User-operated audiometry–an evaluation of expert vs. non-expert headphone placement

Sørensen, Chris Bang; Gyldenlund Pedersen, René; Nielsen, Jacob; Sidiras, Christos; Schmidt, Jesper Hvass; Pedersen, Ellen Raben

Published in:
International Journal of Audiology

DOI:
10.1080/14992027.2022.2106903

Publication date:
2023

Document version:
Accepted manuscript

User-operated audiometry - An evaluation of expert vs. non-expert headphone placement

[1]Chris Bang Sørensen*, [1]René Gyldenlund Pedersen, [1]Jacob Nielsen, [1]Christos Sidiras, [2,3]Jesper Hvass Schmidt and [1]Ellen Raben Pedersen.

Email addresses in order:
chbs@mmmi.sdu.dk
rgp@mmmi.sdu.dk
jani@mmmi.sdu.dk
chsi@mmmi.sdu.dk
jesper.schmidt@rsyd.dk
erpe@mmmi.sdu.dk
Affiliations
1The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark.
2Research Unit for ORL – Head & Neck Surgery and Audiology, Odense University Hospital, Odense, Denmark; University of Southern Denmark, Odense, Denmark.
3OPEN, Odense Patient data Explorative Network, Odense University Hospital, Odense, Denmark.
*Corresponding author.

Abstract

OBJECTIVE:
User-operated audiometry faces multiple barriers. One of these is the concern of audiologists that patients (non-experts) placing headphones by themselves results in invalid hearing thresholds due to greater placement variability.

DESIGN:
Comparative study. Participants took the AMTAS pure-tone air-conduction audiometry under two different conditions, expert and non-expert circumaural headphone placement for five frequencies within the range 250–8000 Hz. Questionnaires were also used to gain insight into the usability of the user-operated audiometry system – as well as the participants' perceived handling of the audiometry headphones.

STUDY SAMPLE:
Thirty participants (mean age 67.5 years).

RESULTS:
No statistically significant mean differences in hearing thresholds between the expert and non-expert conditions were found. The mean system usability scale score was 84.5. Handling the headphones was also rated as being easy (30%) or very easy (60%) by most non-experts.

CONCLUSION:
The conclusion of the study is that non-experts can be trusted to properly equip a pair of circumaural audiometry headphones for the correct conduction of pure-tone audiometry with only a few digital instructions.

Keywords

Audiometry. User-operated. User-placement. Patient-placement. Headphones.

Introduction

User-operated audiometry -psychoacoustic audiometry that ordinary people (non-experts) perform on themselves, has in the last decade been a promising solution for addressing the growing need for audiological services that outpace the growth of traditional audiological services (Margolis and Morgan 2008). The topic of user-operated audiometry and the different facets of the technology has been the subject of several reviews (Bright and Pallawela 2016; Mahomed et al. 2013; Shojaeemend and Ayatollahi 2018; Trecca et al. 2021). The reviews of both literature and implementations of user-operated audiometry find the technology on par with traditional audiometry for measuring hearing thresholds. With enough evidence for the many user-operated audiometry reviews, one would think that the inventions would be ready to change the field of audiology. However, even with technologies that have been proven to meet the clinical standards, practical and human barriers exist facing implementation. An audit of 17 audiological departments regarding ototoxic monitoring practices of children found that nine of the 17 departments had reluctance in the consideration of self-test devices (Brown et al. 2021). Reluctance to change/innovation in healthcare is a known phenomenon (Coiera 2011; Safi, Thiessen, and Schmailzl 2018), and research into how to penetrate barriers and gain acceptance is similarly being conducted (Côté-Boileau et al. 2019; Safi et al. 2018). Part of the solution is to bring in the staff whose workflow will be affected by the technology in order to understand their concerns and gain from their experience. This way the new technology may be perceived as an enabling technology by the staff, rather than being a new barrier for doing their work or worsening the care quality (Kent et al. 2015).

In work done for the UAud project which aims to further realise the implementation and use of user-operated audiometry in clinics (Sidiras et al. 2021), seven audiologists were interviewed following a semi-structured interview guide. These interviews informed the creation of a questionnaire about perceived concerns and potentials of user-operated audiometry. This questionnaire was answered to completion by 69 audiologists and partially answered by an additional 32 audiologists. Multiple concerns were identified, such as patients misunderstanding the test, malingering, having collapsed ear canals during the test, or taking the test with cerumen blockage. But the placement of headphones by patients themselves, rather than by the audiologist, was the second most frequent concern. Cerumen blockage was the most frequent concern.

Variability in hearing thresholds due to headphone placement has been investigated previously to varying results. Zhou and Green (1995) found that a difference in transducer placement would impact higher frequencies (8000–16000 Hz) more than lower frequencies (1000–8000 Hz). Part of their research tested with commercial headphones and measured sound pressure levels via probe tubes in participant's ears. They found higher frequencies were particularly impacted by movement (standard deviation (SD) = 2.67 dB) compared to the non-movement condition (SD = 1.62 dB). Using pink noise and two music excerpts as stimuli, Paquier and Koehl (2015) tested how placement variability affected audibility. Expert and non-expert listeners were tasked with picking the oddball recording in a 3-interval 3-alternative forced choice test under different headphone placement conditions. Listeners were able to significantly detect the recording that corresponded to a different placement, except for the non-expert listeners with one of the placements playing one of the music excerpts. Zhou and Green (1995) and Paquier and Koehl (2015) shows that there is validity to the previously mentioned concern. Their results do, however, not conclude if the non-expert placement of headphones will affect the thresholds of an audiogram.

