

Evaluation of Hearing Aid Manufacturers' Software-Derived Fittings to DSL v5.0 Pediatric Targets

Paula Folkeard¹ Marlene Bagatto^{1,2} Susan Scollie^{1,2}

¹Faculty of Health Sciences, National Centre for Audiology, University of Western Ontario, London, Ontario, Canada

²School of Communication Sciences & Disorders, University of Western Ontario, London, Ontario, Canada

Address for correspondence Paula Folkeard, National Centre for Audiology, The University of Western Ontario, London, ON, Canada N6G 1H1 (e-mail: folkeard@nca.uwo.ca).

J Am Acad Audiol 2020;31:354–362.

Abstract

Background Hearing aid prescriptive methods are a commonly recommended component of evidence-based preferred practice guidelines and are often implemented in the hearing aid programming software. Previous studies evaluating hearing aid manufacturers' software-derived fittings to prescriptions have shown significant deviations from targets. However, few such studies examined the accuracy of software-derived fittings for the Desired Sensation Level (DSL) v5.0 prescription.

Purpose The purpose of this study was to evaluate the accuracy of software-derived fittings to the DSL v5.0 prescription, across a range of hearing aid brands, audiograms, and test levels.

Research Design This study is a prospective chart review with simulated cases.

Data Collection and Analysis A set of software-derived fittings were created for a six-month-old test case, across audiograms ranging from mild to profound. The aided output from each fitting was verified in the test box at 55-, 65-, 75-, and 90-dB SPL, and compared with DSL v5.0 child targets. The deviations from target across frequencies 250-6000 Hz were calculated, together with the root-mean-square error (RMSE) from target. The aided Speech Intelligibility Index (SII) values generated for the speech passages at 55- and 65-dB SPL were compared with published norms.

Study Sample Thirteen behind-the-ear style hearing aids from eight manufacturers were tested.

Results The amount of deviation per frequency was dependent on the test level and degree of hearing loss. Most software-derived fittings for mild-to-moderately severe hearing losses fell within ± 5 dB of the target for most frequencies. RMSE results revealed more than 84% of those hearing aid fittings for the mild-to-moderate hearing losses were within 5 dB at all test levels. Fittings for severe to profound hearing losses had the greatest deviation from target and RMSE. Aided SII values for the mild-to-moderate audiograms fell within the normative range for DSL pediatric fittings, although they fell within the lower portion of the distribution. For more severe losses, SII values for some hearing aids fell below the normative range.

Conclusions In this study, use of the software-derived manufacturers' fittings based on the DSL v5.0 pediatric targets set most hearing aids within a clinically acceptable range around the prescribed target, particularly for mild-to-moderate hearing losses. However, it is likely that clinician adjustment based on verification of hearing aid output would be required to optimize the fit to target, maximize aided SII, and ensure appropriate audibility across all degrees of hearing loss.

Keywords

- ▶ Desired Sensation Level
- ▶ fit to target
- ▶ hearing aid verification
- ▶ hearing aids
- ▶ pediatric
- ▶ Speech Intelligibility Index

Introduction

The Desired Sensation Level (DSL) method (Scollie et al)²⁸ is an evidence-based fitting algorithm with prescribed targets for adults (v5.0 adult) and children (v5.0 pediatric), which is used together with clinical procedures (Bagatto et al.).^{6,21} The DSL algorithm has been implemented in many manufacturers' software systems, allowing clinicians to fit hearing aids using these targets either through prefitting from the hearing aid fitting software or preferably by fine-tuning on the basis of routine verification with a hearing aid analyzer. Verifying that hearing aids match a validated prescriptive target is considered best practice (AAA)¹ and matching targets on the basis of verification to within ± 5 dB of the target is recommended (Ricketts and Mueller;²⁶ Baker and Jenstad).⁸ Several studies have examined hearing aid fittings with children to evaluate the proximity of these fittings to prescriptive targets (McCreery et al)¹⁸ and the overall ability to match targets (Moodie et al).²² In both studies, deviation from the prescriptive target was quantified using the root-mean-square error (RMSE) of the fitting compared with the prescribed DSL targets at 500, 1000, 2000, and 4000 Hz. Aided Speech Intelligibility Index (SII) results were also analyzed in those studies as an indicator of the achieved audibility of amplified speech. Moodie et al (2017) provided evidence that clinical hearing aid fittings can typically be achieved to ≤ 5 -dB RMSE of the DSL v5.0 pediatric targets, although greater deviations at higher frequencies and with more severe hearing losses were noted. Proximity of fit to targets both overall and in the high frequencies has been shown to predict positive long-term outcomes for children who use hearing aids (McCreery et al).^{18,19}

It is recommended that clinicians verify hearing aid fittings to a validated evidence-based hearing aid prescription using real-ear measurement or coupler-based verification corrected with either a measured or age- appropriate real-ear-to-coupler difference (RECD) (AAA)¹; however, it is possible for the clinician to use the manufacturer's software to set the hearing aid's frequency gain response without verification. These software-derived fittings rely on the manufacturer's implementation of the prescriptive method within their software and lack the feedback of measurement of the actual output of the hearing aid for either calibrated speech or high-level signals for maximum output. Typically, the manufacturer offers their own proprietary fitting algorithm in addition to validated prescriptive methods, such as the National Acoustics Laboratories Nonlinear version 2 (Keidser et al)¹³ and DSL version 5, within their fitting software. Seewald et al (2008) compared the fittings obtained from five manufacturers using their proprietary algorithm, or when there was no proprietary algorithm, their default prescriptive method (i.e., [National Acoustics Laboratories Nonlinear version 1] or DSL [i/o]) for a series of nine different hearing losses for a six-month-old child. They found significant variance in fitted output across devices at various speech input levels and up to a 30 dB range of maximum output levels for the mildest audiogram. The authors argued that this 30 dB range of between-manufacturer variation

likely resulted in some hearing aids being fitted with either too much or too little output. Since the time of publication of that study, implementation of generic prescriptions has become more common, and some software systems default to generic prescriptions for pediatric fittings even when proprietary formulae are defaults for adult fittings. However, it is not clear whether implementation of evidence-based generic prescription algorithms has improved consistency in hearing aid output for unverified software-derived fittings. Studies examining the manufacturers' software-derived fit to target (FTT) for various versions of the NAL prescription method have found significant deviations between the software-derived default responses and the prescribed targets (Ricketts and Mueller;²⁶ Leavitt and Flexer;¹⁷ Sanders et al;²⁷ Amlani et al;³ Leavitt et al).¹⁶

