
Electronic Thesis and Dissertation Repository

7-6-2021 10:00 AM

Objective Measurement Optimization in Early Hearing Detection and Intervention

Matthew J. B. Urichuk, *The University of Western Ontario*

Supervisor: Scollie, Susan D., *The University of Western Ontario*

Joint Supervisor: Purcell, David W., *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Health and Rehabilitation Sciences

© Matthew J. B. Urichuk 2021

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Speech and Hearing Science Commons](#), and the [Speech Pathology and Audiology Commons](#)

Recommended Citation

Urichuk, Matthew J. B., "Objective Measurement Optimization in Early Hearing Detection and Intervention" (2021). *Electronic Thesis and Dissertation Repository*. 7958.
<https://ir.lib.uwo.ca/etd/7958>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

Following the fitting of amplification devices, outcomes need to be evaluated to determine benefit of intervention. Objective measurements, which generally require rapid acoustic or electrophysiological measurement without active participation of the individual, are one subset of evaluation methods.

This thesis aimed to improve the reliability, accuracy, and efficiency of a subset of objective outcome measurements: the speech-evoked Envelope Following Response (EFR), and the wideband real-ear-to-coupler-difference (wRECD), which are used in hearing aid validation and verification respectively.

The speech-evoked EFR, a neural response reflecting phase-locked activity to the envelope of a speech stimulus, can be detected using a variety of statistical indicators. Chapter 1 focused on improving speech-evoked EFR detection by comparing the sensitivity and efficiency of statistical indicators in adults and infants. Results show that indicators using phase information tend to outperform those that do not. Accuracy and speed of speech-evoked EFR detection was also found to differ between infants and adults.

The main contribution of this thesis is the proposal and validation of a clinically viable test paradigm for wRECD measurement which does not require a probe-tube microphone and is not affected by reflected wave interference. The projects evaluated the measurement of the integrated pressure level (IPL) wRECD using a Thevenin-equivalent source parameter calibrated transducer. Calibration of the transducer was found to be reliable across time. IPL wRECD improved wRECD reliability, high-frequency performance, and simultaneously assessed middle ear function using wideband acoustic immittance. Below 5 kHz, the IPL wRECD was not clinically significantly different from probe-tube microphone measurements when a generic coupling method was used. An individual's earmold significantly impacts resulting wRECD measurements due to variable lengths of tubing associated with the earmold. The current thesis proposed a method to acoustically determine the tubing length to create accurate and reproducible generic-tip-to-earmold transforms, which improved the estimation of the earmold wRECD.

In summary, IPL wRECD measurement shows promise as an alternative to probe-tube microphone wRECD measurement procedures and is expected to improve validity of hearing aid fittings, especially in the high frequencies. Similarly, speech-evoked EFRs can be used as an objective measurement to validate hearing aid fittings accurately and efficiently.

Keywords

Objective outcome measures, Envelope Following Responses (EFR), Real-ear-to-coupler difference (RECD), real-ear measurement (REM), forward pressure level, integrated pressure level

Summary for Lay Audience

Following hearing aid fitting, outcomes need to be determined to ensure benefit of the intervention. This is done using several outcome measurement tools, such as objective outcome measurements which can be measured by recording brainwaves. Objective outcome measurements are quick and do not require sustained participant cooperation.

This thesis aimed to improve the accuracy and speed of two objective measurements: the speech-evoked Envelope Following Response (EFR) and the wideband real-ear-to-coupler-difference (wRECD). The speech-evoked EFR is used to confirm that the hearing aid is providing enough sound to the individual. The wRECD accounts for individual ear canal effects on sound input to the eardrum, improving test accuracy for growing infants.

The speech-evoked EFR is a brainwave response that can be measured while the participant sleeps. It is detected by various statistical tests which determine if sound is reaching the brainstem. Chapter two compared common statistical tests used for speech-evoked EFR detection in infants and adults. Results show that statistical tests that incorporate response phase outperform those that solely rely on response magnitude.

The main contribution of this thesis is the proposal and validation of a clinically viable wRECD measurement paradigm. Current clinical measurements require a probe-tube microphone placed within millimeters of the eardrum. In contrast, this thesis proposes a new method of measurement, the integrated pressure level (IPL) wRECD. It uses a specially calibrated earpiece which houses a speaker and a microphone. The calibration procedure allows for the calculation of the sound-level at an individual's eardrum using a measurement by the microphone in the earpiece. This thesis found that this calibration procedure was reliable over time. IPL wRECD measurement using a generic tip and the individual's earmold was also validated. Results indicate that IPL wRECD measurements are more reliable than conventional wRECD measurement. Above 4 kHz, IPL wRECD improves on probe-tube microphone measurements. Overall, IPL wRECD is a promising alternative to current wRECD measurements and may improve validity of hearing aid fittings, especially in the high frequencies. The IPL wRECD could be used with speech-evoked EFR to provide accurate, objective measurement of infant brain responses to amplified sound from hearing aids.

List of Abbreviations

AAA = American Academy of Audiology
AEP = Auditory Evoked Potentials
ANSI = American National Standards Institute
B&K = Bruel and Kjaer
BTE = Behind the ear
CI = confidence interval
dB = decibel
DFT = discrete fourier transform
ECLD = ear-canal level difference
ecSPL = ear-canal sound pressure level
EEG = electroencephalography
EFR = envelope following response
EHDI = early hearing detection and intervention
FA = Fourier analyzer
FPL = forward pressure level
HSD = Honest significant difference
Hz = Hertz
ICC = intraclass correlation coefficient
IPL = integrated pressure level
MSC = Magnitude Square Coherence
mm = millimeter
OAE = otoacoustic emissions
peSPL = peak-to-peak equivalent sound pressure level
REAR = real-ear aided response
RECD = real-ear-to-coupler difference
RMS = root-mean square
RPL = reverse pressure level
SD = standard deviation
SE = standard error
SPL = sound pressure level
TEREO = Transform for Estimating Real Ear Output

WAI = wideband acoustic immittance

wRECD = wideband real-ear-to-coupler difference

Co-Authorship Statement

This thesis includes six chapters comprised of an introductory chapter (Chapter 1), four integrated articles (Chapter 2-5), and a concluding chapter (Chapter 6). I, Matthew Urichuk, am responsible for the design, conceptualization, data collection, statistical analyses, and organization of this work. I am the lead author on all manuscripts and chapters. I am the sole author of the introductory chapter (Chapter 1) and the concluding chapter (Chapter 6). Susan Scollie, David Purcell, and Prudence Allen are acknowledged for reviewing drafts of both chapters prior to inclusion in this document. Chapter 2 was co-authored by Vijayalakshmi Easwar, Susan Scollie, and David Purcell. Susan Scollie and David Purcell assisted with analyses, guidance on project design, and edited manuscript drafts prior to submission. David Purcell provided tools for response analysis. Chapters 3-5 were co-authored by David Purcell and Susan Scollie who provided guidance on project design, statistical analyses, and reviewed all three manuscripts prior to submission. David Purcell assisted with analysis code regarding individualized ear canal cross-sectional area estimation. Susan Scollie and David Purcell also provided guidance on obtaining ethics approval.

Acknowledgments

First and foremost, thank you to my supervisors Dr. Susan Scollie and Dr. David Purcell for their continuous encouragement and mentorship. Your support and enthusiasm for my project was incredible and I could not have finished this dissertation without your extensive reviewing, suggestions, and support.

Thank you to my advisory, candidacy, and defense committee members for their feedback on this dissertation. Dr. Ryan McCreery, Dr. Blake Butler, Dr. Ewan Macpherson, Dr. Prudence Allen, and Dr. Daniel Rasetshwane – this work would not be anywhere near the same caliber had you not taken the time out of your busy schedules to be on the committees throughout my PhD, and I am very grateful to you all.

A huge thank you to the lab-mates, honorary lab-mates, and classmates that I have had throughout my years at Western. Negar, Max, Caitlin, Adrienne, Viji, Hasan, Adrian, Sarah, Josh, Nicole: I could not imagine these past four years without your friendship and guidance.

This work was supported by the Ontario Research Fund and Canadian Health and the Natural Sciences and Engineering Research Council of Canada. This work could not have come to fruition without the extensive help and technical support from Jim Jonkman, Martin Jonkman, and everyone I worked with during my time at Audioscan.

Thank you to my family, in particular my parents and siblings, who were always available when I needed a phone call and words of encouragement. I wouldn't be where I am without you all.

Finally, I couldn't have done this without the continued love and support of my fiancée, Taylor, and my soon-to-be in-laws, who took me in during a global pandemic and have been an unbelievable support network. I can't imagine getting through my PhD, or the pandemic, without you all. Taylor, I am so excited for what our next chapter will bring and couldn't imagine doing it with anyone else.

Table of Contents

Abstract.....	ii
Summary for Lay Audience	iv
List of Abbreviations.....	v
Co-Authorship Statement	vii
Acknowledgments	viii
Table of Contents	ix
List of Tables.....	xiv
List of Figures.....	xv
List of Appendices.....	xix
Chapter 1	1
1 Introduction.....	1
1.1 Outcome validation	3
1.2 Accounting for individual ear canal acoustics.....	4
1.3 Measuring individual ear canal SPL during hearing aid fitting	5
1.4 Individual ecSPL during hearing aid fitting: errors and clinical issues.....	7
1.5 Accounting for individual ear canal acoustics: alternatives to probe-tube ecSPL calibration	8
1.6 Purpose of this thesis	11
1.7 Research questions	11
1.8 References	13
Chapter 2	20
2 Performance of statistical indicators in the objective detection of speech-evoked envelope following responses.....	20
2.1 Introduction	20
2.2 Methods	22

2.2.1	Participants	22
2.2.2	Stimulus	22
2.2.3	Procedure	24
2.2.4	Statistical indicators.....	25
2.2.5	Evaluation of test performance.....	27
2.3	Results	27
2.3.1	Comparison between populations.....	33
2.3.2	Minimum testing duration for clinical benefit.....	34
2.4	Discussion.....	35
2.5	Conclusion.....	38
2.6	References	40
Chapter 3	44
3	Thévenin-equivalent source parameter reliability: intra- and inter-sessional effects on stimulus level calibration and wideband reflectance	44
3.1	Introduction	44
3.1.1	Calibration methods: ear canal sound pressure level	45
3.1.2	Calibration methods: alternatives to ear canal SPL.....	46
3.2	Thevenin-equivalent source parameter calibration.....	47
3.2.1	Source parameter derivation.....	47
3.2.2	Ear canal stimulus level measurement	49
3.2.3	Wideband acoustic immittance measurement	50
3.3	Methods	51
3.3.1	Calibration	51
3.3.2	Experimental measurement	52
3.3.3	Analysis	52
3.3.4	Clinical significance of variation.....	53

3.4	Results	54
3.4.1	Inter-session source parameter calibration reliability.....	54
3.4.2	Cross-sectional source parameter calibration reliability	57
3.5	Discussion/Conclusion	58
3.6	References	60
Chapter 4		65
4	Measurement of hearing aid tubing length using a wideband stimulus.....	65
4.1	Introduction	65
4.1.1	Sources of variability.....	67
4.1.2	Potential solution: reflectance-based tubing length estimation	68
4.1.3	Reflectance-based methods to estimate tubing length.....	69
4.1.4	Applications of an estimate of the tubing length of earmolds.....	70
4.2	Methods	70
4.2.1	Instrumentation.....	70
4.2.2	Calibration	71
4.2.3	Experimental measurement	71
4.2.4	Individualized transform	72
4.3	Results	73
4.3.1	Acoustic tubing length measurement	73
4.3.2	Individualized foam-tip-to-earmold correction using tubing length estimation	76
4.4	Discussion.....	78
4.4.1	Limitations.....	80
4.5	Conclusion.....	80
4.6	References	81
Chapter 5		85