Paquier, Koehl, and Jantzem (2016) investigated placement variability by measuring variability due to intra-participant threshold variance. Analysing participants individually, Paquier et al. (2016) observed three types of participants, namely those with: consistent high variability, consistent low variability, and lastly variability only when replacing the headphones. For those participants, variability was seen not only at the higher frequencies measured but also at the middle frequencies. Considering headphone type, Paquier et al. (2016) found statistically significant differences at 2000 and 11000 Hz for circumaural headphones and at 4000 and 6000 Hz for supra-aural headphones respectively. The authors, however, questioned if the differences were due to placement, as no other frequencies were significantly affected by repositioning. The results of Paquier, Koehl, and Jantzem (2016) indicate that the concerns regarding non-expert placement might be isolated to a subset of non-experts, rather than being a pervasive issue for all non-experts.

Zhong et al. (2010) investigated measurement errors in headphone-to-ear-canal transfer functions, with one of the variables being headphone placement. The largest differences between placements were between 8000 and 10000 Hz with an SD of 8 dB. For the middle frequencies, the SD was around 0.8 dB from 1000 to 5000 Hz. The lower frequencies had worse SDs than the middle frequencies, with 250 Hz reaching up towards an SD of 2 dB, making the lower frequencies of potential interest to the current study. While the results of Zhou and Green (1995) indicate that a search for the difference between the expert and non-expert placement should focus on the higher frequencies, the results of Zhong et al. (2010) show that the lower frequencies might also be of interest.

Contrary to the concern about non-expert placement, a 2015 study on the hearing of 300+ workers found that hearing thresholds improved with the non-expert placing the headset themselves rather than the expert (Almeida et al. 2015). However, the non-experts were receiving direct instructions from the expert. This could bias the instructions given with the information the expert has by viewing the actions of the non-expert. It therefore still leaves open the concern of non-experts placing the headphones incorrectly when doing the task without active guidance. Margolis et al. (2016) provided their participants with test kits for home audiometry and found a good comparison with manual audiometry done in the clinic. Insert earphones (modified Etymotic Research mc5) were used in the kit, and participants had to insert these themselves. But the non-expert placement is not directly highlighted in the publication and the 250 Hz tone was not tested as the greatest effect of ambient noise was expected.

While headphone placement has been the subject of some research, concerns of some audiologists remain. Transducer placement influences sound pressure level, but whether this influence will have any effect in an audiogram when the responsibility of placing the circumaural transducers is put in the hands of non-experts has not been determined. The study investigates the following questions:

1. Is there a difference in the audiogram thresholds between digitally instructed user placement compared to the expert placement of headphones?
2. How do users rate their experience with handling the headphones and the usability of the system?

And thereby conclude on whether non-experts can correctly put on and place circumaural audiometry headphones themselves for the successful conduction of audiometry.

Materials and methods

Participants

A total of 30 participants (13 females, 17 males, age 40–81 years, mean age 67.5 years and SD of 10.1 years) were recruited for the study. Twenty-two participants were recruited from two clinics affiliated with the Department of Audiology, University Hospital of Odense. Five of the 22 were from the hearing clinic in Svendborg and 17 of those 22 from the University clinic at Winsløwparken in Odense. Twenty-one were patients recruited directly at the clinics and one was next-of-kin to a patient. The remaining eight participants were recruited externally via a flyer and word of mouth. To ensure that participants were representative of the population at a hearing clinic (Houmøller et al. 2022), an age restriction was applied for inclusion. Inclusion criteria were a) age 40 years or older; b) native Danish speakers; c) capability to read and operate a touchscreen; d) dexterity sufficient to pick up and place the circumaural headphones.

Testing environments

The testing in Svendborg was performed in audiometry booths. The testing site at the University Clinic had not previously been screened for audiometric suitability and so measurements of background noise were performed there. The location fell within acceptable parameters for testing, with no frequencies used in the study measuring above ISO-8253-1:2010 limits. During the 20 min of measuring background noise (with one measurement made every second), the ISO limits were only exceeded at 200 Hz, but by no more than 2 dB and in only 16 readings, which should be of no or a negligible consequence to this study.

Equipment

The study used the DD450 Headphones (RadioEar), a circumaural headphone. The headphones were connected to an Affinity Compact (Interacoustics), which was connected to an HP EliteBook 840 G5 laptop running Windows 10 and the Affinity Compact Suite software. As the laptop does not provide a touch interface, an external touch screen, the ZenScreen™ Touch MB16AMT (ASUSTeK Computer Inc.), was connected to the laptop.

A 6th generation iPad (Apple), model number A1893 was used for measuring background noise. The software used for this was the AudioTools app (Studio Six Digital) with the additional SPL Graph module (newest release as of 02/12/2021). The iPad and software were calibrated in a free-field setup.

Test software

A modified version of the Interacoustics AMTAS software was used. This software is the AMTAS (Margolis et al. 2010) test built into the Affinity Compact Suite, which runs the test via a connected Affinity Compact. With the Interacoustics AMTAS test both air- and bone-conduction pure-tone audiometry, as well as speech audiometry, can be performed by non-experts using the software. However, for this study, only the air-conducted pure-tone audiometry test was used. Modifications to the software were made in the code of the software for research purposes only. The modifications included colour changes for sharper contrasts, font change and text size increase for improved readability, as well as the addition of specific instructions for conducting the headphone placement.

