J Am Acad Audiol 29:520–532 (2018)

Evaluation of the Repeatability and Accuracy of the Wideband Real-Ear-to-Coupler Difference

DOI: 10.3766/jaaa.17007

Jonathan M. Vaisberg*†
Paula Folkeard*
John Pumford‡
Philipp Narten*§
Susan Scollie*¶

Abstract

Background: The real-ear-to-coupler difference (RECD) is an ANSI standardized method for estimating ear canal sound pressure level (SPL) thresholds and assisting in the prediction of real-ear aided responses. It measures the difference in dB between the SPL produced in the ear canal and the SPL produced in an HA-1 2-cc coupler by the same sound source. Recent evidence demonstrates that extended high-frequency bandwidth, beyond the hearing aid bandwidth typically measured, is capable of providing additional clinical benefit. The industry has, in turn, moved toward developing hearing aids and verification equipment capable of producing and measuring extended high-frequency audible output. As a result, a revised RECD procedure conducted using a smaller, 0.4-cc coupler, known as the wideband-RECD (wRECD), has been introduced to facilitate extended high-frequency coupler-based measurements up to 12.5 kHz.

Purpose: This study aimed to (1) compare test–retest repeatability between the RECD and wRECD and (2) measure absolute agreement between the RECD and wRECD when both are referenced to a common coupler.

Research Design: RECDs and wRECDs were measured bilaterally in adult ears by calculating the dB difference in SPL between the ear canal and coupler responses. Real-ear probe microphone measures were completed twice per ear per participant for both foam-tip and customized earmold couplings using the Audioscan Verifit 1 and Verifit 2 fitting systems, followed by measurements in the respective couplers.

Study Sample: Twenty-one adults (mean age = 67 yr, range = 19–78) with typical aural anatomy (as determined by measures of impedance and otoscopy) participated in this study, leading to a sample size of 42 ears.

Data Collection and Analysis: Repeatability within RECD and wRECD was assessed for each coupling configuration using a repeated-measures analysis of variance (ANOVA) with test–retest and frequency as within-participants factors. Repeatability between the RECD and wRECD was assessed within each configuration using a repeated-measures ANOVA with test–retest, frequency, and coupler type as within-participants factors. Agreement between the RECD and wRECD was assessed for each coupling configuration using a repeated-measures ANOVA with RECD value, coupler type, and frequency as within-participants factors. Post hoc comparisons with Bonferroni corrections were used when appropriate to locate the frequencies at which differences occurred. A 3-dB criterion was defined to locate differences of clinical significance.

Results: Average absolute test–retest differences were within ±3 dB within each coupler and coupling configuration, and between the RECD and wRECD. The RECD and wRECD were in absolute agreement

*National Centre for Audiology, Faculty of Health Sciences, University of Western Ontario, London, ON, Canada; †Graduate Program in Health & Rehabilitation Sciences (Hearing Science), Faculty of Health Sciences, University of Western Ontario, London, ON, Canada; ‡Audioscan, a Division of Etymonic Design Incorporated, Dorchester, ON, Canada; §Lübeck University of Applied Sciences, Lübeck, Germany; ¶School of Communication Sciences and Disorders, Faculty of Health Sciences, University of Western Ontario, London, ON, Canada

Corresponding author: Jonathan M. Vaisberg, National Centre for Audiology, University of Western Ontario, London, ON, Canada N6G 1H1; E-mail: jvaisber@uwo.ca

This work was funded with support from Audioscan, a division of Etymonic Design Incorporated.

This work was presented at the 2016 annual meeting of the American Academy of Audiology (AudiologyNOW!) in Phoenix, AZ, April 15, 2016.

following HA-1-referenced transforms, with most frequencies agreeing within ±1 dB, except at 0.2 kHz for the earmold, and 0.2–0.25 kHz for the foam tip, where the average RECD exceeded the average wRECD by slightly >3 dB.

Conclusions: Test–retest repeatability of the RECD (up to 8 kHz) and wRECD (up to 12.5 kHz) is acceptable and similar to previously reported data. The RECD and wRECD are referenced to different couplers, but can be rendered comparable with a simple transform, producing values that are in accordance with the ANSI S3.46-2013 standard.

Key Words: coupler, ear canal, hearing aids, instrumentation, probe tube, RECD, reliability and validity, repeatability, verification

Abbreviations: ANOVA = analysis of variance; CI = confidence interval; RECD = real-ear-to-coupler difference; SD = standard deviation; SPL = sound pressure level; TM = tympanic membrane; VF = Verifit; wRECD = wideband real-ear-to-coupler difference

INTRODUCTION

The real-ear-to-coupler difference (RECD) is a transfer function which measures the difference in dB between the sound pressure level (SPL) produced in the ear canal and a known cavity volume by the same sound source (ANSI, 2013). It is calculated by subtracting a response in a 2-cc coupler from the response in an ear canal produced with the same signal, typically using a foam-tip coupling to the ear (Bentler and Pavlovic, 1989; Hawkins et al, 1990). It is used when converting hearing thresholds from hearing level to SPL and to assist real-ear aided response prediction from coupler-based hearing aid verification (Moodie et al, 1994; Scollie et al, 1998; Munro and Davis, 2003; Bagatto et al, 2005). This procedure improves the accuracy of simulated hearing aid fittings when measures of direct real-ear aided responses are not achievable. The RECD can also improve the accuracy of real-ear measurement, indirectly, by improving the accuracy with which the audiogram is converted to SPL, thereby rendering prescribed targets more accurate, if targets are prescribed from SPL thresholds. The RECD has become standardized as part of ANSI guidelines and has been incorporated into preferred practice guidelines for adult and pediatric populations, internationally (AAA, 2006; 2013; BSA and BAA, 2007).

The ANSI 3.46-2013 standard specifies that the RECD is to be measured as the difference between SPL produced from a foam tip in the ear canal and an HA-1 2-cc coupler using a high-impedance sound source. However, in clinical practice, RECDs are frequently measured using personal earmolds in addition to or as an alternative to foam tips. Although not the standard definition of the RECD, an earmold-based RECD incorporates the unique tubing length present in a patient's own earmold, allowing the clinician to incorporate earmold acoustics into the fitting process. The earmold RECD displays predictable differences attributed to coupling type (and tubing length for earmold), independent of factors such as age (Moodie et al, 2016). These differences have supported the development of earmold–foam-tip transforms which have been applied within modern verification equipment (Audioscan, 2016). Therefore, for the purposes of this article, the RECD will include measures using both foam-tip and earmold couplings.

The RECD, as implemented in some verification systems, may not fully address the requirements for extended high-frequency measurements, which includes frequencies above those typically measured, amplified, and verified in the past. The hearing aid industry appears to be moving toward developing products capable of producing and measuring extended high-frequency audible output. For example, some commercial hearing aids now demonstrate amplification capabilities up to 10 kHz (Kreisman et al, 2010; Kimlinger et al, 2015). High-frequency targets (up to 10 kHz) have also been incorporated into prescriptive formulas, such as CAMEQ2-HF (Füllgrabe et al, 2010; Moore and Sęk, 2013) and integrated into some modern verification systems (Audioscan, 2016).

The need for high-frequency verification has been made evident in a number of clinical applications. High-frequency audibility has been associated with improved speech perception and faster word-learning rates in hearing-impaired children, particularly for syllables with high-frequency energy (Stelmachowicz et al, 2001; Stelmachowicz et al, 2007; Pittman, 2008). Furthermore, frequencies beyond 6 kHz have been associated with improved perceived sound quality for speech and music in adults (Moore and Tan, 2003; Ricketts et al, 2008, Easwar et al, 2015).

