grp device: egen rank_app = rank(rannum)

*  add labels:
label define device_labels 1 "iPhone 3GS" 2 "iPad" 3 "iPhone4S" 4 "iPhone 5"
label values device_rank device_labels

label define app_labels 1 "NoiseSee" 2 "dB Hunter" 3 "SPLnFFT" 4 "Real SPL
Meter" 5 "dB Meter pro" 6 "AdvDbMeter" 7 "SoundMeter" 8"iSPL Pro" 9
"Sound Level Meter" 10 "SPL Pro"
label values rank_app app_labels

label define dBA_level_labels 1 "65 dB" 2 "70 dB" 3 "75 dB" 4 "80 dB" 5 "85
dB" 6 "90 dB" 7 "95 dB"
label values dBA_grp_rank dBA_level_labels

keep block dBA_grp_rank device_rank rank_app
```

Delivering on the Nation's promise: Safety and health at work for all people through research and prevention.

To receive NIOSH documents or other information about occupational safety and health topics, contact NIOSH at

1-800-CDC-INFO (1-800-232-4636)

TTY: 1-888-232-6348

E-mail: cdcinfo@cdc.gov

or visit the NIOSH Web site at www.cdc.gov/niosh

For a monthly update on news at NIOSH, subscribe to NIOSH eNews by visiting www.cdc.gov/niosh/eNews

SAFER • HEALTHIER • PEOPLE