Detection of hearing problems in Aboriginal & Torres Strait Islander children: A comparison between clinician-administered and self-administrated hearing tests

Kiri Mealings[a], Samantha Harkus[b], Brooke Flesher[c], Alea Meyer[d], King Chung[d], and Harvey Dillon[a, e, f]

[a]*National Acoustic Laboratories, Sydney, Australia*

[b]*Australian Hearing, Sydney, Australia*

[c]*Purdue University, United States*

[d]*Northern Illinois University, United States*

[e]*University of Manchester, Manchester, United Kingdom*

[f]*Macquarie University, Sydney, Australia*

Acronyms:

AutoAud Automatic Audiometry

dB HL: Decibel hearing level

LiSN-S: Listening in Spatialized Noise – Sentences test

SPD: Spatial processing disorder

Detection of hearing problems in Aboriginal & Torres Strait Islander children: A comparison between clinician-administered and self-administrated hearing tests

Objective: This study evaluated the agreement of self-administered tests with clinician-administered tests in detecting hearing loss and speech-in-noise deficits in Aboriginal & Torres Strait Islander children.

Design: Children completed clinician-administered audiometry, self-administered automatic audiometry (AutoAud), clinician-administered Listening in Spatialized Noise – Sentences test, and self-administered tablet-based hearing game Sound Scouts. Comparisons were made between tests to determine the agreement of the self-administered tests with clinician-administered tests in detecting hearing loss and speech-in-noise deficits.

Study Sample: 297 Aboriginal and Torres Strait Islander children aged 4-14 years from three schools.

Results: Acceptable threshold differences of ≤ 5 dB between AutoAud and manual audiometry hearing thresholds were found for 88% of thresholds, with greater agreement for older than for younger children. Consistent pass/fail results on the Sound Scouts speech-in-quiet measure and manual audiometry were found for 81% of children. Consistent pass/fail results on the Sound Scouts speech-in-noise measure and LiSN-S high-cue condition were found for 73% of children.

Conclusions: This study shows good potential in using self-administered applications as initial tests for hearing problems in children. These tools may be especially valuable for children in remote locations and those from low socio-economic backgrounds who may not have easy access to healthcare.

Introduction

The Indigenous population of Australia are the Aboriginal and Torres Strait Islander peoples. Aboriginal refers to the people originating from mainland Australia and Tasmania, and Torres Strait Islander refers to those from the Torres Strait Islands between Papua New Guinea and mainland Australia. Aboriginal and Torres Strait Islander children experience significantly higher levels of recurrent and chronic otitis media and associated hearing loss compared to non-Indigenous children (Closing the Gap Clearinghouse, 2014). In remote Australia, an average of 30% of Aboriginal children experience chronic otitis media (Australian Institute of Health and Welfare, 2012) varying to up to 90% in some communities (Closing the Gap Clearinghouse, 2014). The disease can be established by the age of three to six months (Leach et al., 1994). In urban locations, the proportion of Aboriginal children with chronic ear disease is similar (Gunasekera et al., 2018), however, it is usually of a milder form and associated with milder levels of hearing loss.

Chronic ear disease is the main cause of hearing loss for Aboriginal and Torres Strait Islander children. Nine percent of Australian children fitted with hearing aids are Aboriginal and Torres Strait Islander, while they make up only 4% of the Australian child population (Australian Hearing, 2018). While hearing loss is spread equally across the socioeconomic gradient for non-Indigenous Australian children, for Aboriginal and Torres Strait Islander children, it is concentrated in lower socioeconomic locations (Simpson et al, 2017), which are also less likely to have timely access to both specialist otology and hearing services (Gunasekera et al, 2009).

Delays in identification and remediation of peripheral hearing loss in children disrupts language and communication outcomes even for children with mild-to-moderate hearing loss (Briscoe, Bishop, & Norbury, 2001). In all affected populations, high rates of chronic ear infection relates to socially-determined risk factors (Kong and Coates, 2009). Children who have experienced chronic ear infection in early childhood are at increased risk of delayed development of the skills required for school readiness across physical, social, emotional, language, cognition, and communication domains (Bell et al, 2016). Identification and intervention as early as possible in their schooling is required to ensure these children do not fall further behind as they progress through school.

Spatial processing disorder (SPD) is a type of auditory processing disorder that is characterised by reduced ability to segregate target speech from competing signals that come from different directions, despite normal sensitivity to soft sounds (Cameron et al, 2014). There is a strong relationship between SPD and early onset, chronic or recurrent ear disease (Tomlin and Rance, 2014; Graydon et al, 2017). This ability relies on the brain's ability to process binaural sound cues, the learning of which is disrupted by fluctuating hearing in early childhood. Prevalence of SPD among Aboriginal and Torres Strait Islander primary school children in a regional Australian location has previously been shown to be 7% (Cameron et al, 2014).

Modern school classrooms are interactive spaces where small group learning activities are commonplace. These activities lead to increased levels of background noise, presenting a challenge to children who experience difficulty understanding speech in noise, whether because of hearing loss or any type of auditory processing disorder. There are benefits to identifying children with listening-in-noise problems so that hearing-friendly communication and teaching strategies can be implemented, and/or hearing remediation initiated.

People living in low-resourced communities or countries often do not have timely access to hearing services. As a consequence, hearing problems often go undiagnosed and unremediated (Australian Institute of Health and Welfare, 2018). The development and usage of automated hearing test applications, designed to be used by workers who have not had specialised hearing health training, has the potential to enhance access to the hearing assessments required to progress children along appropriate care pathways to specialist treatment and/or hearing rehabilitation (Mahomed, 2013).

Automated audiometry has existed in a variety of forms since Von Bekesy's sweep frequency approach, developed in the 1940s. Currently, automated hearing evaluation applications are available in a variety of formats (pure tone threshold tests, speech in noise tests), on a range of platforms (computer/tablet/phone-based applications and the combination of programs and clinical audiometers). In all service delivery contexts, reliability and validity of automated results are of uppermost concern. In the context of under-served communities, complexity of the application must be balanced against need for administration and interpretation by non-hearing health trained workers. Telehealth practices can partially alleviate this problem.

A study by Whitton, Hancock, Shannon, and Polley (2016) found automatic audiometry to be an accurate measure of tone detection thresholds, frequency discrimination abilities, and word recognition in noise scores in adults. A 2013 meta-analysis of the validity of automated audiometry approaches, evaluated mainly with adults, found that automated approaches had equal test-retest reliability as manual approaches (Mahomed et al, 2013). This study acknowledged, however, that data were limited for children and other more difficult-to-test populations, and for diverse types and degrees of hearing loss.

A 2018 study evaluating validity of an automated audiometry application in school age children found that the particular application yielded results within the clinically acceptable range of variation (10 dB) for 7 to 12-year-olds, but not for 6-year-olds, and that overall, thresholds obtained using automated approaches were on average 6.7 dB higher (poorer) than for manual approaches, risking over-referral for diagnostic audiology (Pereira et al, 2018). A 2013 study using a tablet-based threshold-finding application with 325 6 to10-year-old children showed no difference in referral rates across age groups but a large number of false positives (Kam et al, 2013). Another 2013 study with children aged 3-13 years found no statistically significant difference between warble-tone thresholds obtained by client-controlled audiometry on an iPad compared traditional play audiometry (Beauregard et al., 2013). However, a follow-up study in 2015 revealed that thresholds were not consistent if using uncalibrated headphones, and thresholds at 500 Hz were not reliable if there was noise present (Yeung, Heley, Beauregard, Champagne, & Bromwich, 2015).