5	Validity and reliability of integrated pressure level Real-Ear-to-Coupler Difference measurements	85
5.1	Introduction	85
5.1.1	Accounting for individualized ear-canal acoustics using a source parameter calibrated transducer.....	87
5.1.2	IPL wRECD: Method for foam-tip and earmold wRECD determination.	90
5.2	Methods	92
5.2.1	Participants	92
5.2.2	Procedure.....	92
5.2.3	Stimulus calibration.....	94
5.2.4	Analysis	95
5.2.5	Data exclusion	95
5.3	Results	96
5.3.1	Comparison between IPL wRECD and probe-tube microphone measured foam-tip wRECD.....	96
5.3.2	Comparison between earmold IPL wRECD and probe-tube microphone measured earmold wRECD	98
5.3.3	Validation of individualized hearing tubing length correction.....	101
5.3.4	Test-retest reliability of IPL wRECD measurement	103
5.3.5	Performance of individualized ear canal cross-sectional area estimation	106
5.4	Discussion.....	108
5.4.1	Main findings.....	108
5.4.2	Comparison between IPL wRECD and foam-tip wRECD.....	110
5.4.3	Comparison between earmold IPL wRECD and probe-tube microphone measured earmold wRECD and validation of individualized tubing length corrections	111
5.4.4	Performance of individualized ear canal cross-sectional area estimation	114
5.5	Conclusion.....	115
5.6	Acknowledgements	115

5.8 References	116
Chapter 6	122
6 Contributions, limitations, and conclusions.....	122
6.1 Research aims and summary of findings.....	122
6.2 Limitations.....	126
6.3 Future work	127
6.4 Conclusion.....	128
6.5 References	130
Appendices	133
Curriculum Vitae	135

List of Tables

Table 2-1: Sensitivity of response detection in infant participants for F1, F2+, and fricative stimuli after 10, 20, and 30 minutes using each statistical indicator. The best performing indicator in each condition (row) is bolded. 95% confidence interval of sensitivity estimate is provided in brackets. Pairwise post-hoc comparisons to the best performing test were completed with asterisks denoting a significant decrease in sensitivity ($p < 0.05$).31

Table 2-2: Sensitivity of response detection in adult participants for F1, F2+, and fricative stimuli after 10, 20, and 30 minutes using each statistical indicator. The best performing indicator in each condition (row) is bolded. 95% confidence interval of sensitivity estimate is provided in brackets. Pairwise post-hoc comparisons to the best performing test were completed with asterisks denoting a significant decrease in sensitivity ($p < 0.05$). Determination of significance was completed on log-odds of detection, which asymptote at 100%. As a result, detection chance near 100% can be statistically significantly different even though detection rates may differ by less than 1%.32

Table 2-3: Minimum test duration necessary for better than chance sensitivity using each statistical indicator for both infants and adults across low- (F1), mid- (F2+), and high-frequency (fricative) EFR stimuli. Asterisks denote that the minimum testing duration exceeded the testing duration of the current experiment. Lower durations indicate greater efficiency.35

Table 4-1: Standing wave measurement accuracy across different tubing length conditions.75

Table 5-1: Intraclass correlation coefficients for select 1/3 octave bands. Values within square brackets indicate 95% confidence intervals.106

List of Figures

- Figure 1-1: Framework for identifying sources of error in procedures used in Early Hearing Detection and Intervention. Areas of focus are shaded in grey.3*
- Figure 2-1: Modelled sensitivity of EFR detection in infant participants for F1, F2+, and fricative stimuli across testing duration for each statistical indicator. Shaded area indicates the 95% confidence interval of the sensitivity estimate for each test.30*
- Figure 2-2: Sensitivity of response detection in adult participants for F1, F2+, and fricative stimuli across testing duration for each statistical indicator. Shaded area indicates the 95% confidence interval of the sensitivity estimate for each test.33*
- Figure 2-3: Comparison of modelled response detection sensitivity between populations. Columns are separated by the statistical indicator used, while the rows are split by stimulus type. Shaded area indicates the 95% confidence interval of the sensitivity estimate for the population.34*
- Figure 3-1: Absolute deviation in FPL, IPL, and power reflectance from first day results at all audiometric frequencies. A clinically meaningful difference is indicated by a dotted line every 1 dB for FPL and IPL and every 0.025 unit (2.5%) change for power reflectance.56*
- Figure 3-2: Weekly change for in-ear stimulus measurement (A) and power reflectance (B) caused by source parameter calibration differences. Points indicate the regression coefficient for the weekly change, with 95% confidence intervals ($\pm 2 * SE$) visualized.57*
- Figure 3-3: Distribution of absolute values for intra-session test-retest differences in IPL (A), FPL (B), and power reflectance (C) calculations caused by source parameter calibration differences. Values were averaged across 1/3 octave bands. The transducer was re-inserted between each calibration procedure. The clinically significant cut-off is denoted by a dotted line at 1 dB (A, B) and 2.5% (C). The boxes indicate interquartile range which is the range between 25-75th percentile values with the median marked by a horizontal line. The error-bars extend to 1.5 * IQR and individual outliers are represented by single points.58*

Figure 4-1: (A) Comparison of estimated tubing lengths (y-axis) and the actual physically measured tubing lengths. Points falling on the dotted line indicate perfect agreement. Points above the line indicate overestimation of tubing lengths while points falling below the dotted line indicate underestimation of the tubing length. (B) Difference between estimated tubing length and the physical tubing lengths. Values below 0 indicate underestimation of tubing length, whereas positive values indicate overestimation. All three trials are shown.74

Figure 4-2: (A) Absolute differences (dB) between IPL and SPL measurements and (B) Measured pressure reflectance magnitude for frequencies between 2000-8000 Hz for four hearing aid tubing length (25, 35, 45, 55 mm). Vertical dashed lines are the analytic frequency at which the maximum difference and minimum reflectance values are expected; this is first standing wave frequency corresponding to the half-wave resonance of the tube. 76

Figure 4-3: Plot (A) indicates the individualized foam-tip to earmold RECD corrections across frequencies for 1 mm increments for hearing aid tubing measurements. Averaged transforms calculated by Moodie et al., (2016) are indicated by black asterisks (*). A value of 0 indicates no difference between the earmold RECD and the foam-tip RECD. Plot (B) indicates error caused by using the average foam-tip to earmold RECD rather than the individualized transform. Clinical significance for real-ear measurements is denoted by horizontal dotted lines at ± 2 dB. Averaged foam-tip-to-earmold RECD transforms do not exist for 8 kHz, so a transform value of 0 is used. Positive values indicate that the average transform underestimates the correction factor necessary for a given length of tubing whereas negative values indicate overestimation.78

Figure 5-1: Measurement procedure for determining the insert earphone wRECD and earmold wRECD using (A) probe microphone wRECD measurement procedures and (B) IPL wRECD measurement procedures. Acoustic measurements are taken at the top of each pathway. Each measurement type is referred to in the current study using the underlined term associated with each pathway. The resulting value for use in hearing aid fitting procedures (i.e. the insert-earphone wRECD or earmold wRECD) is indicated in the bottom row.94

Figure 5-2: (A) Average measurement of wRECD values as a function of frequency using either the IPL wRECD (red) or probe-tube microphone SPL wRECD (blue). Average

wRECD values for the normative adult population are overlaid with a dotted line (as implemented in the Audioscan Verifit 2). (b) signed difference in wRECD average between IPL wRECD minus probe-tube microphone foam-tip wRECD. The cut-off for clinical significance is denoted by dotted lines. Positive values indicate increased average wRECD values at a specific frequency for IPL wRECD. Error bars indicate the 95% confidence interval for each estimate. Sub-plots C and D show the distribution of all wRECD measured in the current study. Boxplots represent 25-75th percentile of the distribution, with interior horizontal line at the median and whiskers extending to 1.5 * interquartile range (IQR). Outliers are indicated by black dots.97

Figure 5-3: Individual measurements of wRECD values in all participants' right ears as a function of frequency using IPL wRECD measurement (red; triangles) and probe-tube microphone measured foam-tip wRECD measurement (blue; circles). Average wRECD values for normative adult population (as implemented in the Audioscan Verifit 2) are overlaid with dotted black lines.98

Figure 5-4: Distribution of all earmold wRECD measurements across all tubing length conditions using both earmold IPL wRECD and probe-tube measured earmold wRECD. Width of distribution indicates density of measurements yielding a specific wRECD for a given 1/3 octave band.100

Figure 5-5: (Top row) Average modelled earmold wRECD for earmolds attached to 30, 40, and 50 mm of hearing aid tubing from probe-tube microphone measures. Error bars indicate ± 1 standard error. (Bottom row) Difference between measured and estimated earmold wRECD values for the three methods of earmold wRECD estimation shown in Figure 1. Clinically significant error is indicated by the dotted line. Positive values indicate larger estimated wRECD values than measured by probe-tube microphone.101

Figure 5-6: Differences in estimation of earmold wRECD for two extreme tubing lengths (average 50 mm earmold wRECD minus average 30 mm earmold wRECD) using two tubing-length compensation methods: individualized IPL (red) and average (blue). Clinically significant cut-offs are shown by the dotted lines. Error bars indicate 95% confidence interval for each estimate ($2*SE$).103

*Figure 5-7: Absolute test-retest differences in decibels for each wRECD testing methodology across 1/3 octave bands. Boxes indicate 25-75th percentile; horizontal line indicates median test-retest difference; whiskers represent 1.5 * IQR. Individual outlier points are denoted using a single black dot. (a) Test-retest data taken from 22 individuals, as reported in Vaisberg (2016) for probe-tube microphone wRECD measurements using a generic foam insert. (b) Test-retest data obtained in the current cohort of 21 individuals for IPL wRECD determination. 105*

Figure 5-8: Improvement in agreement between the IPL wRECD and probe-tube microphone foam-tip wRECD when an individualized measurement of characteristic impedance (obtained with individualized cross-sectional area) is used compared to average characteristic impedance. Gray lines show individuals, and the black line is the group average. Error bars indicate 95% confidence interval for each frequency. A dotted line indicates 4000 Hz (see text for discussion). 108

List of Appendices

Appendix A: HSREB Approval	133
----------------------------------	-----

Chapter 1

1 Introduction

With the increasing implementation of early hearing detection and intervention (EHDI) programs, diagnosis and intervention for pediatric hearing loss can occur when an infant is only months old (Sininger et al., 2009; Tomblin et al., 2014; Uus & Bamford, 2006). Early detection and intervention that optimizes the audibility of speech through properly fitted hearing aids has a significant impact on outcomes of infants with hearing loss (McCreery et al., 2015; Moeller & Tomblin, 2015; Tomblin et al., 2015; Yoshinaga-Itano et al., 1998). However, the age of the patient can introduce challenges into the EHDI process. At such a young age, infants are developmentally unable to actively participate in any aspect of the hearing aid fitting process or subsequent outcome evaluation procedures. As such, infants in this age bracket require a modified test battery throughout the EHDI process. Hearing aid output is matched to prescriptive targets calculated from the individual's hearing loss. The prescriptive targets and hearing aid output is calibrated to the sound-pressure level produced in the ear canal. After the fitting, intervention outcomes need to be assessed to determine benefit for the patient (American Academy of Audiology, 2013; Bagatto et al., 2005; The Joint Committee on Infant Hearing, 2019).

