A SMARTPHONE NATIONAL HEARING TEST: PERFORMANCE AND CHARACTERISTICS OF USERS

Karina C. De Sousa a

a. Department of Speech-Language Pathology and Audiology, University of Pretoria, Pretoria, Gauteng, South Africa.

De Wet Swanepoel, Ph.D. a,b,c

a. Department of Speech-Language Pathology and Audiology, University of Pretoria, Pretoria, Gauteng, South Africa
b. Ear Sciences Centre, School of Surgery, University of Western Australia, Nedlands, Australia
c. Ear Science Institute Australia, Subiaco, Australia

David Moore, Ph.D. d,e

d. Communication Sciences Researcher Center, Cincinnati Childrens’ Hospital Medical Center, University of Cincinnati, Ohio, USA
e. Division of Human Communication, Development & Hearing, University of Manchester, United Kingdom

Cas Smits, Ph.D. f

e. Department of Otolaryngology-Head and Neck Surgery, Section Ear & Hearing, and Amsterdam Public Health research institute, VU University Medical Center, Amsterdam, The Netherlands.
ABSTRACT

Purpose: The smartphone digits-in-noise (DIN) hearing test, called hearZA™ was made available as a self-test in South Africa in March 2016. This study determined characteristics and test performance of the listeners who took the test.

Method: A retrospective analysis of 24072 persons who completed a test between March 2016 to August 2017 was conducted. User characteristics, including age, English-speaking competence and self-reported hearing difficulty were analyzed. Regression analyses were conducted to determine predictors of the speech reception threshold (SRT).

Results: Overall referral rate of the hearZA™ test was 22.4%, and 37% of these reported a known hearing difficulty. Age distributions showed that 33.2% of listeners were 30 years and younger, 40.5% were between 31 and 50 years, and 26.4% were older than 50 years. Age, self-reported English-speaking competence and self-reported hearing difficulty were significant predictors of the SRT.

Conclusions: High test uptake, particularly among younger users and high overall referral rate indicates that the hearZA™ App addresses a public health need. The test also reaches target audiences including those with self-reported hearing difficulty and those with normal hearing who should monitor their hearing ability.

Keywords: digits-in-noise; hearing screening; early detection; hearing loss; smartphone hearing screening; national hearing test
INTRODUCTION

The global burden of hearing loss has been increasing steadily with close to half a billion people estimated to suffer from permanent disabling hearing loss (Vos et al., 2015; Hay et al., 2017; Wilson et al., 2017). The rise in recreational noise exposure further places an estimated 1.1 billion young people between the ages of 12 to 35 years at risk for acquiring hearing loss (WHO, 2017). Regardless of high prevalence rates and significant consequences, hearing loss continues to be an undetected and untreated disability that is not adequately positioned or prioritized within many healthcare systems (Mackenzie & Smith, 2013). This is especially true in low-and-middle-income countries (LMICs) where hearing health care is scarce or often unavailable due to the lack of resources and poor awareness amongst the lay public and health professionals (Mackenzie & Smith, 2013). Consequences of unaddressed hearing loss are far-reaching, affecting social participation (Hallam et al., 2008), psychological well-being (Fellinger et al., 2012), cognitive function (Livingston et al., 2017) and employment opportunity (Tucci, Merson & Wilson, 2010; WHO, 2017). With high proportions of occupations dependent on spoken communication (Ruben, 2015), the cumulative effects of hearing loss can have significant economic implications for the individual and society. Recent WHO estimates of the global estimated costs associated with hearing loss equaled 750 billion (WHO, 2017). Emphasis on prevention and treatment of hearing loss is, therefore, important on a global scale with particular focus on LMICs (Wilson et al., 2017).