The purpose of this study was to examine the extent to which self-administered tests agree with clinician-administered tests. More specifically, the study examined the agreement between clinician-administered audiometry and an automatic audiometry program (AutoAud), and clinician-administered audiometry with Sound Scouts (a tablet-based hearing test game) speech-in-quiet test scores. Sound Scouts also measures children’s speech-in-noise abilities, so these were compared with the diagnostic test for SPD: the Listening in Spatialized Noise – Sentences test (LiSN-S).

Method

Aboriginal and Torres Strait Islander children enrolled in three schools with different geographical settings participated in this study. All children completed manual audiometry, Sound Scouts, and the LiSN-S in quiet rooms at the schools (see below for test descriptions).

They were given breaks between testing to avoid fatigue. Children from one of the schools also completed AutoAud.

Participants

A total of 297 children aged 4-14 years from three schools took part in the study. There were 120 children from Campbelltown, Sydney, 118 children from Kuranda, Northern Queensland, and 59 children from Port Augusta, South Australia. Parental consent was obtained from all the children via an opt-out process and permission to conduct this study was obtained from the Australian Hearing Human Research Ethics Committee.

[Insert Table 1 here]

Manual Audiometry

Manual audiometry was administered by the examiner using the software-based Avant MedRx audiometer installed on laptops. Air and bone conduction pure tone thresholds were obtained using the modified Hughson Westlake procedure while children were wearing Peltor sound-attenuating circumaural headphones integrated with D45 transducers. Children were instructed to raise their hand or to perform a play task, for example, drop a block in a bucket, when they heard beeping sounds. Air-conduction thresholds were obtained from 250-8000 Hz in octave intervals.

AutoAud

Automatic audiometry was used only at the school in Kuranda as it was only developed in time for this set of data collection. The AutoAud computer program is a self-administered adaptive hearing application that tests hearing at 1 kHz and 4 kHz. The auditory stimulus consists of three pulsed tones with each tone lasting 290 ms and a 140 ms silence between the tones. AutoAud uses a USB audio device to drive a pair of Howard Leight Sync circumaural headphones. The test was configured to present stimuli up to 60 dB HL and down to 20 dB HL. The participant is asked to press the space bar whenever he or she hears the pulsed tones. The response needs to be made within 2.35 seconds from the onset of tone bursts to be valid.

During the practice phase, the pulsed tones were presented to the children at 60 dB HL. Once the child responded to the tones consistently, the software proceeded to the testing phase in which the left and right ears were tested separately. The starting level for the testing phase was 50 dB HL. The level decreases by 10 dB for each correct response down to 20 dB HL. The application displays a green tick on the screen when a valid response is registered to maintain the children's attention during the test. If the listener responds to 20 dB HL, then his or her threshold is recorded as 20 dB HL and the test at that frequency for that ear ends.

If the listener does not respond to a tone, the level is increased by 10 dB and then in 5 dB steps until a correct response is registered (up to 60 dB). When a correct response is recorded, the level decreases in 5 dB steps. The final level at which a correct response is received is deemed to be the threshold. If the listener does not record a correct response at 60 dB HL, his or her threshold is recorded as being > 60 dB HL. A pass on the test is considered to be thresholds of 20 dB HL at both frequencies in both ears.

Sound Scouts

Sound Scouts is a self-administered tablet game-based hearing test that can be completed without a trained audiologist or calibrated audiometers. The test utilizes three interactive games to test the child's ability to 1) detect tones in background noise, 2) understand speech in quiet, and 3) understand speech in background noise. Sound Scouts reliably detects four-frequency average hearing thresholds (4FAHL; average of 500, 1000, 2000 and 4000 Hz) greater than 30 dB HL in children aged 4.5 years and older (Dillon, Mee, Moreno, and Seymour, 2018). Z scores are calculated by the app using age-appropriate normative data for each of these measures. There are two versions of Sound Scouts – the long version which takes 20 minutes to complete (which was used in Campbelltown and Port Augusta) and a short version with reduced narration and different speech-in-quiet and tone-in-noise stimuli, that takes 10 minutes to complete (this new version was trialed in Kuranda). The short version was designed to require a lower level of English proficiency, allow for more children to be tested more quickly at school, and be less likely to be a challenge for those children with attention difficulties due to its shorter timeframe.

The children played the game in a quiet room at their school. Before testing each child, the examiner showed him/her the pictures on the iPad of the words used for the speech-in-quiet task (colour words for the long version and spondees for the short version). In the case of the short version, if a child was unfamiliar with any of the spondees, other spondees with which they were familiar were selected. The child needed to answer several questions correctly in the practice trials in order to proceed to the test. Both the practice and the test were administered via Sennheiser HD215 circumaural headphones (Hanover, Germany) using a Microsoft Surface Pro 3 touchscreen computer (Microsoft, China) or iPad Air 2 (Apple Inc., California).

The speech-in-quiet calibration game identifies the softest speech sound that the child can understand compared to an adult with normal hearing, an approach taken to avoid the need to have calibrated headphones or tablet device. For this study, a researcher with normal hearing first completed the calibration game and then the child proceeded with the game themselves. The speech-in-noise game asks the child to identify objects spoken by a target talker who sounds like they are in front of the child while there are two competing talkers that are spatially separated from the target talker (Dillon et al., 2018). The masking sounds are presented at a pre-specified level relative to the child's speech reception thresholds in quiet. The examiner monitored the test environment and if it became noisy, the instructor could choose to pause or restart the test.

Listening in Spatialized Noise – Sentences Test (LiSN-S)

The Listening in Spatialized Noise – Sentences Test (LiSN-S) (Cameron and Dillon, 2009) measures a listener's ability to utilise inter-aural time and level cues to understand a target talker in the presence of distracting talkers coming from the same or different spatial angles. It compares the test scores in four different conditions to assess whether the listener has SPD (Cameron and Dillon, 2007a, 2007b, 2008). The first (high-cue) condition, which is also known as the Different Voices ± 90° condition, is initially treated as a screener. Children with z-scores ≥ -1.5 would be considered as passing the test and were not tested further with LiSN-S. Children with z scores poorer than -1.5 in the high cue condition proceeded to the other three conditions (Same Voice ± 90°, Different Voices 0°, Same Voice 0°).

During the high-cue condition, the child heard target sentences through headphones and was asked to repeat the sentences back to the examiner. The target voice sounded like it was coming from the front of the child, while two different distracting voices apparently came from +90° and -90° azimuths (i.e. on either side of the listener). This condition was called the high-cue condition because the child could use both spatial and talker cues to differentiate the target voice from the distractors. The signal-to-noise ratio of the speech signal and background noise adapted throughout the test depending on how many words in the sentence the child got correctly. The test was administered by the examiner with the child wearing Sennheiser HD215 circumaural headphones (Hanover, Germany) using a Microsoft Surface Pro 3 touchscreen computer (Microsoft, China) or iPad Air 2 (Apple Inc., California).