Modified pediatric screening and diagnostic procedures consist of multiple independent testing procedures and implement the cross-check principle, which states that no one test result should be accepted during the intervention until it is confirmed by another independent procedure (Jerger & Hayes, 1976; Norrix, 2015). In a pediatric hearing aid fitting, this means that objective response measurement is confirmed by behavioural testing or vice versa. Behavioural testing requires active participation, with responses acquired from the individuals' reactions to stimuli, whether that be turning their head, placing a toy in a box, or pressing a button. Objective response measurement does not rely on behavioural responses, making it especially helpful for interventions where the child is developmentally unable to participate. Objective responses are determined based on acoustic or neural responses to a stimulus (Hall, 2016). Objective measures are used to estimate hearing thresholds, fit and verify hearing aids, and assess intervention outcomes

(Hall, 2016). Later, these measures are followed up with behavioural testing when the child is developmentally able to participate (The Joint Committee on Infant Hearing, 2019). For this reason, objective measures are especially important for pediatric hearing aid intervention.

Objective measurements, whether used for diagnosis or for outcome evaluation, share two common factors:

1. They need to be accurately measured – generally using either an acoustic measurement or neural responses, sometimes employing a statistical detection paradigm.
2. They require stimulus presentation to the auditory system.

Sources of variability in stimulus calibration or in response, if not accounted for, will introduce error into objective response measurement, whether it be auditory evoked potentials, measurements of the real-ear aided response, or otoacoustic emissions (OAE). The main sources of interpersonal variability come from differing residual ear-canal dimensions, tympanic membrane impedance, and measurement location in the canal (Voss & Herrmann, 2005). These sources of variability, if unaccounted for, have the potential to result in misdiagnosis and mismanagement of an infant's hearing loss throughout the EHDI process (Figure 1-1).

The overarching aim of this thesis is to develop new methods to improve objectivity in EDHI. More specifically, the aim is to investigate and improve key areas linked to objective response detection and/or effects of ear-canal acoustics on stimulus presentation levels. Improvements in these two areas can impact all processes in EHDI due to the importance of the underlying procedure, objective outcome measurement. Objective outcome measurement combines the assessment of ear-canal acoustic alterations and response detection, making it a key sub-procedure of EHDI as well as the main focus of this thesis. This introductory chapter briefly reviews objective measurements for hearing aid outcome validation and current approaches used to account for ear canal acoustics throughout the EHDI process.

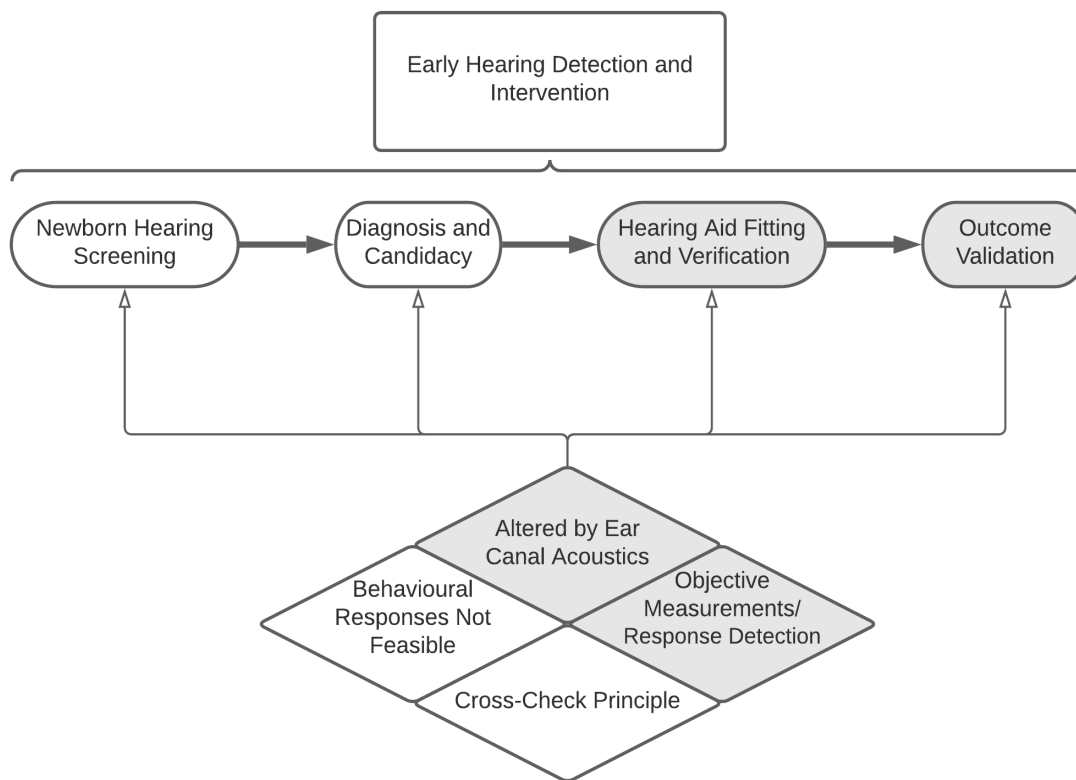


Figure 1-1: Framework for identifying sources of error in procedures used in Early Hearing Detection and Intervention. Areas of focus are shaded in grey.

1.1 Outcome validation

Following the fitting of amplification devices, outcomes need to be evaluated to determine the benefit of the intervention (American Academy of Audiology, 2013). Typically, outcome evaluation is completed using a combination of subjective or objective outcome measurements, as both have important advantages and disadvantages (Easwar, 2014). Subjective outcomes are generally not assessed directly from the child; they are often based on observations from the child's caregiver. In contrast, objective outcome measurements, such as auditory evoked potentials (AEP), do not require active patient or caregiver participation.

Auditory evoked potentials (AEP) minimize the bias in the responses, require less time between the fitting of amplification and outcome evaluation, and provide greater

ecological validity compared to subjective outcome measures (Easwar, 2014). Given these advantages, there has been increased interest in using AEP to assess pediatric hearing aid outcomes and speech audibility (Anderson & Kraus, 2013; Easwar, Beamish, et al., 2015; Easwar, Birstler, et al., 2020; Glista et al., 2012; Golding et al., 2007; Jenkins et al., 2018; Purdy et al., 2010).

One AEP of particular interest is the envelope-following response (EFR), a neural response with a dominant brainstem generator (Bidelman, 2018; Herdman & Stapells, 2001) that reflects phase-locked neural activity to the envelope of the stimulus (Aiken & Picton, 2008). The EFR is a promising technique for verifying speech audibility in infants because, unlike many other auditory evoked potentials, speech stimuli can be used for elicitation (referred to as the speech-evoked EFR; (Aiken & Picton, 2006; Choi et al., 2013) and the patient can sleep during testing (Jeng et al., 2011). Detection of a speech-evoked EFR indicates that the child's brainstem is being activated by the sound-level presented to their auditory system, with smaller cortical contributions present (Bidelman, 2018; Easwar et al., 2021). Additionally, using speech stimuli more accurately represents real-world hearing-aid amplification, especially when non-linear hearing aids are used (Easwar et al., 2012; Henning & Bentler, 2005; Scollie & Seewald, 2002).

The speech-evoked EFR can be detected rapidly (Easwar, Purcell, et al., 2015) and has excellent test-retest reliability (Easwar, Scollie, et al., 2020). Given the small amplitude of the response, the EFR requires sufficient averaging over time and is not always detected, even when an individual reports hearing a sound (Easwar, Purcell, et al., 2015; Easwar, Scollie, et al., 2020). Extensive research on EFR detection accuracy shows a dependence on a variety of testing parameters: sound-level, statistical indicator used, frequency of the elicitor, and subject noisiness (Aiken & Picton, 2006; Easwar, Birstler, et al., 2020; Vanheusden et al., 2019).

1.2 Accounting for individual ear canal acoustics

Traditionally, the amplitude of a stimulus is calibrated using the sound-pressure level (SPL) response as measured inside a reference hard-walled coupler or simulated human ear (Scollie et al., 1998; Valente et al., 1994). The amplitude of the sound-level, however,

depends on the characteristics of the acoustic cavity itself: the larger the cavity, the lower the stimulus level. Human ear canals are not all the same – they vary by shape, size, and impedance. Consequently, the same stimulus can produce ear canal SPL measurements that vary as much as 40 dB between individuals (Munro & Hatton, 2000; Munro & Salisbury, 2002; Saunders & Morgan, 2003). Unfortunately, the differences in sound-level vary in a manner that is difficult to predict across frequencies and is dependent on the transducer used (Bagatto et al., 2005; Munro & Hatton, 2000; Munro & Salisbury, 2002) and the individual's ear canal geometry (Stinson & Lawton, 1989). To account for this variability, acoustic measurements of the individual's ear canal need to be made. Differing approaches to measuring the ear canal sound-level have been proposed and implemented during the hearing aid fitting/verification process and during diagnostic procedures. These approaches are discussed below.

1.3 Measuring individual ear canal SPL during hearing aid fitting

The aided sound level presented to the individual's auditory system needs to be verified to ensure that it is adequate and appropriate (American Academy of Audiology, 2013; Moodie, Pietrobon, et al., 2016). Given the large effect of ear canal acoustics on the presented sound-level, the individual's ear canal SPL (ecSPL) should be measured to accurately fit the hearing aid to prescriptive targets, also calibrated in ecSPL. Due to hearing aid non-linearity, the ecSPL should be measured for multiple input levels to ensure both adequate access to speech and that the maximum hearing aid output does not exceed prescribed levels (American Academy of Audiology, 2013). The individualized ecSPL can be determined using one of two approaches: (1) directly measuring the sound-level that the hearing aid produces in the ear canal, known as the real-ear aided response (REAR) or (2) quantifying the individual's acoustic-transform and comparing it to that of a reference coupler, known as real-ear-to-coupler difference (RECD; American Academy of Audiology, 2013).

To measure the individual's REAR, a flexible probe microphone is placed in the ear canal close (<4 mm) to the eardrum by a highly trained professional (Vaisberg et al., 2016). The REAR is measured in real-time as a stimulus, typically speech, is played and

processed by the hearing aid. The individual's REAR is then directly compared to the ecSPL-referenced prescriptive hearing aid targets. This approach to accounting for the individual ear canal acoustics is the gold-standard for most adult hearing aid fittings and is recommended for vented hearing aid fittings and for children with long (>35 mm) earmold tubing (Gustafson et al., 2013; Moodie, Pietrobon, et al., 2016). In pediatric fittings, however, obtaining direct, accurate REAR measurements can be difficult because they require sustained cooperation from the patient and are negatively impacted by head movement. These requirements introduce practicality concerns for working in a pediatric population, where these conditions may not be reliably met, requiring an alternative approach for measuring the individual's REAR.

The RECD can be used as a level-independent acoustic transform used to accurately convert the SPL produced in a reference coupler to the individual's unique ecSPL (Bagatto et al., 2005; Moodie et al., 1994). The RECD is obtained by taking the difference between the SPL produced in a reference coupler and the SPL produced in the individual's ear canal (Bagatto et al., 2005). The resultant RECD can then convert SPL produced in the reference coupler to the estimated ecSPL produced in the individual. The real-ear portion of this measurement, like REAR measurement, requires the placement of a probe-tube microphone near the eardrum (Vaisberg et al., 2018). After the tube is in place, the RECD requires only seconds of cooperation and attention from the patient, making it ideal for the pediatric population (Scollie et al., 1998). After the RECD is measured, the rest of the hearing aid fitting can be done in the reference coupler because the hearing aid output can be accurately fit to estimated ecSPL prescriptive target data. During the hearing aid fitting, the RECD is used twice: first to convert the ecSPL hearing thresholds to equivalent adult dB HL thresholds for prescriptive target calculation and then to convert reference coupler measurements to equivalent ecSPL measures (Bagatto et al., 2005; Gustafson et al., 2013; Scollie et al., 1998).