Procedure

The study participants were given a brief introduction to the purpose of the experiment and then signed a consent form to participate. Following this, participants were asked about their experience with hearing aids and audiological testing, experience with tinnitus, vertigo and fluctuating or asymmetrical hearing loss. Afterward, otoscopy was performed. The expert (a trained audiologist) collecting these data was present in the room with the participants throughout the entire data collection session. However, the expert could only guide the patients' understanding during specific parts of the questionnaire. If questions arose during the user-operated audiometry, responses from the expert would be in the manner of "Just try your best to do as the system tells you," so as to not influence the tests beyond being in the room. Name and social security number were input by the expert into the software and were then automatically migrated to the study's online research database provided by Odense Patient Data Explorative Network (OPEN).

Participants took the user-operated pure-tone test twice. Once with the user placing the headphones themselves by following digital instructions (non-expert placement), and once with an expert placing the headphones. The order of non-expert and expert placement was alternated between participants. The orientation of the headphones on the desk was also alternated right-left/left-right between participants to counter any coincidental correct orientation. The user-operated audiometry tested the frequencies 1000, 6000, 8000, 250 and 500 Hz. This limited number of tones was selected to keep the test sessions as short as possible and not cause participant fatigue, while still testing the frequencies found to be most relevant by previous literature. After completing the user-operated audiometry twice, the participant was allowed a 5-min break before answering a questionnaire. Presenting the questionnaire after the audiometries ensured that all participants completed it.

Placing headphones instructions

For the participants to place the headphones themselves, instructions on headphone placement were given by the software. The software only presented instructions during non-expert placement, so instructions had not been viewed twice by the non-experts receiving expert placement first. Instructions for headphone placement were presented in two sets, with the user going from the first set of instructions to the second set by clicking the "Fortsaet" (Continue) button. The first set of instructions focussed on describing the composition of the headset and the right-left properties (colour differentiated). The first set consisted of 388 characters (spaces excluded) divided into 79 words. The second set of instructions focussed on the actions and sensations the non-expert were to use to achieve a placement acceptable for audiometry. The second set consisted of 386 characters (spaces excluded) divided into 77 words. Screenshots of the placement instruction screens, as well as the Danish text and a english translation, can be found in the online supplemental material at https://www.tandfonline.com/doi/10.1080/14992027.2022.2106903.

Questionnaire

The questionnaire consisted of 29 questions, 10 of which were the SUSDK questionnaire (Hvidt et al. 2020) SUSDK is a Danish translation of the System Usability Scale (SUS) (Brooke 1995). Besides the 29 questions, two conditional questions and three free comment fields were included for the participant to further elaborate their answers. The 19 questions that were not part of SUSDK ranged from documenting IT capability, usage of glasses, and awareness of any tinnitus the participant may have, to questions regarding the subjective experience with handling the headphones and operating the software, as to investigate if some of these aspects could be potential predictors for the difference in non-expert and expert placement.

Statistical approach

Data storage was done using the Research Electronic Data Capture (REDCap) tool developed by Vanderbilt University, Nashville, Tennessee, United States (Harris et al. 2019; Harris et al. 2009), supplied by OPEN. Stata 17.0 (Revision 17 Jan 2022) was the primary statistical tool used for the analysis of the result. The analysis consisted of Shapiro-Wilk normality tests, calculating means and SD of differences and paired t-tests ($p < 0.05$).

Due to safety limitations in the software, the system could not measure hearing thresholds above 80 dB HL at 250 Hz, 90 dB HL at 500 Hz, 95 dB HL at 1000 Hz, 80 dB HL at 6000 Hz, and 75 dB HL at 8000 Hz, and any hearing threshold above this was reported as "not registered" by the system and are not included in the data analysis.

Outliers

Potential outliers were decided as any hearing threshold absolute difference at or greater than two times the expanded uncertainty, with expanded uncertainty being as defined by ISO 8253:2010 annexe A.

Results

For each test (five frequencies on both ears) the average test time was 7 min and 43 seconds (SD of 1 min and 48 seconds). Mean and SD of real and absolute differences between hearing thresholds obtained by expert and non-expert placing the headphones are shown in Table 1. The means of real differences fell between −0.3 and 1.7 dB and the means of absolute differences were between 2.0 and 4.3 dB. The SDs of the real differences ranged from 3.1 to 6.5 dB and SDs of absolute differences ranged from 2.5 to 5.8 dB.

Table 1. Mean and SD of threshold differences between expert and non-expert placement for all data and with potential outliers (n=4) omitted. For the real differences, non-expert placement was subtrahend.