Hearing aid performance characteristics are specified to be measured within a 2-cc coupler (ANSI, 2014). Although the measurement band used for hearing aid specification and tolerance tests ends at 5 kHz (ANSI, 2014), other types of hearing aid analyses (fitting to target, analysis of signal processing) can include frequencies above 5 kHz. The 2-cc coupler is limited to measuring hearing aid output levels up to ~8 kHz due to the microphone noise floor in the high-frequency region (IEC, 2006). This creates some challenges in verifying performance characteristics of extended-bandwidth hearing aids, especially >8 kHz. As a result, a 0.4-cc coupler was recently developed (IEC, 2016). This smaller volume is advantageous because it provides higher measured SPL relative to a larger volume for a device generating the same output. For example, Figure 1 shows the spectrum of the Audioscan Verifit (VF) 2 RECD transducer calibration stimulus as presented in an HA-1 2-cc coupler (black curve) and 0.4-cc coupler (gray curve) measured in 1/12th octave bands. In both cases, the transducer is coupled via a standard foam tip with 26-mm tubing. Averaged across frequencies (0.19–16 kHz), the 0.4-cc coupler SPL levels are ~14 dB greater than the HA-1 2-cc coupler. This coupler volume difference improves the ratio of the signal to microphone noise floor, especially for high frequencies when hearing aid outputs are low.

An RECD using a 0.4-cc coupler, or wideband RECD (wRECD), has now been implemented to support measurement of the electroacoustic performance of hearing aids and amplification up to 12.5 kHz (Audioscan, 2016; Glista et al, 2016). Like a standard RECD, the wRECD is measured by comparing the dB SPL produced by a known signal using a foam-tip or earmold coupling in an occluded ear canal to the dB SPL produced in a coupler, but the 0.4-cc coupler is used rather than the HA-1 2-cc coupler that is specified by ANSI for the RECD. Due to the volume differences between the 2- and 0.4-cc coupler, RECDs of the same signal are inherently different based on which coupler they were measured in. As a result, system-specific coupler differences per frequency have been developed to transform between the 2- and 0.4-cc coupler. This transform was developed by measuring the spectral difference between couplers for the same output (Figure 1, dotted curve). It may be possible, therefore, to transform the wRECD to the ANSI-standard RECD reporting format, which is referenced to the HA-1 2-cc coupler (ANSI, 2013). This transformation would facilitate the use of the measured wRECD in systems requiring the standard HA-1 2-cc reporting format (e.g., hearing aid fitting software, other verification systems). Assessment of this transformation is one of the purposes of this article.

A second purpose of this article is to assess the test–retest repeatability of the wRECD procedure. Repeatability is defined as the absolute test–retest differences between repeated measurements. This assessment is already well-established for 2-cc RECD measurements. Specifically, repeatability assessments of the RECD have revealed test–retest differences of approximately ±2 dB or less for most frequencies between 0.2 and 6 kHz (Sinclair et al, 1996, Scollie et al, 1998; Munro and Davis, 2003; Scollie et al, 2011), with larger test–retest differences observed at the lowest and highest frequencies. Low-frequency variability is often attributed to slit leak venting during recordings during real-ear measurement (Bagatto, 2001; Bagatto et al, 2002), whereas high-frequency variability is often attributed to interactions between probe tube microphone depth and standing waves in the ear canal (Mueller, 2001; Vaisberg et al, 2016). Note that these sources of variability occur due to the real-ear portion of the RECD measurement. The introduction of the wRECD warrants the need to assess its repeatability, particularly in the extended high-frequency region.

While the extended high-frequency limitation of the 2-cc coupler is overcome by using a 0.4-cc coupler, there are still extended high-frequency concerns to consider during the real-ear portion of the wRECD measurement. Real-ear measurement is a procedure which seeks to measure signal output levels at the tympanic membrane (TM) using a probe tube microphone. However,

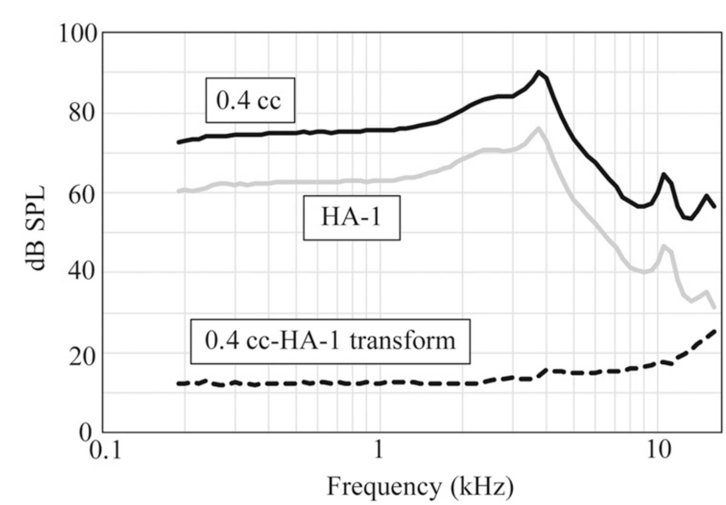

Figure 1. Spectral responses of the Audioscan VF 2 RECD transducer calibration stimulus as measured using 1/12th octave bands in the 0.4-cc (black line) and HA-1 2-cc (gray line) couplers. The dotted line represents the difference between spectra in dB, which is used to transform wRECD values to agreement with HA-1 2-cc RECD values.

to prevent physical contact at the TM, the probe tube is often placed distally. This exposes the measurement to standing waves, which occur due to interference between a sound wave traveling toward the TM and its reflection off of the TM. Standing waves cause a reduction in sound intensity relative to the intensity at the TM and can introduce error into real-ear measurements made at a distance from the TM (Gilman and Dirks, 1986; Dirks and Kincaid, 1987; Chan and Geisler, 1990). These reductions become more apparent, and occur at lower frequencies, as the probe tube moves further from the TM. As a result, shallow probe tube placement measurements may not reflect the true level at the TM.

To reduce the impact of standing wave attenuation on the typically measured frequency range of interest, ANSI (2013) recommends placing the probe tube within 6 mm of the TM. Strategies for verifying correct probe tube placement include the 6-kHz notch method, in which a 6-kHz acoustic notch can be used to determine the location of the probe tube relative to the TM (Storey and Dillon, 2001). Other methods include visually assisted probe tube placement using an otoscope (Hellstrom and Axelsson, 1993), and/or placing the end of the probe tube a fixed distance beyond the intratragal notch, depending on the age and gender of the patient (Moodie et al, 1994). In a clinic, a combination of all these strategies, plus clinical judgment, is recommended for correct probe tube placement (Pumford and Sinclair, 2001). Doing so permits up to ±2 dB of error relative to the level at the TM up to 6 kHz. More error typically occurs in the 6- to 8-kHz range for the RECD due to interactions between standing waves and the probe tube. For example, Vaisberg et al (2016) measured real-ear white noise responses in female adult ear canals at four insertion depths from the intratragal notch: 24, 26, 28, and 30 mm. Relative to the 30-mm insertion, all the shallower insertions revealed attenuations in the 6- to 8-kHz range, with the shallowest (24 mm) insertion demonstrating the most error. Given that reference measurements were made using the 30-mm insertion depth, there are likely attenuations of greater magnitude in the 6- to 8-kHz range if measuring relative to the TM. Therefore, there is possibly more error when considering extended high-frequency measurements. While this error may theoretically be overcome by placing the probe tube closer the TM, placements beyond recommended insertion depths can cause patient discomfort, preventing exceptionally deep probe tube placement measurements from being conducted (Vaisberg et al, 2016). It is therefore important to consider high-frequency measurement in the ear canal as a legitimate source of variability in wRECD measurement.