Data Analysis

Data were analyzed using Statistica version 10. Pearson correlations were used to assess the strength of the relationships between manual audiometry versus Sound Scouts speech-in-quiet results and the LiSN-S high-cue condition versus Sound Scouts speech-in-noise results. Sensitivity and specificity are reported for AutoAud versus manual audiometry results, Sound Scouts speech-in-quiet versus manual audiometry results, and Sound Scouts speech-in-quiet versus LiSN-S results.

Results

The following section provides a cross-validation of the self-administered tests with the clinician-administered tests. More in depth details of the individual test results by school and the prevalence of children’s hearing and speech-in-noise deficits can be found in Mealings, Hwang, Fragoso, Chung, Harkus, and Dillon (2020).

A. Hearing Assessment Cross-Validation: Manual Audiometry versus AutoAud

Figure 1 shows the difference between the manual audiometry and AutoAud results as a function of the child’s age. The size of the marker is directly proportional to the number of child with a particular threshold (i.e. larger circles indicate more children). A negative difference indicates that the manual audiometry results gave a lower (i.e. better) threshold than the AutoAud results and vice versa for a positive difference. For this analysis, we converted the manual audiometry scores that were < 20 dB HL to be at 20 dB HL so that there was no difference between the manual and AutoAud scores because the latter, by design, gave thresholds down to only 20 dB HL.

Averaged across test frequencies and ears, the thresholds differed by at most 5 dB for 88% of children. At every combination of frequency and ear, the percentage of children for whom there was a difference of 10 dB or greater was larger for children less than 10 years old than for older children. Across test frequencies and ears, 31% of children less than 10 years old had differences of 10 dB or greater for a given frequency/ear compared to 16% for children 10 years or older, however this difference did not reach significance ($\chi^2(1) = 3.31$, $p = 0.07$).

[Insert Figure 1 here]

Table 2 compares the percentage of children who passed/failed manual audiometry compared to AutoAud from all schools. Failing manual audiometry and AutoAud was defined as a threshold > 20 dB HL at each particular frequency. Consistent results on both tests were found for 82-87% of children, depending on the frequency. Where there were inconsistent results, children more often failed AutoAud and passed manual audiometry than vice versa. Overall sensitivity when a hearing loss was defined as > 20 dB HL was 29% and specificity was 87%. When hearing loss was defined as > 30 dB HL, sensitivity was still 29% and specificity was 91%.

[Insert Table 2 here]

B. Hearing Assessment Cross-Validation: Manual Audiometry versus Sound Scouts

A Pearson's correlation analysis was carried out to determine how well the self-administered Sound Scouts speech-in-quiet z scores (which are based on the poorer of the two speech reception thresholds measured) correlated with the manual audiometry results using the child's worst ear 4FAHL as well as the threshold at each of 500, 1000, 2000, and 4000 Hz in the ear with the worst 4FAHL. For all schools combined, a moderate correlation was found for the 4FAHL in the child's worse ear ($r(238) = -0.50$, $p < 0.01$ indicating that lower Sound Scouts speech-in-quiet z scores were associated with higher thresholds in the child's worse ear (Figure 2a). Moderate correlations were also found with thresholds at all individual frequencies: 500 Hz ($r(238) = -0.50$, $p < 0.01$); 1000 Hz ($r(238) = -0.48$, $p < 0.01$), 2000 Hz ($r(238) = -0.41$, $p < 0.01$), and 4000 Hz ($r(238) = -0.35$, $p < 0.01$). Note that many children had their thresholds measured down to only 15 dB HL which is the reason for a cluster of responses around this threshold. When taking into account only those children with hearing loss, a stronger correlation was found ($r(40) = -0.77$, $p < 0.01$, Figure 2b). Since two different versions of Sound Scouts were used in the projects at different schools, we also measured the correlations between the children's 4FAHL vs. the long version used in Campbelltown and Port Augusta, and the short version used in Kuranda. For the long version, a moderate correlation was found ($r(144) = -0.51$, $p < 0.01$, Figure 2c). For the short version, a similar moderate correlation was also found ($r(94) = -0.50$, $p < 0.01$, Figure 2d).

Sound Scouts detected (i.e. indicated a result outside the normal range) eight out of nine children with hearing loss greater than 30 dB HL in their worse ear. Of these eight children, three had unilateral conductive hearing loss, four had bilateral conductive hearing loss, and one child had a bilateral sensorineural hearing loss (4FAHL in worse ear of 60 dB HL). The child that Sound Scouts did not detect had a mild bilateral conductive hearing loss (4FAHL in worse ear of 31.25 dB HL).

[Insert Figure 2 here]

Table 3 compares the percentage of children who passed/failed manual audiometry compared to Sound Scouts speech-in-quiet test from all schools. Failing manual audiometry was defined as 4FAHL > 20 dB HL for the first analysis and 4FAHL > 30 dB HL for the second analysis. Failing Sound Scouts was defined by a z score of < -2. Consistent results on both tests were found for 81% of children for a manual audiometry 4FAHL of > 20 dB HL, and consistent results were found for 87% of children for a manual audiometry 4FAHL of > 30 dB HL. Inconsistent results were evenly split for those who passed manual audiometry and failed Sound Scouts and vice versa when a manual audiometry fail was classed at a 4FAHL > 20 dB HL. When classed at a 4FAHL > 30 dB HL, inconsistent results were due to children passing manual audiometry but failing Sound Scouts. The sensitivity of Sound Scouts for a 4FAHL of > 20 dB HL was 41% and specificity was 89%. Sensitivity for a 4FAHL of > 30 dB HL was 88% and specificity was 88%.

[Insert Table 3 here]

C. Speech-in-Noise Cross-Validation: Sound Scouts versus LiSN-S

Sound Scouts speech-in-noise SRTs are plotted relative to the LiSN-S high-cue SRTs (both expressed as a SNR in dB) in Figure 3. The Pearson's correlation coefficient between the two measures was $r(239) = 0.53$ ($p < 0.01$). Based on the simple correlation alone, 71% of the variance in LiSN-S is therefore not explained by Sound Scouts scores (and vice versa). Both measures, of course, have intrinsic random measurement errors, which have previously been estimated based on test-retest studies. For Sound Scouts, the standard deviation due to random measurement error of speech-in-noise test scores has been calculated to be 1.9 dB (unpublished data). These data are based on repeat measurements, one month after the first administration of Sound Scouts, on 34 of the children described in Dillon et al. (2018). Speech SNR improved by 1.52 dB on retest, with a SD around that mean of 2.74 dB. As the test-retest difference reflects the random error component of both the test and the retest, the random error component of either test or retest alone is estimated as 2.74 dB/√2, or 1.93 dB. For the LiSN-S high-cue condition, the corresponding value is 1.0 dB (Cameron & Dillon, 2007b). When the effect of random errors in each test are allowed for, this proportion reduces to 56% of the variance in LiSN-S not accounted for by Sound Scouts, and 46% of the variance in Sound Scouts not accounted for by LiSN-S.