Practice surveys indicate that only a small subset of hearing health-care providers (30–44%) regularly use verification measures to cross-check the accuracy of their fittings (Mueller;²³ Kirkwood;¹⁴ Mueller and Picou).²⁵ In a more recent report, Leavitt et al (2017) reported that the 30–40% overestimated the use of real-ear measurement for their adult population because hearing aids tested in their study suggested only 2.3% had been verified. The overall limited adoption of the use of real-ear verification over the past 30 years (Mueller)²⁷ suggests that many hearing health-care professionals continue to rely on the manufacturers' implementation of the hearing aid prescription to generate the expected output by frequency and level. For these reasons, evaluation of the accuracy of these systems may help inform whether fitting method is an important factor in hearing aid fitting consistency.

Most studies discussed earlier used the manufacturers' implementation of the NAL family of targets for adult hearing aid fittings. It is, therefore, of interest to consider clinical usage of verification in the pediatric context. The results of several studies suggest that the use of best practice procedures, including verification of hearing aid fittings, is higher with pediatric audiologists. Moodie et al (2016) surveyed more than 300 North American audiologists who self-identified as delivering pediatric audiology services. They found 70% reported using best-practice protocols. McCreery et al (2013) also found that most pediatric audiologists surveyed in their study reported using methods consistent with pediatric verification guidelines. Although these numbers are higher than those found for general audiology practitioners, results still indicate 14–30% of pediatric audiologists did not follow recommended guidelines for verification. McCreery et al (2013) also emphasized that many children with hearing loss are not serviced by pediatric audiologists, suggesting the number of unverified pediatric fittings may be higher. Krishnan and Simpson (2018) surveyed their Purdue University Audiology students about their experience with real-ear measures at clinical placements in eight American states. Results from their study showed that students reported 37% of pediatric fittings and 46% of adult fittings were not verified. Overall, these results indicate that an evaluation of the success of software-derived fittings to the DSL v5.0 pediatric targets may be important to describe the quality of unverified pediatric hearing aid fittings.

Recent advances have produced better tools for use in developing software-derived fittings. These include industry-standard speech test signals (Holube et al)¹¹, industry-standard audiograms (Bisgaard et al)⁹, a published index of fit to targets using a rootmean-square error measure (McCreery et al;¹⁸ Moodie et al)²², updated standards for electroacoustic evaluations of hearing aids (ANSI/ASA S3.22-2009), normative data for the aided SII (Moodie et al)²², and an increased awareness of the need to improve accuracy in software-derived fittings for validated prescriptive methods in general. It was the purpose of this study, therefore, to provide a contemporary evaluation of fitting accuracy for the DSL method to update our knowledge of software-derived hearing aid fitting performance.

In the present study, we examined the accuracy of software-derived hearing aid fittings to the DSL v5.0 pediatric targets by verifying the output of the hearing aid from a range of major hearing aid manufacturers. The objectives of this study were to (a) evaluate the software-derived fit to targets using frequency-specific measures to determine which audiograms were fitted within ± 5 dB per frequency as recommended by Moodie et al (2017); (b) evaluate the fit to targets using the RMSE index to determine whether software-derived fittings fall within a commonly recommended ≤ 5 dB value (Baker and Jenstad;⁸ Moodie et al;²² McCreery et al;)^{18,19} and (c) report the aided SII value against normative data to determine whether software-derived fittings fall within the same range as verified hearing aids reported by Moodie et al (2017).

Methods

Thirteen behind-the-ear style hearing aids were chosen from eight manufacturers. Evaluation was conducted at the National Centre for Audiology at Western University in London, Canada. Approval from the board of research ethics was not required because the study did not involve human participants. All hearing aids were commercially available at the time of testing (2016) and were in the mid-to-high end range of technology. Only hearing aids that offered the option of fitting to the DSL v5.0 pediatric targets in their software were included in this study. If a manufacturer had two hearing aids with the same reported fitting range on different software platforms, the hearing aid on the newer platform was included in the study. A manufacturer may have had more than one hearing aid per audiogram included in the study if the fitting ranges for the two hearing aids were reportedly different in the software.

Hearing aids were fitted to the standard "N" audiograms described by Bisgaard et al (2010). These gently sloping audiograms range from mild to profound in severity and were developed to facilitate device evaluation studies. ► **Table 1** describes the number of manufacturers and number of hearing aids per standard audiogram included in this study.

ANSI tests (ANSI 3.22)⁴ were completed to ensure that each hearing aid was functioning according to the manufacturer's specifications. One hearing aid failed ANSI 3.22 criteria and another would not connect to the manufacturer's software. Both devices were replaced with the same model that met specifications.

Table 1 The Number of Hearing Aids and Manufacturers Tested for Each of the Standard Audiograms

Audiogram	Number of Hearing Aids	Number of Manufacturers
N1	7	6
N2	9	8
N3	10	8
N4	12	8
N5	13	8
N6	8	7
N7	4	4

Using a client age of six months, client files were created in the NOAH 4 (Hearing Instrument Manufacturers' Software Association, Denmark) for each of the standard N-audiograms (Bisgaard et al)⁹, which range in severity from mild to profound, and in configuration from flat to moderately sloping. This resulted in seven client files. The hearing aids were programmed using the most current version of each manufacturer's software installed on NOAH 4 as of July 2016. Using the automatic fit option in the manufacturers' software, the aids were fitted monaurally to the left ear for a six-month-old child, using an insert phone audiogram with a foam tip coupling to the ear and a predicted foam tip RECD. This was repeated for each audiogram, using a hearing aid that fell within the recommended fitting range for each audiogram.