The RECD should be measured with the same transducer used for hearing threshold determination (Gustafson et al., 2013) because RECD values are dependent on transducer characteristics (Munro & Hatton, 2000; Munro & Salisbury, 2002). The transducer dependence is due to differences in the tubing length associated with the transducer,

which has an impact on the transducer's acoustic impedance (Munro & Salisbury, 2002). As a result, an RECD measured using a foam-tip insert is going to differ significantly, as much as 7-8dB at 4000 Hz, from an RECD of the same individual measured using their hearing aid tubing and ear-mold (Gustafson et al., 2013; Moodie, Pietrobon, et al., 2016). The systematic error introduced by this transducer mismatch can be avoided by using the same transducer. In cases where the same transducer cannot be used for RECD and threshold determination, average foam-tip-to-earmold transforms can be used (Scollie et al., 1998; Valente et al., 1994). Even if the transducer mismatch is accounted for, however, there are other sources of errors and clinical issues with the measurement of RECD that affect all probe-tube ecSPL measurements.

1.4 Individual ecSPL during hearing aid fitting: errors and clinical issues

Measurements of ecSPL, completed with a probe tube for in situ measurement or to determine the individual's RECD, result in more accurate determinations of auditory input than coupler-calibrated stimuli (American Academy of Audiology, 2013; The Joint Committee on Infant Hearing, 2019). However, even when these measurements are implemented in the hearing aid verification process, they can introduce significant errors into the calibration procedure. Probe-tube ecSPL measurements are contaminated by standing wave errors as large as ± 20 dB in the audiometric range (Dreisbach & Siegel, 2001; Siegel, 1994; Siegel & Hirohata, 1994; Stinson et al., 1982; Stinson & Lawton, 1989). The standing wave errors are caused by the destructive interference between the stimulus and the sound-wave reflected by the eardrum (Chan & Geisler, 1990; Gilman & Dirks, 1986; Stinson et al., 1982; Stinson & Lawton, 1989). The reflected sound wave interference results in an erroneous sound-level reading at the location of the probe-tube that is not present at the individual's eardrum. The largest effect of the reflected sound wave interference is found near the frequency of the $\frac{3}{4}$ -wavelength pressure node. However, the amplitude, width, and exact frequency location of the reflected wave interference are difficult to predict as they are dependent on transducer depth, probe-tube depth, ear canal geometry, and many other factors (Feigin et al., 1989, 1990; Stinson & Lawton, 1989; Vaisberg et al., 2016). Furthermore, as the frequency of the stimulus

increases, the wavelength decreases, resulting in more complex interactions between the forward-moving and reflected sound waves. As a result, probe-tube microphone measurements of ecSPL are increasingly unreliable and inaccurate as stimulus frequency increases (McCreery et al., 2009; Vaisberg et al., 2016). Probe-tube ecSPL measurements also require precise probe-tube placement by a highly trained individual, which is not possible in every audiological setting (i.e. teleaudiology, screening procedures).

Even with these limitations, measuring the ecSPL using a probe-tube improves the accuracy of hearing aid measurement over methods that do not account for individual ear canal acoustics. Probe-tube measures are recommended practice for hearing aid fitting for both adults and children (American Academy of Audiology, 2013). However, 25% of North American pediatric audiologists self-report that they never or seldom measure the RECD during hearing aid verification and almost two-thirds rarely or never use the RECD during the diagnostic stage (Moodie, Rall, et al., 2016). The numbers are significantly worse when looking at hearing aid validation in adults, with less than half of all audiologists self-reporting that they use probe-tube microphone measures on the day of the fitting (Mueller & Picou, 2010). The most commonly reported reasons for not completing RECD measures in clinic are due to a lack of clinical time and self-doubts about the ability to make the RECD measurement (Moodie, Pietrobon, et al., 2016). These results indicate a clear need for improved access to training procedures (Koch et al., 2018, 2020) and/or new validation and verification procedures to decrease the amount of clinical time and technical skill needed to complete such measurements.

1.5 Accounting for individual ear canal acoustics: alternatives to probe-tube ecSPL calibration

Recently, there has been interest in alternative methods of ear canal sound-level measurement that do not require precise placement of a probe-tube microphone. These calibration approaches require a specially calibrated transducer that houses both a sound-source and microphone (Lewis et al., 2009; Scheperle et al., 2008; Souza et al., 2014). These alternative approaches utilize an analysis of the eardrum reflectance to negate the impact of the reflected sound-wave interference. The distal sound-level measurement is completed using a microphone that is flush with the sound-source, allowing the

transducer/ear canal coupling to be represented as a Thevenin electrical circuit (Allen, 1986; Keefe et al., 1993; Scheperle et al., 2008). By accounting for the transducer's source pressure and impedance using a special Thevenin-equivalent source parameter calibration procedure (reviewed in Chapter 3), the sound-level in the ear canal can be decomposed into the forward-moving pressure level (FPL) and the reflected sound-level (RPL; Souza et al., 2014). The total ear canal sound level can be represented in dB FPL or as the integrated pressure level (IPL), which is equivalent to the eardrum SPL without any reflected wave interferences (Lewis et al., 2009; Scheperle et al., 2011; Souza et al., 2014; Withnell et al., 2014). Ear canal sound level measurements calibrated in FPL or IPL require only seconds of patient cooperation, do not require probe-tube placement, and simultaneously measure the individual's wideband acoustic immittance and ear canal sound levels (WAI; Souza et al., 2014).

Early work on the development of FPL has focused on potential benefits in diagnostic settings – mainly for the purpose of in-ear calibration of stimuli used for otoacoustic emission (OAE) testing and audiometry. FPL-based calibration is beneficial for OAE testing because, along with the improved accuracy over traditional SPL approaches, the FPL-referenced stimulus retains its phase information which is useful for some applications (Allen, 1986; Keefe et al., 1993; Scheperle et al., 2008). Furthermore, OAE probes are already constructed with the microphone in the same sound-source, meaning that such calibrations can be completed with currently available OAE transducers.

Early testing with OAE stimuli demonstrated mixed results regarding the benefit of FPL-calibration on testing accuracy. Some studies found significant improvements (Kirby et al., 2011; Scheperle et al., 2008, 2011) while some found no discernible benefit over traditional SPL approaches (Burke et al., 2010; Rogers et al., 2010). These mixed results may be due to testing at discrete frequencies, which minimizes the ability to measure the effects of the standing wave errors, and the frequency of the standing wave may fall between discrete frequencies (Scheperle et al., 2011).

Work on behavioural threshold testing with FPL and IPL calibrated stimuli, however, has consistently shown improved accuracy (Lapsley Miller et al., 2018; Lewis et al., 2009;

McCreery et al., 2009; Withnell et al., 2009, 2014). Further accuracy improvements have been made using individualized in-situ acoustic estimations of ear canal cross-sectional area, which can be completed during the FPL/IPL measurement (Rasetshwane & Neely, 2011). By individualizing the cross-sectional area estimation for use in FPL and IPL calculation using an in-situ time-domain reflectance measurement, a more accurate measurement of ear canal sound level (on average ~ 1 dB) can be obtained (Souza et al., 2014).

While the calibration of stimuli referenced in FPL or IPL may be advantageous for diagnostic purposes, it is difficult to use FPL-referenced hearing thresholds in a hearing aid fitting, where prescriptive targets are referenced to ecSPL. FPL-to-ecSPL transforms are complex and vary between individuals, especially in the mid- and high frequencies, where the measurements differ on average 12.2 dB at the individual's standing wave frequency (mean = 5909 Hz, SD = 1943 Hz; McCreery et al., 2009). For use in current hearing aid intervention, FPL-referenced normative values or an easy to measure/interpret FPL-to-ecSPL transform would be necessary.

Unlike the FPL, the integrated pressure level (IPL) is theoretically equivalent to the SPL present at the termination of the ear canal without reflected-wave interactions (Lewis et al., 2009; Souza et al., 2014). IPL calculation, however, requires the determination of FPL and shows similar advantages when compared to ecSPL measurement (Souza et al., 2014). IPL has been shown to be relatively independent of transducer depth, with ± 3 dB accuracy in determining behavioural thresholds up to 16 kHz in adults (McCreery et al., 2009; Souza et al., 2014). Given the theoretical equivalence between IPL and SPL at the eardrum, direct comparison to ecSPL measurements or prescriptive targets used with conventional probe-tube microphone verification systems is possible. Being able to quickly and efficiently measure the IPL-referenced RECD (IPL RECD), for example, would allow for consistent calibration across the pediatric hearing workflow, with possible applications at both the hearing assessment (Scollie et al., 1998) and coupler-based hearing aid verification (Moodie et al., 1994) stages, as well as in the calculation of prescriptive targets that apply transforms for each of these stages (Bagatto et al., 2005). Measurement of the IPL RECD is expected to improve accuracy of individual ear canal

acoustic transforms and allow for such corrections to be completed in any audiological setting, even if a highly trained practitioner is not present.

1.6 Purpose of this thesis

The purpose of this thesis is to improve the accuracy of objective measurements throughout the pediatric hearing aid verification and validation procedures. To meet such aims, Chapter 2 is based on improving the clinical efficacy of speech-evoked EFR as an objective outcome measurement. In a clinical setting, objective outcome validation will be restricted to the duration of time that the infant is able to stay asleep. Previous studies have analyzed statistical indicator efficacy in speech-evoked EFR, however, they have not completed such an analysis on an infant population. Furthermore, many previous comparisons of statistical indicators have employed fixed testing durations (e.g. 30 minutes regardless of response detection or absence), which limits real-world generalization where test-durations may vary. In Chapter 2, different statistical indicators used to detect auditory evoked potentials are assessed to determine the speed and accuracy of response detection in infants and adults. By evaluating the speed and accuracy of detection, it is hoped that the incorporation of such an objective outcome validation may be incorporated in clinical settings.

The following three chapters (Chapter 3-5) focus on the development and subsequent proposal of an IPL-calibrated, probe-tube free RECD measurement that occurs concurrently with an analysis of the middle ear. The rationale for these final chapters was to (1) decrease potential clinical barriers to the introduction of IPL RECD in clinic, (2) improve RECD measurement accuracy for foam-tip and ear-mold RECD determination, and (3) to decrease the time and difficulty to accurately measure the child's RECD.

1.7 Research questions

The research questions, stated in the order of appearance throughout chapters 2-5, are:

- a. Is the accuracy of speech-evoked EFR detection dependent on the statistical indicator used? Can the sensitivity of speech-evoked EFR detection be calculated given a variable test duration?

- b. What is the impact of intra- and inter-session reliability of Thevenin-equivalent source parameter calibration on stimulus calibration and wideband reflectance accuracy?
- c. Can valid and reliable measurements of individual hearing aid tubing length be determined acoustically using a source parameter calibrated transducer?
- d. Is IPL RECD measurement more reliable and accurate than probe-tube microphone measured RECD? Can IPL RECD measurement be a direct replacement to probe-tube microphone measured foam-tip and/or earmold RECD for use in hearing aid workflows for ecSPL-derived hearing aid targets?

