International Journal of Audiology

ISSN: 1499-2027 (Print) 1708-8186 (Online) Journal homepage: https://www.tandfonline.com/loi/iija20

Evaluation of an internet-based speech-in-noise screening test for school-age children

Marya Sheikh Rashid, Wouter A. Dreschler & Jan A. P. M. de Laat

To cite this article: Marya Sheikh Rashid, Wouter A. Dreschler & Jan A. P. M. de Laat (2017) Evaluation of an internet-based speech-in-noise screening test for school-age children, International Journal of Audiology, 56:12, 967-975, DOI: 10.1080/14992027.2017.1378932

To link to this article: https://doi.org/10.1080/14992027.2017.1378932

Taylor & Francis
Taylor & Francis Group

International Journal of Audiology 2017; 56: 967–975

International Journal of Audiology

Original Article

Evaluation of an internet-based speech-in-noise screening test for school-age children

Marya Sheikh Rashid[1], Wouter A. Dreschler[1], and Jan A. P. M. de Laat[2]

[1]Clinical and Experimental Audiology, Amsterdam Public Health Research Institute, Academic Medical Center (AMC) Amsterdam, Amsterdam, The Netherlands and [2]Department of Audiology, Leiden University Medical Center, Leiden, The Netherlands

The British Society of Audiology

The International Society of Audiology

NAS NORDIC AUDIOLOGICAL SOCIETY

Abstract

Objective: To evaluate a Dutch online speech-in-noise screening test (in Dutch: ''Kinderhoortest'') in normal-hearing school-age children. Sub-aims were to study test–retest reliability, and the effects of presentation type and age on test results. *Design*: An observational cross-sectional study at school. Speech reception thresholds (SRTs) were obtained through the online test in a training condition, and two test conditions: on a desktop computer and smartphone. The order of the test conditions was counterbalanced. *Study sample*: Ninety-four children participated (5–12 years), of which 75 children were normal-hearing ($\leq$25 dB HL at 0.5 kHz, $\leq$20 dB HL at 1–4 kHz). *Results*: There was a significant effect for test order for the two test conditions (first or second test), but not for presentation type (desktop computer or smartphone) (repeated measures analyses, $F(1,75) = 12.48$, $p < 0.001$; $F(1,75) = 0.01$, $p = 0.982$). SRT significantly improved by age year (first test: 0.25 dB SNR, 95% CI: –0.43 to –0.08, $p = 0.004$. Second test: 0.29 dB SNR, 95% CI: –0.46 to –0.11; $p = 0.002$). *Conclusions*: The online test shows potential for routine-hearing screening of school-age children, and can be presented on either a desktop computer or smartphone. The test should be evaluated further in order to establish sensitivity and specificity for hearing loss in children.

Key Words: Hearing screening, online speech-in-noise test, school-age children, speech understanding in noise, maturation, smartphone

Introduction

Untreated mild to severe childhood hearing loss may have serious negative consequences for speech and language, educational and socio-emotional development (Davis et al. 1986; Brookhouser, Worthington, and Kelly 1991; Bess, Dodd-Murphy, and Parker 1998; Yoshinaga-Itano et al. 1998). Therefore, early identification is of great importance. Well-established neonatal hearing screening programmes in European countries, including otoacoustic emissions (OAE) or automated auditory brainstem response (AABR) screening, identify permanent congenital hearing losses, but they do not detect delayed-onset or acquired sensorineural losses (Skarzynski and Piotrowska 2012; Winston-Gerson and Sabo 2016). In the USA, the prevalence of mild permanent sensorineural hearing loss at 6 kHz in children aged 6–19 years is 12.5%, and in children aged seven years is 6% (Niskar et al. 2001). Up to 90%, more children are diagnosed with hearing loss before the age of nine years than are diagnosed as newborns (Fortnum et al. 2001). Early diagnosis of hearing loss can be achieved with hearing screening during pre-school, and primary school years, reducing the impact on speech and language development (Lu et al. 2014; Prieve et al. 2015). There are several screening methods to identify delayed-onset or acquired sensorineural hearing loss in pre-school- and school-age children. OAE and pure-tone screening are the most reliable and commonly used tools, though pure-tone screening is considered to be the preferred reference standard (Prieve et al. 2015). In the Netherlands, childhood hearing assessment is performed in all children between the age of four and six years. The assessment is performed at school by a youth health care nurse through pure-tone threshold screening at regular contact sessions.

When hearing screening is performed in the school setting, a large number of children can be reached (Winston-Gerson and Sabo 2016). However, pure-tone screening in remote settings, such as schools, is often performed in less than optimal test conditions. High ambient noise levels, but also calibration issues and examiner's and examinee's training, experience and motivation, negatively influence the accuracy of screening results, making pure-tone

---

Correspondence: Marya Sheikh Rashid, Clinical and Experimental Audiology, Amsterdam Public Health Research Institute, Academic Medical Center (AMC) Amsterdam, P.O. Box 22660, Amsterdam, the Netherlands. E-mail: m.sheikhrashid@amc.uva.nl

(Received 26 January 2017; revised 7 September 2017; accepted 7 September 2017)
ISSN 1499-2027 print/ISSN 1708-8186 online
DOI: 10.1080/14992027.2017.1378932

Abbreviation lists

AABR	automated auditory brainstem response
CI	confidence interval
CVC	consonant-vowel-consonant
HTML	hypertext markup language
OAE	otoacoustic emissions
PTA_{5124}	pure-tone average of 0.5, 1, 2, and 4 kHz
RMS	root mean square
SD	standard deviation
SNR	signal-to-noise ratio
SRT	speech-reception threshold

screening less reliable for detecting hearing losses (Bamford et al. 2007; Schlauch and Carney 2012; Kam et al. 2013; Prieve et al. 2015). Therefore, there is a need for appropriate, effective and efficient periodic hearing screening that can be performed accurately and reliably in school or other remote settings, to identify suspected mild to severe sensorineural hearing losses in pre-school and school-age children.