Because of age-based default hearing aid settings within some manufacturers' fitting software, certain features were changed before verification. Specifically, microphone mode was set to omnidirectional and noise reduction, frequency compression, feedback cancellation, and other adaptive features were turned off or set to minimum for testing. Once programmed to one of the standard audiograms, the hearing aid was coupled to the 0.4 cc coupler with behind-the-ear adaptor and placed in an Audioscan VF2 hearing aid test box (software version 4.4.4; Dorchester, ON, Canada). The thresholds for the appropriate standard audiogram were entered into the VF2, and DSL v5.0 pediatric targets for a six-month-old child with an average foam tip RECD were generated within the VF2. The software functions within the VF2 transformed the coupler-based measures to simulated real-ear-aided responses. Simulated real-ear-aided responses were measured for the International Speech Test Signal at 55-, 65-, and 75-dB SPL, and for the maximum power output (MPO) stimulus from the VF2. Some hearing aids were able to fit more than one of the standard audiograms. In total, 63 fittings were created across the manufacturer and audiogram type (► **Table 1**).

Session files from the VF2 were exported (.XML format), saved, and imported into a spreadsheet for further analysis. Session files contained both the prescriptive targets and the measured aided responses and aided SII values from the verification measurements. From these data, the differences between the DSL v5.0 pediatric target and the measured output for each of the four signal levels (55-, 65-, 75-, and 90-dB SPL) were calculated to find the deviation from target at

250, 500, 1000, 2000, 3000, 4000, and 6000 Hz. The overall proximity of the fitting to the DSL target was calculated using the average RMSE of the software-derived fitting compared with the DSL v5.0 pediatric target for 500, 1000, 2000, and 4000 Hz (McCreery et al).¹⁸ The aided SII for each hearing aid was extracted from each VF2 session file for further analysis, and the audiometric pure-tone average for 500, 1000, and 2000 Hz was calculated for each of the standard audiograms and logged for later evaluation of the aided SII values.

Results

Frequency-Specific Deviation from Target

The deviations from target across frequency were computed from 250 to 6000 Hz by subtracting the target from the measured response. Computed in this way, positive values indicate that the hearing aid was over target. A repeated measures analysis of variance (ANOVA) was completed using frequency as the within-subject variable, test level and standard audiogram type as between-subject variables, and deviations from target per frequency as the dependent variable (SPSS v24; IBM Corporation, Chicago, IL). The effect of test frequency was significant, indicating that more deviation was observed at some frequencies than others [$F_{(3,590, 879,492)} = 287.875$, $p < 0.001$, $\eta^2 = 0.540$]. Test level was significant, indicating that match to target may be leveldependent [$F_{(3, 245)} = 18.347$, $p = < 0.001$, $\eta^2 = 0.183$]. The effect of audiogram type was also significant, indicating that match to target may be more successful for some hearing losses than others [$F_{(6, 245)} = 48.753$, $p = < 0.001$, $\eta^2 = 0.544$]. Furthermore, there were interactions between frequency and standard audiogram [$F_{(21,539, 879,492)} = 26.935$, $p = < 0.001$, $\eta^2 = 0.397$] and frequency and test level [$F_{(10,769, 879,492)} = 6.204$, $p = < 0.001$, $\eta^2 = 0.071$].

Pairwise comparisons indicated that the fittings for the mild-to-moderate N1, N2, and N3 audiograms were not significantly different from each other but were significantly different from the fittings for the N4-N7 audiograms. Fittings for N4 and N5 were not significantly different from each other but were significantly different than the others, and similarly, N6 and N7 were not significantly different from each other but were significantly different from N1 to N5. Descriptive statistics of the signed mean difference, standard deviation (SD), and proportion of fittings that fell within a 65-dB deviation from DSL v5.0 target at each frequency for the four levels tested are provided for each of these groups in **Table 2**. Positive and negative values of the signed means in this table indicate fittings that were on average over or under target, respectively.

The minimum and maximum deviations from target across frequency for 55-, 65-, and 75-dB SPL speech signals for each of the standard audiograms are shown in **Figure 1**.

In general, these results indicate that the fitting accuracy declined as the severity of the audiogram increased.

RMSE: Effects of Audiogram Type and Test Level

Recall that the RMSE value is used to index the overall fit to targets and has been reported across the frequencies 500, 1000, 2000, and 4000 Hz (McCreery et al;¹⁸ Baker and Jenstad;⁸ Moodie et al).²² Fittings evaluated using this method are considered to have an acceptable proximity to targets if the RMSE value is < 5 dB. The mean RMSE and SD for each fitting in this study are reported in **Table 3**.

Consistent with results from Moodie et al (2017), when averaged across fittings, an RMSE of ≤ 5 dB was achievable for all levels of the mild-to-moderately severe (N1-N5) audiograms. Examining the RMSE of individual hearing aid fittings

