Evaluation of smartphone sound measurement applications[a)]

Chucri A. Kardous[b)] and Peter B. Shaw
National Institute for Occupational Safety and Health, 4676 Columbia Parkway, Cincinnati, Ohio 45226
ckardous@cdc.gov, pbs3@cdc.gov

Abstract: This study reports on the accuracy of smartphone sound measurement applications (apps) and whether they can be appropriately employed for occupational noise measurements. A representative sample of smartphones and tablets on various platforms were acquired, more than 130 iOS apps were evaluated but only 10 apps met our selection criteria. Only 4 out of 62 Android apps were tested. The results showed two apps with mean differences of 0.07 dB (unweighted) and −0.52 dB (A-weighted) from the reference values. Two other apps had mean differences within ±2 dB. The study suggests that certain apps may be appropriate for use in occupational noise measurements.

PACS numbers: 43.50.Qp, 43.50.Rq, 43.50.Yw, 43.58.Fm [SS]
Date Received: August 29, 2013 **Date Accepted:** January 16, 2014

1. Introduction

As of June 2013, smartphone penetration in the U.S. market has reached more than 60% of all mobile subscribers with more than 140×10^6 devices. Apple iOS and Google Android platforms account for 93% of those devices (Nielsen, 2013). Worldwide adoption is expected to hit 2×10^9 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 apps). Interest in such sound measurement apps 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 and Matthias, 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.

The National Institute for Occupational Safety and Health (NIOSH) conducts scientific research and provides guidance to reduce workplace injury and illness. Workplace noise surveillance efforts in the past 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 sophistication of current sound measurement apps present a great opportunity to revolutionize current data collection and surveillance practices for noise. Through the use of crowdsourcing techniques, workers around the world can collect and share workplace (or task-based) noise exposure data using their smartphones. Scientists and occupational safety and health professionals

[a)]The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the National Institute for Occupational Safety and Health. Mention of any company or product does not constitute endorsement by NIOSH.
[b)]Author to whom correspondence should be addressed.

can 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 workers' awareness about their ambient noise environment and allows them to make informed decisions about the hazard to their hearing.

Occupational and general purpose sound level measurements are conducted using type 1 or type 2 sound measurement instruments that must meet the requirements of 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.

This paper describes a pilot study to assess the functionality and accuracy of smartphone sound measurement apps, examine the variability of device hardware on the accuracy of the measurements, and determine whether these apps can be relied on to conduct participatory noise monitoring studies in the workplace.

2. Methods

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 or fast response, and (4) equivalent continuous average sound level (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 the reference sound level and were tested with their original calibration settings to simulate a typical user experience that may not have access to a calibrated sound source or equipment. Ten iOS apps out of more than 130 apps were examined and downloaded from the iTunes store. The list of the ten iOS apps tested and examined in this paper is shown in Table 1.

A total of 62 Android apps were examined and downloaded from the Google Play store but only 4 apps *(SPL Meter by AudioControl, deciBel Pro by BSB Mobile Solutions, dB Sound Meter by Darren Gates*, and *Noise Meter by JINASYS*) partially met our selection criteria. 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,

Table 1. List of iOS sound measurement apps.

App	Developer	Features
Adv Decibel Meter 2.0	Amanda Gates	A/C weighting, Int/Ext mic, Calibration
Decibel Meter Pro 2.0.5	Performance Audio	A/C/Z weighting, Calibration
iSPL Pro 1.1.4	Colours Lab	A/C/SPL weighting, Calibration
Noise Hunter 1.0.1	Inter.net2day	A/C/SPL weighting, Int/Ext mic, TWA, Calibration
NoiSee 1.0	IMS Merilni Sistemi	A/C/Z weighting, ISO/OSHA, Dose, Calibration
Sound Level Meter 1.5	Mint Muse	A/C/SPL weighting, Calibration
SoundMeter 3.3.1	Faber Acoustical	A/C/SPL weighting, Leq, Int/Ext mic, Calibration
(Real) SPL Meter 1.0	BahnTech	A/C/SPL weighting, Calibration
SPL Pro 3.6	Andrew Smith	A/C weighting, Leq, Int/Ext mic, Calibration
SPLnFFT 4.0	Fabien Lefebvre	A/C/SPL weighting, Leq, Int/Ext mic, Calibration

and *NoiseWatch* published by the European Environment Agency. 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.