Although it is well established that timely identification and management could substantially reduce the functional impairment of hearing loss (Cacciatore et al., 1999), most cases remain undiagnosed and untreated, especially in older adults (Pronk et al., 2011). Hearing screening programmes have a critical role in monitoring hearing and early detection of hearing difficulty (WHO, 2017). However, implementation and success of these programmes are reliant on specific human and technological resources like trained screening personnel, audiological equipment, and optimal quiet test environments. Also, within LMICs, individuals may be required to travel substantial distances for hearing screening as health care providers are severely limited and unequally distributed (Mulwafu et al., 2017). The high costs involved and limited access to population-based hearing screening has led to various initiatives to use telephone and internet based speech in noise (SIN) screening tests (Smits et al., 2004;
Although all these tests measure speech recognition in noise, they vary in terms of speech stimuli, type of background noise and test procedure. Compared to conventional pure tone audiometry, SIN tests do not require calibrated equipment and also have the advantage of being less sensitive to ambient noise and transducer type (Jansen et al., 2010). Furthermore, SIN tests measure the functional deficits related to hearing loss as opposed to hearing sensitivity (Smits & Houtgast, 2005). The digits-in-noise (DIN) test is a type of SIN test that measures the speech reception threshold (SRT) where a listener can correctly identify 50% of digit triplets (e.g., 3-2-7) presented in speech noise. Compared to SIN tests that use sentences, English digits are easily understood, even in multilingual populations making it less dependent on the listener’s linguistic skills and suitable for a wide range of users (Smits et al., 2013). The DIN was first developed for landline telephone use as Netherland’s national hearing test in 2004 (Smits et al., 2004). Several telephone and online versions after that arose in countries like the USA, Australia, Germany, France and Poland (Folmer et al., 2017; Jansen et al. 2010; Smits et al. 2006a; Watson et al. 2012; Zokoll et al. 2012). A challenge related to telephone DIN tests, however, is insufficient landline coverage in LMICs. In South Africa for example, only 13.9% of people have access to a landline telephone (STATSSA, 2013). Capitalizing on rapid dispersal of low-cost smartphones with estimated 80% worldwide access by the year 2020 (The Economist, 2015), offering the DIN test as a smartphone application was a suitable alternative to landline testing. The first smartphone DIN test was made available in 2016 as South Africa’s national hearing test, called hearZA™ (Potgieter et al., 2016; Potgieter et al., 2017). The application can be completed within 3 minutes, correlates highly with pure tone average (0.5,1,2,4 kHz) and has sensitivity and specificity over 90%, sufficient for population-based screening (Potgieter et al., 2016; Potgieter et al., 2017).

As a free downloadable application on Android and iOS app stores hearZA™ has been marketed through sponsorships and endorsements by South African celebrities serving as hearing health ambassadors. The hearZA™ campaign and platform are utilized for several different purposes, of which hearing screening is only one. These purposes include serving as a 1) strategic public awareness tool for hearing health; 2) screening tool for hearing loss; 3) providing a personal hearing profile for tracking
hearing health; 4) decision support tool encouraging action on hearing loss developed in collaboration with the Ida Institute; 5) location-based referral system to connect persons to their closest hearing health providers in partnership with national audiological societies (Swanepoel, 2017).

Analysis of large-scale hearing screening programs, like the hearZA™ National Hearing Test, can provide indications for test-use, considerations for interpretation and measures of reach. The purpose of the study was therefore to determine characteristics and test performance of individuals that have tested themselves with the hearZA™ national hearing test App.

METHOD

The study was reviewed and received ethical approval from the Faculty of Humanities Research Ethics Committee, University of Pretoria.

Participants
A retrospective analysis of 30321 DIN tests completed from March 2016 to August 2017 was conducted. There were 297 listeners who indicated ages of 100 years and older, of which 283 incorrectly indicated ages of 2017 or 2018 years. The remaining 14 subjects had ages 100 up to 117 years. Their data, together with listeners who kept the application’s default setting of 27 years, were excluded as incorrect data entries (n=2175). Furthermore, 4075 listeners who completed the test without headphones or earbuds were excluded to prevent confounding variables. This resulted in a total of 24072 tests for analysis. The ages of the remaining sample ranged between 5 to 99 years (Average: 39 years; SD: 16.6 years; n=24072). Listeners were further divided into two groups based on their self-rated level of English speaking competence. English speaking competence is rated on a non-standardized visual analogue scale of 1 to 10, a lower score indicating poorer competence and a higher score, better competence. The two groups were 1) native (N) and non-native (NN) ≥6 and 2) NN ≤5 (Potgieter et al., 2017). Pass and fail criteria were based on that of Potgieter et al. (2017), N and NN ≥6 with a cut-off of -9.55 dB and N≤ 5 with a cut-off of -7.5 dB.