[Insert Figure 3 here]

The correlation between the Sound Scouts and LiSN-S scores is partly the result of there being a range of ages in the participants, as older children on average to do better on both tests. To control for the effects of age, Pearson's correlation was repeated using the z scores for each test as z scores have the same distribution around the mean at all ages, which makes results comparable across age groups. A moderate, but slightly lower correlation was found (r (240) = 0.42, $p < 0.01$). As seen in Figure 4a there was an outlier where the LiSN-S z score = -18. A correlation analysis was run again excluding this outlier, resulting in $r(239) = 0.36$, $p < 0.01$. Again, as two different versions of Sound Scouts were used in the projects at different schools, we also measured the correlations between the children's LiSN-S high-cue z scores vs. the long version used in Campbelltown and Port Augusta, and the short version used in Kuranda. For the long version, a moderate correlation was found ($r(152) = -0.43$, $p < 0.01$, Figure 4b). For the short version, a moderate correlation was also found ($r(89) = -0.38$, $p < 0.01$, Figure 4c).

[Insert Figure 4 here]

Table 4 compares the percentage of children who passed/failed the LiSN-S and/or Sound Scouts speech-in-noise test from all schools (defined by a z score of < -2). Consistent results on both tests were found for 73% of children. Interestingly, 23% of children who failed the LiSN-S passed Sound Scouts. Possible reasons for this are raised in the discussion. Sensitivity for Sound Scouts speech-in-noise test was 30% and specificity was 94%.

[Insert Table 4 here]

D. Presence of Hearing Loss versus Speech-in-Noise Difficulties

A comparison of the percentage of children who passed/failed manual audiometry and/or the LiSN-S from all schools was also made. Failing manual audiometry was again classified as a 4FAHL > 20 dB HL. A fail on the LiSN-S was defined as a z score < -2. Fifty-eight percent of children passed both tests. Thirteen percent of children failed both tests, i.e. had hearing loss and speech-in-noise issues. Twenty-one percent of children did not have a hearing loss but did have speech-in-noise issues. Seven percent of children had a hearing loss without any speech-in-noise issues.

Discussion

The purpose of this study was to examine the extent to which self-administered tests agree with clinician-administered tests. The study examined the agreement between manual audiometry and AutoAud, manual audiometry and Sound Scouts speech-in-quiet test scores, and the LiSN-S and Sound Scouts speech-in-noise scores. As Aboriginal and Torres Strait Islander children can often be in remote locations, self-administered tests provide a way for all children to be tested for hearing problems, and then those who show problems can be referred for a more complete assessment. Additionally, as many communities that are more remote have a lower socio-economic status, these people may be unable to afford to travel to a healthcare professional to get their child's hearing checked, or pay for the appointment. Therefore, self-administered tests have the potential to make services more accessible.

Hearing Assessment Cross-Validation: Manual Audiometry versus AutoAud

AutoAud measured the children's hearing threshold down to 20 dB HL at 1 kHz and 4 kHz for their left and right ears. Acceptable threshold differences of ≤ 5 dB between the AutoAud and manual audiometry results were found for 86-91% of the children depending on the frequency and ear. The children who had threshold differences of ≥ 10 dB were mostly less than 10 years old. Generally, when there were discrepancies, it was the clinician-administered audiometry that gave lower (i.e. better) thresholds than AutoAud. Overall sensitivity when a hearing loss was defined as > 20 dB HL was 29% and specificity was 87%. When hearing loss was defined as > 30 dB HL, sensitivity was still 29% and specificity was 91%. This sensitivity is much less than that reported by Margolis, Frisina, and Walton (2011) who detected the inaccurate audiograms with a sensitivity (i.e. hearing loss agreement with the gold standard pure tone audiometry) of 71% and a specificity of 91%. Their study, however, combined children and adults, so the adult data as well as the test's inherent properties may have improved the agreement with the gold standard.

Hearing Assessment Cross-Validation: Manual Audiometry versus Sound Scouts

Sound Scouts measured children's speech reception thresholds in quiet which can then be compared to their hearing threshold measure by manual audiometry. If a speech-in-quiet z-score poorer than -2 were to be taken as in indication of hearing loss, then it is evident from Figure 2a that Sound Scouts did not detect hearing loss for more than half the children with 4FAHL between 20 and 30 dB HL (assuming that the manual audiometry results are the gold standard). For losses greater than 30 dB HL, however, z-scores poorer than -2 (usually markedly so) were obtained for 8 out of 9 children with this degree of loss. The one mis-classified child in this category had a loss of 32 dB 4FAHL in the poorer ear. The sensitivity of Sound Scouts for a 4FAHL of > 20 dB HL was 41% and specificity was 89%. Sensitivity was improved when using a hearing loss criteria 4FAHL of > 30 dB HL, with sensitivity at 88% and specificity at 88%. This result is consistent with the very high sensitivity previously found for Sound Scouts (Dillon et al., 2018) for 4FA hearing thresholds greater than 30 dB HL.

Speech-in-Noise Cross-Validation: Sound Scouts versus LiSN-S

Sound Scouts also measured children's speech-in-noise hearing ability which could be compared to the children's performance on the clinician-administered LiSN-S. A child was considered to have a speech-in-noise deficit on these tests if their z score was < -2. Sensitivity for Sound Scouts speech-in-noise test was 30% and specificity was 94%. Consistent pass/fail results on the Sound Scouts speech-in-noise measure and the LiSN-S high-cue condition were found for 73% of children. Interestingly, however, 23% of children who passed the Sound Scouts speech-in-noise test failed the LiSN-S. A possible reason for this is due to the number of items in each test. The LiSN-S has up to 30 sentences, each with four, five, or six scored items. Consequently, scoring on up to 150 items gives a more precise score than Sound Scouts, which has only 24 items for the speech-in-noise task. This is demonstrated by the random error standard deviation of 1.0 dB for LiSN-S versus 1.9 dB for Sound Scouts. Because of this greater precision, the normative range on the LiSN-S will be smaller, so a smaller deficit from average (in dB) is more likely to be outside normal limits. Consequently, the LiSN-S should pick up more cases of children with speech-in-noise deficits, though Sound Scouts should pick up the most severe cases. The speech-in-noise test in Sound Scouts could be made longer to increase its precision, however, if it is too long then the child will lose concentration which will also make their scores unreliable. As Sound Scouts also has the advantage of detecting hearing loss, and being self-administered, it is a trade-off between the number of things the application can test for, combined with the greater range of circumstances in which the test can be used, versus the precision that the game is able to provide for each test. Disagreement among test scores is not simply a case of the relative precision of the two tests. Even when the impact of the expected random test errors were taken into account, approximately half of the variance in scores in each test remained unaccounted for when using the other test as a predictor. However, as the study population differs from the general population in various ways, the test scores may have a greater random component than has been assumed.

There are several potential sources of different speech-in-noise test results. First, attention and engagement are achieved by game play for Sound Scouts versus by encouragement from a clinician for LiSN-S. Either of these could potentially lead to better performance, depending on the child. Second, there is a greater need to use semantic and syntactic language knowledge in LiSN-S than in Sound Scouts. This may be an issue for Indigenous children who often have languages other than English as their first language. Third, there is a need to associate visual symbols with words in Sound Scouts, but not in LiSN-S. Consequently, it cannot be assumed that in each case of disagreement the LiSN-S result is the more representative test of real-life listening in noise ability.