Table 2 Descriptive Statistics of Deviation from Target, Including Signed Mean Difference, SD of the Signed Difference, and Percentage of the Fittings within a ± 5 -dB Deviation from Target for Each Frequency Per Level for the Groupings of N1-N3; N4-N5; and N6-N7

dB SPL	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	6000 Hz
N1-N3						
55, mean (SD) %	-2.1 (2.3) 85	-3.7 (2.2) 73	-2.5 (2.3) 88	1.44 (3.8) 73	-1.9 (2.8) 92	-1.6 (4.0) 77
65, mean (SD) %	-1.6 (3.3) 88	-3.3 (1.8) 81	-2.1 (2.3) 92	2.4 (3.9) 73	-2 (3.1) 88	-1.4 (3.9) 81
75, mean (SD) %	-0.1 (3.2) 92	-1.4 (1.8) 100	0.4 (2.4) 100	2.8 (3.9) 77	-2.6 (3.8) 69	-3.5 (3.3) 73
N4-N5						
55, mean (SD) %	-2.2 (4.5) 80	-2.0 (3.4) 88	-1.6 (3.5) 84	2.9 (4.1) 72	-1.8 (3.0) 88	-5.4 (5.4) 48
65, mean (SD) %	-1.9 (4.4) 72	-2.5 (3.1) 88	-2.2 (3.6) 76	0.5 (3.8) 72	-4.5 (2.5) 72	-8.9 (4.9) 24
75, mean (SD) %	-2.0 (4.5) 72	-1.1 (3.5) 84	0.0 (2.9) 100	0.1 (4.1) 76	-5.6 (3.5) 40	-11.5 (4.6) 8
MPO, mean (SD) %	-0.5 (3.8) 84	-0.3 (3.6) 84	1.8 (3.4) 80	2.9 (4.1) 68	0.0 (4.6) 72	-8.9 (8.3) 32
N6-N7						
55, mean (SD) %	-1.0 (6.0) 75	-1.4 (2.8) 100	-1.4 (3.3) 100	3.7 (2.5) 83	-13.2 (3.6) 0	-20.8 (6.6) 0
65, mean (SD) %	-4.7 (5.7) 58	-3.6 (3.7) 58	-3.4 (3.6) 58	-0.7 (2.3) 92	-15.2 (3.4) 0	-24.1 (7.8) 0
75, mean (SD) %	-6.8 (5.7) 50	-6.0 (5.1) 33	-4.8 (3.8) 50	-3.5 (3.2) 58	-14.1 (3.6) 0	-28.2 (8.1) 0
MPO, mean (SD) %	-1.8 (3.2) 83	-0.7 (3.4) 92	0.7 (3.5) 92	1.9 (3.9) 75	-9.0 (4.7) 25	-23.8 (14.7) 0

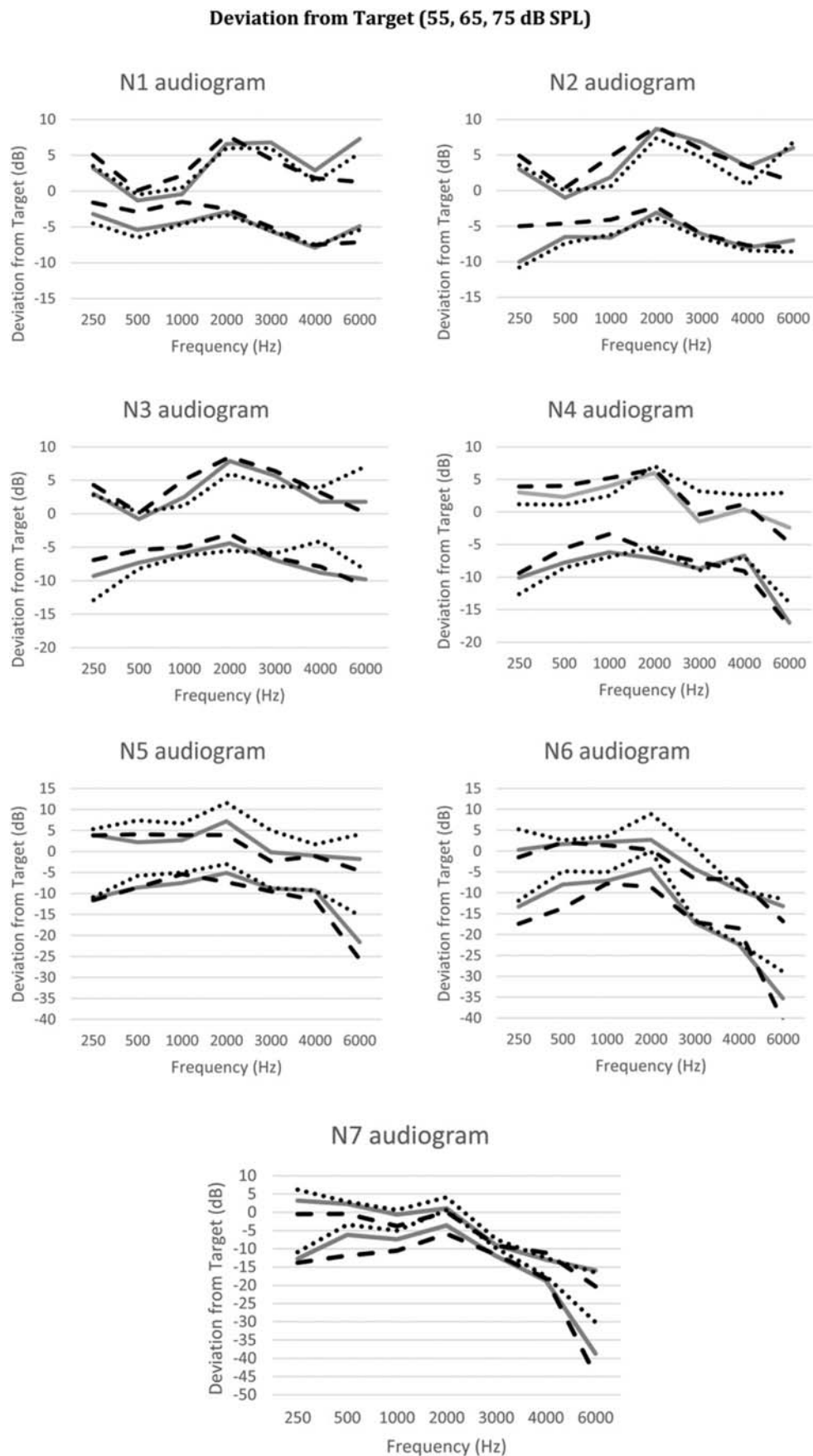


Fig. 1 For each standard audiogram, the minimum and maximum deviation from target values from all hearing aid fittings are plotted for each frequency for 55 dB (... ..), 65 dB (—), and 75 dB (---).