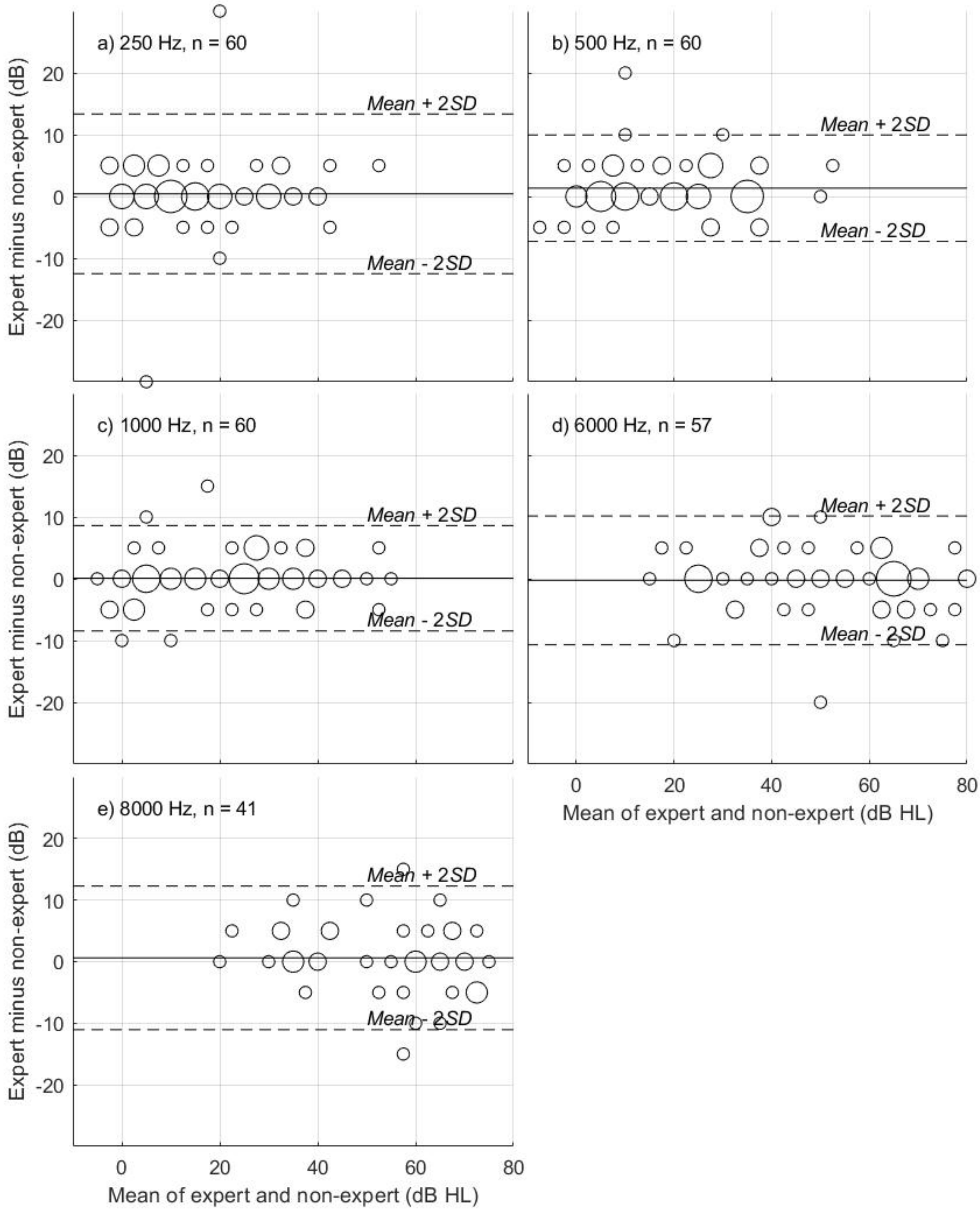
	250 right	250 left	500 right	500 left	1000 right	1000 left	6000 right	6000 left	8000 right	8000 left
Mean and (SD) of real differences	1.2 (6.5)	-0.3 (6.4)	1.7 (5.3)	1.0 (3.1)	0.2 (4.0)	0.0 (4.6)	-0.2 (5.0)	-0.3 (5.5)	0.5 (5.9)	0.8 (5.9)
Mean and (SD) of absolute differences	3.2 (5.8)	3.0 (5.7)	3.3 (4.4)	2.0 (2.5)	2.2 (3.4)	3.0 (3.4)	3.0 (3.9)	3.5 (4.3)	3.8 (4.5)	4.3 (4.1)
n	30	30	30	30	30	30	28	29	21	20
Mean and (SD) of real differences without outliers	0.2 (3.7)	0.7 (3.2)	1.0 (4.1)			-0.5 (3.6)				
Mean and (SD) of absolute differences without outliers	2.2 (2.9)	2.5 (2.5)	2.8 (3.2)			2.6 (2.5)				
n without outliers	29	29	29			29				

The paired t-tests for hearing thresholds measured at each frequency for each ear found no statistically significant difference between the expert and non-expert placement. Across the 10 paired t-tests, the degrees of freedom ranged from 29 to 19 with t-values from −1.80 to 0.34 (four positive values and six negative values). The p-values were all greater than 0.05, ranging from 0.08 to 1.00, meaning there was no evidence for non-random differences between measurements in our sample.

Four hearing threshold differences were found as potential outliers. If the potential outliers are removed from the means and SD calculations, the means of differences fall within −0.5 to 1.0 dB and the means of absolute differences were all less than 3.8 dB. SD of real differences across the frequencies ranged from 3.1 to 5.9 dB and SDs of absolute differences ranged from 2.5 to 4.5 dB. See

Table 1 for the means and SD of each frequency per ear with potential outliers removed. For visualisation of the potential outliers, see Figure 1 (Bland-Altman plots). The Bland-Altman plots in Figure 1 show no disproportionality in terms of disagreement between hearing threshold obtained by expert and non-experts placing the headphones.