1.8 References

- Aiken, S. J., & Picton, T. W. (2006). Envelope Following Responses to Natural Vowels. *Audiology and Neurotology*, *11*(4), 213–232. <https://doi.org/10.1159/000092589>
- Aiken, S. J., & Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hearing Research*, *245*(1–2), 35–47. <https://doi.org/10.1016/j.heares.2008.08.004>
- Allen, J. B. (1986). Measurement of Eardrum Acoustic Impedance. In J. B. Allen, J. L. Hall, A. E. Hubbard, S. T. Neely, & A. Tubis (Eds.), *Peripheral Auditory Mechanisms* (Vol. 64, pp. 44–51). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-50038-1_6
- American Academy of Audiology. (2013). *American Academy of Audiology Clinical Practice Guidelines Pediatric Amplification*.
- Anderson, S., & Kraus, N. (2013). The Potential Role of the cABR in Assessment and Management of Hearing Impairment. *International Journal of Otolaryngology*, *2013*, 1–10. <https://doi.org/10.1155/2013/604729>
- Bagatto, Moodie, S., Scollie, S., Seewald, R., Moodie, S., Pumford, J., & Liu, K. P. R. (2005). Clinical Protocols for Hearing Instrument Fitting in the Desired Sensation Level Method. *Trends in Amplification*, *9*(4), 199–226. <https://doi.org/10.1177/108471380500900404>
- Bidelman, G. M. (2018). Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. *NeuroImage*, *175*, 56–69. <https://doi.org/10.1016/j.neuroimage.2018.03.060>
- Burke, S. R., Rogers, A. R., Neely, S. T., Kopun, J. G., Tan, H., & Gorga, M. P. (2010). Influence of Calibration Method on Distortion-Product Otoacoustic Emission Measurements: I. Test Performance. *Ear and Hearing*, *31*(4), 533–545. <https://doi.org/10.1097/AUD.0b013e3181d86b3d>
- Chan, J. C. K., & Geisler, C. D. (1990). Estimation of eardrum acoustic pressure and of ear canal length from remote points in the canal. *The Journal of the Acoustical Society of America*, *87*(3), 1237–1247. <https://doi.org/10.1121/1.398799>
- Choi, J. M., Purcell, D. W., Coyne, J.-A. M., & Aiken, S. J. (2013). Envelope Following Responses Elicited by English Sentences. *Ear and Hearing*, *34*(5), 637–650. <https://doi.org/10.1097/AUD.0b013e31828e4dad>
- Dreisbach, L. E., & Siegel, J. H. (2001). Distortion-product otoacoustic emissions measured at high frequencies in humans. *The Journal of the Acoustical Society of America*, *110*(5), 2456–2469. <https://doi.org/10.1121/1.1406497>
- Easwar, V. (2014). *Evaluation of Auditory Evoked Potentials as a Hearing aid Outcome Measure*. University of Western Ontario.
- Easwar, V., Beamish, L., Aiken, S., Choi, J. M., Scollie, S., & Purcell, D. (2015). Sensitivity of envelope following responses to vowel polarity. *Hearing Research*, *320*, 38–50. <https://doi.org/10.1016/j.heares.2014.11.008>

- Easwar, V., Birstler, J., Harrison, A., Scollie, S., & Purcell, D. (2020). The accuracy of envelope following responses in predicting speech audibility. *Ear & Hearing, Publish Ahead of Print*. <https://doi.org/10.1097/AUD.0000000000000892>
- Easwar, V., Scollie, S., Lasarev, Urichuk, M., Aiken, S., Purcell, D. (2021). Characteristic of speech-evoked envelope following responses in infancy. *Trends in Hearing, 25, 1-17*. DOI: 10.1177/23312165211004331..
- Easwar, V., Purcell, D., Aiken, S. J., Parsa, V., & Scollie, S. (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 and Hearing, 36*(6), 18.
- Easwar, V., Purcell, D., & Scollie, S. (2012). Electroacoustic Comparison of Hearing Aid Output of Phonemes in Running Speech versus Isolation: Implications for Aided Cortical Auditory Evoked Potentials Testing. *International Journal of Otolaryngology, 2012*, 1–10. <https://doi.org/10.1155/2012/518202>
- Easwar, V., Scollie, S., Aiken, S., & Purcell, D. (2020). Test-Retest Variability in the Characteristics of Envelope Following Responses Evoked by Speech Stimuli: *Ear and Hearing, 41*(1), 150–164. <https://doi.org/10.1097/AUD.0000000000000739>
- Feigin, J. A., Barlow, N. L. N., & Stelmachowicz, P. G. (1990). The Effect of Reference Microphone Placement on Sound Pressure Levels at an Ear Level Hearing Aid Microphone: *Ear and Hearing, 11*(5), 321–326. <https://doi.org/10.1097/00003446-199010000-00001>
- Feigin, J. A., Kopun, J., Stelmachowicz, P., & Gorga, M. (1989). Probe-tube microphone measures of ear-canal sound pressure levels in infants and children. *Ear and Hearing, 10*(4), 254–258.
- Gilman, S., & Dirks, D. D. (1986). Acoustics of ear canal measurement of eardrum SPL in simulators. *The Journal of the Acoustical Society of America, 80*(3), 783–793. <https://doi.org/10.1121/1.393953>
- Glista, D., Easwar, V., Purcell, D. W., & Scollie, S. (2012). A Pilot Study on Cortical Auditory Evoked Potentials in Children: Aided CAEPs Reflect Improved High-Frequency Audibility with Frequency Compression Hearing Aid Technology. *International Journal of Otolaryngology, 2012*, 1–12. <https://doi.org/10.1155/2012/982894>
- Golding, M., Pearce, W., Seymour, J., Cooper, A., Ching, T., & Dillon, H. (2007). The Relationship between Obligatory Cortical Auditory Evoked Potentials (CAEPs) and Functional Measures in Young Infants. *Journal of the American Academy of Audiology, 18*(2), 9.
- Gustafson, S., Pittman, A., & Fanning, R. (2013). Effects of Tubing Length and Coupling Method on Hearing Threshold and Real-Ear to Coupler Difference Measures. *American Journal of Audiology, 22*(1), 190–199. [https://doi.org/10.1044/1059-0889\(2012/12-0046\)](https://doi.org/10.1044/1059-0889(2012/12-0046))

- Hall, J. W. (2016). Crosscheck Principle in Pediatric Audiology Today: A 40-Year Perspective. *Journal of Audiology and Otology*, 20(2), 59–67. <https://doi.org/10.7874/jao.2016.20.2.59>
- Henning, R. W., & Bentler, R. (2005). Compression-Dependent Differences in Hearing Aid Gain Between Speech and Nonspeech Input Signals: *Ear and Hearing*, 26(4), 409–422. <https://doi.org/10.1097/00003446-200508000-00004>
- Herdman, A. T., & Stapells, D. R. (2001). Thresholds determined using the monotic and dichotic multiple auditory steady-state response technique in normal-hearing subjects. *Scandinavian Audiology*, 30(1), 41–49. <https://doi.org/10.1080/010503901750069563>
- Jeng, F.-C., Hu, J., Dickman, B., Montgomery-Reagan, K., Tong, M., Wu, G., & Lin, C.-D. (2011). Cross-Linguistic Comparison of Frequency-Following Responses to Voice Pitch in American and Chinese Neonates and Adults: *Ear and Hearing*, 32(6), 699–707. <https://doi.org/10.1097/AUD.0b013e31821cc0df>
- Jenkins, K. A., Fodor, C., Presacco, A., & Anderson, S. (2018). Effects of Amplification on Neural Phase Locking, Amplitude, and Latency to a Speech Syllable: *Ear and Hearing*, 39(4), 810–824. <https://doi.org/10.1097/AUD.0000000000000538>
- Jerger, J., & Hayes, D. (1976). The Cross-Check Principle in Pediatric Audiometry. *Archives of Otolaryngology - Head and Neck Surgery*, 102, 614–620.
- Keefe, D. H., Bulen, J. C., Arehart, K. H., & Burns, E. M. (1993). Ear-canal impedance and reflection coefficient in human infants and adults. *The Journal of the Acoustical Society of America*, 94(5), 2617–2638. <https://doi.org/10.1121/1.407347>
- Kirby, B. J., Kopun, J. G., Tan, H., Neely, S. T., & Gorga, M. P. (2011). Do “Optimal” Conditions Improve Distortion Product Otoacoustic Emission Test Performance?: *Ear and Hearing*, 32(2), 230–237. <https://doi.org/10.1097/AUD.0b013e3181fa5da2>
- Koch, R. W., Moodie, S., Folkeard, P., Scollie, S., Janeteas, C., Agrawal, S. K., & Ladak, H. M. (2018). Face and Content Validity of a Probe Tube Placement Training Simulator. *Journal of the American Academy of Audiology*. <https://doi.org/10.3766/jaaa.17114>
- Koch, R. W., Saleh, H., Folkeard, P., Moodie, S., Janeteas, C., Agrawal, S. K., Ladak, H. M., & Scollie, S. (2020). Skill Transference of a Probe-Tube Placement Training Simulator. *Journal of the American Academy of Audiology*, 31(01), 040–049. <https://doi.org/10.3766/jaaa.18054>
- Lapsley Miller, J. A., Reed, C. M., Robinson, S. R., & Perez, Z. D. (2018). Pure-Tone Audiometry With Forward Pressure Level Calibration Leads to Clinically-Relevant Improvements in Test–Retest Reliability: *Ear and Hearing*, 39(5), 946–957. <https://doi.org/10.1097/AUD.0000000000000555>
- Lewis, J. D., McCreery, R. W., Neely, S. T., & Stelmachowicz, P. G. (2009). Comparison of *in-situ* calibration methods for quantifying input to the middle ear. *The Journal*

of the Acoustical Society of America, 126(6), 3114–3124.
<https://doi.org/10.1121/1.3243310>

- McCreery, R. W., Pittman, A., Lewis, J., Neely, S. T., & Stelmachowicz, P. G. (2009). Use of forward pressure level to minimize the influence of acoustic standing waves during probe-microphone hearing-aid verification. *The Journal of the Acoustical Society of America*, 126(1), 15–24. <https://doi.org/10.1121/1.3143142>
- McCreery, R. W., Walker, E. A., Spratford, M., Oleson, J., Bentler, R., Holte, L., & Roush, P. (2015). Speech Recognition and Parent Ratings From Auditory Development Questionnaires in Children Who Are Hard of Hearing: *Ear and Hearing*, 36, 60S-75S. <https://doi.org/10.1097/AUD.0000000000000213>
- Moeller, M. P., & Tomblin, J. B. (2015). Epilogue: Conclusions and Implications for Research and Practice. *Ear and Hearing*, 36, 92S-98S. <https://doi.org/10.1097/AUD.0000000000000214>
- Moodie, Shane, Seewald, R. C., & Sinclair, S. (1994). Procedure for Predicting Real-Ear Hearing Aid Performance in Young Children. *American Journal of Audiology*, 3(1), 23–31. <https://doi.org/10.1044/1059-0889.0301.23>
- Moodie, Sheila, Pietrobon, J., Rall, E., Lindley, G., Eiten, L., Gordey, D., Davidson, L., Moodie, K. S., Bagatto, M., Haluschak, M. M., 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. *Journal of the American Academy of Audiology*, 27(03), 264–275. <https://doi.org/10.3766/jaaa.15086>
- Moodie, Sheila, Rall, E., Eiten, L., Lindley, G., Gordey, D., Davidson, L., Bagatto, M., & Scollie, S. (2016). Pediatric Audiology in North America: Current Clinical Practice and How It Relates to the American Academy of Audiology Pediatric Amplification Guideline. *Journal of the American Academy of Audiology*, 27(03), 166–187. <https://doi.org/10.3766/jaaa.15064>
- Mueller, H. G., & Picou, E. M. (2010). Survey examines popularity of real-ear probe-tube microphone measures: *The Hearing Journal*, 63(5), 27–28. <https://doi.org/10.1097/01.HJ.0000373447.52956.25>
- Munro, K. J., & Hatton, N. (2000). Customized Acoustic Transform Functions and Their Accuracy at Predicting Real-Ear Hearing Aid Performance. *Ear and Hearing*, 21(1), 59–69.
- Munro, K. J., & Salisbury, V. (2002). Is the real-ear to coupler difference independent of the measurement earphone. *International Journal of Audiology*, 41(7), 408–413.
- Norrix, L. W. (2015). Hearing Thresholds, Minimum Response Levels, and Cross-Check Measures in Pediatric Audiology. *American Journal of Audiology*, 24(2), 137–144. https://doi.org/10.1044/2015_AJA-14-0095
- Purdy, S. C., Katsch, R., Dillon, H., Storey, L., Sharma, M., & Agung, K. (2010). Aided orlital Auditory Evoked Potentials for Hearing Instrument Evaluation in Infants. In *A Sound Foundation Through Early Amplification*.