Table 3 Mean RMSE from DSL v5.0 Pediatric Target for Each of the Fittings Per Bisgaard Standard Audiogram Calculated Using 500, 1000, 2000, and 4000 Hz with (SD)

Audiogram	55 dB	65 dB	75 dB	90 dB—MPO
N1	3.8 (0.6)	3.6 (0.7)	3.2 (1.1)	2.9 (0.9)
N2	3.8 (1.2)	3.6 (1.3)	3.5 (1.4)	3.0 (1.9)
N3	3.4 (1.8)	3.6 (1.3)	3.5 (1.7)	2.8 (1.5)
N4	3.8 (1.5)	3.8 (1.2)	4.0 (1.5)	3.6 (1.2)
N5	3.9 (1.5)	4.3 (1.4)	4.5 (1.3)	3.8 (1.1)
N6	7.2 (1.8)	8.3 (1.7)	8.4 (2.0)	5.4 (2.3)
N7	7.7 (1.1)	9.0 (1.5)	9.8 (1.7)	6.8 (2.2)

for this range, results showed that most fittings achieved an RMSE of ≤ 5 . RMSE was larger for more severe audiograms with only 8% of the hearing aid fittings for N6 and N7 achieving an RMSE of ≤ 5 at 55-, 65-, and 75-dB SPL (► **Table 4**).

A repeated measures ANOVA was used to evaluate differences between RMSE values across audiograms and test levels. The Greenhouse-Geisser epsilon was used to adjust the degrees of freedom for nonsphericity in the data. RMSE at 55-, 65-, 75-, and 90-dB SPL were included in the initial analysis. With all standard audiograms included in the analysis, the test level was significant [$F_{(1,992, 111,55)} = 10.24, p \leq 0.001, \eta^2 = 0.155$] as was the audiogram type [$F_{(6, 56)} = 31.44, p \leq 0.001, \eta^2 = 0.771$]. Post hoc paired comparisons were completed with Bonferroni corrections between the seven audiogram types. The RMSE values for the fittings to the N1, N2, N3, N4, and N5 audiograms were not significantly different from one another, nor were the RMSE values for fittings to N6 and N7. However, the N1-N5 group was significantly different than the N6-N7 group. Because of this finding, the data were separated into two groups, (a) N1-N5 and (b) N6-N7, and ANOVAs were run on each subgroup to examine the effects of test level. Within the N1-N5 group, there were no significant effects of audiogram type or test level, nor any interaction between audiogram type and test level ($p > 0.05$). For the N6-N7 group of fittings, test level was significant [$F_{(1,898, 18,979)} = 11.172, p = 0.001, \eta^2 = 0.528$], but there were no effects of audiogram type and no interaction between test level and audiogram.

Effects of Manufacturer Type: RMSE

Recall that hearing aids were fitted only to the standard audiograms that fell within the manufacturer's reported

Table 4 Percentage of Fittings with an RMSE of ≤ 5 dB with (maximum RMSE if > 5) Per Level

dB SPL	N1–N5	N6–N7
55	88 (6)	8(11)
65	84 (6)	8(11)
75	88 (7)	8(11)
90—MPO	92 (6)	58 (10)

fitting range. As a result, not all manufacturers were included in all the standard audiogram types. However, the N5 audiogram was fitted by all of the 13 hearing aids from all the manufacturers. Therefore, the subset of fittings for the N5 audiogram was selected to evaluate whether significant differences between software-derived fit to targets were dependent on the manufacturer type. A repeated measures ANOVA, with test level as a repeated factor, indicated that there was no significant difference between the different manufacturers' RMSE at any test level ($p > 0.05$).

Effects of Manufacturer Type: Frequency Specific Deviation from Target

To examine differences between manufacturers across the entire frequency range, the measured output at 250-6000 Hz for 55-, 65-, and 75-dB SPL and MPO for each hearing aid fitted to N5 was entered into SPSS, together with the DSL targets. A repeated measures ANOVA comparing aided outputs with targets showed that the "manufacturer" was a significant factor in differences from target [$F_{(8, 20)} = 11.782, p < 0.001, \eta = 0.825$]. Manufacturer type also interacted with test frequency [$F_{(18,10, 45,25)} = 5.73, p < 0.001, \eta = 0.696$]. There was no three-way interaction between the manufacturer, frequency, and level. Pairwise comparisons of the DSL targets to aided output for each manufacturer showed no significant overall output difference for six of the eight manufacturers; however, two manufacturers, each having two aids fit within the N5 range, were significantly different from DSL ($p = 0.014$ and $p \leq 0.001$).

Pairwise comparisons of the DSL target output and each manufacturer's measured output at each frequency (collapsed across test level) found that most manufacturers provided fits to targets within 5 dB for most frequencies; however, most fittings at 6000 Hz were outside of this range (► **Table 5**). In addition, review of the individual fittings in this group revealed some that fell within a recommended < 5 -dB RMSE but had larger frequency-specific deviations

Table 5 Number of Manufacturers (Out of Eight Total) Whose Output (Collapsed across Test Levels) is > 5.0 dB from the DSL Target Output when Programmed to the N5 Audiogram

Frequency (Hz)	Number of Manufacturers with > 5.0 -dB Difference
250	3
500	1
1000	0
2000	0
3000	1
4000	3
6000	6
Collapsed across all frequencies	1

Note: Each frequency is presented as well as the result of overall output difference collapsed across frequencies.