The purposes of this study were twofold: (a) to evaluate and compare the test–retest repeatability and accuracy between the wRECD and RECD using both earmold and foam-tip couplings and (b) to evaluate coupler transformations for converting the wRECD to an HA-1 2-cc coupler reference, for comparison to HA-1 2-cc RECD values. Repeatability was defined as the absolute test–retest differences between repeated measurements and accuracy was defined as the wRECD or RECD values obtained per measurement. We hypothesized that repeatability would not differ on a per-frequency basis between the wRECD and RECD from 0.2 to 8 kHz. Furthermore, we hypothesized that a coupler transformation would allow conversion of the wRECD from a 0.4-cc reference to a standard HA-1 2-cc reference.

METHODS

Participants

Participants were recruited from the University of Western Ontario Translational Research Unit participant database. A total of 21 adults between ages 19 and 78 (mean age = 66.5 yr, standard deviation [SD] = 16.2, males = 12, females = 9) participated in this study. All participants presented with normal middle ear status, as verified by impedance results that were consistent with normal adult tympanometric data (Margolis and Heller, 1987). All participants had sensorineural hearing loss sufficient to wear hearing aids. There was a range of hearing aid experience, from infrequent users to long-term experienced users. Since each participant presented with typical anatomy in both ears, each ear was treated as an independent participant. This yielded a sample total of 42 ears. This study was approved by the University of Western Ontario Health Sciences Research Ethics Board. All participants completed informed consent and were compensated for their time.

Procedure

Testing took place in a quiet carpeted laboratory at the National Centre for Audiology at the University of Western Ontario. RECD and wRECD measurements were completed using the Audioscan VF 1 (version 3.12.2) and VF 2 (version 4.2.2) hearing instrument fitting systems, respectively. The probe microphone and coupler microphone of the VF 1 and VF 2 were calibrated using the standard Audioscan VF calibration stimulus and calibration procedure. Responses in the coupler were measured before each participant by presenting the calibration stimulus from the RECD Audioscan RE770 transducer into the HA-1 2-cc coupler (for the VF 1) or 0.4-cc coupler (for the VF 2).

To measure real-ear responses, the probe tube was inserted into the ear canal until it was within 5 mm of the eardrum as judged by an experienced audiologist using the visually assisted placement technique. The audiologist measured 31 mm between the marker and the tip of the probe tube, inserted the probe tube into the ear canal, and placed the tip of the probe tube on the ridge just in front of the TM. Correct placement and probe tube orientation (i.e., the probe tube lay flat within the ear canal) was verified using otoscopy.

The ear was then occluded using either foam-tip or earmold couplings. The foam tip was an ER3A yellow or beige foam insert. The outer surface of the tip sat flush with the opening of the ear canal when inserted correctly. The majority of earmolds were custom made specifically for this study. Two participants had recent shell-style earmolds with 13 tubing and so study-specific custom earmolds were not required for them. Any earmold venting was plugged medially using putty. Tubing length for earmolds in the current study ranged from 38 to 56 mm (mean = 49 mm, SD = 0.41).

The foam tip or earmold tubing was then attached to the RECD RE770 transducer. The probe tube was repositioned if it moved during foam-tip or earmold insertion, as observed via the tube position marker. The real-ear response was measured by presenting the same VF calibration noise from the RECD RE770 transducer previously used for the coupler measurement to the ear canal via the attached foam tip or custom earmold. If slit leaks occurred, the coupling to the ear was reseated and the response was remeasured. The difference in dB between the real-ear response and coupler response using the same test signal was defined as the RECD or wRECD (depending on the fitting system used) across frequencies for each ear. The real-ear portion of the measurements was taken twice on both the VF 1 and VF 2 systems back to back for the same ear for the same participant. The coupling was reinserted into the ear canal for each measurement. Once the measurement was completed, the foam tip or earmold and the probe tube were removed from the ear. The procedure was then repeated using the participant's other ear. After measurements in both ears were completed, the procedure was repeated using the same coupling type. The procedure was then repeated again using the alternate coupling type. The ear that was measured first as well as coupling strategy was counterbalanced between participants.

There were a total of 16 measurements per participant (2 ears × 2 test–retest × 2 couplings × 2 couplers). Recordings were analyzed using 1/12th octave band analysis at 17 1/3rd octave band center frequencies (0.2, 0.25, 0.315, 0.4, 0.5, 0.63, 0.8, 1, 1.25, 1.6, 2, 2.5, 3.15, 4, 5, 6.3, and 8 kHz) for the RECD and 19 1/3rd octave band center frequencies (0.2, 0.25, 0.315, 0.4, 0.5, 0.63, 0.8, 1, 1.25, 1.6, 2, 2.5, 3.15, 4, 5, 6.3, 8, 10, and 12.5 kHz) for the wRECD.

Analysis

The data were analyzed as follows: to ensure recordings were reliably measured, test–retest repeatability for each coupling within each coupler was assessed using raw data. Statistically comparable results between the first and second measurement would suggest recordings were repeatable within configurations and that there was minimal human error. Repeatability was then assessed between coupler types for each coupling using absolute test–retest differences between recordings. Absolute values were selected so that the results would maximize the ability to observe test–retest error, rather than errors being cancelled out by averaging positive and negative differences together from the raw data. To compare the accuracy between RECD and wRECD responses, each ear's frequency response was averaged across the test and retest for each coupling.

The raw data and average coupler responses were inspected for normality using the Shapiro–Wilk test and inspection of descriptive statistics. However, absolute test–retest differences inherently yield an L-shaped distribution with substantial positive skewness, thus violating the normality assumption for parametric statistics. Therefore, to approximate a normal distribution, the data set was transformed using a logarithmic transformation with zero values (Tabachnik and Fidell, 1996; Howell, 2002). Transformed absolute test–retest differences were then inspected for normality in the same manner as the other data sets. If a condition continued to deviate from normality and led to an effect of interest, results were accepted if two conditions were satisfied: (a) equal sample sizes across groups and (b) error degrees of freedom were $\geq$20 (Mardia, 1971; Tabachnik and Fidell, 1996). A summary of conditions which violated the normality assumption can be found in Table 1. As seen in the following analysis, all conditions consisted of equal sample sizes across groups and error degrees of freedom that were $>$20.

Raw data were analyzed between test and retest within each coupler and coupling configuration using a two-factor (test–retest [2] and frequency [17] or [19] for RECD or wRECD, respectively) repeated-measures analysis of variance (ANOVA). Provided that there were no significant interactions between recordings, a two-factor (coupler type [2] and frequency [17]) repeated-measures ANOVA was used to assess repeatability of absolute test–retest differences between couplers within each coupling. For assessment of accuracy, a two-factor (coupler type [2] and frequency [17]) repeated-measures ANOVA was used to compare RECD values between couplers within each coupling. For repeatability and accuracy assessments between coupler types, responses at 10 and 12.5 kHz for the wRECD were omitted since direct pairwise comparisons at these frequencies were unavailable for the RECD. All data sets were inspected for sphericity using the Mauchly test. To protect against deviations from sphericity, degrees of freedom for all analyses were modified using

Table 1. Proportion of Conditions Which Were Statistically Significant Using the Shapiro–Wilk Test for Each Configuration within Each Analysis

Analysis	Configuration	Number of Conditions (Factors/Levels)	Count of Conditions with $p < 0.05$	Percentage of Conditions with $p < 0.05$
Repeatability within each coupling and coupler	Foam-tip RECD	34 (frequency [17] × session [2])	14	41.1
	Foam-tip wRECD	38 (frequency [19] × session [2])	15	39.5
	Earmold RECD	34 (frequency [17] × session [2])	16	47.1
	Earmold wRECD	38 (frequency [19] × session [2])	13	34.2
Repeatability between couplers (transformed values)	Foam tip	34 (frequency [17] × coupler [2])	25	73.5
	Earmold	34 (frequency [17] × coupler [2])	24	70.6
Accuracy between couplers	Foam tip	34 (frequency [17] × coupler [2])	15	44.1
	Earmold	34 (frequency [17] × coupler [2])	9	26.5

Note: $p < 0.05$ was considered statistically significant.

the Greenhouse–Geisser correction (Gray and Kinnear, 1999). Post hoc analyses were completed when appropriate using the Bonferroni method.