Procedures
Listeners completed the DIN test by downloading the application on an Android or iOS operated smartphone. Individuals were made aware of the App due to marketing
campaigns or word-of-mouth recommendations. Before the test, listeners were indicated to complete the test with either earbuds or headphones. Earlier versions of the App did not prevent test execution when earbuds or headphones are not plugged in, which was subsequently changed. The application requires listeners to provide their date of birth, home language and to rate their English speaking competence on a non-standardized scale of 1-10. Users were required to indicate whether they experience hearing difficulty by selecting “Yes” or “No” on the application. After that, listeners self-select a comfortable listening intensity. Based on the comfort level selected, the application uses a fixed overall level with an adaptive signal-to-noise ratio (SNR) (see Potgieter et al. 2016 for details). Masking noise is formed to match the long-term average speech spectrum of the digits. Speech signals are presented diotically. Listeners are expected to enter the digits heard on the smartphone, where a correct response will prompt the application to present the next digit at 2 dB lower SNR. Where the listener is uncertain of the digits heard, they are instructed guess. When the response is incorrect, the application presents the next triplet at 2 dB higher SNR. The result is calculated by averaging the last 19 SNRs (Potgieter et al., 2016; Potgieter et al., 2017).

Statistics
Data analysis was completed using Statistical Package of the Social Sciences (SPSS v23.0). Results were analyzed using one-way analysis of variance (ANOVA) and Analysis of Covariance (ANCOVA). Post hoc comparisons were made using the Bonferroni adjustment. A p-value of 0.05 was used to indicate significance. Multiple regression was performed for continuous and categorical variables to determine predictors of the SRT result. The relationships between percentile SRTs and age were determined by following the procedure as described in Smits et al. (2013). Briefly, percentile values were calculated for 11-year age groups. These age groups were shifted in 1-year steps from 20 to 94 years. The percentile values were calculated for each age group and a group-size weighted fit to an exponential growth function was performed.
RESULTS

Approximately one in four persons (22.4%) who took the hearZATM test ($n=24072$) failed based on normative data by Potgieter et al. (2017). Of persons for whom self-report of hearing status was available ($n=17611$), 37% reported a hearing difficulty of which 30.2% failed the test. Of those who did not self-report a hearing difficulty, 19.7% failed. The age distribution of persons taking the hearZATM test (figure 1) indicated that 33.2% were 30 years of age and younger, 40.5% were between 31 and 50 years of age while 26.4% were older than 50 years of age. Over 56 years, SRTs became worse and referral rates rapidly increased (Figure 2).

![Figure 1. Age distribution of persons taking the hearZATM test ($n=24072$)](image)
A one way ANOVA indicated a significant effect of age groups on the SRT results ($F[5,24066]=182; p<0.001$). Listeners were further evaluated according to their level of English speaking competence ($n=24072$). Most listeners selected English as their first language ($n=17832$). When divided into $N \& NN \geq 6$ ($n=22737$) and $NN \leq 5$ ($n=1335$) groups based on the categories used by Potgieter et al. (2017), significant mean SRT differences between the two groups were found. $NN \leq 5$ listeners performed significantly worse than $N \& NN \geq 6$ listeners (95% CI, 1.1 to 1.6; $t(1410.9)=-11.3; p<0.001$). Referral rate according to cut-off criteria by Potgieter et al. (2017) increased with age group, the highest referral rates in the 51-60 and 61-99-year groups (Table 1). After adjustment for age and level of English-speaking competence, there was a statistically significant difference ($F(2,24067)=132.89; p<0.001$; partial $\eta^2=0.011$) in SRTs between listeners without self-reported hearing difficulty ($n=11089$) and with a
self-reported hearing difficulty \((n=6522)\). SRTs in listeners with a self-reported hearing difficulty were significantly higher \((-9.5 \pm 0.04\) dB SNR) than in listeners without a reported hearing difficulty \((-10.2 \pm 0.02\) dB SNR), with a mean difference of 0.7 (95% CI, 0.6 to 0.8) dB SNR, \(p<0.001\).