Figure 1. Bland-Altman plots (difference plots) of the hearing thresholds for both ears were obtained with experts and non-experts placing the headphones. Plots a) through e) shows the different frequencies. The x- and y-axis of all the plots are on the same scale. The size of the circles indicates the number of values at a given position; the larger the circles the more overlapping values.



Measurement uncertainty

An uncertainty budget, combined uncertainty and expanded uncertainty are calculated in accordance with ISO 8253:2010. The first part of the uncertainty budget is repeated measurements, denoted as L'ht,est. This uncertainty is for measurements under the same conditions, which is what we wish to evaluate our findings against. For air-conducted stimuli, L'ht,est is 2.5 dB up to 4000 Hz and 4 dB above 4000 Hz. Audiometric equipment is denoted as δeq. For this uncertainty measure, the step size is also considered, with the step size here being 5 dB (Eikelboom et al. 2013). For the equipment and step size in this study, δeq is 2.3 dB up to 4000 Hz and 3.2 dB above 4000 Hz. The next relevant uncertainty measure is transducers and their fitting denoted δtr. The contribution of δtr is 1.5 dB up to 4000 Hz and 2.5 dB above 4000 Hz. Masking is applied by the AMTAS method, and if masking is suspected to be non-optimized (not following ISO 8253:2010) an uncertainty measure of 2 dB can be added. However, this measure was omitted, as the masking level on the non-test ear at −20 dB relative to the signal level on the test ear of the AMTAS test is not disruptive. The combined uncertainty is denoted as u and can be calculated as the square root of the sum of the squared uncertainty measures.

$u = \sqrt{(L'^2_{ht,est} + \delta^2_{eq} + \delta^2_{tr})}$

Resulting in u = 3.7 dB up to 4000 Hz and 5.7 dB above 4000 Hz.

The expanded uncertainty, denoted U, is u times a coverage factor k. For coverage probability of 95% and a normal distribution k = 2;

$U = u * 2$

Resulting in U = 7.4 dB up to 4000 Hz and 11.4 dB above 4000 Hz. As all SDs of real differences fall within the expanded uncertainty, threshold measurements are within the expected variance for pure-tone audiometries as specified by ISO 8253:2010.

Questionnaire results

The average SUS score, calculated as described by Brooke (1995), was 84.5 out of 100 with an SD of 13.2, a very favourable score. The questions were in Danish and are translated here to English. On the question “How was it to put on the headphones?” with the five-point answering scale going from “Very hard” to “Very easy,” three answered “Neither hard nor easy,” nine answered “Easy,” and 18 answered “Very Easy.”

On the question “How did it feel to have the headphones on?” with the five-point answering scale going from “Very uncomfortable” to “Very comfortable,” one answered “Uncomfortable,” eight answered “Neither uncomfortable or comfortable,” 16 answered “Comfortable,” and five answered “Very comfortable.”

On the question “Could you feel a difference between your own placement of the headphones and the professional's placement?” with the five-point answering scale going from “His placement felt noticeably better” to “My placement felt noticeably better,” one answered “His placement felt noticeably better,” nine answered “His placement felt a little better,” 17 answered “No difference,” three answered “My placement felt a little better,” and zero answered “My placement felt noticeably better.”

The left-right orientation of the headphones on the table was alternated between participants to ensure that headphone placement instructions were not followed accidentally, but with intent by the participant. One participant out of the 30 participants failed to place the headphones correctly regarding left-right orientation on their first placement attempt.

Discussion

Mean difference and standard deviation

In this study, hearing thresholds via expert and non-expert placement were collected to ascertain whether non-experts performing headphone placement for the conduction of audiometry has any effect diagnostically, i.e. on the audiogram. The audiologists' concerns regarding inaccurate thresholds due to non-expert placement were found to be resolved. No statistically significant difference was found between hearing thresholds on the tested frequencies. The frequencies chosen were based on literature which, in the 250–8000 Hz frequency range, found the lower and higher ends to have the most audibility and SPL variance due to placement variance. The 1000 Hz frequency was measured specifically to be used as a low variance baseline (Zhong et al. 2010) for comparison.

The results of this study do not suggest any difference between the expert and non-expert placement when the non-expert user is provided with digital text and image instructions for placing the circumaural headphones. This is based on the means being close to 0 dB and not systemically biased, with SDs in the range normally expected for pure-tone audiometry and less than the expected uncertainty for the specific setup used in this study (ISO 8253:2010).

While the issue of headphone placement variance has been studied for some time with varying results (Almeida et al. 2015; Paquier, Koehl, and Jantzem 2016; Zhong et al. 2010; Zhou and Green 1995) the current study is to our knowledge the first to explicitly and intentionally leave non-experts with the full responsibility of achieving correct headphone placement. Placement variability between experts and non-experts was of concern to the validity of the resulting audiogram, an important concern to address as the field of audiology looks poised to employ user-operated audiometry even more. The reliability of non-expert placement is a necessity for user-operated audiometry to be employed in ways more scalable than the reach of today's traditional manual audiometry.

The SDs between non-expert and expert placement in this study also falls in line with SDs found by Margolis et al. (2010) who compared the AMTAS test to manual audiometry on 30 participants. Eikelboom et al. (2013) performed test-retests (two tests with equal conditions) of the AMTAS test on 10 participants. They found similar means of real differences, but higher SDs and means of absolute differences, which may be because they examined the AMTAS test in a non-sound-treated room with a maximum ambient noise of 70 dB. While the present study also used non-sound-treated rooms, these were within ISO (8253-1:2010) limits when accounting for the attenuation provided by the headphones. While the SDs of Eikelboom et al. (2013) may have been higher due to a noise condition not present in this study, the fact that this study found similar means to a study that performed test-retests, further supports that non-experts can perform adequate headphone placement on themselves. Table 2 shows the results of this study alongside the results from the Margolis et al. (2010) and Eikelboom et al. (2013) studies.