To determine statistical differences of clinical interest, we estimated 95% confidence intervals (CIs) for the difference between real-ear measurement and simulated real-ear measurement data reported by Munro and Davis (2003). The size of the CI was computed by adding 2 SDs to the mean difference (Munro and Davis, 2003, Table 6) and averaging values across frequencies. For the RE770 transducer, this value was 2.6 dB for measures made with the earmold and 2.9 dB for measures made with the foam tip, referenced to the HA-1 coupler. This is similar to the average 2.3 dB 95% CI reported by Scollie et al (1998) for a HA-2-referenced foam-tip RECD and less than the 5.4 dB and 4.4 dB 95% CIs reported for HA-2-referenced earmold and foam-tip RECDs, respectively, by Bagatto et al (2002). We therefore adopted a 3-dB criterion for the purposes of identifying measures that differed more than would be expected based on typical test–retest variance alone. Therefore, if a test result was statistically significant, but its comparison's mean difference descriptively fell below 3 dB, then the results were not of significant clinical interest.

For measures of repeatability within each coupler and for measures of accuracy between couplers, we used RECD values as the data for analysis. Therefore, following an ANOVA, we anticipated main effects of frequency for both measures. However, these effects are likely driven by magnitude differences of different frequencies and resonant properties in the ear canal and coupler.

RESULTS

Repeatability within Each Coupler

There were no significant interactions observed between test–retest and frequency per measurement configuration (Table 2), indicating that the data were measured reliably up to 8 kHz for the RECD and 12.5 kHz for the wRECD, using both foam-tip and earmold couplings. Although there was a main effect of test–retest for the foam-tip wRECD configuration, its mean difference was 0.34 dB, falling below the 3-dB criterion and therefore not clinically significant. Given that there were no main effects of clinical interest between test–retest, or interactions of test–retest differences with

Table 2. Summary of Repeated-Measures ANOVAs for Raw Data within Each Configuration

Configuration	Factors	df	F Ratio	*p* Value
Foam tip RECD	Frequency	3.9, 160.2	154.8	<0.01
	Test–retest	1,41	0.002	0.97
	Frequency × test–retest	3.0, 122	0.49	0.69
Foam tip wRECD	Frequency	4.5, 177.2	123.2	<0.01
	Test–retest	1, 41	4.54	<0.05
	Frequency × test–retest	4.5,175.0	0.43	0.81
Earmold RECD	Frequency	3.6, 149.2	152.3	<0.01
	Test–retest	1, 41	0.31	0.58
	Frequency × test–retest	11.8, 178.7	0.27	0.81
Earmold wRECD	Frequency	4.5, 184.9	156.6	<0.01
	Test–retest	1, 41	0.342	0.56
	Frequency × test–retest	2.7, 112.1	0.77	0.50

Notes: The main effect for test–retest using the foam-tip wRECD was not considered clinically significant because the mean difference between recordings was 0.34 dB. $\alpha = 0.05$.

frequency, absolute test–retest differences were used for assessing repeatability and average RECD values across test–retest were used for assessing accuracy.

Repeatability between RECD Types

This analysis compared whether the RECD and wRECD (coupler [2] and frequency [17]) differed in repeatability using the foam tip or the earmold. Untransformed group and individual data are displayed in Figure 2.

For the foam-tip analysis, the ANOVA revealed a main effect of frequency [$F_{(4.4,182.2)} = 31.7$, $p < 0.01$, $\eta^2 = 0.436$]. This effect is described below in the context of the 3-dB criterion. The interaction of coupler type by frequency was nonsignificant [$F_{(5.2,215.1)} = 0.72$, $p = 0.62$, $\eta^2 = 0.017$]. Although there was a main effect of coupler type [$F_{(1,41)} = 5.318$, $p < 0.05$, $\eta^2 = 0.115$], the mean difference between the RECD and wRECD was 0.41 dB, falling below the 3-dB criterion and therefore not clinically significant.

At the group level, as seen in Figure 2, mean absolute test–retest differences fell below 3 dB from 0.2 to 6.3 kHz for both couplers. The mean absolute test–retest difference exceeded 3 dB at 8 kHz for both couplers, as well as 10 and 12.5 kHz for the wRECD. At the individual level, absolute test–retest differences >3 dB were observed in up to 20% of ears measured across center frequencies from 0.315 to 5 kHz for the wRECD, and up to 12% of ears measured across center frequencies from 0.25 to 5 kHz for the RECD. Between 23% and 52% of ears elicited absolute test–retest differences >3 dB at 0.2, 2.5, and 6.3–12.5 kHz for the wRECD, and 0.2, 6.3, and 8 kHz for the RECD. Descriptively, there was a trend for more ears to elicit absolute test–retest differences >3 dB as the frequency increased from 6.3 to 8 kHz and as the frequency decreased from 2.5 to 2 kHz for both couplers. This trend continued increasing to and peaked at 10 kHz for the wRECD. Slightly fewer ears elicited absolute test–retest differences >3 dB at 12.5 kHz, relative to 10 kHz.

For the earmold analysis, the ANOVA revealed a main effect of frequency [$F_{(4.6,188.6)} = 50$, $p < 0.01$, $\eta^2 = 0.53$]. This effect is described below in the context of the 3-dB criterion. The effect of coupler type was nonsignificant [$F_{(1,41)} = 0.61$, $p = 0.44$, $\eta^2 = 0.015$] as was the interaction of coupler by frequency [$F_{(6.3,257.2)} = 1.1$, $p = 0.39$, $\eta^2 = 0.025$].

At the group level, as seen in Figure 2, mean absolute test–retest differences fell below 3 dB from 3.15 to 6.3 kHz and 10 to 12.5 kHz for the wRECD and from 0.25 to 6.3 kHz for the RECD. The mean absolute test–retest difference exceeded 3 dB at 0.2 and 8 kHz for both couplers, and 0.25 kHz for the wRECD. At the individual level,