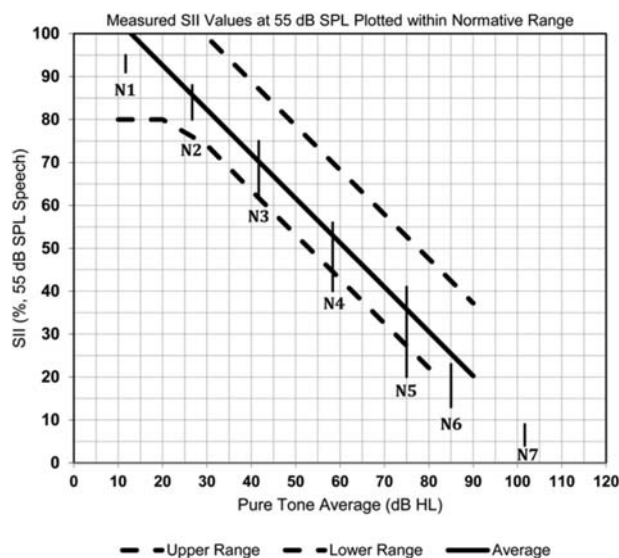


Fig. 2. Ranges of SII values obtained across hearing aids at 55-dB SPL. Results are plotted against the typical range for DSL pediatric fittings from Moodie et al (2017).

that were located as either test level specific or were at frequencies not included in the RMSE.

Aided SII

In clinical practice, aided SII values from hearing aid verification may be plotted against three frequency pure-tone average (500, 1000, and 2000 Hz) on a score sheet if the clinician intends to evaluate the aided SII value against typical audiologist-verified fittings (Bagatto et al).⁷ An updated score sheet for this process has recently been developed (Moodie et al).²² We plotted the observed aided SII values, for speech at 55- and 65-dB SPL, from the present study, on this updated score sheet (► **Figures 2 and 3**). Results indicate that for the 55- and 65-dB signals, SII values from all

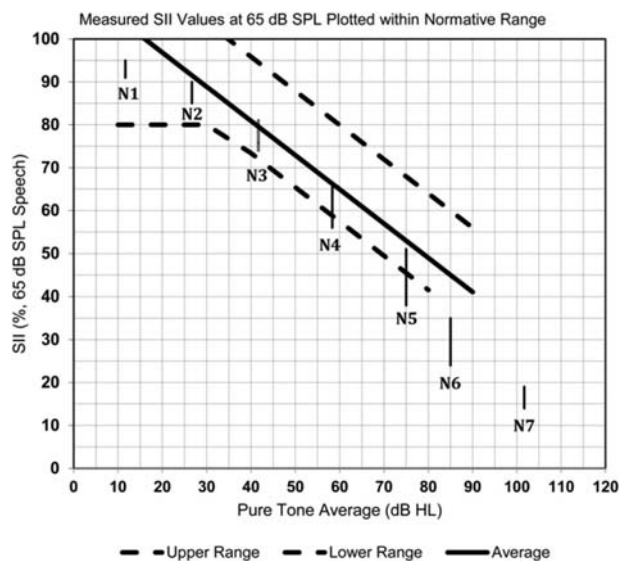


Fig. 3. SII values obtained for each hearing aid fitting at 65- dB SPL plotted within the normative range for DSL pediatric fittings from Moodie et al (2017).

hearing aid fittings fell within the typical, albeit lower portion of, the normative range for the N1-N3 audiograms. However, as the hearing losses became more severe, more of the hearing aids fell below the lower cutoff of the acceptable range. More hearing aids fell within the mid-to-upper area of the range for the 55-dB SPL test level, compared with the 65-dB SPL test level.

Discussion

In the present study, we evaluated deviation from target for hearing aid fittings that were derived from manufacturers' software-derived fittings, without any further verification and fine-tuning. Fittings were based on the fitting software's automated fit to DSL v5.0 pediatric targets, using a six-month-old infant test case, repeated across a set of standardized audiograms. Hearing aids were chosen from a range of manufacturers and were selected for inclusion only if the audiogram fell within the fitting range for the specific hearing aid. All fittings were tested using speech at three test levels, and with a test of maximum output, and the resulting descriptive verification was compared against targets to determine if the software-assisted fittings met target.

A variety of FTT methods are used in the literature, including measures of raw and/or absolute deviation from targets, summary measures based on root-meansquare deviation at selected frequencies, and evaluation of whether the aided SII falls within a typical range. In this study, we presented all three types of measures and evaluated whether consistent and accurate fittings were obtained from software-derived fittings.

For deviation from targets, it is recommended that hearing aids be verified and fine-tuned to meet prescribed targets within ± 5 dB, across as broad a frequency range as possible (Baker and Jenstad;⁸ Moodie et al).²² Previous studies found that most hearing aids measured on ear were 7-10 dB below target above 2000 Hz when programmed using the software-derived National Acoustics Laboratories Nonlinear version 2 (Sanders et al).²⁷ In our study, most hearing aids measured in the coupler using the software-derived DSL v5.0 fell within ± 5 dB for N1-N3 (► **Table 2**). Deviations from target did increase with increasing severity of hearing loss in the high frequencies and at the 75-dB SPL test level for speech. Fittings to the N4 to N7 audiograms progressively showed increased deviation in the high frequencies, with most fittings falling outside of the recommended 6.5 dB at 6000 Hz for the N4-N5 audiograms and at 4000 and 6000 Hz for the N6-N7 audiograms. The negative-signed means in ► **Table 2** and plotted range of deviation (► **Figure 1**) demonstrate that deviations resulted from fits that were under target. Because this effect was located mainly in high frequencies, significant under-target responses would be expected to limit the audible bandwidth provided by the fitting. Overall, the trend to provide fittings that were under target increased as the severity of the hearing loss increased. McCreery et al (2017) examined the impact of meeting or not meeting the recommendation to meet frequency-specific targets within ± 5 dB. They compared aided speech recognition in children

whose hearing aids met targets within ± 5 dB at all levels, at some levels, or at no test levels during verification. The children in the group who had under-fitted hearing aids at all test levels had significantly poorer aided speech recognition scores than did children in the other groups, and this effect was most pronounced for children who were underfitted in the high frequencies (McCreery et al).¹⁹ As summarized earlier, we observed that although most manufacturers' fittings met this ± 5 -dB criterion, some did not. This results in a pattern of under-fitting in the high frequencies that is similar to that observed by McCreery et al (2017), which was speculated to relate to a lack of routine verification. The results of this study support this speculation.