Table 2. Mean and SD values of this study and other studies using the AMTAS test. Margolis et al. (2010) did not collect hearing thresholds on 6000 Hz nor did Eikelboom et al. (2013). Eikelboom et al. (2013) also combined right and left ear measurements.

Mean and (SD) of real differences	250 right	250 left	500 right	500 left	1000 right	1000 left	6000 right	6000 left	8000 right	8000 left

This Study	0.2 (3.7)	0.7 (3.2)	1.0 (4.1)	1.0 (3.0)	0.2 (4.0)	0.0 (4.6)	-0.2 (5.0)	-0.3 (5.5)	0.5 (5.9)	0.8 (5.9)
Margolis et al (2010)	0.4 (6.4)	0.4 (3.5)	–0.2 (4.7)	–0.6 (4.2)	–1.6 (4.0)	–1.4 (6.4)			3.2 (8.8)	1.4 (4.7)
Eikelboom et al (2013)	-1.5 (5.4)		1.3 (6.3)		0.3 (7.9)				1.8 (12.2)	
Mean and (SD) of absolute differences	250 right	250 left	500 right	500 left	1000 right	1000 left	6000 right	6000 left	8000 right	8000 left
This Study	2.2 (2.9)	2.1 (2.5)	2.8 (3.2)	2.0 (2.5)	2.2 (3.4)	3.0 (3.4)	3.0 (3.9)	3.5 (4.3)	3.8 (4.5)	4.3 (4.1)
Margolis et al (2010)	4.0 (5.0)	2.4 (2.5)	3.4 (3.1)	2.6 (3.3)	2.8 (3.3)	3.8 (5.3)			6.4 (6.7)	2.6 (4.1)
Eikelboom et al (2013)	3.5 (4.3)		4.8 (4.1)		5.3 (5.7)				4.7 (4.9)	

System usability and headphone experience

Non-experts rated their experience both with placing the headphones as well as wearing them as mostly positive. That it does not present as a negative experience is of importance to the implementation of user-operated audiometry technology on a wider scale than traditional manual audiometry can currently cover, as non-experts having to handle the headphones by themselves is necessary for the implementation of user-operated audiometry in places where no expert is going to be available to do the placement. However, even though they rated their experience with the headphones mostly positive, a third of participants (n = 10) still preferred the placement conducted by the expert to their own, with only 10% of participants (n = 3) preferring their own placement, and the rest of the participants (n = 17) being neutral about either placement. Whether this could be attributed to the expert placing it more comfortably than the non-experts did themselves, or if this is because having an expert perform the placement instils a greater trust in the placement for the non-expert, is unclear.

The SUSDK questions were used in this study to gauge the usability of the user-operated audiometry from the point of view of the non-experts. The SUS provides a rough insight into any type of system regardless of context, being described as a "quick and dirty" usability scale by its author (Brooke 1995). While being context-free is a great strength of the scale in terms of wide-scale applicability, this also means it is not perfect for any one system. This was also the experience of using it in this study, where especially the first question "I think that I would like to use this system frequently" presented issues. This type of system is not meant to be owned by the respondent and is only something they would use occasionally or due to a medical need. Therefore, as expected, this question scored lower on average than the other SUS questions.

While the SUS has been widely applied in publications in many fields, the authors of this study are only aware of one other publication that used the SUS within audiology. Philips et al. (2018) used the SUS to measure the usability of a tablet application for cochlear implant users, which allowed the users to access relevant information as well as perform user-operated hearing tests at home. Philips et al. (2018) found an average SUS score of 75.6 with an SD of 8.6, and while this is slightly lower than the average SUS score found in this study, both scores would be considered good (Bangor, Kortum, and Miller 2008). although it might just reflect the specific implementations, it still bodes well for user-operated audiometry as a technology that favourable SUS scores are achieved by said implementations, more usability studies of user-operated audiometry would do the field well.

Furthermore, non-experts were found to place headphones with the correct left-right orientation, as instructed by the system, regardless of how the headphones were situated before testing. This was seen in all but one participant. Said participant was stopped by the left-right orientation check in the system that comes after headphone placement and was directed back to the headphone placement instructions by the system. This prompted the participant to check their headphone placement and correct their orientation. Correct headphone orientation is vital for measuring correct hearing thresholds and providing high-quality treatment, especially in participants with asymmetrical hearing loss, and user-operated audiometry systems should implement some strategy to ensure correct orientation and placement. This result also addresses the concern of audiologists regarding headphone placement by proving the quality of the audiograms produced is reliable. However, it is recommended to also have the system run an orientation check to avoid any errors of wrongfully reported hearing thresholds.

Potential outliers

Four threshold differences stood out in this study as potential outliers. A threshold difference of negative 30 dB on 250 Hz left ear tone between expert (−10 dBHL) and non-expert (20 dBHL) belonged to a participant that had been observed to have very narrow ear canals. This aroused the suspicion that an ear canal collapse might have happened due to the user placement, as this had been a suspected potential cause of erroneous hearing thresholds during the non-expert conditions, as the participants are not aware of this being a possibility when they place the headphones. However, there was no difference between the expert and non-expert hearing thresholds at 500 Hz left ear, making it less likely that it was ear canal collapse as that would have created a more uniform hearing threshold increase across frequencies (Bardsiri 2018; Mohamad and Algarni 2016) and more likely to simply be an accidental outlier with no clear explanation in the gathered data.