- Rasetshwane, D. M., & Neely, S. T. (2011). Inverse solution of ear-canal area function from reflectance. *The Journal of the Acoustical Society of America*, *130*(6), 3873–3881. <https://doi.org/10.1121/1.3654019>
- Rogers, A. R., Burke, S. R., Kopun, J. G., Tan, H., Neely, S. T., & Gorga, M. P. (2010). Influence of Calibration Method on Distortion-Product Otoacoustic Emission Measurements: II. Threshold Prediction: *Ear and Hearing*, *31*(4), 546–554. <https://doi.org/10.1097/AUD.0b013e3181d86b59>
- Saunders, G. H., & Morgan, D. E. (2003). Impact on hearing aid targets of measuring thresholds in dB HL versus dB SPL: El impacto en la medición de los umbrales en dB HL o en dB SPL, en las metas de un auxiliar auditivo. *International Journal of Audiology*, *42*(6), 319–326.
- Scheperle, R. A., Goodman, S. S., & Neely, S. T. (2011). Further assessment of forward pressure level for *in situ* calibration. *The Journal of the Acoustical Society of America*, *130*(6), 3882–3892. <https://doi.org/10.1121/1.3655878>
- Scheperle, R. A., Neely, S. T., Kopun, J. G., & Gorga, M. P. (2008). Influence of *in situ*, sound-level calibration on distortion-product otoacoustic emission variability. *The Journal of the Acoustical Society of America*, *124*(1), 288–300. <https://doi.org/10.1121/1.2931953>
- Scollie, S., & Seewald, R. C. (2002). Evaluation of Electroacoustic Test Signals I: Comparison with Amplified Speech: *Ear and Hearing*, *23*(5), 477–487. <https://doi.org/10.1097/00003446-200210000-00009>
- Scollie, S., Seewald, Richard, Cornelisse, Leonard, & Jenstad, Lorianne. (1998). Validity and Repeatability of Level-Independent HL to SPL Transforms. *Ear and Hearing*, *19*, 407–413.
- Siegel, J. H. (1994). Ear-canal standing waves and high-frequency sound calibration using otoacoustic emission probes. *The Journal of the Acoustical Society of America*, *95*(5), 2589–2597. <https://doi.org/10.1121/1.409829>
- Siegel, J. H., & Hirohata, E. T. (1994). Sound calibration and distortion product otoacoustic emissions at high frequencies. *Hearing Research*, *80*(2), 146–152. [https://doi.org/10.1016/0378-5955\(94\)90106-6](https://doi.org/10.1016/0378-5955(94)90106-6)
- Sininger, Y. S., Martinez, A., Eisenberg, L., Christensen, E., Grimes, A., & Hu, J. (2009). Newborn Hearing Screening Speeds Diagnosis and Access to Intervention by 20–25 Months. *Journal of the American Academy of Audiology*, *20*(1), 49–57. <https://doi.org/10.3766/jaaa.20.1.5>
- Souza, N. N., Dhar, S., Neely, S. T., & Siegel, J. H. (2014). Comparison of nine methods to estimate ear-canal stimulus levels. *The Journal of the Acoustical Society of America*, *136*(4), 1768–1787. <https://doi.org/10.1121/1.4894787>
- Stinson, M., & Lawton, B. W. (1989). Specification of the geometry of the human ear canal for the prediction of sound-pressure level distribution. *The Journal of the Acoustical Society of America*, *85*(6), 2492–2503. <https://doi.org/10.1121/1.397744>

- Stinson, M., Shaw, E., & Lawton, B. W. (1982). Estimation of acoustical energy reflectance at the eardrum from measurements of pressure distribution in the human ear canal. *The Journal of the Acoustical Society of America*, 72(3), 766–773. <https://doi.org/10.1121/1.388257>
- The Joint Committee on Infant Hearing. (2019). Year 2019 Position Statement: Principles and Guidelines for Early Hearing Detection and Intervention Programs. *The Journal of Early Hearing Detection and Intervention*, 4(2), 1–44.
- Tomblin, J. B., Harrison, M., Ambrose, S. E., Walker, E. A., Oleson, J. J., & Moeller, M. P. (2015). Language Outcomes in Young Children with Mild to Severe Hearing Loss: *Ear and Hearing*, 36, 76S-91S. <https://doi.org/10.1097/AUD.0000000000000219>
- Tomblin, J. B., Oleson, J. J., Ambrose, S. E., Walker, E., & Moeller, M. P. (2014). The Influence of Hearing Aids on the Speech and Language Development of Children With Hearing Loss. *JAMA Otolaryngology–Head & Neck Surgery*, 140(5), 403. <https://doi.org/10.1001/jamaoto.2014.267>
- Uus, K., & Bamford, J. (2006). Effectiveness of Population-Based Newborn Hearing Screening in England: Ages of Interventions and Profile of Cases. *PEDIATRICS*, 117(5), e887–e893. <https://doi.org/10.1542/peds.2005-1064>
- Vaisberg, J. M., Folkeard, P., Pumford, J., Narten, P., & Scollie, S. (2018). Evaluation of the Repeatability and Accuracy of the Wideband Real-Ear-to-Coupler Difference. *Journal of the American Academy of Audiology*, 29(06), 520–532. <https://doi.org/10.3766/jaaa.17007>
- Vaisberg, J. M., Macpherson, E. A., & Scollie, S. (2016). Extended bandwidth real-ear measurement accuracy and repeatability to 10 kHz. *International Journal of Audiology*, 55(10), 580–586. <https://doi.org/10.1080/14992027.2016.1197427>
- Valente, M., Potts, L. G., Valentet, M., Vass, W., & Goebel, J. (1994). Intersubject Variability of Real-Ear Sound Pressure Level: Conventional and Insert Earphones. *Journal of the American Academy of Audiology*, 5(6), 9.
- Vanheusden, F., Chesnaye, M., Simpson, D., & Bell, S. (2019). Envelope frequency following responses are stronger for high-pass than low-pass filtered vowels. *International Journal of Audiology*, 58, 1–8. <https://doi.org/10.1080/14992027.2018.1562243>
- Voss, S. E., & Herrmann, B. S. (2005). How Does the Sound Pressure Generated by Circumaural, Supra-aural, and Insert Earphones Differ for Adult and Infant Ears?: *Ear and Hearing*, 26(6), 636–650. <https://doi.org/10.1097/01.aud.0000189717.83661.57>
- Withnell, R. H., Jeng, P. S., Parent, P., & Levitt, H. (2014). The clinical utility of expressing hearing thresholds in terms of the forward-going sound pressure wave. *International Journal of Audiology*, 53(8), 522–530. <https://doi.org/10.3109/14992027.2014.898122>

- Withnell, R. H., Jeng, P. S., Waldvogel, K., Morgenstein, K., & Allen, J. B. (2009). An *in situ* calibration for hearing thresholds. *The Journal of the Acoustical Society of America*, *125*(3), 1605–1611. <https://doi.org/10.1121/1.3075551>
- Yoshinaga-Itano, C., Sedey, A. L., Coulter, D. K., & Mehl, A. L. (1998). Language of Early- and Later-identified Children With Hearing Loss. *Pediatrics*, *102*(5), 1161–1171. <https://doi.org/10.1542/peds.102.5.1161>

Chapter 2

2 Performance of statistical indicators in the objective detection of speech-evoked envelope following responses¹

2.1 Introduction

Early hearing detection and intervention programs are being implemented in numerous countries with the aim of completing hearing loss diagnoses and intervention at an early age, often before the 6-month mark (Sininger et al., 2009; The Joint Committee on Infant Hearing, 2019; Uus & Bamford, 2006). During this developmental period, children cannot participate in conventional behavioural testing, making it difficult to ensure adequate speech audibility throughout intervention. Recently, there has been interest in using electrophysiological responses to objectively assess neural responses to auditory input in children (Anderson & Kraus, 2013; Easwar et al., 2015; Easwar, Scollie, et al., 2020; Jenkins et al., 2018). One such electrophysiological measure is the envelope-following response (EFR), a neural response with a dominant brainstem generator for rates greater than 150 Hz (Bidelman, 2018; Herdman & Stapells, 2001) that reflects phase-locked neural activity to the envelope of the stimulus (Aiken & Picton, 2008). Unlike many other auditory evoked potentials, the EFR can be measured in response to running speech while the individual sleeps (Choi et al., 2013; Jeng et al., 2011). A speech stimulus is advantageous because it more accurately represents real-world hearing-aid amplification when compared to tonal stimuli, particularly when the hearing aids are non-linear (Henning & Bentler, 2005; Scollie & Seewald, 2002). These characteristics make speech-evoked EFRs a promising technique for objective hearing aid outcome validation of speech audibility in infants.

¹ A version of this chapter has been submitted: Urichuk, M., Easwar, V., Scollie, S., Purcell, D. Performance of statistical indicators in the objective detection of speech-evoked envelope following responses. Submitted to *Ear and Hearing*.

Although there is evidence that speech-evoked EFRs can be useful clinically, there are several factors that can influence the accuracy of the measure that must be considered. Numerous recent studies on speech-evoked EFR detection have been completed in adult populations, showing excellent test-retest reliability, a rapid detection paradigm, and high sensitivity of various statistical indicators (Aiken & Picton, 2006; Easwar, Birstler, et al., 2020; Easwar et al., 2015; Vanheusden et al., 2018). Some statistical indicators rely on the amplitude of the EFR, such as the F-test, while others may rely on the phase of the EFR, like the Rayleigh test. Additionally, there are statistical indicators that use a combination of both the amplitude and phase information such as the Hotelling T^2 , Magnitude Square Coherence (MSC), and the Rayleigh-Moore test. The choice of statistical indicator can significantly impact the accuracy of response detection, with indicators that use both amplitude and phase often exhibiting equivalent or superior performance to statistics that use either phase or amplitude alone (Easwar et al., 2020; Mijares et al., 2013; Vanheusden et al., 2019).

The age of the individual can also significantly impact speech-evoked EFR detection. Various characteristics of speech-evoked EFRs, as well as response detection, differ significantly in infant and adult populations (Easwar et al., 2021; Savio et al., 2001; Van Dyke et al., 2017). Both amplitude, and to a lesser extent phase coherence, may be lower in an infant population. It is possible that decreased detectability is due to spectral magnitude differences rather than decreased phase locking in the infant population (Easwar et al., 2021; Lins et al., 1996; Van Dyke et al., 2017). If this is the case, it is expected that statistical indicators that heavily rely on phase information will perform more accurately for speech-evoked EFR detection in infants than amplitude-based indicators. The current study assesses the performance of various statistical indicators that use phase coherence, amplitude, or a combination of both for speech-evoked EFR detection in an infant population as a function of testing duration.

Easwar and colleagues (2020) compared the sensitivity of three common statistical indicators in an adult population. The researchers found that after approximately 30 minutes of testing, the statistical indicators that incorporated phase-information explicitly had comparable performance. Using an identical stimulus, the current work builds on

these results in two important ways. First, sensitivity analyses are extended to an infant population and two additional statistical indicators are assessed. Secondly, the sensitivity of each statistical indicator is assessed continuously throughout testing. This will account for the variable testing durations that often occur in a clinical setting and will enable the comparison of both sensitivity and efficiency of the statistical indicators between infants and adults. We hypothesize that the infants will require a longer test duration when compared to adults and will have comparatively low detection sensitivities. Age-based effects are expected to be largest when using amplitude incorporating statistical indicators.