We also computed the summary RMSE measure to quantify overall FTT, using methods recommended by McCreery et al.^{19,20} It is typically recommended that this value be <5 -dB RMSE. We observed values of ≤ 4.5 dB for the five mildest audiograms at all test levels (**►Table 3**). However, examination of the deviation across the individual frequencies tested (250-6000 Hz) revealed that for each audiogram, at least one hearing aid fell outside the ± 5 dB tolerance for at least one frequency (**►Figure 1**), despite the fitting having acceptable RMSE. We performed an across-manufacturer comparison for one commonly fitted audiogram to test consistency across brands without the additional factor of degree of hearing loss. This analysis examined both RMSE and frequency-specific deviation from target. Results revealed that seven of the eight manufacturers matched the DSL targets within 65 dB when collapsed across frequency and level. However, some manufacturers had frequency-specific deviations from targets that occurred at only some test levels, or that occurred at frequencies not included in the summary measure of FTTs (RMSE). Without routine verification, these deviations would not be detected by the clinician and are, therefore, more likely to be unresolved. Verification curves provide data at many frequencies and use graphical depictions that allow the clinician to examine frequency response smoothness and FTTs in detail, and make fine-tuning decisions in response. Overall, these results indicate that whereas large (poor) RMSE values successfully index deviation from target, small (good) RMSE values may miss frequency-specific deviations from target. This may be merely a function of the selected frequencies that are included in the RMSE calculation (i.e., 500, 1000, 2000, and 4000 Hz). Future studies could examine the impact of including more frequencies in this measure to improve sensitivity.

The final measure that was evaluated in this study was the aided SII, and we specifically evaluated whether aided SII values fell within the recommended ranges proposed by Moodie et al (2017). In this study, all SII values from the software-derived hearing aid fittings to the three mildest audiograms fell within the recommended range for 55- and 65-dB SPL speech. It should be noted, however, that these values fell mainly in the lower portion of the typical range, indicating that the software-derived fittings provided somewhat less audibility, on average, than would be expected for clinician-driven fittings that use verification and fine-tuning. As the hearing losses fell within the moderately severe to

severe range (N4-N7), SII values continued to fall within the lower end of the normative range, and some fittings fell outside of the lower boundary. This result, as well as the inability to match target for some hearing aids particularly in the high frequencies, may be attributable to either (a) the details of the manufacturer's implementation of software-derived fitting or (b) limitations in the power of the hearing aids to match prescribed targets. As mentioned previously, the hearing aids were fitted to audiograms that fell within the manufacturer's fitting range of the device. However, we noted during study implementation that fitting ranges typically do not change when the fitting formula is changed, despite differences in gain requirements across prescriptions. We also note that between-target differences increase as the hearing loss becomes more severe (Johnson and Dillon;¹² Ching et al).¹⁰ It is, therefore, possible that some recommended fitting ranges were not specific to DSL targets, and that this may have impacted whether the fitted devices in fact had enough power to provide a good match to targets. Verification of the hearing aid, possibly at the preselection stage, would be required to confirm whether a specific make and model of hearing aid had the required power to meet prescribed targets. Routine verification of hearing aids would allow clinician insight into this limitation, which could in turn affect selection decisions to base choices on expected FTTs rather than relying only on the suggested fitting range. This speculation could underlie the differences between the clinician-driven aided SII range and the observed SII values in the present study. Further investigation would be required to examine whether clinician-driven fittings depart from suggested fitting ranges and whether this is more prevalent as degree of hearing loss increases.

Summary

Overall, the software-derived fittings examined in this study provided a good starting point for most audiograms tested in this study, particularly for mild-to-moderate hearing losses. Considering the absolute magnitude of fitting deviations (**►Figure 1**), the fittings in this study exhibit higher between-manufacturer consistency compared with previously reported results by Seewald et al (2008). This is an encouraging improvement, which provides evidence that manufacturer implementation of DSL prescription in hearing aid fitting software may help contribute to fitting consistency. However, the authors acknowledge that several aspects of the study design, including the use of a single-age case with predicted rather than measured RECD values, reduced the variability that occurs across children within the same age range and as they grow. As a result, findings presented may represent the best-case scenario for audibility and proximity to target, particularly for children with greater degrees of hearing loss or devices that are already near their functional output limits at an early age. Routine verification ensures that individual ear acoustics that affect the amount of gain required from the hearing aid as the child grows are considered, and the importance of this for individual children was not assessed in this study. The need to remeasure ear canal

acoustics and retune the hearing aid to accommodate for changes is a component of practice guidelines and recent research on outcomes (AAA;¹ McCreery et al).²⁰

The results reinforce the importance of the clinician's role in the preselection of hearing aids that are in an appropriate fitting range for the chosen prescriptive target, as well as in verification and fine-tuning of the fittings. Fine-tuning can ensure that fittings fall within 5 dB of target, consider bandwidth and smoothness of fit to target, and take into account the aided SII to maximize the audibility provided by the fitting across test levels.

Abbreviations

ANOVA	analysis of variance
DSL	Desired Sensation Level
FTT	fit to target
MPO	maximum power output
NAL	National Acoustic Laboratories
RECD	real-ear-to-coupler difference
REAR	real-ear-aided response
RMSE	root-mean-square error
SD	standard deviation
SII	Speech Intelligibility Index

Funding

This project was supported in part by the Ontario Research Fund RE08-072.

Acknowledgments

The authors wish to thank Nicole Willis and Usha Choudhary for their contributions to data analysis.