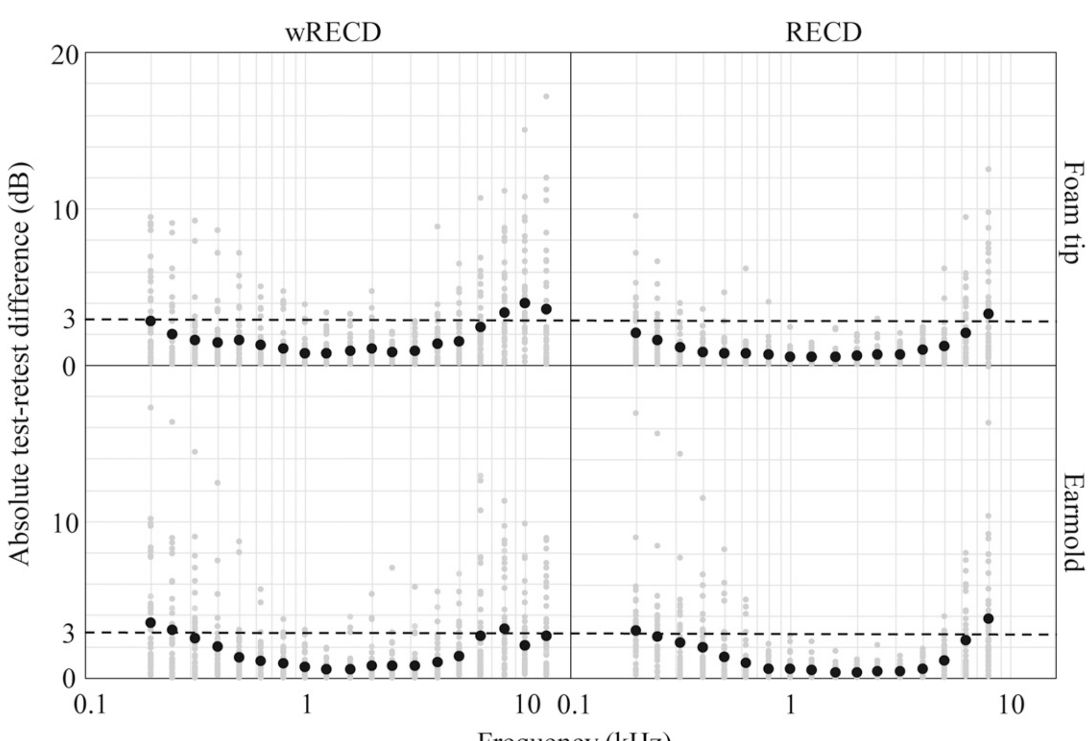

Figure 2. Test–retest repeatability measured as an absolute difference in dB between test and retest for the wRECD from 0.2 to 12.5 kHz (left column) and RECD from 0.2 to 8 kHz (right column). The top row represents measurements using the foam-tip coupling and the bottom row represents measurements using the earmold coupling. Group measurements are represented by the black circles and individual measurements are represented by the gray circles. The dotted black line represents the 3-dB criterion.

absolute test–retest differences above 3 dB were observed in up to 19% of ears across center frequencies from 0.4 to 5 kHz for the wRECD and up to 19% of ears across center frequencies from 0.4 to 5 kHz for the RECD. Between 21% and 43% of ears elicited absolute test–retest differences above 3 dB at 0.2–0.315 kHz and 8 kHz, for both couplers, as well as 10–12.5 kHz for the wRECD. Descriptively, there was a trend for more ears to elicit absolute test–retest differences above 3 dB as the frequency increased to 8 kHz and as the frequency decreased from 0.315 to 0.2 kHz for both couplers. For the wRECD, fewer ears elicited absolute test–retest differences above 3 dB at 10 kHz relative to 12.5 kHz, and even more so relative to 8 kHz.

Accuracy of RECD Types

This statistical analysis compared whether the RECD and wRECD differed following transformation to a common coupler. Examination of the untransformed RECD and wRECD values (coupler [2] and frequency [17]) shows clear differences in the magnitude of the frequency response between couplers within each coupling, as would be expected. The wRECD frequency responses were corrected to HA-1 RECD values using the transform values obtained in Figure 1. Transformed wRECD and RECD values were then averaged across the first and second run before analysis. Group data are described in Figure 3.

For the foam-tip wRECDs, the ANOVA revealed a main effect of coupler [$F_{(1,41)} = 21.8$, $p < 0.01$, $\eta^2 = 0.347$]. Across frequencies, the RECD was greater than the transformed wRECD by ~1 dB, falling below the 3-dB criterion. Therefore, this main effect was not considered clinically significant. The ANOVA also revealed a significant coupler type by frequency interaction [$F_{(3,122.9)} = 24.1$, $p < 0.01$, $\eta^2 = 0.37$]. Post hoc comparisons were used to locate the frequencies and coupler types that varied significantly from one another, comparing adjacent pairs of frequency-specific measurements across RECD types, with all RECDs transformed to the HA-1 2-cc reference (Table 3). Results revealed significant effects for frequencies 0.2–2 kHz, and 0.315–5

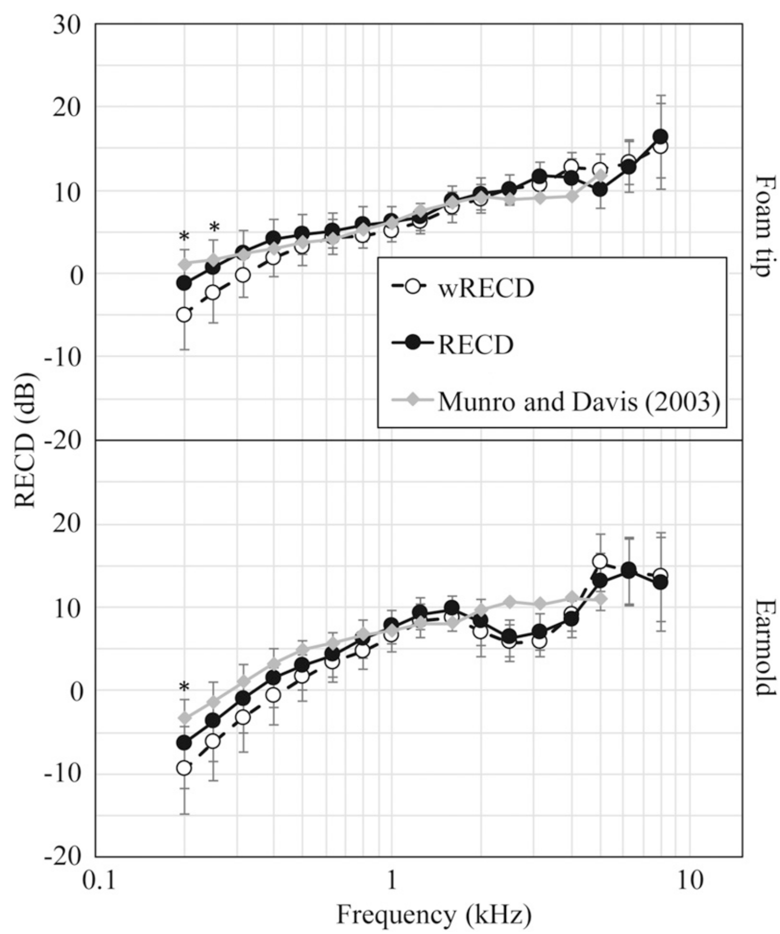

Figure 3. Mean RECDs (dark circles) and HA-1-referenced wRECDs (white circles). The top panel represents RECDs measured using the foam-tip coupling and the bottom panel represents RECDs measured using the earmold couplings. Error bars represent 1 SD of the mean. Asterisks indicate responses at frequencies in which the RECD and wRECD differed by >3 dB. Values at 10 and 12.5 kHz for the wRECD are not plotted because there is no correction factor scaling them to the RECD measurement as the RECD measures to 8 kHz. Gray diamonds represent mean RECDs plotted in Figure 3 of Munro and Davis (2003).

Table 3. Post Hoc Comparisons with Bonferroni Corrections for the Interaction of Coupler Type (RECD and wRECD) by Frequency (0.2–8 kHz at 1/3rd Octave Bands) for Measures of Accuracy

	Foam Tip			Earmold		
Frequency (kHz)	Adjusted *p* Value	Mean Difference (dB)	Clinically Significant	Adjusted *p* Value	Mean Difference (dB)	Clinically Significant
0.2	<0.01	3.739	Yes	<0.01	3.101	Yes
0.25	<0.01	3.084	Yes	<0.01	2.462	No
0.315	<0.01	2.763	No	<0.01	2.293	No
0.4	<0.01	2.348	No	<0.01	2.201	No
0.5	<0.01	1.421	No	<0.01	1.443	No
0.63	<0.01	0.793	No	<0.01	0.838	No
0.8	<0.01	1.404	No	<0.01	1.519	No
1	<0.01	0.955	No	<0.01	1.146	No
1.25	<0.01	0.585	No	<0.01	0.807	No
1.6	<0.01	0.777	No	<0.01	1.001	No
2	<0.01	0.626	No	<0.01	1.243	No
2.5	0.678	0.076	No	<0.05	0.525	No
3.15	<0.01	1.046	No	<0.01	1.229	No
4	<0.01	1.199	No	0.075	0.410	No
5	<0.01	2.388	No	<0.01	2.296	No
6.3	0.179	0.620	No	0.824	0.113	No
8	0.087	1.151	No	0.139	0.860	No

Notes: The mean difference is the RECD value relative to the HA-1-referenced wRECD value. Effects were considered clinically significant if the mean difference was >3 dB. $\alpha = 0.05$.

kHz. All mean differences fell below the 3-dB criterion, except for 0.2 and 0.25 kHz, in which the RECD was greater than the wRECD by 3.7 and 3.1 dB, respectively.