Presence of Hearing Loss versus Speech-in-Noise Difficulties

A comparison of the percentage of children who passed/failed manual audiometry and/or the LiSN-S from all schools was made to determine which children have either a hearing loss or speech-in-noise difficulties, neither, or both. Fifty-eight percent of children did not have either issue, and 13% of children had both hearing loss and speech-in-noise issues. Interestingly, 21% of children did not have a hearing loss but did have speech-in-noise issues and 7% of children had a hearing loss without any speech-in-noise issues. This shows the importance of testing for both hearing loss and speech-in-noise difficulties as one does not necessarily mean the other.

Limitations of the Study

While this study provided valuable insight into the effectiveness of using automated testing procedures in identifying children with hearing loss and/or speech-in-noise difficulties, there were a couple of limitations.

First, both the LiSN-S and Sound Scouts tests are administered in English. For some Aboriginal and Torres Strait Islander children, English is their second or even third language as they grow up learning the traditional language(s) of their community. Although the examiners administering these tests were careful in making sure the children had sufficient English to be able to complete the testing, and were more lenient when scoring the LiSN-S by allowing for grammatical errors, it is still possible that the language barrier impacted on the children’s results. To address this, we are currently developing a language-independent version of the LiSN-S to avoid a person’s language background influencing their test results. The Sound Scouts short version also uses simplified English compared to the long version, so this may be the better app to use for this population.

Second, the validation of both AutoAud and Sound Scouts relied on the clinician-administered audiometry and LiSN-S results being sufficiently accurate to be regarded as gold standard tests. However, both of these tests are still behavioural tests and rely on the compliance and motivation of the child, and the skill of the clinician in maintaining motivation, to yield accurate results. Therefore, differences in results between the clinician-administered and self-administered results may not necessarily mean that the self-administered test is the inaccurate test. Since this study was conducted, Sound Scouts has undergone significant improvements so the results of this study apply to the version that existed at that time.

Conclusions

The results of this study show the potential in using self-administered applications as initial tests for hearing problems in children. Self-administered tests provide a way for all children to be tested for hearing problems, and then those who show problems can be referred for a more complete assessment. Self-administered hearing test apps such as AutoAud and Sound Scouts have the advantage of being able to be completed without a clinician or clinical audiology equipment, so are more accessible and have a lower cost. The speech-in-quiet component of Sound Scouts in particular gave good sensitivity and specificity results when using a hearing loss criterion of > 30 dB HL. As shown in this study, self-administered tests are not always 100% accurate though, so future research is needed to continue to validate these tests, especially for AutoAud and the speech-in-noise component of Sound Scouts as these tests had low sensitivity. However, these tests do provide a valuable starting point in identifying a greater number of children with hearing loss more efficiently.

Acknowledgements. We would like to thank the participating schools and health services in the three communities, Mark Mitchell from the Queensland Aboriginal and Islander Health Council, the audiology students from the University of Northern Illinois, Meagan Ward and Belinda Lesina from Australian Hearing, and Pam Jackson from the National Acoustic Laboratories, for their help with collecting data for this study. The study was supported by the Prime Minister and Cabinet Department's Indigenous Advancement Scheme. Harvey Dillon acknowledges the support of the Australian Department of Health, Macquarie University and the NIHR Manchester Biomedical Research Centre.

Professional meeting details. Portions of this study were presented at the Audiology Australia conference in Sydney, May 20-23, 2018, and the Otitis Media Australia conference in Darwin, August 14-16, 2018.

Disclosure statement. This study was funded by the Indigenous Advancement Strategy Project 4-1M7F2PW, through the Department of Prime Minister and Cabinet, Australia.

References

Australian Hearing. (2018). Demographic details of young Australians aged less than 26 years with a hearing loss, who have been fitted with a hearing aid or cochlear implant at 31 December 2017. Sydney, Australia: Australian Hearing. https://www.deafnessforum.org.au/wp-content/uploads/2019/08/Demographic-Details-Australians-aged-less-than-26-with-a-hearing-loss-who-have-been-fitted-with-a-hearing-aid-or-cochlear-implant-at-31-December-2018.pdf

Australian Institute of Health and Welfare. (2012). Northern Territory emergency response child health check initiative: Follow-up services for oral and ear health, final report, August 2007–June 2012. Cat. no. DEN 223. Canberra: AIHW.

Australian Institute of Health and Welfare. (2018). Australia's health 2018. Australia's health series no. 16. AUS 221. Canberra: AIHW.

Beauregard, Y., Yeung, J., Heley, S., Javidnia, H., Bromwich, M., & Champagne, S. (2013). The new age of play audiometry: prospective validation testing of an iPad-based play audiometer. Journal of Otolaryngology - Head and Neck Surgery, 42(1), 21. doi:10.1186/1916-0216-42-21

Bell, M.F., Bayliss, D.M., Glauert, R., Harrison, A., & Ohan, J.L. (2016). Chronic illness and developmental vulnerability at school entry. *Pediatrics*, 137(5): e20152475

Briscoe, J., Bishop, D.V.M., & Norbury, C.F. (2001). Phonological processing, language, and literacy: A comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *Journal of Child Psychology and Psychiatry, 42*(3), 329-340.

Cameron, S., & Dillon, H. (2007a). Development of the Listening in Spatialized Noise-Sentences Test (LiSN-S*). Ear and Hearing*, 28(2): 196-211. doi:10.1097/AUD.0b013e318031267f

Cameron, S., & Dillon, H. (2007b). The listening in spatialized noise-sentences test (LiSN-S): Test-retest reliability study. *International Journal of Audiology*, 46: 145-153. doi:10.1080/14992020601164170

Cameron, S., & Dillon, H. (2008). The Listening in Spatialized Noise–Sentences Test (LiSN-S): Comparison to the prototype LISN and results from children with either a suspected (central) auditory processing disorder or a confirmed language disorder. *Journal of the American Academy of Audiology*, 19(5): 377-391.

Cameron, S., & Dillon, H. (2009). Listening in Spatialized Noise – Sentences test (LiSN-S) (Version 2.4). Murten, Switzerland: Phonak Communications AG.

Cameron, S., Dillon, H., Glyde, H., Kantha, S., and Kania, A. (2014). Prevalence and remediation of spatial processing disorder (SPD) in Indigenous children in regional Australia. *International Journal of Audiology*, 53(5): 326-335. doi:10.3109/14992027.2013.871388

Ching, T.Y.C., Dillon, H., Leigh, G., and Cupples, L. (2018). Learning from the Longitudinal Outcomes of Children with Hearing Impairment (LOCHI) study: summary of 5-year findings and implications. *International Journal of Audiology*, 57(2), S105-S111.

Closing the Gap Clearinghouse (AIHW & AIFS). (2014). Ear disease in Aboriginal and Torres Strait Islander children. Resource sheet no. 35. Produced by the Closing the Gap Clearinghouse. Canberra: Australian Institute of Health and Welfare & Melbourne: Australian Institute of Family Studies.