The last three of the outlying threshold differences belonged to the same participant. Their first potential outlying difference was 30 dB on the 250 Hz right ear tone between expert (35 dBHL) and non-expert (5 dBHL). The second was a difference of 20 dB on the 500 Hz right ear tone between expert (20 dBHL) and non-expert (0 dBHL). The third and last difference was 15 dB on the 1000 Hz left ear between expert (25 dBHL) and non-expert (10 dBHL) This participant received the expert placement condition first, but as they had read the test instructions and pressed to proceed to the hearing test, the system started loading but never finished. (Test instructions are different from headphone placement instructions and were received under both expert and non-expert placement conditions, headphone placement instructions were not present during expert placement conditions). After a period of loading and no progress, the expert stepped in and had to reset the software, upon which the participant could continue unimpeded. It was noted that the participant did not pay as much attention immediately after the restart of the software, going quickly through the previously witnessed screens to reach the hearing test section. This may have resulted in them performing worse due to a less clear understanding of the test and a concentration disturbed by the IT issues experienced.

As the authors cannot be sure that an ear collapse had happened with one participant, and since IT issues will always be a possibility, the study has not excluded these potential outliers in its conclusions. While the findings of Paquier, Koehl, and Jantzem (2016) suggested that there might be a subset of individuals that would have high variability due to placement, this was not the case in the present study.

Conclusion

This study investigated the validity of having non-experts perform headphone placement for the conduction of user-operated audiometry. This was done by intra-subject comparison of audiograms conducted with non-expert and expert headphone placement. A SUS score, as well as ratings on the experience of the participants regarding handling of the circumaural headphones, was also collected.

Threshold differences between expert and non-expert placement were compared with a paired t-test for each ear and frequency, and no statistically significant difference was found between expert and non-expert placement conditions. The mean difference between conditions was also quite low, spanning a range of −0.3 to 1.7 dB across the different frequencies and ears. A mean SUS score of 84.5 out of 100 was achieved, which is considered a good score. Participants rated their experience with handling the headphones positively, but still showed a preference for the placement of the expert over their own in a third of all cases.

A concern about non-expert placement has been expressed by audiologists, and previous research has shown that headphone placement variability changes sound pressure levels in the ear canal, giving validity to the concern of the audiologists. But whether placement variability introduces a significant change at the measurement level of the audiogram was not previously investigated, and this study shows that in a sample of 30 participants, non-expert placement of headphones was valid for achieving a correct audiogram. This is important to know if user-operated audiometry is to be widely implemented and possibly in ways that allow new workflows at clinics to better meet the demands for hearing health care.

Ethical approval

The necessity for approval by the ethical committee was waived by the Research Ethics Committee of the University of Southern Denmark.

Data availability statement

Data is not readily available for sharing due to privacy laws.

Funding

This work was supported by Innovation Fund Denmark Grand Solutions 9090-00089B and the William Demant Foundation case number 19-1251.

Acknowledgements

Carsten Daugaard of Force Technology (Odense) provided knowledge, equipment, and environment for calibration of the iPad that was used to take background noise measurements. Louise Thygesen Schmidt created and conducted the interviews with audiologists and patients whose results in large parts motivated the creation of this study. Data storage was facilitated through OPEN, Open Patient data Explorative Network, Odense University Hospital, Region of Southern Denmark, which provided a REDCap solution with secure servers for GDPR-compliant data storage.

References

Almeida, B. P., Menezes, P. d. L., de Andrade, K. C. L., & Teixeira, C. F. (2015). Positioning of earphones and variations in auditory thresholds. *Brazilian Journal of Otorhinolaryngology, 81*(6), 642-646. doi:10.1016/J.BJORL.2015.08.016

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human–Computer Interaction, 24*(6), 574-594. doi:10.1080/10447310802205776

Bardsiri, M. M. (2018). False Unilateral Conductive Hearing Loss Due to Collapsed Ear Canals in a 78 Years Old Male. *Global Journal of Otolaryngology, 18*(4). doi:10.19080/GJO.2018.18.555991

Bright, T., & Pallawela, D. (2016). Validated Smartphone-Based Apps for Ear and Hearing Assessments: A Review. *JMIR rehabilitation and assistive technologies, 3*(2). doi:10.2196/REHAB.6074

Brooke, J. (1995). SUS - A quick and dirty usability scale. *Usability Eval. Ind., 189*. Retrieved from https://www.researchgate.net/publication/228593520_SUS_A_quick_and_dirty_usability_scale

Brown, E. C. M., Caimino, C., Benton, C. L., & Baguley, D. M. (2021). An audit of UK audiological practice in specialist paediatric oncology centres regarding hearing assessment of children at risk of ototoxicity due to chemotherapy. *The Journal of laryngology and otology, 135*(1), 14-20. doi:10.1017/S0022215121000025

Coiera, E. (2011). Why system inertia makes health reform so difficult. *BMJ, 342*. doi:10.1136/BMJ.D3693

Côté-Boileau, É., Denis, J. L., Callery, B., & Sabean, M. (2019). The unpredictable journeys of spreading, sustaining and scaling healthcare innovations: a scoping review. *Health Research Policy and Systems, 17*(1), 84-84. doi:10.1186/S12961-019-0482-6