One of the early signs of hearing impairment is the difficulty experienced in understanding speech in background noise in daily situations (Smoorenburg 1992; Kramer, Kapteyn, and Festen 1998). Therefore, one potential approach to identify hearing loss is speech-in-noise testing. Advanced time-efficient online self-administered and automated speech-in-noise tests have been developed, that focus on the detection of sensorineural hearing losses (Leensen et al. 2011; Jansen 2013; Smits, Goverts, and Festen 2013). The main advantages of such tests are that the tests are easily accessible and less susceptible to environmental noise (Smits, Kapteyn, and Houtgast 2004; Culling, Zhao, and Stephens 2005). A Dutch online speech-in-noise hearing screening test for children has been developed by the Leiden University Medical Center and the Academic Medical Center in the Netherlands, and was implemented online in January 2007 (in Dutch: ''Kinderhoortest''). The test was developed with the aim of allowing the evaluation of children's speech perception in noise in an easy and accessible way at a remote setting, such as the school environment. The goal of such testing would be the early detection of perceptive sensorineural hearing loss in school-age children. The relatively simple test with suitable speech material may be useful for children aged five years and older. An important limitation is that this test may not be assumed to be sensitive to conductive hearing losses caused by external or middle ear pathologies, such as otitis media. According to the underlying model by Plomp and Mimpen (1979), speech-in-noise results do not lead to higher critical SNR's for pure conductive hearing losses. Conductive hearing losses are one of the potential forms of hearing losses in school-age children, though they are more common in pre-school-age children (Samelli et al. 2012). Most of the children experience temporary conductive hearing losses, which can be treated medically. The main objective of this study was to evaluate the suitability of the Dutch online speech-in-noise screening test for use in primary school children. The sub-aims were to evaluate the test–retest reliability of the test, the effect of the presentation type: on a desktop computer or smartphone, and to assess age effects on test results.

Methods

Subjects

This study was performed in 94 primary school children. Recruitment took place at the Koningin Wilhelminaschool in Rijnsburg, the Netherlands. Information letters, informed consent forms and short questionnaires were sent to the parents. All children were native speakers of the Dutch language. Speech-in-noise data were collected from 94 children. The results of 19 children were excluded from the analyses, because these children were younger than five years old ($N = 1$), had poor-hearing thresholds at one or more octave frequencies ($\geq$25 dB HL at 0.5 kHz and/or $\geq$20 dB HL at 1–4 kHz) for at least one ear ($N = 10$), had missing data on the speech-in-noise tests for at least one condition ($N = 2$), had instable SRT measurements for at least one test ($N = 5$) and had a floor score for at least one test ($N = 1$) (the definitions of an instable measurement and of a floor score are explained in the section ''Statistical analyses''). The data of the remaining 75 normal-hearing children were analysed further.

Measurement procedures

This cross-sectional study was approved by the medical ethics committee of Leiden University Medical Center (project number P11-108). Informed consent was given by parents and the school's board of directors. For every child, information of concerning age, gender and grade was collected. All tests took place during school hours in quiet rooms at the school.

Pure-tone audiometry

Pure-tone audiometry was performed as a reference standard. The five-year olds performed play audiometry. Hearing thresholds were measured for both ears at the frequencies 0.5, 1, 2 and 4 kHz. Bone-conduction thresholds were assessed as well when a hearing threshold was 20 dB HL or worse at one frequency. Pure-tone audiometry was performed in the teachers' office room, with an average background noise level of 43 dB(A). Because there was no soundproof cabin, and audiometric tests were potentially subject to environmental noise, a hearing threshold of 25 dB HL or better at 0.5 kHz was defined as normal. Between 1 and 4 kHz, a hearing threshold of 20 dB HL or better was defined as normal. Children with poorer thresholds were referred for further investigation, and parents or caregivers were informed. For the pure-tone audiometry measurements the Interacoustics AD229b audiometer was used with Telephonics TDH−39P headphones with Amplivox Audiocups to attenuate ambient sound, and a Radioear B71 bone conductor.

Speech-in-noise testing

Children's perception of speech in noise was assessed by means of the online speech-in-noise test for children. First, a training condition was performed in a group session in class, i.e., all children belonging to one grade performed the test in the same computer classroom at the same time, but each performed the test individually on a personal desktop computer. Spoken instructions on the test procedure were given by the research assistant before the training test started. The children were instructed to identify the presented words by clicking on the corresponding pictures on the screen. They were also instructed to click on the picture depicting a question mark if a presented word could not be identified. Then, in two test conditions, all children were tested with a desktop computer, and with a smartphone. For these two test conditions, children performed the test one by one, separate from the other children, in the teachers' office room. The order of the type of presentation was counterbalanced. The computer classroom in which the children underwent the training condition had an average background noise level of 48 dB(A). The two test conditions took place in the same teachers' office room in which the pure-tone audiometry took place, with an average background noise level of 43 dB(A). The online speech in noise tests was presented on a standard desktop computer and DKT Eduline of Philips SHP2000 headphones, and on the smartphones Nokia Lumia 625 or the Huawei G6, with Ewent headphones.