Multiple regression was conducted to determine if English speaking competence, self-reported hearing difficulty, and age were significant predictors of the SRT result. Data of 6461 listeners who did not select whether they had a known hearing problem were excluded. In the remaining sample \((n=17611)\), all the variables contributed significantly to the prediction \((F(3,17067)=219.55, p<0.001)\); adj. \(R^2=0.04\). Linear regression analysis for N & NN ≥6 \((n=19430)\) and NN ≤5 \((n=1413)\) listeners were conducted for the same variables. For N & NN ≥6 listeners, the model was significant \((F(3,16705)=146.79, p<0.001)\); adj. \(R^2=0.02\) with age (beta 0.11; 95% CI 0.01 to 0.02; \(p<0.001\)) and self-reported hearing difficulty (beta -0.09; 95% CI -0.62 to -0.44; \(p<0.001\)) contributing significantly to the prediction. Overall the model was also significant for NN ≤5 listeners \((F(3,898)=13.66; p<0.001)\); adj. \(R^2=0.04\) with age (beta

<table>
<thead>
<tr>
<th>Age Category in years ((n))</th>
<th>Group mean SRT in dB SNR (SD)</th>
<th>Number of N/ NN English Listeners</th>
<th>Mean SRT in dB SNR (SD)</th>
<th>Referral Rate (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-15 (1441)</td>
<td>-8.5 (4.6)</td>
<td>N &amp; NN ≥6: 1275 NN ≤5: 166</td>
<td>-8.7 (4.4)</td>
<td>No Norms</td>
</tr>
<tr>
<td>16-30 (6543)</td>
<td>-10.2 (2.9)</td>
<td>N &amp; NN ≥6: 6217 NN ≤5: 326</td>
<td>-10.3 (2.8)</td>
<td>17.3% (1076)</td>
</tr>
<tr>
<td>31-40 (5893)</td>
<td>-10 (2.8)</td>
<td>N &amp; NN ≥6: 5629 NN ≤5: 264</td>
<td>-10.2 (2.8)</td>
<td>18.3% (1030)</td>
</tr>
<tr>
<td>41-50 (3840)</td>
<td>-9.8 (2.8)</td>
<td>N &amp; NN ≥6: 3662 NN ≤5: 178</td>
<td>-10.1 (2.7)</td>
<td>19.9% (729)</td>
</tr>
<tr>
<td>51-60 (3279)</td>
<td>-9.8 (2.8)</td>
<td>N &amp; NN ≥6: 3082 NN ≤5: 197</td>
<td>-9.8 (2.7)</td>
<td>26.4% (814)</td>
</tr>
<tr>
<td>61-99 (3670)</td>
<td>-8.7 (3.7)</td>
<td>N &amp; NN ≥6: 2872 NN ≤5: 204</td>
<td>-8.8 (3.6)</td>
<td>45.6% (1309)</td>
</tr>
</tbody>
</table>

\(N\); native, \(NN\); non-native SRT; speech reception threshold, SD; standard deviation

Table 1. Mean SRTs and referral rates of N&NN ≥6 and NN ≤5 English speakers according to age (referral according to Potgieter et al. In Press norms)
0.09; 95% CI 0.01 to 0.03; \( p<0.001 \)), self-reported English speaking competence rating (beta -0.15; 95% CI -0.67 to -0.28; \( p<0.001 \)) and self-reported hearing difficulty (beta -0.11; 95% CI -1.46 to -0.35) contributing significantly to the prediction.

DISCUSSION
The hearZA\textsuperscript{TM} national hearing test App had 30321 persons tested from 3 March 2016 to 14 August 2017. Overall the test had a referral rate of 22.4%. Compared to the 81% referral rate of the US national hearing test (Watson et al., 2015), the referral rate of hearZA\textsuperscript{TM} was low. It is possible that hearZA\textsuperscript{TM} targets a younger population group, where the prevalence of hearing loss is still low, explaining the overall low SRTs across age groups. The tests for which self-report of a hearing problem was available, 37% reported a known hearing difficulty. Most persons who took the test (40.5%) were between 31 and 50 years of age, and 27.2% were 30 years of age and younger. The median age of the listeners was 37 years. Compared to the Dutch national hearing test (Smits et al., 2006\textsuperscript{a}), the median age in this analysis was lower than both internet (40 years) and telephone version (54 years) of the test, suggesting the hearZA\textsuperscript{TM} App is reaching a younger population. Over 50 years of age, test uptake dropped with less than a third (26.4%) of the sample in this category. The same pattern was seen for the Dutch National Hearing Test (Smits et al., 2006\textsuperscript{a}). A study by Moore et al. (2015) linked poorer computer literacy with age among the elderly population. Although computer literacy does not necessarily relate to smartphone use, it reflects general digital literacy and limited ability and perhaps willingness to take self-testing using internet technology. Physical restrictions such as visual impairment and limited upper extremity dexterity, typical in the aging population, were also suggested contributors to limited use of Internet-based health provision (Or et al., 2011; Moore et al., 2015).