At the group level, as seen in Table 3, mean differences between RECD types were less than 3 dB, except for slightly higher differences in the low frequencies. At the individual level, at least 90% or more of RECDs measured had posttransformation differences of ≤3 dB from 0.5 to 6.3 kHz; 74% of RECDs had posttransformation differences of ≤3 dB at 8 kHz. These percentages decreased as frequency decreased in the low-frequency range, with 76%, 64%, 52%, and 43% of ears eliciting posttransformation differences below 3 dB at 0.4, 0.315, 0.25, and 0.2 kHz, respectively.

For the earmold analysis, the ANOVA revealed a main effect of coupler [$F_{(1,41)} = 31.9$, $p < 0.01$, $\eta^2 = 0.437$]. However, the RECD was greater than the wRECD by ~1 dB, falling below the 3-dB criterion. Therefore, this main effect was not considered clinically significant. The ANOVA also revealed a significant coupler by frequency interaction [$F_{(3,122.9)} = 24.1$, $p < 0.01$, $\eta^2 = 0.37$]. Post hoc comparisons were used to locate the frequencies and coupler types that varied significantly from each another, comparing adjacent pairs of frequency-specific measurements across coupler depths (Table 3). Results revealed significant effects for frequencies 0.2–3.15 kHz, and 5 kHz. All mean differences fell below the 3-dB criterion, except for 0.2 kHz, in which the RECD was greater than the wRECD by 3.1 dB.

At the group level, as seen in Table 3, mean differences between couplers fell below 3 dB for all frequencies except at 0.2 kHz, for which the mean difference was 3.1 dB. At the individual level, 83% or more of RECDs had mean differences of ≤3 dB from 0.5 to 8 kHz and 92% or more of RECDs had mean differences of ≤3 dB from 0.63 to 5 kHz. These percentages decreased in the low-frequency range, with 69%, 59%, 52%, and 48% of ears eliciting between-coupler differences below 3 dB at 0.4, 0.315, 0.25, and 0.2 kHz, respectively.

DISCUSSION

Repeatability between Couplers

Small test–retest differences were observed in the majority of individual ear canals, with greater differences occurring toward the lowest and highest frequencies. This U-shaped trend was observed for all coupling and coupler configurations, suggesting that the wRECD is equally as reliable as the RECD. The observed U-shaped trend is also consistent with previous evaluations of RECD repeatability (Sinclair et al, 1996; Scollie et al, 1998; Munro and Davis, 2003). The majority of test–retest differences fell within the 3-dB criterion, at least over the frequency range for which 95% CIs were determined (0.25–4 kHz in Scollie et al [1998] and 0.25–6 kHz in Munro and Davis [2003]).

As expected, a substantial proportion of test–retest differences were greater than the 3-dB criterion in frequencies 0.25 kHz and below for the foam-tip RECDs and 0.5 kHz and below for the earmold RECDs. This variability is likely attributed to slit leak venting during both foam-tip and earmold measurements (Bagatto et al, 2002). Recall that the 3-dB criterion was calculated on the average of all frequencies. Frequency-specific 95% CIs from these previous studies in the lower frequency range are greater than the average. The larger CIs better capture the distributions seen in the present study.

The extended-bandwidth range (≥6.3 kHz) also elicited more test–retest differences exceeding the 3-dB criterion. This variability was expected due to standing wave interactions with the probe tube microphone during the real-ear-measurement portion of the RECD (Mueller, 2001; Vaisberg et al, 2016). However, it was unknown if the 3-dB criterion was appropriate for assessing the clinical significance of the extended-bandwidth region, since the criterion was only calculated on the basis of CIs in the 0.25- to 6-kHz range. The extended-bandwidth individual test–retest differences up to 8 kHz appear to be in relative agreement with individual test–retest differences displayed in Figure 2A in Scollie et al (2011). For the wRECD, individual test–retest differences at 10 and 12.5 kHz were either in agreement or less than those at 8 kHz for both the RECD and wRECD. This suggests that the high-frequency range newly available using the 0.4-cc coupler is at least no less or more reliable than the upper limit of the frequency range available using the 2-cc coupler.

While the repeatability in the 10- to 12.5-kHz range of the wRECD appears to be comparable to the highest frequencies measured using the RECD, there is still more variability in the extended-bandwidth range relative to the 0.5- to 6-kHz range. Since this variability is attributed to standing wave interactions with the probe tube in the ear canal and not the coupler, it is of interest for future research to investigate methods which minimize these standing wave attenuations in the ear canal. One such method is the forward pressure level calibration technique. This technique isolates the incident wave from the reflected wave in the ear canal and measures the level of the incident wave only. As a result, it minimizes the effects of standing waves even when real-ear measurements are completed away from the TM (Lewis et al, 2009; McCreery et al, 2009; Richmond et al, 2011; Scheperle et al, 2011) and outperforms other methods of high-frequency measurement in the ear canal (Souza et al, 2014).

Accuracy between Couplers

The expected effect was that HA-1-referenced wRECD values made in the 0.4-cc coupler would be in agreement with HA-1 RECD values made in the 2-cc coupler. The present results are essentially consistent with that hypothesis. For the foam-tip coupling, statistically significant differences were observed between couplers at all frequencies except 2.5, 6.3, and 8 kHz. However, most differences fell below the 3-dB criterion for clinical significance, with the majority of differences falling below 1.5 dB. For the earmold coupling, statistically significant differences were observed between couplers at all frequencies except 4, 6.3, and 8 kHz. However, most differences also fell below the 3-dB criterion for clinical significance, with the majority of differences falling below 1.5 dB.

There were differences between couplers which exceeded the 3-dB criterion in the low-frequency range, which could therefore be considered clinically significant. The wRECD and RECD differed by >3 dB at 0.2 kHz for both earmold (mean difference = 3.1 dB, SD = 3.9) and foam-tip (mean difference = 3.7 dB, SD = 4.5) couplings, and at 0.25 kHz for the foam-tip coupling (mean difference = 3.1 dB, SD = 3.7). However, these differences require further consideration. First, The BSA and BAA (2007) recommends that successful hearing aid gain matches within ±5 dB of prescriptive targets at 0.5, 1, and 2 kHz during real-ear measurement. This is similar to the guidelines recommended by the AAA (2006). Therefore, the error associated with the wRECD-HA-1 correction at the lower frequencies can likely be captured within this target-matching tolerance. Second, recall that the clinical criterion was defined by calculating the average 95% CI over the 0.25- to 6-kHz frequency range for the mean difference between the measured and derived HA-1 2-cc coupler real-ear SPL in Munro and Davis (2003). The magnitude of this average is driven up by the 95% CI at the lowest frequency, 0.25 kHz, which is 5 dB for foam-tip coupling and 4.3 dB for the earmold coupling. If a clinical criterion is defined on a per-frequency basis, then the between-coupler differences would no longer be considered clinically significant.