The speech material consisted of a closed set of eight Dutch monosyllable consonant-vowel-consonant (CVC) words. The words were all nouns, and highly familiar for young children, as they were selected from the Dutch word lists used for diagnostic speech audiometry in children (Bosman 1989). The response buttons on the screen were pictures accompanied by written words. The written words were: ''lion'', ''goat'', ''book'', ''rose'', ''moon'', ''thumb'', ''fire'' and ''chicken'' (in Dutch), and were all represented by easily recognisable pictures. To prevent guessing, a ninth response button with a question mark and the text ''not understood'' was added. The response screen is shown in Figure 1. In order to enhance test reliability, the words were perceptually homogenised (Sheikh Rashid and Dreschler 2014). To achieve equal intelligibility of the words, the presentation levels of the specific words were adjusted. These level corrections, based on the average SRTs for the individual words, were derived from the slopes of word-specific psychometric functions according to the method described in Leensen et al. (2011). The word-specific psychometric functions were based on online results of tests that were performed by children, from January 2007 to August 2014 ($N = 46{,}742$). Perceptually difficult words were amplified, and perceptually simple words were attenuated (the level corrections ranged between 1.51 and –2.55 dB). The words were presented in a masking noise, which was a broadband continuous noise, with a spectrum that corresponded with the long-term average speech spectrum of the homogenised word material. The test was diotic (binaural); i.e., both ears were measured at the same time. The volume level of the stimuli could be set by means of the volume scale to a comfortable level. The minimum and maximum volume levels were controlled for in the clinical setting, and set at 15 dBA (i.e., whisper level) to 85 dBA (without distortion of the sounds). To familiarise listeners with the test and the response buttons, the words were presented in a masking noise prior to the test. The test consisted of 20 stimuli. All words were randomised, and each word was presented two or three times. The test started with a signal-to-noise ratio (SNR) of –1 dB and the intensity of the word was being varied by means of an up–down procedure in steps of 2 dB SNR. The intensity of the masking noise was fixed. The speech reception threshold (SRT in dB SNR) at which 50% of the material is correctly understood, was based on the mean SNR values of the last 10 presentations. The test was presented in an HyperText Markup Language (HTML) format, and could, therefore, be performed on any electronic device that supported the

Figure 1. Response screen of the Dutch online speech-in-noise test for school-age children (In Dutch: ''Kinderhoortest'').

format, such as a desktop computer, tablet, or smartphone. Test duration was approximately 3 min in all age groups.

Statistical analyses

Data were analysed using SPSS (IBM SPSS Statistics 20 and 22; IBM Corp, Armonk, NY, USA). Results of normal-hearing children between the age of 5 and 12 years old were analysed. The data of children with incomplete or invalid test results, due to instable measurements or a floor effect, were excluded from the analyses. An instable measurement refers to an intra-individual standard deviation of 3 dB or larger. A floor effect refers to a minimal SRT score. A floor effect can be the result of consecutive incorrect responses due to a hearing loss or not understanding the test procedure.

Descriptive analyses were performed on hearing thresholds and speech SRTs of the subjects. The normality assumption was assessed by means of Q-Q plots and goodness of fit tests. SRT data showed normal distributions. Therefore, General Linear Model Repeated measures analyses were performed on the two test conditions to analyse the effect of the type of presentation: desktop computer or smartphone (within-subject factor), the order: the first or second test (within-subject factor) and age in categories (between-subject factor) on SRT (in dB SNR). Post hoc analyses with Bonferroni corrections were performed when significant effects were found. To analyse SRT (in dB SNR) as a function of age (in years) and test (training condition, first and second tests), multiple regression analyses were performed. In order to assess the consistency of the first test and second results, a measurement error was calculated by taking the quadratic mean of the within-subject standard deviations of the repeated measurements. Finally, in order to assess age-related differences, a regression analysis was performed on SRT scores (in dB SNR) of the first and second test, as a function of age (in years).

Table 1. Age group (in years), number of participants per age group, and mean pure-tone average (PTA) for the frequencies 0.5, 1, 2, and 4 kHz (PTA_{5124}) (in dB HL) (SD) for right and left ear.

		PTA_{5124}	
Age group	N *(75)*	*Right ear*	*Left ear*
5	7	12.9 (1.2)	13.6 (2.3)
6	13	8.8 (5.3)	9.2 (4.6)
7	10	7.5 (5.0)	6.3 (3.8)
8	11	5.7 (5.4)	5.6 (4.7)
9	13	5.7 (5.3)	7.8 (5.5)
10	11	5.7 (5.1)	5.6 (6.0)
≥11	10	8.8 (3.8)	8.0 (2.8)

Results

Table 1 shows the pure-tone average (PTA) thresholds for the octave frequencies 0.5, 1, 2 and 4 kHz (PTA_{5124}). The mean results are given per age group and per ear. At least 10 children participated per age group, except for the youngest age group ($N = 7$). The 11–12-year olds ($N = 8$ and $N = 2$, respectively) are clustered in the oldest age group (≥11 years).

SRT scores for all test conditions

The mean SRT scores (in dB SNR) for the training condition and the two counterbalanced test conditions (first and second tests, and on desktop computer and smartphone) were calculated for each age group (Table 2). Children performed better on both test conditions (desktop computer and smartphone) as compared to the training condition, with a significant difference in mean SRT of –1.5 dB SNR ($F(2,75) = 17.64$, $p < 0.001$). For the two-test conditions, there was no significant main effect for type of presentation (desktop computer or smartphone) ($F(1,75) = 0.01$, $p = 0.982$), but there was a significant main effect for test order (first or second test) ($F(1,75) = 12.48$, $p < 0.001$). The mean SRT scores of the second test were significantly better than the SRT scores on the first test with a difference of 0.7 dB SNR (95% CI: 0.3–1.1; $p = 0.001$). The main effect of age was significant as well ($F(6,75) = 3.09$, $p = 0.01$). Post hoc analyses showed that the 10–11-year olds performed better as compared to the 5–6-year olds (with a difference of 1.9 dB SNR, 95% CI: –3.7 to –0.1; $p = 0.023$), and to the 6–7-year olds (with a difference of 1.6 dB SNR, 95% CI: –3.1 to –0.1; $p = 0.028$). There were no significant interaction effects between age and type of presentation ($F(6,75) = 0.67$, $p = 0.674$), or age and test order ($F(6,75) = 1.06$, $p = 0.393$).