For our experimental setup, we generated pink noise with a 20 Hz to 20 kHz frequency range, at levels from 65 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. The experimental design 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. 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 randomization schedule required six total replications (blocks) to achieve adequate power (>0.924) for this study. We examined the accuracy of the A-weighted and unweighted (or flat) sound levels for each device over the 65 to 95 dB test range.

The measurements were conducted in a diffuse sound field at a reverberant noise chamber at the NIOSH acoustic testing laboratory. The diffuse sound field ensured that the location and size of the smartphones did not influence the results of the study. 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 $\frac{1}{2}$-in. 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 (SPLs). The microphone and sound level meter were calibrated before and after each measurement using G.R.A.S. (Holte, Denmark) model 42AP piston phone. All the reference measurement instrumentation used in this study underwent annual calibration at a National Institute of Standards and Technology accredited laboratory. Smartphones were set up on a stand in the middle of the chamber at a height of 4 ft and approximately 6 in. from the reference microphone as shown in Fig. 1.

To analyze the data, we generated a randomization sampling schedule and employed analysis of variance using both SAS (Cary, NC) and Stata software (College Station, TX). We used the difference between the actual sound level (as measured by the reference microphone) and the app measurement (the outcome variable), and then determined the effects of noise level, device, and app 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.

Fig. 1. (Color online) The SoundMeter app on the iPhone 5 (left) and iPhone 4S (right) compared to $\frac{1}{2}$-in. Larson-Davis 2559 random incidence type 1 microphone (center).

Table 2. Means of differences in unweighted and A-weighted sound levels using Tukey multiple pairwise comparisons. Means that share the same letter designation (A, B, C, or D) are not significantly different.

App	N	Mean (dB)	S. E. (dB)	Mean (dBA)	S. E. (dBA)
Adv Decibel Meter	168	3.78	0.25	−5.04	0.27
Decibel Meter Pro	168	−8.65	0.32	-13.17^{A}	0.27
iSPL Pro	168	−7.42	0.27	-2.57^{C}	0.25
Noise Hunter	168	−12.21	0.33	-1.92^{B}	0.27
NoiSee	168	1.97^{D}	0.29	−1.12	0.25
Sound Level Meter	168	6.76	0.29	3.60	0.27
SoundMeter	168	1.75^{D}	0.23	−0.51	0.12
(Real) SPL Meter	168	−5.58	0.30	-13.13^{A}	0.27
SPL Pro	168	2.78	0.23	2.48	0.11
SPLnFFT	168	0.06	0.35	$-2.27^{B,C}$	0.25

3. Results

The effects of noise level, device, and app were all highly significant ($p < 0.0001$), as were the interactions between the three variables. For both the unweighted and A-weighted sound level measurements, the fit of the split-split-plot design model to the data was good with R^2 values of 0.983 and 0.977, respectively.

The effect of primary interest, the app, was highly significant with a p-value much less than 0.0001 for both the unweighted and A-weighted sound level measurements. In order to see which apps provided measurements closest to the actual reference sound levels, Tukey multiple pairwise comparisons were used to compare the means of the differences between the app measurements and the reference sound levels. Use of the Tukey approach ensured an overall significance level of 0.05. A total of 168 sample combinations of device and noise levels were used to calculate the means of the differences for each app. The results of the comparisons are shown in Table 2.

Figure 2 shows box plots of the distribution of differences between reference and app measurements of the unweighted and A-weighted sound levels. Any data points greater than the upper adjacent value or less than the lower adjacent value are shown as dots directly above or below the adjacent value.

The effect of the 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. The results are shown in Table 3.

Figure 3 shows box plots of the distribution of differences between the app measurements and the reference microphone in unweighted and A-weighted sound levels for the four iOS devices tested.