Dillon H., Mee C., Moreno J.C., and Seymour J. (2018). Hearing tests are just child's play: The Sound Scouts game for children entering school. *International Journal of Audiology*, 57(7): 529-537. doi: 10.1080/14992027.2018.1463464

Graydon, K., Rance, G., Dowell, R., and Van Dun, B. (2017). Consequences of early conductive hearing loss on long-term binaural processing. *Ear and Hearing*, (38)5: 621-627. doi:10.1097/ AUD.0000000000000431

Gunasekera, H., Morris, P.S., Daniels, J., Couzos, S., and Craig, J.C. (2009). Otitis media in Aboriginal children: The discordance between burden of illness and access to services in rural/remote and urban Australia. *Journal of Paediatrics and Child Health*, 45: 425-430.

Gunasekera, H., Miller, H.M., Burgess, L., Chando, S., Sheriff, S.L., Tsembis, J.D., Kong, K.M., Coates, H.L.C., Curotta, J., Falster, K., McIntyre, P.B., Banks, E., Peter, N.J., and Craig, J.C. (2018). Agreement between diagnoses of otitis media by audiologists and otolaryngologists in Aboriginal Australian children. *Medical Journal of Australia*, 209(1): 29-35.

Kam, A.C.S., Gao, H., Li, L.K.C., Zhao, H., Qiu, S., and Tong, M.C.F. (2013). Automated hearing screening for children: A pilot study in China. *International Journal of Audiology*, 52: 855-860.

Kong, K. and Coates, H.L.C. (2009). Natural history, definitions, risk factors and burden of otitis media. *Medical Journal of Australia*, 191(9): S39.

Leach, A., Boswell, J., Asche, V., Nienhuys, T.G., and Mathews, J.D. (1994). Bacterial colonisation of the nasopharynx predicts very early onset and persistence of otitis media in Australian Aboriginal infants. *Paediatric Infectious Diseases Journal*, 13: 983-989.

Mahomed, F., Swanepoel, D., Eikelboom, R.H. and Soer, M. (2013). Validity of automated threshold audiometry: A systematic review and meta-analysis. *Ear and Hearing*, 34: 745-752.

Margolis, R.H., Frisina, R. & Walton, J.P. (2011) AMTAS®: Automated method for testing auditory sensitivity: II. Air conduction audiograms in children and adults, *International Journal of Audiology*, *50*(7), 434-439, doi: 10.3109/14992027.2011.553206

Mealings, K., Hwang, J., Fragoso, J., Chung, K., Harkus, S., and Dillon, H. (2020). Hearing loss and auditory processing in Aboriginal and Torres Strait Islander children from different locations and varying socio-educational advantage. *International Journal Pediatric Otorhynolaryngology*. doi: 10.1016/j.ijporl.2019.109741

Pereira, O., Pasko, L.E., Supinski, J., Hammond, M., Morleta, T., and Nagao, K. (2018). Is there a clinical application for tablet-based automated audiometry in children? *International Journal of Pediatric Otorhinolaryngology*, 11: 87-92.

Simpson, A., Enticott, J.C., and Douglas, J. (2017). Socioeconomic status as a factor in Indigenous and non-Indigenous children with hearing loss: analysis of national survey data. *Australian Journal of Primary Health*, 23: 202-207. doi: 10.1071/PY16088.

Tomlin, D. and Rance, G. (2014). Long-term hearing deficits after childhood middle ear disease. *Ear and Hearing*, 35(6): e233–e242. doi:10.1097/AUD.0000000000000065.

Whitton, J. P., Hancock, K. E., Shannon, J. M., & Polley, D. B. (2016). Validation of a Self-Administered Audiometry Application: An Equivalence Study. Laryngoscope, 126(10), 2382–2388. doi:10.1002/lary.25988

Yeung, J. C., Heley, S., Beauregard, Y., Champagne, S., & Bromwich, M. A. (2015). Self-administered hearing loss screening using an interactive, tablet play audiometer with ear bud headphones. International Journal of Pediatric Otorhinolaryngology, 79(8), 1248–1252. doi:10.1016/j.ijporl.2015.05.021

Tables

Table 1: Number of children who completed each test at each school.

	Port Augusta	Campbelltown	Kuranda
Manual Audiometry	58	115	115
Automatic Audiometry	-	-	114
Sound Scouts	48 (Long version)	119 (Long version)	110 (Short version)
LiSN-S	44	119	97

Table 2: Percentage of children who i) passed both manual audiometry and AutoAud; ii) failed manual audiometry but passed AutoAud; iii) passed manual audiometry but failed AutoAud; and iv) failed both manual audiometry and AutoAud for different frequencies and ears. Bold values show agreement between the two tests.

			Manual Audiometry	
			Pass	Fail
AutoAud	1 kHz Left	Pass	**82%**	2%
		Fail	13%	**4%**
	4 kHz Left	Pass	**83%**	9%
		Fail	11%	**4%**
	1 kHz Right	Pass	**82%**	5%
		Fail	13%	**0%**
	4 kHz Right	Pass	**83%**	4%
		Fail	12%	**1%**
	Overall	Pass	**83%**	5%
		Fail	12%	**2%**

Table 3: Percentage of children who i) passed both manual audiometry and Sound Scouts speech-in-quiet test; ii) failed manual audiometry but passed Sound Scouts speech-in-quiet test; iii) passed manual audiometry but failed Sound Scouts speech-in-quiet test; and iv) failed both manual audiometry and Sound Scouts speech-in-quiet test for manual audiometry 4FAHLs of > 20 dB HL and > 30 dB HL. Bold values show agreement between the two tests.

		Manual Audiometry 4FAHL > 20		Manual Audiometry 4FAHL > 30	
		Pass	Fail	Pass	Fail
Sound Scouts	Pass	**74%**	10%	**84%**	0.4%
Speech-in-Quiet	Fail	9%	**7%**	12%	**3%**

Table 4: Percentage of children who i) passed both LiSN-S and Sound Scouts speech-in-noise test; ii) failed LiSN-S but passed Sound Scouts speech-in-noise test; iii) passed LiSN-S but failed Sound Scouts speech-in-noise test; and iv) failed both LiSN-S and Sound Scouts speech-in-noise test. Bold values show agreement between the two tests.

		LiSN-S	
		Pass	Fail
Sound Scouts	Pass	**63%**	23%
Speech-in-Noise	Fail	4%	**10%**