The results of the multiple regression analysis with SRT (in dB SNR) as a function of age (in years) and test (training condition, first test and second test) are shown in Table 3. According to the model, the mean SRT score for a five-year-old child in the first test session was –12.6 dB SNR. There was a significant improvement (decrease) in mean SRT score of 0.3 dB SNR per age year. Performance on the first test was 1.2 dB SNR better than on the training condition. Performance on the second test was 0.7 dB SNR better as compared to the first test. These differences were

Table 2. Mean SRT (in dB SNR) (SD) per age group (in years), for training and two test conditions.

	Mean SRT (dB SNR) (SD)				
		Test order		*Presentation type*	
Age group	*Training*	*First*	*Second*	*Desktop computer*	*Smartphone*
5	–11.8 (2.4)	–12.6 (1.2)	–13.8 (1.8)	–13.3 (0.6)	–13.1 (2.2)
6	–11.2 (2.6)	–13.2 (1.2)	–13.8 (1.1)	–13.9 (1.0)	–13.1 (1.3)
7	–12.0 (2.4)	–13.8 (1.3)	–14.2 (1.1)	–14.0 (1.1)	–14.0 (1.3)
8	–12.6 (1.6)	–13.9 (1.3)	–14.1 (1.6)	–13.6 (1.8)	–14.3 (1.0)
9	–12.9 (1.8)	–13.9 (1.6)	–14.0 (1.5)	–14.0 (1.5)	–13.9 (1.5)
10	–14.0 (1.3)	–14.7 (1.1)	–15.5 (1.8)	–15.1 (1.5)	–15.2 (1.6)
≥11	–13.6 (1.5)	–13.8 (2.2)	–15.5 (1.4)	–14.6 (2.5)	–14.8 (1.4)
Total	–12.6 (2.1)	–13.8 (1.5)	–14.4 (1.6)	–14.1 (1.6)	–14.1 (1.6)

statistically significant. The measurement error between the first and second tests was 1.3 dB.

Age-related differences

In Figure 2, the SRT score by age in percentiles is given for the first test. To analyse age-related differences in SRT, a regression analysis was performed, with SRT (in dB SNR) of the first test as outcome measure and age (5–11 years) as an explaining factor. According to this model, there is a significant improvement of 0.3 dB SNR in mean SRT score per age year ($\beta = 0.25$, 95% CI: –0.43 to –0.08; $p = 0.004$, $R^2 = 0.11$). A comparable age effect was observed in the second test ($\beta = 0.29$, 95% CI: –0.46 to –0.11; $p = 0.002$, $R^2 = 0.13$).

To establish age-corrected cut-off values for pass-refer criteria for the screening test, the 90th percentile of the SRT results of the test for six-year olds was used as a starting point. The beta-value of –0.25 dB SNR per age year for the first test was then used to correct for age. The age-corrected cut-off values are presented in Table 4. These cut-off values were then applied to the results of the first test of the 75 children. Based on this categorisation, 92% ($N = 69$) of the normal-hearing children passed the test.

Table 3. Multiple regression analysis with SRT (in dB SNR) as a function of age (in years) and condition (training, first and second test) (reference = first test).

	β	*p*	*95% CI*	
Constant	–12.58	<0.001	–13.13	–12.03
Age (in years)	–0.33	<0.001	–0.44	–0.21
Condition				
Training	1.16	<0.001	0.63	1.68
Second test	–0.68	0.013	–1.20	–0.15

Explained variance $R^2 = 0.27$.

Table 4. Age-corrected cut-off values.

Age category (years)	*Cut-off value (score positive result) (dB SNR)*
5–6	>–11.00
6–7	>–11.25
7–8	>–11.50
8–9	>–11.75
9–10	>–12.00
10–11	>–12.25
11–12	>–12.50

Discussion

This study focussed on the practical evaluation of the Dutch online speech-in-noise screening test in normal-hearing school-age children of 5–12 years old. To assess the reliability of the test, the test was performed at a primary school: first, a group training session was performed on a desktop computer, than two test conditions were performed, on a personal desktop computer and on a smartphone. The order of the test conditions was counter-balanced. The two tests were performed better as compared to the training, with a difference between the average of both tests, and the training test of 1.5 dB SNR. There were no significant differences in SRT score by type of presentation. There was an effect of test order between the two test conditions, indicating an additional learning effect after training of 0.7 dB SNR. A measurement error of 1.3 dB was found, indicating reasonable test–retest reliability. The standard deviations on the test and retest conditions were smaller than the stepsize that was used in the adaptive procedure (2 dB), indicating the homogeneity of the participant's results. Furthermore, according to the regression analysis, the oldest children had better SRT scores as compared to the youngest children, with a difference in the order of 1.5 dB SNR for the first test.