A similar examination of Android apps and devices was not performed because of the low number of apps with similar functionality, and the lack of conformity of features between devices. Testing conducted with the four Android smartphones also revealed a high variance in measurements of similar apps between different devices.

4. Discussion

The results reported in Table 2 show that the *SPLnFFT* app had the best agreement, in unweighted SPLs, with a mean difference of 0.07 dB from the actual reference values. The *SoundMeter* app had the best agreement, in A-weighted sound levels, with a mean difference of −0.52 dBA from the reference values. For unweighted sound level measurements, *NoiSee*, *SoundMeter*, and *SPLnFFT* had mean differences within the ±2 dB of the reference measurement. For A-weighted sound level measurements, *Noise Hunter*, *NoiSee*, and *SoundMeter* had mean differences within ±2 dBA of the reference

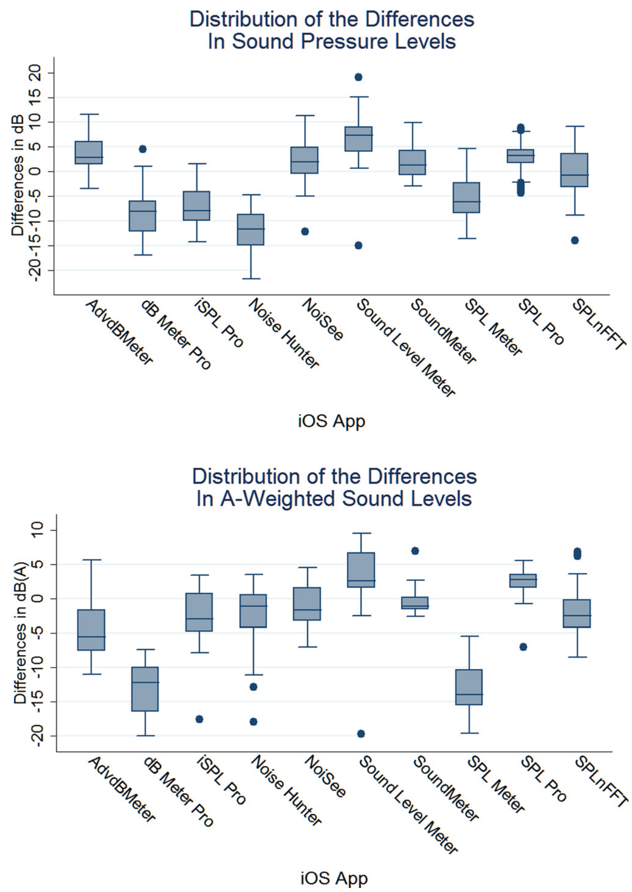

Fig. 2. (Color online) (a) Box plots of differences in unweighted SPLs between reference microphone and app measurements by app. (b) Box plots of differences in A-weighted sound levels between reference microphone and app measurements by app.

measurements. 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 the narrowest distribution of differences, as shown by the box plot [Fig. 2(b)]. The apps with differences outside the ± 2 dB/2 dBA are considered not to be in good agreement with unweighted and A-weighted measurements.

Table 3. Means of differences in unweighted and A-weighted sound levels using Tukey multiple pairwise comparisons.

App	N	Mean (dB)	Mean (dBA)
iPhone 3Gs	420	0.44	−0.70
iPhone 4s	420	−0.83	−2.57
iPad 4thGen	420	−2.67	−5.38
iPhone 5	420	−3.62	−4.80