The wRECD and RECD values observed in the present study are generally consistent with those seen in previous literature when using a similar RECD configuration. A uniform configuration was desired for comparison so that additional correction factors would not need to be applied. Munro and Davis (2003) measured RECDs in an HA-1 2-cc coupler using an RE770 transducer coupled to a foam tip and custom

Table 4. Mean Differences (1 SD) between Earmold and Foam-Tip Configurations for RECDs and wRECDs

	Frequency (kHz)					
	0.25	0.5	1	2	4	8
wRECD	−3.9 (5.4)	−1.7 (3.1)	1.2 (1.5)	−2.1 (2.1)	−3.8 (2.7)	−1.6 (6.0)
RECD	−4.6 (5.1)	−1.7 (3.3)	1.4 (1.7)	−1.5 (1.8)	−3.0 (2.8)	−3.7 (6.9)

earmold. Their RECD values were obtained for comparison purposes with the current study by digitizing both plots in Figure 3 from their article using Web Plot Digitizer (Rohatgi, 2016). The values were then plotted on Figure 3 of the current article at 0.2, 0.25, 0.315, 0.4, 0.5, 0.63, 0.8, 1, 1.25, 1.6, 2, 2.5, 3.15, 4, and 5 kHz. These frequencies were selected because they were the only frequencies available in both studies with direct comparisons. Mean RECDs at most frequencies measured by Munro and Davis (2003) fell within 1 SD of RECDs measured using the HA-1 2-cc coupler in the present study and within 2 SDs of HA-1-referenced wRECDs measured using the 0.4-cc coupler. Mean RECDs in the present study underestimated those by Munro and Davis (2003) by an average 3.2 dB from frequencies 2.5–4 kHz for the earmold, and overestimated them by 2.6 dB at 4 kHz for the foam tip. This difference may be attributed to differences in tubing length. As tubing length increases, SPL decreases, therefore affecting the value of the RECD particularly in the 2- to 4-kHz range (Gustafson et al, 2013). Mean earmold tubing length in the current study was 49 mm, which is greater than the 45 mm used by Munro and Davis (2003). In addition, the participants' mean age in the current study was 67 yr of age and each individual was an experienced hearing aid user. This is an older population relative to the average 27-yr-of-age normal-hearing adult (n = 16) evaluated by Munro and Davis (2003). This age difference may introduce variability between studies, and in addition raises the question of whether an average adult RECD should be defined for both younger and older adults. Another possible source of error may be related to the accuracy of the Web Plot Digitizer software. As previously mentioned, the previous RECD values were not explicitly provided by Munro and Davis (2003). The values were extracted from a digital image using the software. The error associated with this software has not been investigated, and may therefore be partially responsible for some of the differences seen for both the earmold and foam-tip configurations. However, despite these potential sources of variability, on average, the wRECD and RECD measured here were in agreement with those measured in a previous study using a similar configuration.

Finally, although not statistically evaluated, mean foam tip–earmold differences in measured RECDs were calculated by subtracting the average foam-tip RECD per ear across both sessions from the average earmold RECD per ear for both the HA-1 2- and 0.4-cc couplers. The values are plotted in Table 4 as a function of frequency. The mean difference values obtained were comparable across both couplers, suggesting that between-coupling differences are independent of coupler size. These values are also consistent with those plotted in Figure 3 (middle panel) of Moodie et al (2016), providing further evidence for coupling corrections derived from that study.

CONCLUSIONS AND FUTURE DIRECTIONS

The wRECD is both an accurate and repeatable clinical measurement which can be used to capture the acoustic properties of the individual ear canal, and to verify electroacoustic hearing aid performance in a simulated ear canal up to 12.5 kHz. The mean absolute test–retest difference, averaged across frequency, was 1.7 dB for the foam-tip coupling and 1.6 dB for the earmold coupling. For most frequencies, half or more of individual ears tested exhibited absolute test–retest differences of ≤3 dB, with almost all ears below this criterion in the mid-frequency range. The greatest variability occurred toward the low- and high-frequency ranges. When referenced to the HA-1 2-cc coupler, these results were in agreement with absolute test–retest differences measured using the HA-1-based RECD (at least in the 0.2- to 8-kHz range). The extended high-frequency test–retest repeatability measured using the wRECD (8–12.5 kHz) was comparable to the test–retest repeatability measured at 8 kHz using the RECD. Furthermore, wRECD accuracy values were in agreement with those measured using the RECD when measured on a per-frequency basis for both the earmold and foam-tip couplings. The accuracy and repeatability of wRECDs appear to be comparable to that of HA-1 2-cc RECDs, provided that the correct protocol has been administered. Overall, these findings support the perspective that the wRECD is a useful tool for dB HL to SPL audiometric conversions, including their use in defining real-ear measurement targets, and simulated real-ear measurement in the extended high-frequency band. However, it is currently unknown if the present findings generalize to children and populations presenting with atypical anatomy. Future studies should pursue RECD–wRECD comparisons using these populations. Furthermore, future studies could investigate comparisons between predicted and actual high-frequency real-ear measurement, as well as predicted versus measured high-frequency SPL threshold conversions.

Acknowledgments. The authors thank Jim Jonkman and Jonathan Pietrobon of Audioscan, a Division of Etymonic Design Incorporated, for their technical support and comments in preparation of this manuscript.

REFERENCES

American Academy of Audiology (AAA). (2006) *Guidelines for the Audiologic Management of Adult Hearing Impairment.* http://audiology-web.s3.amazonaws.com/migrated/haguidelines.pdf_53994876e92e42.70908344.pdf. Accessed August 31, 2015.

American Academy of Audiology (AAA). (2013) *Clinical Practice Guidelines Pediatric Amplification.* http://galster.net/wp-content/uploads/2013/07/AAA-2013-Pediatric-Amp-Guidelines.pdf. Accessed August 31, 2015.

American National Standards Institute (ANSI). (2013) *Methods of Measurement of Real-Ear Performance Characteristics of Hearing Aids. ANSI S3.46-2013*. New York, NY: Acoustical Society of America.

American National Standards Institute (ANSI). (2014) *Specification of Hearing Aid Characteristics. ANSI S3.22-2014*. New York, NY: Acoustical Society of America.

Audioscan. (2016) *Audioscan Verifit® User's Guide 4.8*. Dorchester, Canada: Etymonic Design Inc. www.audioscan.com/Docs/vf2manual.pdf. Accessed December 5, 2016.

British Society of Audiology (BSA) and British Academy of Audiology (BAA). (2007) Guidance on the use of real ear measurement to verify the fitting of digital signal processing hearing aids. www.thebsa.org.uk/wp-content/uploads/2014/04/REM.pdf. Accessed August 31, 2015.

Bagatto MP. (2001) Optimizing your RECD measurements. *Hear J* 54(9):32–36.

Bagatto M, Moodie S, Scollie S, Moodie S, Pumford J, Liu KP. (2005) Clinical protocols for hearing instrument fitting in the Desired Sensation Level method. *Trends Amplif* 9(4):199–226.

Bagatto MP, Scollie SD, Seewald RC, Moodie KS, Hoover BM. (2002) Real-ear-to-coupler difference predictions as a function of age for two coupling procedures. *J Am Acad Audiol* 13(8):407–415.

Bentler RA, Pavlovic CV. (1989) Transfer functions and correction factors used in hearing aid evaluation and research. *Ear Hear* 10(1):58–63.

Chan JC, Geisler CD. (1990) Estimation of eardrum acoustic pressure and of ear canal length from remote points in the canal. *J Acoust Soc Am* 87(3):1237–1247.