Figures

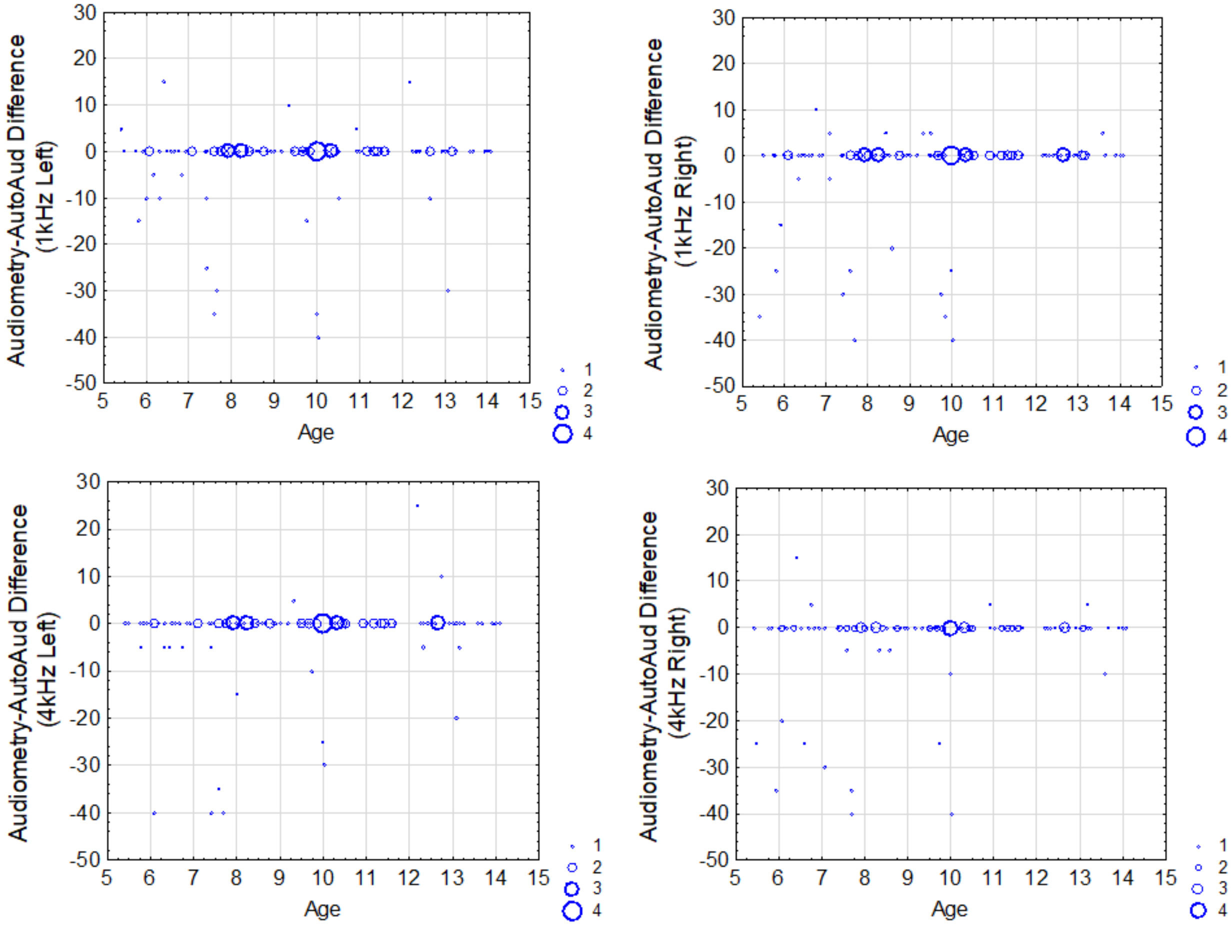

Figure 1: Difference between clinician-administered audiometry and self-administered automatic audiometry results at 1 kHz and 4 kHz for the child's left and right ears.

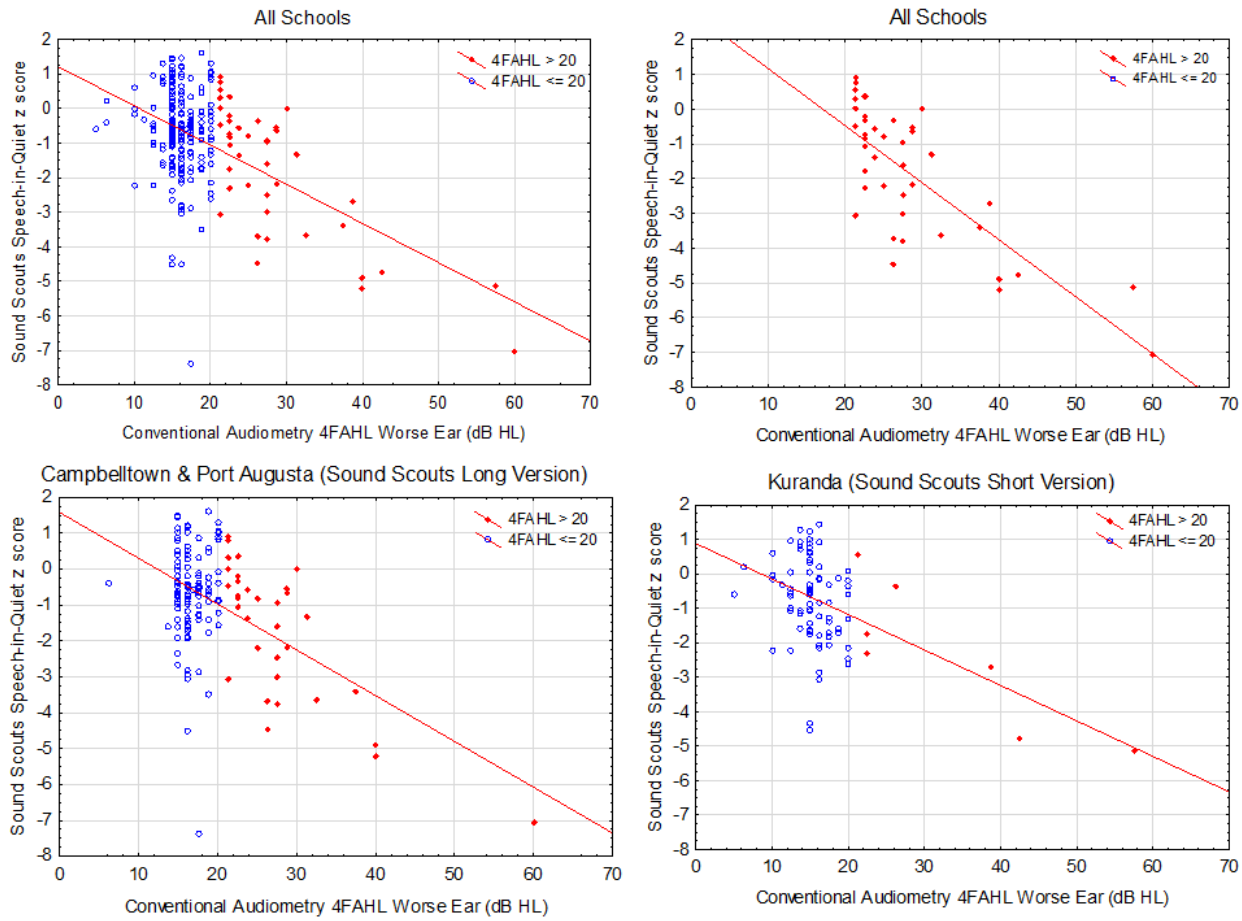

Figure 2: Correlations between manual audiometry and Sound Scouts speech-in-quiet results for a) all children at all schools; b) only children with hearing loss at all schools; c) Sound Scouts Long Version at Campbelltown and Port Augusta; d) Sound Scouts Short Version at Kuranda only.