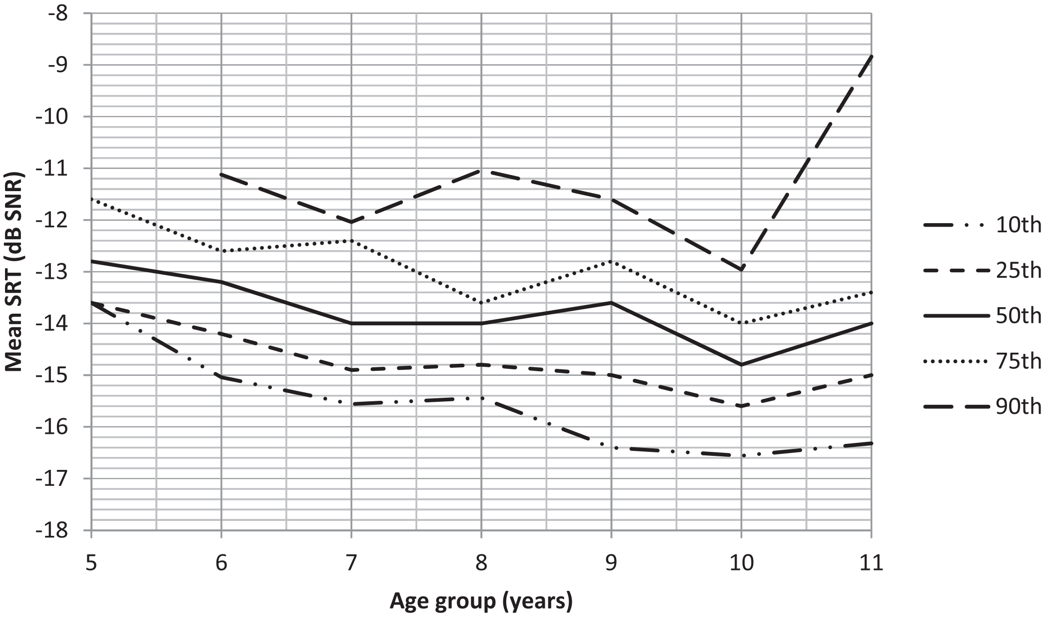

Figure 2. Reference values. Mean SRT in dB SNR by age for the first test. Distribution in percentiles (10th, 25th, 50th, 75th, and 90th).

Presentation mode

According to this study, the type of presentation (i.e., electronic device) did not influence SRT score. The test can be performed either on a desktop computer or on a smartphone combined with commercial headphones. The expectation that the test can be delivered on all types of electronic devices that support HTML applications is supported by the results of this study. Also, the children did not experience any difficulties in using the different electronic devices. According to a study by Culling, Zhao, and Stephens (2005), variations in equipment and listening environment do not present any significant obstacles to the development of a self-administered screening test based on speech in noise. There are several hearing screening tests delivered on different types of electronic devices (Leensen et al. 2011; Jansen 2013; Smits, Goverts, and Festen 2013, Potgieter et al. 2015). Studies on these computer- and smartphone-based speech-in-noise screening tests have shown that there are indeed no significant effects of transducer type on test outcome. Jansen (2013) showed that there is no significant effect of transducer type (headphones, built-in laptop speakers, in-ear phones and external speakers) on SRT in uncontrolled circumstances, for the Flemish computer-based digit triplet test. Recently, a smartphone-based digits-in-noise hearing test in South African English has been developed and validated (Potgieter et al. 2015). It was investigated whether different types and quality of headphones, including standard smartphone headphones and clinical headphones, would influence SRT. Statistically significant effects were not found. The South African smartphone-based screening test is based on the digit triplet test, developed by Smits, Goverts, and Festen (2013). Although the current test uses CVC words in noise, it is based on the same principles of speech-in-noise hearing testing.

Test–retest reliability

To assess the test–retest reliability of the speech-in-noise test, the first test was compared to the second test. Children performed significantly better on the second test as compared to the first test, indicating a learning effect of 0.7 dB SNR. It is important to note that the children were already familiar with the test procedure and the word material, because of the training condition that was performed prior to the two test conditions. For this reason, it cannot be ruled out that the actual learning effect is even greater than the learning effect found between the first and the second tests. It is unclear to what extent the training condition may have influenced test results. The difference of 1.2 dB SNR found between the training condition and the first test indicates an initial learning effect, but part of this difference could be ascribed to other factors related to testing together in one classroom, such as distraction.

Test–retest reliability of speech-in-noise tests in children has been studied earlier. Schafer et al. (2012) assessed the test–retest reliability of the Phrases in Noise Test (PINT) in normal-hearing children, a speech-recognition test for use in a clinical or educational setting. The PINT seemed fairly reliable as differences between two lists were within 3 dB SNR for 90% of the children. A smaller learning effect was found in the current study. This may be due to the use of a closed set of highly familiar words instead of sentences that have higher linguistic demands (Smits, Goverts, and Festen 2013). Jansen (2013) investigated the feasibility of the digit triplet test as an automated self-test in school-age children, and found a measurement error of 0.5–0.7 dB for different age groups. The smaller measurement error may be a result of the use of digit triplets as speech material instead of single CVC words, leading to more reliable estimates of the SRT (Jansen et al. 2014).

Age-related effects

In this study, age-related effects were present; the older children outperformed the younger children in all test conditions. There were no significant interactions with presentation type or order; age effects were consistently the same in all conditions, and were also present in the first and second tests. Several studies have been focussing on (school-age) children's ability to recognise speech in noise, and age-effects in auditory processing abilities (Elliott 1979; Elliott et al. 1979; Fallon, Trehub, and Schneider 2000; Johnson 2000; Talarico et al. 2007; Vaillancourt et al. 2008; Schafer et al. 2012; Jansen 2013; Koopmans, Goverts, and Smits 2014). In these studies, several auditory tasks and speech-in-noise tests were performed in different noise conditions. The majority of these studies has demonstrated maturation of the auditory system of normal-hearing children. Speech-in-noise recognition tends to improve with age and adult-like performance is reached in adolescence, depending on the speech-in-noise listening condition (Fallon, Trehub, and Schneider 2000; Johnson 2000; Talarico et al. 2007; Vaillancourt et al. 2008). Fallon, Trehub, and Schneider (2000) found that five-year old children required SNRs that were 5 dB more favourable than those of adults to obtain comparable performance on low-context sentences presented in background babble. Talarico et al. (2007) investigated the effect of age and cognition in 6- to 16-year olds with a task that included (non-)words in noise, varying in confusability and difficulty. Mean SNR scores decreased across all age groups, indicating better speech-in-noise recognition in the older children (up to 3 dB SNR). No correlations were found between speech in noise conditions and IQ scores. According to Elliot (1979), there are developmental changes in SRTs of young children up to 16 years of age. Age effects on SRT performance are mainly explained by developing auditory processing abilities, associated with developing linguistic skills, and cognition-related abilities such as memory capacities, experience and attention (Elliott 1979; Elliott et al. 1979; Boothroyd 1997; Hnath-Chisolm, Laipply, and Boothroyd 1998; Eisenberg et al. 2000; Fallon, Trehub, and Schneider 2000; Vaillancourt et al. 2008).