Dirks DD, Kincaid GE. (1987) Basic acoustic considerations of ear canal probe measurements. *Ear Hear* 8(5, Suppl):60S–67S.

Easwar V, Purcell DW, Aiken SJ, Parsa V, Scollie SD. (2015) Evaluation of speech-evoked envelope following responses as an objective aided outcome measure: effect of stimulus level, bandwidth, and amplification in adults with hearing loss. *Ear Hear* 36(6):635–652.

Füllgrabe C, Baer T, Stone MA, Moore BCJ. (2010) Preliminary evaluation of a method for fitting hearing aids with extended bandwidth. *Int J Audiol* 49(10):741–753.

Gilman S, Dirks DD. (1986) Acoustics of ear canal measurement of eardrum SPL in simulators. *J Acoust Soc Am* 80(3):783–793.

Glista D, Hawkins M, Moodie S, Scollie S. (2016) Understanding how RECDs fit into the "big picture" of hearing aid fitting. *Hear Rev* 23(3):26.

Gray CD, Kinnear PR. (1999) *SPSS for Windows Made Simple*. Hove, UK: Psychology Press.

Gustafson S, Pittman A, Fanning R. (2013) Effects of tubing length and coupling method on hearing threshold and real-ear to coupler difference measures. *Am J Audiol* 22(1):190–199.

Hawkins DB, Cooper WA, Thompson DJ. (1990) Comparisons among SPLs in real ears, 2 cm^3 and 6 cm^3 couplers. *J Am Acad Audiol* 1(3):154–161.

Hellstrom PA, Axelsson A. (1993) Miniature microphone probe tube measurements in the external auditory canal. *J Acoust Soc Am* 93(2):907–919.

Howell DC. (2002) *Statistical Methods for Psychology*. 5th ed. Pacific Grove, CA: Duxbury.

International Electrotechnical Commission (IEC). (2006) *Electroacoustics—Simulators of human head and ear—Part 5: 2 cm^3 coupler for the measurement of hearing aids and earphones coupled to the ear by means of ear inserts. IEC International Standard 60318*-5. Geneva, Switzerland: IEC.

International Electrotechnical Commission (IEC). (2016) *Electrocacoustics—Hearing aids—Method for measuring electroacoustic performance up to 16 kHz. IEC Technical Specification 62886*. Geneva, Switzerland: IEC.

Kimlinger C, McCreery R, Lewis D. (2015) High-frequency audibility: the effects of audiometric configuration, stimulus type, and device. *J Am Acad Audiol* 26(2):128–137.

Kreisman BM, Mazevski AG, Schum DJ, Sockalingam R. (2010) Improvements in speech understanding with wireless binaural broadband digital hearing instruments in adults with sensorineural hearing loss. *Trends Amplif* 14(1):3–11.

Lewis JD, McCreery RW, Neely ST, Stelmachowicz PG. (2009) Comparison of in-situ calibration methods for quantifying input to the middle ear. *J Acoust Soc Am* 126(6):3114–3124.

Mardia KV. (1971) The effect of nonnormality on some multivariate tests and robustness to nonnornality in the linear model. *Biometrika* 58(1):105–121.

Margolis RH, Heller JW. (1987) Screening tympanometry: criteria for medical referral. *Audiology* 26(4):197–208.

McCreery RW, Pittman A, Lewis J, Neely ST, Stelmachowicz PG. (2009) Use of forward pressure level to minimize the influence of acoustic standing waves during probe-microphone hearing-aid verification. *J Acoust Soc Am* 126(1):15–24.

Moodie S, Pietrobon J, Rall E, Lindley G, Eiten L, Gordey D, Davidson L, Moodie KS, Bagatto M, Haluschak MM, Folkeard P, Scollie S. (2016) Using the real-ear-to-coupler difference within the American Academy of Audiology Pediatric Amplification Guideline: protocols for applying and predicting earmold RECDs. *J Am Acad Audiol* 27(3):264–275.

Moodie K, Seewald R, Sinclair S. (1994) Procedure for predicting real-ear hearing aid performance in young children. *Am J Audiol* 3:23–31.

Moore BCJ, Sęk A. (2013) Comparison of the CAM2 and NAL-NL2 hearing aid fitting methods. *Ear Hear* 34(1):83–95.

Moore BCJ, Tan C-T. (2003) Perceived naturalness of spectrally distorted speech and music. *J Acoust Soc Am* 114(1):408–419.

Mueller HG. (2001) Probe microphone measurements: 20 years of progress. *Trends Amplif* 5(2):35–68.

Munro KJ, Davis J. (2003) Deriving the real-ear SPL of audiometric data using the "coupler to dial difference" and the "real ear to coupler difference". *Ear Hear* 24(2):100–110.

Pittman AL. (2008) Short-term word-learning rate in children with normal hearing and children with hearing loss in limited and extended high-frequency bandwidths. *J Speech Lang Hear Res* 51(3):785–797.

Pumford J, Sinclair S. (2001) Real-ear measurement: basic terminology and procedures. *AudiologyOnline* (Article 1229). www.audiologyonline.com/articles/real-ear-measurement-basic-terminology-1229. Accessed August 5, 2016.

Richmond SA, Kopun JG, Neely ST, Tan H, Gorga MP. (2011) Distribution of standing-wave errors in real-ear sound-level measurements. *J Acoust Soc Am* 129(5):3134–3140.

Ricketts TA, Dittberner AB, Johnson EE. (2008) High-frequency amplification and sound quality in listeners with normal through moderate hearing loss. *J Speech Lang Hear Res* 51(1):160–172.

Rohatgi A. (2016) WebPlotDigitizer. http://arohatgi.info/WebPlotDigitizer. Accessed November 7, 2016.

Scheperle RA, Goodman SS, Neely ST. (2011) Further assessment of forward pressure level for in situ calibration. *J Acoust Soc Am* 130(6):3882–3892.

Scollie S, Bagatto M, Moodie S, Crukley J. (2011) Accuracy and reliability of a real-ear-to-coupler difference measurement procedure implemented within a behind-the-ear hearing aid. *J Am Acad Audiol* 22(9):612–622.

Scollie SD, Seewald RC, Cornelisse LE, Jenstad LM. (1998) Validity and repeatability of level-independent HL to SPL transforms. *Ear Hear* 19(5):407–413.

Sinclair ST, Beauchaine KL, Moodie SK, Feigin JA, Seewald RC, Stelmachowicz PG. (1996) Repeatability of a real-ear-to-coupler difference measurement as a function of age. *Am J Audiol* 5:52–56.

Souza NN, Dhar S, Neely ST, Siegel JH. (2014) Comparison of nine methods to estimate ear-canal stimulus levels. *J Acoust Soc Am* 136(4):1768–1787.

Stelmachowicz PG, Lewis DE, Choi S, Hoover B. (2007) Effect of stimulus bandwidth on auditory skills in normal-hearing and hearing-impaired children. *Ear Hear* 28(4):483–494.

Stelmachowicz PG, Pittman AL, Hoover BM, Lewis DE. (2001) Effect of stimulus bandwidth on the perception of /s/ in normal- and hearing-impaired children and adults. *J Acoust Soc Am* 110(4):2183–2190.

Storey L, Dillon H. (2001) Estimating the location of probe microphones relative to the tympanic membrane. *J Am Acad Audiol* 12 (3):150–154.

Tabachnik BG, Fidell LS. (1996) Multivariate normality. In: Woods C, ed. *Using Multivariate Statistics*. 3rd ed. California State University, Northbridge: HarperCollins College Publisher, 381.

Vaisberg JM, Macpherson EA, Scollie SD. (2016) Extended bandwidth real-ear measurement accuracy and repeatability to 10 kHz. *Int J Audiol* 55(10):580–586.