References

- American Academy of Audiology (AAA). American Academy of Audiology Clinical Practice Guidelines on Pediatric Amplification. 2013 <https://galster.net/wp-content/uploads/2013/07/AAA-2013-Pediatric-Amp-Guidelines.pdf>. Accessed March 16, 2019
- American National Standards Institute (ANSI). Specification of Hearing Aid Characteristics. ANSI 3.22. New York, NY: ANSI; 2009
- Amlani AM, Pumford J, Gessling E. Real-ear measurement and its impact on aided audibility and patient loyalty. *Hearing Review*. 2017 <http://www.hearingreview.com/2017/09/real-ear-measurement-impact-aided-audibility-patient-loyalty/>. Accessed April 6, 2018
- American National Standards Institute (ANSI). Specification of hearing aid characteristics. ANSI 3.22. New York, NY: ANSI; 2009
- Bagatto M, Moodie S, Brown C, Malandrino A, Richert F, Clench D, Scollie S. Prescribing and verifying hearing aids applying the American Academy of Audiology pediatric amplification guideline: protocols and outcomes from the Ontario Infant Hearing Program. *J Am Acad Audiol* 2016;27(03):188–203
- Bagatto M, Moodie S, Scollie S, Seewald R, Moodie K, Pumford J, Liu KPR. Clinical protocols for hearing instrument fitting in the Desired Sensation Level method. *Trends Amplif* 2005;9(04):199–226
- Bagatto MP, Moodie ST, Malandrino AC, Richert FM, Clench DA, Scollie SD. The University of Western Ontario pediatric audiological monitoring protocol. *Trends Amplif* 2011;15(01):57–76
- Baker S, Jenstad L. Matching real-ear targets for adult hearing aid fittings: NAL-NL1 and DSL v5.0 prescriptive formulae. *Can J Speech Lang Pathol* 2017;41(02):227–235
- Bisgaard N, Vlaming M, Dahlquist M. Standard audiograms for the IEC 60118-15 measurement procedure. *Trends Amplif* 2010;14(12):113–120
- Ching T, Quar T, Johnson E, Newall P, Sharma M. Comparing NAL-NL1 and DSL v5 in hearing aids fit to children with severe or profound hearing loss: goodness of fit-to-targets, impacts on predicted loudness and speech intelligibility. *J Am Acad Audiol* 2015;26(03):260–274
- Holube I, Fredelake S, Vlaming M, Kollmeier B. Development and analysis of an international speech test signal (ISTS). *Int J Audiol* 2010;49(12):891–903
- Johnson EE, Dillon H. A comparison of gain for adults from generic hearing aid prescriptive methods: impacts on predicted loudness, frequency bandwidth, and speech intelligibility. *J Am Acad Audiol* 2011;22:441–459
- Keidser G, Dillon H, Flax M, Ching T, Brewer S. The NALNL2 prescription procedure. *Audiol Res* 2011;1(01):e24
- Kirkwood DH. Survey: dispensers fitted more hearing aids in 2005 at higher prices. *Hear J* 2006;59(04):40–50
- Krishnan LA, Simpson JM. Stop the Madness: Verify Hearing Aid Fittings! Hope for the Future? 2018 <https://www.audiology.org/novemberdecember-2018/stop-madness-verify-hearing-aid-fittings-hope-future-title-inspired-roeser-and>. Accessed April 4, 2019
- Leavitt R, Bentler R, Flexer C. Hearing aid programming practices in Oregon: fitting errors and real ear measures. *Hear Rev* 2017;24(06):30–33
- Leavitt R, Flexer C. The importance of audibility in successful amplification of hearing loss. *Hearing Review*. 2012 <http://www.hearingreview.com/2012/12/the-importance-of-audibility-in-successful-amplification-of-hearing-loss/>. Accessed April 4, 2019
- McCreery R, Bentler RA, Roush P. The characteristics of hearing aid fittings in infants and young children. *Ear Hear* 2013;34(06):701–710
- McCreery R, Brennan M, Walker EA, Spratford M. Perceptual implications of level- and frequency-specific deviations from hearing aid prescription in children. *J Am Acad Audiol* 2017;28(09):861–875
- McCreery R, Walker EA, Spratford M, Bentler R, Holte L, Roush P, Oleson J, Van Buren J, Moeller MP. Longitudinal predictors of aided speech audibility in infants and children. *Ear Hear* 2015;36(01):24S–37S
- Moodie ST, Rall E, Eiten L, Lindley G, Gordey D, Davidson L, Bagatto M, Scollie S. Pediatric audiology in North America: current clinical practice and how it relates to the AAA Pediatric Amplification Guideline. *J Am Acad Audiol* 2016;27(03):166–187
- Moodie ST, Scollie S, Bagatto M, Keene K. The Network of Pediatric Audiologists of Canada. Fit to targets for the DSL v5.0 hearing aid prescription method for children. *Am J Audiol* 2017;26(03):251–258
- Mueller HG. Probe-mic measures: hearing aid fitting's most neglected element. *Hear J* 2005;57(10):33–41
- Mueller HG. 20Q: Today's Use of Validated Prescriptive Methods for Fitting Hearing Aids—What Would Denis Say? *Article* 14101. 2015 <http://www.audiologyonline.com>. Accessed March 8, 2018
- Mueller HG, Picou EM. Survey examines the popularity of real-ear probe-microphone measures. *Hear J* 2010;63(05):27–32
- Ricketts TA, Mueller HG. Whose NAL-NL fitting method are you using? *Hear J* 2009;62(08):10–17
- Sanders J, Stody T, Weber J, Mueller HG. Manufacturers' NAL-NL2 fittings fail real-ear verification. *Hear Rev* 2015;21(03):24
- Scollie S, Seewald R, Cornelisse L, Moodie S, Bagatto M, Lurnagaray D, Beaulac S, Pumford J. The desired sensation level multistage input/output algorithm. *Trends Amplif* 2005;9(04):159–197
- Seewald R, Mills J, Bagatto M, Scollie S, Moodie S. A comparison of manufacturer-specific prescriptive procedures for infants. *Hear J* 2008;61(11):26–34