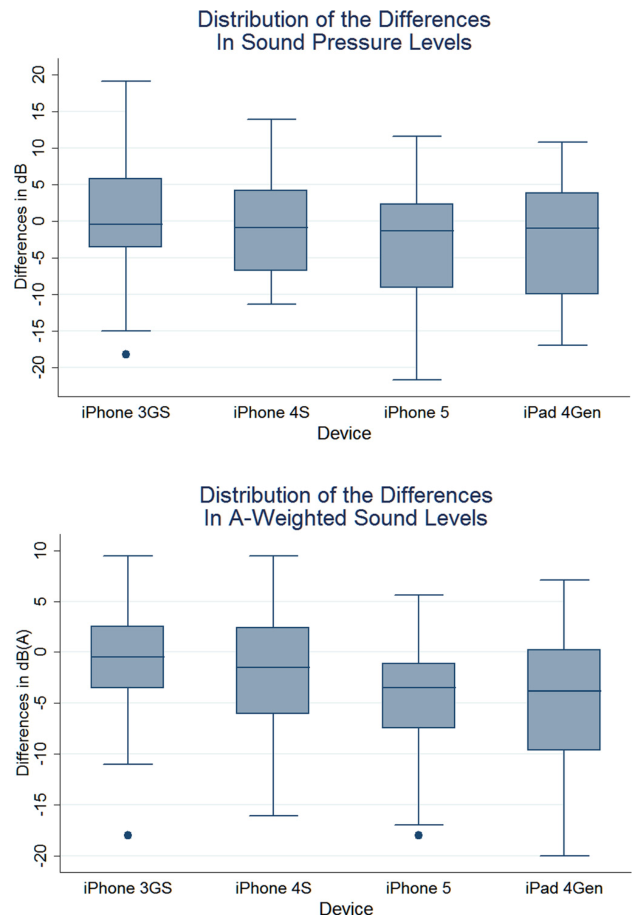

Fig. 3. (Color online) (a) Box plots of differences in unweighted SPLs between reference microphone and app measurements by device. (b) Box plots of differences in A-weighted sound levels between reference microphone and app measurements by device.

The effect of the four different iOS devices used in this study on sound level measurements as demonstrated in Table 3 and Figs. 3(a) and 3(b) show that the older iPhone 3Gs model produced the best agreement for all the apps and noise levels tested (420 samples), with mean differences of 0.44 dB and −0.71 dBA between the apps and the 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 and 17.8 mV/Pa and can capture signals as low as 30 dB SPL and as high as 120 to 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.

The Android-based apps did not have features and functionality similar to the 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 apps, bad or corrupted data, and mechanisms for storing and accessing such data. Most of these issues are being carefully studied and addressed (Maisonneuve *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.

5. Conclusions

This study showed that certain sound measurement apps for Apple smartphones and tablets may be considered accurate and reliable to be used to assess occupational noise exposures. Android and Windows developers do not offer apps that meet the functionality needed for occupational noise assessments. Recent developments in the use of crowdsourcing and participatory noise monitoring techniques of environmental noise suggest that these techniques may also be appropriate for use in the occupational environment to improve awareness of workplace noise and help advance the hearing health of workers.

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the National Institute for Occupational Safety and Health. Mention of any company or product does not constitute endorsement by NIOSH.

References and links

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

European Environment Agency. (**2013**). eyeonearth.org. http://eyeonearth.org/map/NoiseWatch/ (Last viewed June 23, 2013).

Kanhere, S. S. (**2013**). "Participatory sensing: Crowdsourcing data from mobile smartphones in urban spaces," in *Distributed Computing and Internet Technology* (Springer, Berlin), pp. 19–26.

Kanjo, E. (**2010**). "NoiseSPY: A real-time mobile phone platform for urban noise monitoring and mapping," Mobile Networks Appl. **15**(4), 562–574.

Maisonneuve, N., and Matthias, N. (**2010**). "Participatory noise pollution monitoring using mobile phones" Inf. Polity **15**, 51–71.

Maisonneuve, N., Mattias, S., Nissen, M. E., Hanappe, P., and Steels, L. (**2009**). "Citizen noise pollution monitoring," in *Proceedings of the 10th Annual International Conference on Digital Government Research: Social Networks: Making Connections between Citizens, Data and Government*, Digital Government Society of North America, pp. 96–103.

Nielsen. (**2013**). Mobile Majority: U.S. smartphone ownership tops 60%. http://www.nielsen.com/us/en/newswire/2013/mobile-majority–u-s–smartphone-ownership-tops-60-.html (Last viewed June 23, 2013).

Sieber, W. K. (**1991**). "Development, use, and availability of a job exposure matrix based on national occupational hazard survey data," Am. J. Ind. Med. **20**(2), 163–174.