Eikelboom, R. H., Swanepoel, D. W., Motakef, S., & Upson, G. S. (2013). Clinical validation of the AMTAS automated audiometer. *International journal of audiology, 52*(5), 342-349. doi:10.3109/14992027.2013.769065

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., . . . Duda, S. N. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics, 95*, 103208-103208. doi:10.1016/J.JBI.2019.103208

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*(2), 377-381. doi:10.1016/J.JBI.2008.08.010

Houmøller, S. S., Wolff, A., Möller, S., Narne, V. K., Narayanan, S. K., Godballe, C., . . . Schmidt, J. H. (2022). Prediction of successful hearing aid treatment in first-time and experienced hearing aid users: Using the International Outcome Inventory for Hearing Aids. *International journal of audiology, 61*(2), 119-129. doi:10.1080/14992027.2021.1916632

Hvidt, J. C. S., Christensen, L. F., Sibbersen, C., Helweg-Jørgensen, S., Hansen, J. P., & Lichtenstein, M. B. (2019). Translation and Validation of the System Usability Scale in a Danish Mental Health Setting Using Digital Technologies in Treatment Interventions. *https://doi.org/10.1080/10447318.2019.1680922, 36*(8), 709-716. doi:10.1080/10447318.2019.1680922

Kent, B., Redley, B., Wickramasinghe, N., Nguyen, L., Taylor, N. J., Moghimi, H., & Botti, M. (2015). Exploring nurses' reactions to a novel technology to support acute health care delivery. *Journal of clinical nursing, 24*(15-16), 2340-2351. doi:10.1111/JOCN.12881

Mahomed, F., Swanepoel, D. W., Eikelboom, R. H., & Soer, M. (2013). Validity of automated threshold audiometry: A systematic review and meta-analysis. *Ear and Hearing, 34*(6), 745-752. doi:10.1097/01.aud.0000436255.53747.a4

Margolis, R. H., Glasberg, B. R., Creeke, S., & Moore, B. C. J. (2010). AMTAS: automated method for testing auditory sensitivity: validation studies. *International journal of audiology, 49*(3), 185-194. doi:10.3109/14992020903092608

Margolis, R. H., Killion, M. C., Bratt, G. W., & Saly, G. L. (2016). Validation of the Home Hearing Test™. *Journal of the American Academy of Audiology, 27*(5), 416-420. doi:10.3766/jaaa.15102

Margolis, R. H., & Morgan, D. E. (2008). Automated pure-tone audiometry: an analysis of capacity, need, and benefit. *American journal of audiology, 17*(2), 109-113. doi:10.1044/1059-0889(2008/07-0047)

Mohamad, A., & Algarni, M. (2016). Bilateral conductive hearing loss due to collapsed ear canals in a 35 years old female. *Journal of Otolaryngology-ENT Research, 4*(4). doi:10.15406/joentr.2016.04.00110

Paquier, M., & Koehl, V. (2015). Discriminability of the placement of supra-aural and circumaural headphones. *Applied Acoustics, 93*, 130-139. doi:10.1016/J.APACOUST.2015.01.023

Paquier, M., Koehl, V., & Jantzem, B. (2016). Effect of headphone position on absolute threshold measurements. *Applied Acoustics, 105*, 179-185. doi:10.1016/j.apacoust.2015.12.003ï

Philips, B., Smits, C., Govaerts Paul, J., Doorn, I., & Vanpoucke, F. (2018). Empowering Senior Cochlear Implant Users at Home via a Tablet Computer Application. *American journal of audiology, 27*(3S), 417-430. doi:10.1044/2018_AJA-IMIA3-18-0014

Safi, S., Thiessen, T., & Schmailzl, K. J. G. (2018). Acceptance and Resistance of New Digital Technologies in Medicine: Qualitative Study. *JMIR Research Protocols, 7*(12). doi:10.2196/11072

Shojaeemend, H., & Ayatollahi, H. (2018). Automated Audiometry: A Review of the Implementation and Evaluation Methods. *Healthcare informatics research, 24*(4), 263-275. doi:10.4258/HIR.2018.24.4.263

Sidiras, C., Sanchez-Lopez, R., Pedersen, E. R., Sørensen, C. B., Nielsen, J., & Schmidt, J. H. (2021). User-Operated Audiometry Project (UAud) – Introducing an Automated User-Operated System for Audiometric Testing Into Everyday Clinic Practice. *Frontiers in Digital Health, 0*, 141-141. doi:10.3389/FDGTH.2021.724748

Trecca, E. M. C., Lonigro, A., Gelardi, M., Kim, B., & Cassano, M. (2021). Mobile Applications in Otolaryngology: A Systematic Review of the Literature, Apple App Store and the Google Play Store. *Annals of Otology, Rhinology and Laryngology, 130*(1), 78-91. doi:10.1177/0003489420940350

Zhong, X., Liu, Y., Xiang, N., & Xie, B. (2010). *Errors in the measurements of individual headphone-to-ear-canal transfer function*. Retrieved from https://pub.dega-akustik.de/DAGA_2010/data/articles/000096.pdf

Zhou, B., & Green, D. M. (1998). Reliability of pure-tone thresholds at high frequencies. *The Journal of the Acoustical Society of America, 98*(2), 828-828. doi:10.1121/1.413509