2.2 Methods

2.2.1 Participants

23 adults (21-27 years old, 20 females) and 21 infants (0.16-1.15 years; 11 females) were tested at Western University. All participants had no reported hearing loss, ear disease, or concerns about hearing status. A case-history and pure-tone audiometry screening procedure was completed on all adult participants to confirm normal hearing (< 20 dB HL) at octave and half-octave frequencies between 250 Hz-8000 Hz. All adult participants passed tympanometric testing to confirm normal middle-ear function. Infants underwent otoacoustic emissions screening and had previously passed a newborn hearing screening. All participants were monetarily compensated \$10/hr for their time. Data reported in this study are a sub-set of the data reported in a study by Easwar and colleagues (2021).

2.2.2 Stimulus

A modified male-spoken speech token /susa^hi/ was used to elicit EFRs with eight carriers of low-, mid-, and high-frequency (Easwar et al., 2015; Easwar, Scollie, et al., 2020). The EFR carriers stimulate a frequency range that is functional for assessing objective hearing aid outcomes, spanning a frequency range similar to that of the commonly used Speech Intelligibility Index and Ling-6 sound tests (Easwar et al, 2020). The fricatives (/ʃ/, /s/) were both 100% amplitude modulated at 93.02 Hz and high-pass filtered at 3 kHz and 4 kHz, respectively. Both fricatives were maintained at their original root-mean-square

(RMS) level relative to the vowels and to each other. The vowels in the /susaʃi/ stimulus were altered to create two different fundamental frequencies in the same vowel. The natural fundamental frequency (f_0) was unmanipulated for frequency regions above the first formant (F2+; mid-frequency). The fundamental frequency for harmonics within and below the first formant (F1; low-frequency) were lowered by ~ 8.5 Hz relative to the natural f_0 of each vowel. The alteration of the fundamental frequency allowed for increased frequency specificity in responses while using the broadband /susaʃi/ signal (Easwar et al., 2020). A sound level meter (Bruel and Kjaer [B & K] Type 2250) in an ear simulator (B&K Type 4157) was used to calibrate the stimulus at an overall level of 65 dB SPL. The stimulus was presented monaurally with no further individual ear-canal level corrections in the adult population. Each epoch was 1.003125 seconds in duration with each sweep consisting of 4 sequential epochs for a total sweep duration of 4.1045 seconds. The first half of the sweep was presented in one polarity and the second half was presented in the opposite polarity, where the stimulus was multiplied by -1. To account for the systematic increase in sound-level presented to the infant ears caused by their smaller occluded ear-canal volume relative to adults, the stimulus was corrected for the individual infant's ear-canal acoustics using a level-independent acoustic transform known as the real-ear-to-coupler difference (RECD). The RECD is a transfer function used to convert the sound-pressure level (SPL) produced by the stimulus in a reference coupler to the SPL produced by the same stimulus in the individual's ear canal (Bagatto et al., 2005; Moodie et al., 1994; Vaisberg et al., 2018). The RECD is most often completed in a frequency specific manner, however, for the current experiment, the broadband stimulus level was adjusted to present 65 dB SPL in the infant's ear using a Tucker Davis Technologies PA-5 attenuator (Alachua, FL). The in-ear stimulus level was determined using the average level of three stimulus repetitions as measured by an ER-7C probe microphone system (Etymotic Research IL, USA). The probe-tube extended 3-4 mm beyond the termination of the pediatric foam tip (Bagatto et al., 2002; Sininger et al., 1997).

2.2.3 Procedure

Speech-evoked EFRs were measured for all participants using the following procedures. Following skin preparation, three disposable Medi-Trace Ag/AgCl electrodes were placed on the participants to record EEG responses to the monaurally presented stimulus. Ear selection was randomized (12 right ears in infants, 10 right ears in adults). EEG responses were recorded using a single-channel electrode montage with the non-inverting electrode placed on the forehead (Fz), inverting electrode on the ipsilateral mastoid and ground electrode placed on the lateral forehead. Electrode impedances measured before and after the EFR recording were below 5 k Ω with less than 2 k Ω between individual electrode impedances.

Adult participants were seated in a reclining chair in a double-walled sound booth and were instructed to relax, minimize movement as much as possible, and sleep if they were able. Infant participants were placed in a crib or in their parent's arms in a double-walled sound booth. Testing in infants only began after the baby fell asleep, while adult testing began as soon as they were relaxed and resting. Both infant and adult testing sessions aimed for 450 sweeps (30.8 minutes). All adult sessions ended after 450 sweeps. The number of sweeps collected from infant participants varied, as testing was terminated if the participant woke up and was unable to sleep again. Analysis was completed offline using MATLAB (MathWorks, Natick, MA), and R (R Core Team, 2020). Rejection of muscle artifacts was completed on an individual basis using a noise metric computed as the average EEG amplitude between 80-240 Hz for each epoch, which was one-quarter of a sweep in duration. All epochs exceeding two standard deviations above the mean noise metric for each participant were rejected prior to the use of the statistical detection methods. The sweep containing a rejected epoch was also discarded, resulting in a variable number of usable sweeps across both adult (mean = 415, SD = 20, range = 360-435 sweeps) and infant (mean = 401, SD = 43.7, range = 285-435 sweeps) participants.

After artifact rejection and correcting for a 10 ms brainstem delay (Choi et al., 2013), EEG was evaluated using discrete Fourier transforms (DFT; Easwar et al., 2015; Easwar et al., 2020) and a Fourier analyzer (FA) for fricative and vowel elicited EFRs, respectively. Amplitude and phase parameters obtained for each sweep were used to

evaluate response presence at an alpha level of 0.05 for each of the statistical tests: Hotelling's T^2 , Magnitude-Square Coherence (MSC), Rayleigh, Rayleigh-Moore, and the F-test. Each participant was expected to have an EFR present for every carrier in the stimulus due to the suprathreshold presentation level. True positives were operationally defined as a statistically significant signal at the expected f_0 frequency of the carrier. Sensitivity, the proportion of correctly classified significant responses, was calculated using the ratio of significant EFR detections to the total number of expected true positives.

2.2.4 Statistical indicators

F-test: The F-test compares the power of the EEG at a stimulus' f_0 to the average power of neighboring frequencies (Dobie & Wilson, 1996; Lins et al., 1996; Picton et al., 2003). The accuracy of the noise component of this measure is improved with an increasing number of frequency bins (Picton et al., 2003). However, EEG noise is not distributed evenly across frequencies. For an estimate of EEG noise to act as a valid representation of the noise component at the frequency of the EFR, it must be adequately close in frequency to the f_0 of the stimulus (Picton et al., 2003). The frequency resolution of our analysis depended on the reciprocal of stimulus duration, so EEG noise was estimated using 14 neighboring frequencies for the relatively long duration vowels, eight for /s/, and six for the relatively short duration /j/. A response was considered present when its F-ratio exceeded the critical value for F with 2,2x degrees of freedom, where x is the number of adjacent frequencies used in the noise estimate.

Hotelling's T^2 : The Hotelling's T^2 is a multivariate analog of Student's t test, where the expected value of all means is zero (Hotelling, 1931). The Hotelling's T^2 , like the Student's t test, is a parametric statistic that assumes that both the amplitude and phase information follow a multivariate normal distribution. The Hotelling's T^2 test has been used frequently in the literature for detection of EFRs, and other auditory evoked potentials (Cebulla et al., 2006; Easwar, Scollie, et al., 2020; Valdes, Jorge et al., 1997; Vanheusden et al., 2018; Victor, Jonathon & Mast, Joelle, 1991). The Hotelling's T^2 statistic can be transformed into an F-ratio with 2 and n-2 degrees of freedom by

multiplication of the T^2 statistic with $(n - 2)/(2n - 2)$, where n is the number of sweeps (Picton et al., 2003).

Magnitude-Square Coherence: The MSC is mathematically equivalent to the Circular T^2 test (Dobie & Wilson, 1993), a variant of the Hotelling's T^2 , that assumes the variance in the real and imaginary components of Fourier analysis are equivalent (Victor, Jonathon & Mast, Joelle, 1991). Using this assumption, the Circular T^2 , and therefore MSC, is transformed as an F-ratio with 2 and $2n-2$ degrees of freedom by multiplying the Circular T^2 statistic by n , where n is the number of sweeps collected (Picton et al., 2003).

Rayleigh Test: The Rayleigh test is a non-parametric statistic that determines the distribution of phase estimated from each sweep, assessing whether it is uniform (Moore, 1980). The R-statistic is a measure of uniformity in the phase distribution bound between 0 and 1. As the variance in the response phase approaches 0 (i.e. all response phase measurements are the same), R approaches 1, indicating a stronger signal in the noise (Mardia & Jupp, 2009). The R statistic can be distributed as a Chi-square statistic with 2 degrees of freedom given n sweeps where $\chi^2 = 2nR^2$ (Mardia & Jupp, 2009; Stapells et al., 1987).

Rayleigh-Moore Test: The Rayleigh-Moore test is an analogue of the Rayleigh test developed for use with weighted vector data to assess uniformity of the phase distribution (Moore, 1980). Each vector is a complex representation of the amplitude rank and phase estimated from a sweep. For EFRs, this means that phase values are weighted based on the rank of the vector length (spectral magnitude) of the response corresponding to each sweep. Unlike the other statistics being analyzed, to the knowledge of the authors it is not possible to transform the Rayleigh-Moore statistic into an F-ratio. Therefore, the significance of the Rayleigh-Moore statistic is instead tested against a critical value of the modified R statistic (R^*), which can be approximated by $R/n^{3/2}$ as the number of sweeps n increases (Moore, 1980).

2.2.5 Evaluation of test performance

Using R (R Core Team, 2020) and lme4 (Bates et al., 2015), we constructed generalized logistic mixed-effects models to assess the relationship between statistical detection paradigm accuracy, testing duration, and EFR stimulus. Separate logistic regression models were constructed for infant and adult participants to simplify interpretation.

Effect of testing duration (continuous, minutes), EFR carrier (categorical: F1, F2+, fricative), and statistical indicator (categorical: F-test, Rayleigh-Moore test, Rayleigh test, Hotelling's T^2 , and MSC) with three-way interactions, were coded as fixed-effects. Random intercepts for participants and random slopes for both frequency of the stimulus and testing duration were incorporated into the model. All reported p-values were obtained using Wald Chi-square tests. Post-hoc testing was completed by comparing mean response accuracy for each factor adjusted for other variables' estimated marginal means. Estimated marginal mean comparison was completed using the Emmeans package (Lenth, 2020) with Tukey honest significant difference (HSD) multiple comparison corrected p-values as necessary. Estimated marginal means are tested against the standard normal distribution, which is equivalent to obtaining p-values from a t-distribution with infinite degrees of freedom. Overall model fit was also assessed using Nakagawa's pseudo- R^2 (Nakagawa & Schielzeth, 2013) in both populations.

2.3 Results

In infants, there were significant effects of statistical indicator, stimulus frequency, and test duration on speech-evoked EFR detection. These effects are visualized using the estimated sensitivity of each statistical indicator/frequency pairing across testing durations (Figure 2-1). The choice of statistical indicator significantly altered observed sensitivity of response detection ($\chi^2 = 217.82$, $df = 4$, $p < 0.001$). A significant three-way interaction between statistical indicator, test duration, and stimulus frequency was observed ($\chi^2 = 24.73$, $df = 8$, $p = 0.001$). This interaction indicates that the effects of statistical indicator on EFR detection is modulated by the frequency of the carrier and the duration of testing. That is, the benefit of additional testing duration is dependent on the statistical indicator used, the frequency of the stimulus being tested, and the interaction

between these variables. The effect of the statistical indicator interacted with the frequency of the stimulus ($\chi^2 = 74.81$, $df = 8$, $p < 0.001$), with the largest differences between statistical indicators being observed in vowel-elicited EFR detection. There was a significant difference in detection rates between EFR stimulus frequencies and testing duration ($\chi^2 = 33.19$, $df = 4$, $p < 0.001$), indicating that lower frequency stimuli require more time to be detected than higher frequency stimuli. Finally, there was a significant impact of stimulus frequency on detection sensitivity ($\chi^2 = 20.54$, $df = 2$, $p < 0.001$), with higher frequency stimuli having more sensitive and efficient detection of responses when compared to lower frequency stimuli.