For the test session in this study, five-year olds required SNRs that were 1.5 dB more favourable than those of 11–12-year olds to achieve comparable 50% of correct performance. Jansen (2013) assessed the reference SRT for normal-hearing listeners for the digit triplet test and found that the SRTs of the 5th graders are 0.6 dB SNR worser as compared to those of the 7th graders. In the present study, a comparably small age-effect was found. This may be due to the relatively simple test procedure and the use of a closed-set of highly familiar, short and context-free monosyllabic words, supported by visual response buttons with pictograms and written words. The influence of linguistic abilities is expected to be small in this type of task (Fallon, Trehub, and Schneider 2000; Jansen 2013). The age-effect found in our study may be mainly a result of immature auditory perceptual abilities, combined with the influence of attentional limitations (Schafer et al. 2012; Jansen 2013), and the difficulty experienced in understanding test instructions in younger children.

Study limitations

This research has some limitations. First, the inclusion of normal-hearing children was based on hearing thresholds measured by means of pure-tone audiometry, which was not performed in an sound-isolated booth, but in the teachers' room. Although the room was considered quiet, as confirmed by the ambient noise-level measurements that were performed, environmental noise could not be completely avoided, and this may have influenced the pure-tone measurements. To attenuate the ambient noise, audio cups were used in combination with the headphones. Also, to assure normal-hearing, the criteria for normal hearing (at the lower octave frequencies) were adjusted. Looser threshold criteria are reasonable in school settings (Kam et al. 2013). However, children may have performed worse than they would have under optimal test conditions, which implies that possibly some normal-hearing children may have been excluded.

According to the pure-tone screening, 10 children had poor-hearing thresholds at one or more octave frequencies (≥ 25 dB HL at 0.5 kHz and/or ≥ 20 dB HL at 1–4 kHz) for at least one ear, and their results were, therefore, excluded from the analyses. Based on the established age-corrected pass-refer criteria, only one of them failed the online speech-in-noise test (age =5 years, SRT score for the first test= –10.8 dB SNR). This child had a PTA_{512} of 20 dB HL for the right ear, and 23 dB HL for the left ear. The large number of false-negatives, however, could possibly be explained by the less reliable test environment of the pure-tone screening (i.e., the false-negatives could actually be true-negatives). Another explanation could be that, since the test was binaural, children with an unilateral hearing loss were still able to pass the test (four out of the remaining nine children had an unilateral hearing loss). This may be an important limitation of the test. Also, the majority of the children with a bilateral hearing loss had a relatively small, and educationally insignificant, hearing loss, and probably were, therefore, still able to perform well on the online speech-in-noise test. In order to assess an optimal cut-off point for a dichotomous pass/fail outcome with a proper trade-off between sensitivity and specificity for clinically relevant hearing losses, it is necessary to include a large representative sample of children showing a wide range of hearing thresholds. This cut-off point would probably correspond to a higher degree of hearing loss than the relatively strict criteria that were proposed in this study.

Another limitation is that the youngest age group (the five-year olds) was underrepresented in this study as compared to the other age groups. Results of this age group may be less reliable. Also, the five-year olds had some trouble understanding the test instructions. For these young children, it was difficult to understand the goal of the test and the procedures. As the suitability of the test is still unclear for young children, reference values are only established for children older than five years. It is important to evaluate the established reference SRT values in larger populations. Also, it is important to have simple, clear and understandable instructions for the youngest children.

Finally, due to the setup of this research, it is difficult to distinguish a learning effect from the effect of test condition for the training condition versus the first test. Testing in a classroom setting may be less reliable than separate testing in a teacher's office room, due to distractions in the classroom that may hinder children's listening and focussing abilities (Knecht et al. 2002). According to Culling, Zhao, and Stephens (2005), group presentation of speech-in-noise tests in classrooms should be discouraged, mainly because of the potential negative effect of high levels of room reverberation. The training effect as well as simultaneous group testing versus separate testing in different settings need to be explored further. To assess test–retest reliability in more detail, tests and (multiple) retests should be performed under the same conditions, and with different time intervals.

Implications for practice and future research

The speech-in-noise test has important implications for hearing loss screening purposes in school-age children. In the Netherlands, online-speech-in-noise tests are already being used frequently to raise awareness in teenagers and adolescents (De Laat, Van Deelen, and Wiefferink 2016), but not yet for screening purposes. The current test appears to be suitable to be used in a national hearing screening programme, as it is a simple test, appropriate for small children, which can be performed in 3 min when performed binaurally. Due to the type of speech material, the influence of cognition, attention and linguistic demands is minimal. The independency of the test for soundproof test rooms and type of presentation creates opportunities for time-efficient simultaneous group testing and screening in remote settings. However, before the test can be implemented as a screening test, it is important to assess its sensitivity and specificity for detecting clinically relevant or educationally significant degrees and types of hearing losses in children. The test, therefore, needs to be evaluated in a larger representative sample of school-age children, including hearing-impaired children with a large range of hearing losses, in a realistic testing environment such as a school setting. The hypothesis is that children with hearing loss will have higher SNRs as compared to normal-hearing children; however, differences in performance should be investigated. In addition, the current test was conducted binaurally. However, in order to detect unilateral hearing losses, the test should be evaluated when conducted monaurally, as well. Furthermore, to establish sensitivity and specificity, it is of great importance to compare the speech-in-noise test with a reliable reference standard. Therefore, pure-tone audiometry should be performed in better test conditions as compared to the test conditions in the current study.