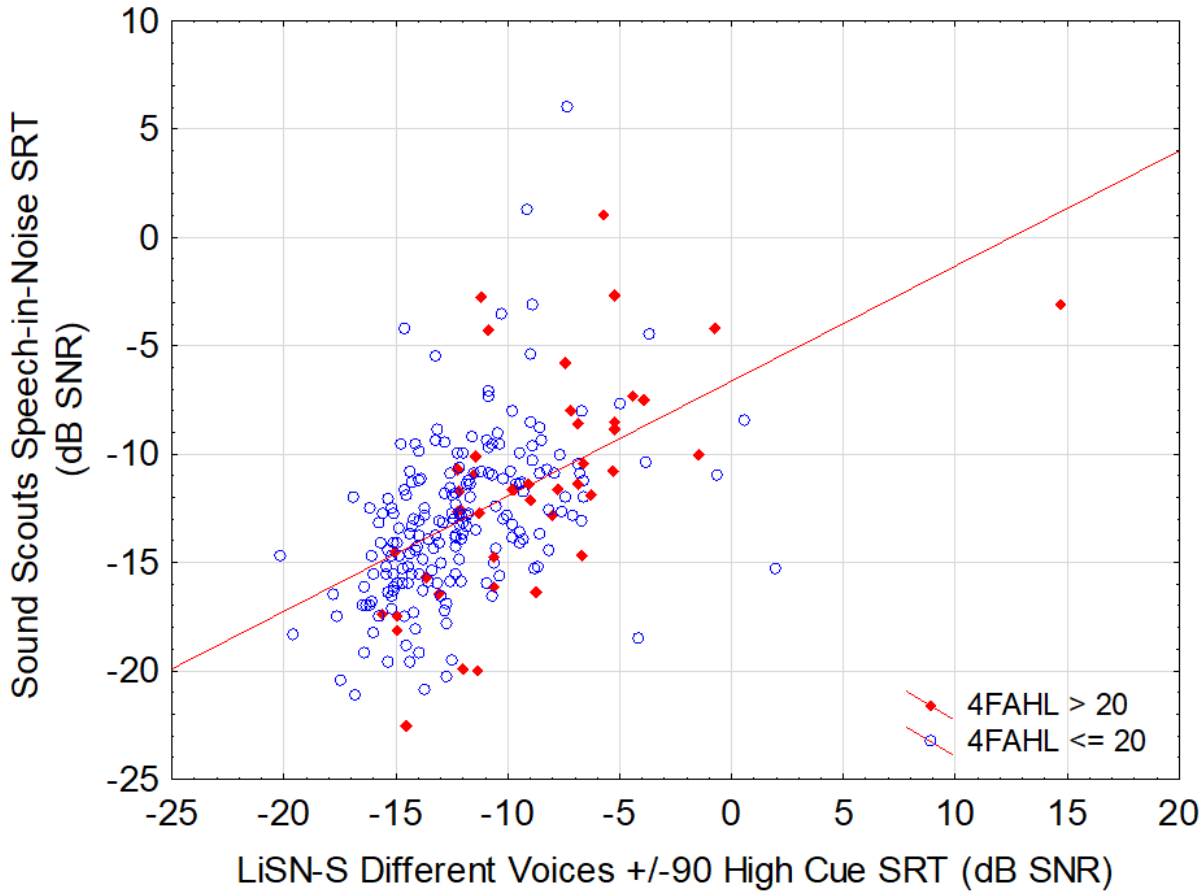

Figure 3: The relationship between LiSN-S high-cue condition speech reception threshold signal-to-noise ratios and Sound Scouts speech-in-noise speech reception threshold signal-to-noise ratios at all schools.

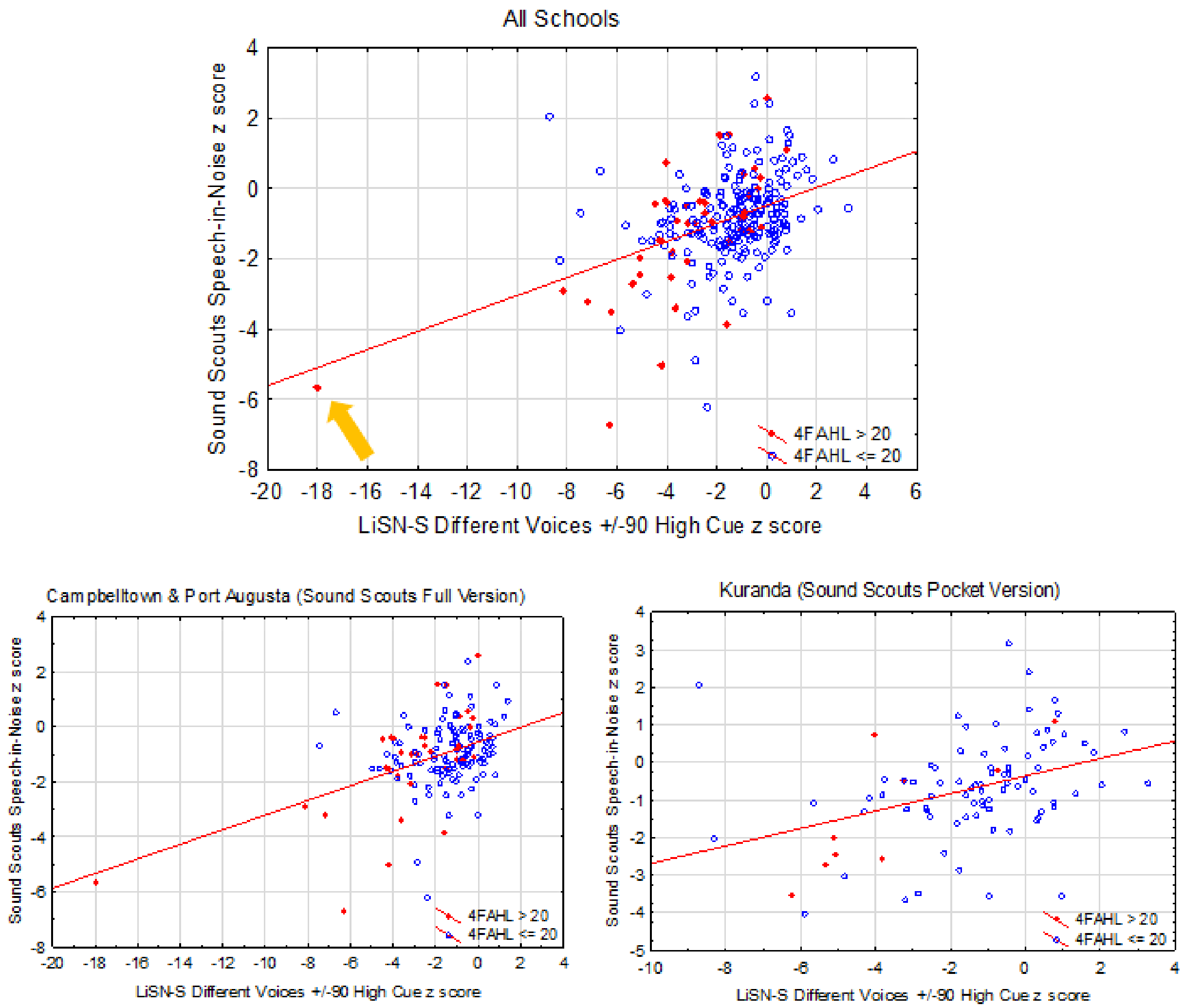

Figure 4: Correlations between LiSN-S high-cue condition z scores and Sound Scouts speech-in-noise z scores for a) all children at all schools; b) Sound Scouts Long Version at Campbelltown and Port Augusta; c) Sound Scouts Short Version at Kuranda only. Arrow shows outlier.