Of course, the hearZA\textsuperscript{TM} marketing efforts, mostly on digital media, also biases the sample to younger populations.

Referral rates remained constant for native and non-native English-speaking listeners between the ages of 16-50 years. Over the age of 50 years, however, hearing deteriorates quickly (Smits et al., 2006\textsuperscript{b}) and referral rates rapidly increase. In South Africa, a multilingual country with 11 official languages, native English speakers comprise only 9.6% of the population (STATSSA 2011). It was therefore interesting that the majority of listeners in this analysis selected English as their first language.
One apparent reason for the high rate of listeners designated as native English speakers in the data set is the fact that it was the default setting on the hearZA™ App (changed subsequently). Mantonakis et al. (2009) investigated how the order of choices affects selection and found that most individuals are likely to select the first option in a sequence. Another possibility could be attributed to the fact that the test is reaching mostly native English speakers because marketing campaigns were conducted and distributed in an English medium. In line with the study of Potgieter et al. (2017), self-rated English-speaking competence was a significant predictor of the SRT. Even though some listeners may have kept the default selection of English as first language, native and non-native listeners with English speaking competence equal or better than 6 still performed significantly better than non-native listeners with ratings equal or lower than 5. The self-reported English-speaking competence rating was also a significant predictor in non-native listeners with English speaking competence ≤ 5, poorer scores showing poorer SRTs.

The high test uptake, especially amongst younger population groups, and the high overall referral rate (22.4%) indicates that the hearZA™ App is addressing a public health need. Although developed for adult users, there was a substantial sample (n=1275) of listeners under 15 years of age. The ability to understand speech in noise is an intrinsic attribute of the auditory system that matures with age (Talarico et al., 2007). Koopmans et al. (2018) demonstrated the effect of age on the Dutch DIN test SRT for a group of more than 100 normal hearing children. It will, therefore, be essential to determine age-specific norms of the test for the users between the ages of 5-15 years since it is clear that parents want to use it for this purpose. Age, self-reported English-speaking competence and hearing difficulty were all significant factors influencing test outcomes. A limitation of this analysis, however, was that results assumed that listeners correctly entered personal information such as date of birth, self-reported hearing problem, native language, non-native language English speaking competence rating, over which researchers had no control. Furthermore, between the two groups with and without self-reported listening difficulty, although statistically different, the difference was small (0.7 dB). Large distributions of SRTs in both groups likely shifted the means of the two groups, another limitation of field testing. Moreover, Pronk, Deeg and Kramer (2018) found that self-report of hearing disability is influenced by other factors such as demographics, personality, mood and
social situation, causing discrepancies between self-report and DIN test results. This could also possibly account for the small difference between the two groups. Lastly, considering the effect of self-reported English-speaking competence, the SRT cut-off criteria set out for native and non-native speakers in Potgieter et al. (2017) are appropriate.

Overall, this study demonstrated widespread uptake of the hearZA™ App across age groups with a substantial number of persons self-reporting hearing problems (37%) and failing the test (22.4%). This means it is reaching an important target audience, those who think they have hearing loss. Conversely, it also reaches a high proportion of persons not yet presenting with a clear hearing problem but having taken the test are aware of their hearing status and can then track it through the App’s personal profile (Swanepoel, 2017). Users are also reminded annually (via in-app notifications) to conduct follow-up tests allowing for longitudinal tracking and the possibility of early detection of hearing problems. A failed DIN test result has demonstrated positive influence on uptake of interventions (Smits et al., 2005; Watson et al., 2015). This method of self-testing should therefore, promote increased self-efficacy and accessible hearing health behaviors among users.

ACKNOWLEDGEMENTS
Portions of this paper were presented at the 3rd International Internet & Audiology Meeting, Louisville, KY, July 2017. Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number 5R21DC016241-02.

REFERENCES