This research shows an age-dependency for SRT in normal-hearing children (an amelioration of –0.25 dB SNR per age year). The test result can be misleading if this is not corrected for. Therefore, the suggestion is to use age-corrected cut-off values in order to prevent false interpretations of positive test results of young children. The proposed age-corrected SRT cut-off values need to be validated in hearing-impaired children as well. Furthermore, this research indicated a learning effect. Learning effects may be accounted for by training or repeated conditional testing, i.e., introducing an automatic retest for children who failed the test. The possible influence of a learning effect on screening test outcomes should be studied further.

Conclusions

The online-speech-in-noise test with simple word material was shown to be appropriate for use in school-age children, and shows potential for a routine-hearing screening test. The test can be conducted simultaneously in a classroom setting, and can be delivered on either a desktop computer or on a smartphone in combination with commonly available headphones. When testing, age and learning effects should be considered. Age-corrected SRT cut-off values for pass/refer categories are proposed for screening purposes. A learning effect exists which could be reduced by training and/or conditional repeated testing. The test should be evaluated further in a larger representative population of school-age children, including hearing-impaired children, in order to evaluate its sensitivity and specificity for identifying childhood-hearing loss in realistic screening settings.

Acknowledgements

The authors would like to thank the pupils and tutors of the Koningin Wilhelminaschool in Rijnsburg for their cooperation. The authors would also like to thank M. Oomen and E. Heuninck of KU-Leuven, J. Braakman, K. Heiligenberg and M. Meijering of Windesheim in Zwolle, for performing the measurements, and Zicht Online for implementing and adapting the online tests. This study was facilitated by the Netherlands Hearing Health Foundation and funded by the Dutch Ministry of Social Affairs.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Bamford, J., H. Fortnum, K. Bristow, J. Smith, G. Vamvakas, L. Davies, R. Taylor, et al. 2007. ‘‘Current Practice, Accuracy, Effectiveness and Cost-effectiveness of the School Entry Hearing Screen’’. *Health Technology Assessment* 11 (32).

Bess, F. H., J. Dodd-Murphy, and R. A. Parker. 1998. ‘‘Children with Minimal Sensorineural Hearing Loss: Prevalence, Educational Performance, and Functional Status.’’ *Ear & Hearing* 19: 339–354.

Boothroyd, A. 1997. ‘‘Auditory Development of the Hearing Child.’’ *Scandinavian Audiology* 26: 9–16.

Bosman, A. J. 1989. ‘‘*Speech Perception by the Hearing Impaired.’’ Ph.D. diss*. [in Dutch]. University of Utrecht.

Brookhouser, P. E., D. W. Worthington, and W. J. Kelly. 1991. ‘‘Unilateral Hearing Loss in Children.’’ *The Laryngoscope* 101: 1264–1272.

Culling, J. F., F. Zhao, and D. Stephens. 2005. ‘‘The Viability of Speech-in-Noise Audiometric Screening Using Domestic Audio Equipment.’’ *International Journal of Audiology* 44: 691–700.

Davis, J. M., J. Elfenbein, R. Schum, and R. A. Bentler. 1986. ‘‘Effects of Mild and Moderate Hearing Impairments on Language, Educational, and Psychosocial Behavior of Children.’’ *Journal of Speech and Hearing Disorders* 51: 53–62.

De Laat, J. A., L. Van Deelen, and K. Wiefferink. 2016. ‘‘Hearing Screening and Prevention of Hearing Loss in Adolescents.’’ *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine* 59: 243–245.

Eisenberg, L. S., R. V. Shannon, A. S. Martinez, J. Wygonski, and A. Boothroyd. 2000. ‘‘Speech Recognition with Reduced Spectral Cues as a Function of Age.’’ *The Journal of the Acoustical Society of America* 107: 2704–2710.

Elliott, L. L. 1979. ‘‘Performance of Children Aged 9 to 17 Years on a Test of Speech Intelligibility in Noise Using Sentence Material with Controlled Word Predictability.’’ *The Journal of the Acoustical Society of America* 66: 651–653.

Elliott, L. L., S. Connors, E. Kille, S. Levin, K. Ball, and D. Katz.1979. ‘‘Children’s Understanding of Monosyllabic Nouns in Quiet and in Noise.’’ *The Journal of the Acoustical Society of America* 66: 12–21.

Fallon, M., S. E. Trehub, and B. A. Schneider.2000. ‘‘Children’s Perception of Speech in Multitalker Babble.’’ *The Journal of the Acoustical Society of America* 108: 3023–3029.

Fortnum, H. M., A. Q. Summerfield, D. H. Marshall, A. C. Davis, J. M. Bamford, A. Davis, C. Yoshinaga-Itano, and S. Hind. 2001. ‘‘Prevalence of Permanent Childhood Hearing Impairment in the United Kingdom and Implications for Universal Neonatal Hearing Screening: Questionnaire Based Ascertainment Study.’’ *BMJ* 323: 536–540.

Hnath-Chisolm, T. E., E. Laipply, and A. Boothroyd. 1998. ‘‘Age-Related Changes on a Children’s Test of Sensory-Level Speech Perception Capacity.’’ *Journal of Speech, Language, and Hearing Research* 41: 94–106.

Jansen, S. 2013. *Efficient and Sensitive Hearing Assessment Based on Speech Perception. Ph.D. diss*. KU Leuven.

Jansen, S., H. Luts, P. Dejonckere, A. Van Wieringen, and J. Wouters. 2014. ‘‘Exploring the Sensitivity of Speech-in-Noise Tests for Noise-Induced Hearing Loss.’’ *International Journal of Audiology* 53: 199–205.

Johnson, C. E. 2000. ‘‘Children’s Phoneme Identification in Reverberation and Noise.’’ *Journal of Speech, Language, and Hearing Research* 43: 144–157.

Kam, A. C. S., H. Gao, L. K. C. Li, H. Zhao, S. Qiu, and M. C. F. Tong. 2013. ‘‘Automated Hearing Screening for Children: A Pilot Study in China.’’ *International Journal of Audiology* 52: 855–860.

Knecht, H. A., P. B. Nelson, G. M. Whitelaw, and L. L. Feth. 2002. ‘‘Background Noise Levels and Reverberation Times in Unoccupied Classrooms: Predictions and Measurements.’’ *American Journal of Audiology* 11: 65–71.

Koopmans, W., T. Goverts, and C. Smits. 2014. *Spraakverstaan in complexe luistersituaties bij basisschoolkinderen* (in Dutch). Amsterdam, the Netherlands: VU Medisch Centrum.

Kramer, S. E., T. S. Kapteyn, and J. M. Festen. 1998. ‘‘The Self-Reported Handicapping Effect of Hearing Disabilities.’’ *Audiology: Official Organ of the International Society of Audiology* 37: 302–312.

Leensen, M. C., J. A. De Laat, A. F. Snik, and W. A. Dreschler. 2011. ‘‘Speech-in-Noise Screening Tests by Internet, Part 2: Improving Test Sensitivity for Noise-Induced Hearing Loss.’’ *International Journal of Audiology* 50: 835–848.

Lu, J., Z. Huang, Y. Ma, Y. Li, L. Mei, G. Yao, Y. Wang, X. Shen, and H. Wu. 2014. ‘‘Comparison Between Hearing Screening-Detected Cases and Sporadic Cases of Delayed-Onset Hearing Loss in Preschool-Age Children.’’ *International Journal of Audiology* 53: 229–234.

Niskar, A. S., S. M. Kieszak, A. E. Holmes, E. Esteban, C. Rubin, and D. J. Brody. 2001. ‘‘Estimated Prevalence of Noise-Induced Hearing Threshold Shifts Among Children 6 to 19 Years of Age: The Third National Health and Nutrition Examination Survey, 1988–1994, United States.’’ *Pediatrics* 108: 40–43.

Plomp, R., and A. M. Mimpen. 1979. ‘‘Improving the Reliability of Testing the Speech Reception Threshold for Sentences.’’ *International Journal of Audiology* 18: 43–52.

Potgieter, J. M., W. Swanepoel de, H. C. Myburgh, T. C. Hopper, and C. Smits. 2015. ‘‘Development and Validation of a Smartphone-Based Digits-in-Noise Hearing Test in South African English.’’ *International Journal of Audiology* 55: 405–411.

Prieve, B. A., T. Schooling, R. Venediktov, and N. Franceschini. 2015. ‘‘An Evidence-Based Systematic Review on the Diagnostic Accuracy of Hearing Screening Instruments for Preschool- and School-Age Children.’’ *American Journal of Audiology* 24: 250–267.

Samelli, A. G., C. M. Rabelo, M. B. Pereira, M. N. Portela, S. G. G. Sanches, and I. F. Neves-Lobo. 2012. ‘‘Comparison of Screening Methods for Conductive Hearing Loss Identification in Children: Low-Cost Proposal.’’ *Journal of Medical Screening* 19: 1–7.

Schafer, E. C., S. Beeler, H. Ramos, M. Morais, J. Monzingo, and K. Algier. 2012. ‘‘Developmental Effects and Spatial Hearing in Young Children with Normal-Hearing Sensitivity.’’ *Ear and Hearing* 33: e32–e43.

Schlauch, R. S., and E. Carney. 2012. ‘‘The Challenge of Detecting Minimal Hearing Loss in Audiometric Surveys.’’ *American Journal of Audiology* 21: 106–119.

Sheikh Rashid, M., and W. A. Dreschler. 2014. *Evaluatie Kinderhoortest* (in Dutch). The Hague, The Netherlands: Nationale Hoorstichting.

Skarzynski, H., and A. Piotrowska. 2012. ‘‘Screening for Pre-School and School-Age Hearing Problems: European Consensus Statement.’’ *International Journal of Pediatric Otorhinolaryngology* 76: 120–121.

Smits, C., T. S. Kapteyn, and T. Houtgast. 2004. ‘‘Development and Validation of an Automatic Speech-in-Noise Screening Test by Telephone.’’ *International Journal of Audiology* 43: 15–28.

Smits, C., S. T. Goverts, and J. M. Festen. 2013. ‘‘The Digits-in-Noise Test: Assessing Auditory Speech Recognition Abilities in Noise.’’ *The Journal of the Acoustical Society of America* 133: 1693–1706.

Smoorenburg, G. F. 1992. ‘‘Speech Reception in Quiet and in Noisy Conditions by Individuals with Noise-Induced Hearing Loss in Relation to Their Tone Audiogram.’’ *The Journal of the Acoustical Society of America* 91: 421–437.

Talarico, M., G. Abdilla, M. Aliferis, I. Balazic, I. Giaprakis, T. Stefanakis, K. Foenander, D. B. Grayden, and A. G. Paolini. 2007. ‘‘Effect of Age and Cognition on Childhood Speech in Noise Perception Abilities.’’ *Audiology and Neurotology* 12: 13–19.

Vaillancourt, V., C. Laroche, C. Giguere, and S. D. Soli. 2008. ‘‘Establishment of Age-Specific Normative Data for the Canadian French Version of the Hearing in Noise Test for Children.’’ *Ear and Hearing* 29: 453–466.

Winston-Gerson, R., and D. L. Sabo. 2016. ‘‘Hearing Loss Detection in Schools and Early Child Care Settings: An Overview of School-Age Hearing Screening Practices.’’ *NASN School Nurse* 31: 257–262.

Yoshinaga-Itano, C., A. L. Sedey, D. K. Coulter, and A. L. Mehl. 1998. ‘‘Language of Early- and Later-Identified Children with Hearing Loss.’’ *Pediatrics* 102: 1161–1171.