The comparative sensitivities of different statistical indicators were assessed after 10, 20, and 30 minutes of usable (i.e. considering only retained sweeps) testing time (Table 2-1). The Rayleigh-Moore test, the ranked, non-parametric analogue of the Rayleigh test, had the greatest sensitivity across all EFR stimuli at 10, 20, and 30 minutes (Table 2-1) except for F1 stimuli after ten minutes of testing, where it performed insignificantly worse than the Rayleigh test. The Rayleigh test performed similarly to the Rayleigh-Moore statistic across all vowel-evoked responses and only marginally worse for fricative-evoked responses. In all conditions with greater than chance sensitivity, the improvement of the Rayleigh-Moore test was significant over the MSC for all stimuli, and over the F-test for all vowel stimuli. Rayleigh-Moore tests were also significantly better than the Hotelling's T^2 statistic for all conditions except for F2+ stimuli at 10 and 30 minutes. The F-test performed poorly for vowel-elicited response detection, failing to significantly exceed chance detection for vowel stimuli after 30 minutes of testing.

Analyzing the model's fit to the infant data, the Nakagawa Pseudo- R^2 for the fixed effects (test duration, frequency of stimulus, and statistical indicator) was 0.49. When accounting for the random effects of the model, the conditional pseudo- R^2 rose to 0.68, indicating a reasonably good fit of the total model to the experimental data.

In adults, the statistical indicator had a significant effect on EFR sensitivity ($\chi^2 = 2011.2$, $df = 4$, $p < 0.001$). A significant interaction was observed between the duration of testing and statistical indicator ($\chi^2 = 144.6$, $df = 4$, $p < 0.001$) and a three-way interaction

between the test duration, statistical indicator, and frequency of stimulus ($\chi^2 = 24.57$, $df = 8$, $p = 0.002$). These results are indicative of a complex relationship between the statistical indicator used, stimulus frequency and the benefit of increased test duration. In contrast to the infant group, the stimulus frequency was not found to be a significant predictor of overall detection sensitivity ($\chi^2 = 3.50$, $df = 2$, $p = 0.17$), nor was the interaction between stimulus frequency and duration of testing ($\chi^2 = 1.97$, $df = 2$, $p = 0.37$). Although both were not statistically significant, a modest descriptive trend of improved accuracy with higher frequency stimuli was observed, especially in phase-incorporating statistical indicators (Figure 2-2).

Statistical indicator sensitivity values were assessed for short (10 minute), mid (20 minute), and long (30 minute) testing durations (Table 2-2). Of note is the improved sensitivity of response detection using the F-test across all stimuli. This trend is particularly notable in vowel-elicited responses, where the F-test had the highest sensitivity during the first 10 minutes of testing. The Hotelling's T^2 and MSC tests performed comparatively poorly in the adult population, with significantly lower sensitivity than the F-test throughout the duration of testing.

Analyzing the model's fit to the adult data, the Nakagawa Pseudo- R^2 for the fixed effects (test duration, frequency of stimulus, and statistical indicator) was 0.33, indicating a modest fit of the fixed effects. When random effects were incorporated, the pseudo- R^2 value rose to 0.83, indicating a very good fit of the model to the data. The comparatively large improvement in model fit with the incorporation of random effects indicates a larger variation in expected response detection sensitivity based on the individual in adults than in infants.

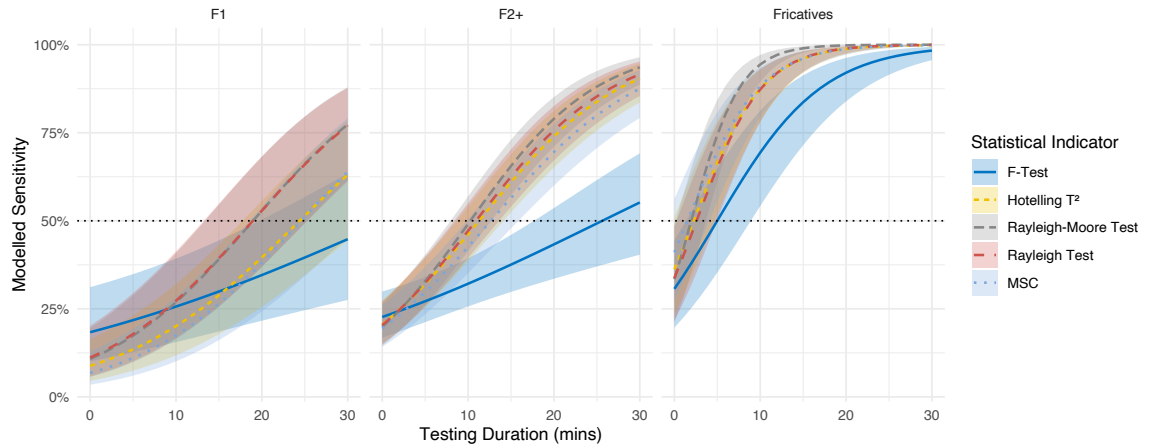


Figure 2-1: Modelled sensitivity of EFR detection in infant participants for F1, F2+, and fricative stimuli across testing duration for each statistical indicator. Shaded area indicates the 95% confidence interval of the sensitivity estimate for each test.

Table 2-1: Sensitivity of response detection in infant participants for F1, F2+, and fricative stimuli after 10, 20, and 30 minutes using each statistical indicator. The best performing indicator in each condition (row) is bolded. 95% confidence interval of sensitivity estimate is provided in brackets. Pairwise post-hoc comparisons to the best performing test were completed with asterisks denoting a significant decrease in sensitivity ($p < 0.05$).

Infant EFR Detection Sensitivity					
As function of testing time and elicitor					
	F-Test	Hotelling T ²	Rayleigh-Moore Test	Rayleigh Test	MSC
10 minutes					
F1	26% [16-39]	20% [12-32]*	27% [16-41]	27% [17-41]	17% [10-28]*
F2+	32% [26-39]*	46% [39-54]	49% [41-57]	47% [39-55]	43% [35-50]*
Fricatives	69% [54-81]*	87% [78-93]*	94% [89-97]	87% [78-93]*	88% [80-94]*
20 minutes					
F1	35% [22-50]*	40% [26-56]*	53% [37-68]	53% [37-68]	38% [24-54]*
F2+	43% [34-54]*	74% [65-81]*	79% [71-85]	76% [67-83]	70% [60-78]*
Fricatives	92% [84-96]	99% [97-100]*	100% [100-100]	99% [98-100]*	99% [97-99]*
30 minutes					
F1	45% [28-63]*	63% [44-78]*	77% [62-88]	77% [61-88]	64% [45-79]*
F2+	55% [40-69]*	90% [84-95]	94% [89-96]	92% [85-95]	88% [79-93]*
Fricatives	98% [96-99]*	100% [100-100]*	100% [100-100]	100% [100-100]*	100% [100-100]*
* Significantly worse than the best performing indicator for this condition					

Adult EFR Detection Sensitivity					
As function of testing time and elicitor					
	F-Test	Hotelling T ²	Rayleigh-Moore Test	Rayleigh Test	MSC
10 minutes					
F1	77% [55-90]	42% [20-66]*	64% [39-83]*	60% [35-81]*	39% [19-64]*
F2+	82% [66-91]	56% [36-75]*	74% [55-86]*	69% [49-83]*	53% [33-72]*
Fricatives	94% [85-98]*	92% [81-97]*	96% [91-99]	97% [93-99]	89% [76-96]*
20 minutes					
F1	93% [74-98]	77% [41-94]*	91% [69-98]	91% [68-98]	75% [39-94]*
F2+	95% [83-99]*	89% [66-97]*	96% [87-99]	95% [83-99]	86% [62-96]*
Fricatives	100% [98-100]*	100% [98-100]*	100% [99-100]*	100% [100-100]	99% [97-100]*
30 minutes					
F1	98% [86-100]	94% [63-99]*	99% [88-100]	99% [89-100]	94% [63-99]*
F2+	99% [91-100]*	98% [86-100]*	100% [97-100]	99% [96-100]	97% [83-100]*
Fricatives	100% [100-100]	100% [100-100]	100% [100-100]	100% [100-100]	100% [100-100]
* Significantly worse than the best performing indicator for this condition					

Table 2-2: Sensitivity of response detection in adult participants for F1, F2+, and fricative stimuli after 10, 20, and 30 minutes using each statistical indicator. The best performing indicator in each condition (row) is bolded. 95% confidence interval of sensitivity estimate is provided in brackets. Pairwise post-hoc comparisons to the best performing test were completed with asterisks denoting a significant decrease in sensitivity ($p < 0.05$). Determination of significance was completed on log-odds of detection, which asymptote at 100%. As a result, detection chance near 100% can be statistically significantly different even though detection rates may differ by less than 1%.

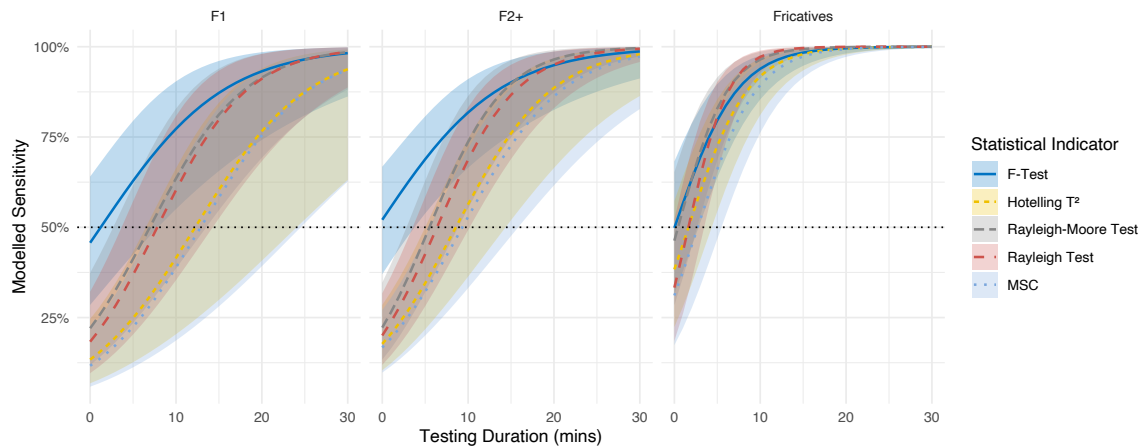


Figure 2-2: Sensitivity of response detection in adult participants for F1, F2+, and fricative stimuli across testing duration for each statistical indicator. Shaded area indicates the 95% confidence interval of the sensitivity estimate for each test.

2.3.1 Comparison between populations

As hypothesized, infants had comparable or lower overall sensitivity to response detection across all conditions (Figure 2-3). The magnitude of the differences in sensitivities was largest for the vowel elicited EFRs, which was expected given the relatively accurate detection of responses elicited by the fricatives which were higher frequency stimuli. The F-test exhibited the largest differences between adult and infant detection sensitivities, with generally less than chance detection rates in the infant population. Notably, to approach comparative sensitivities to those for adults, longer testing durations were necessary in the infant population for vowel stimuli regardless of statistical indicator used.

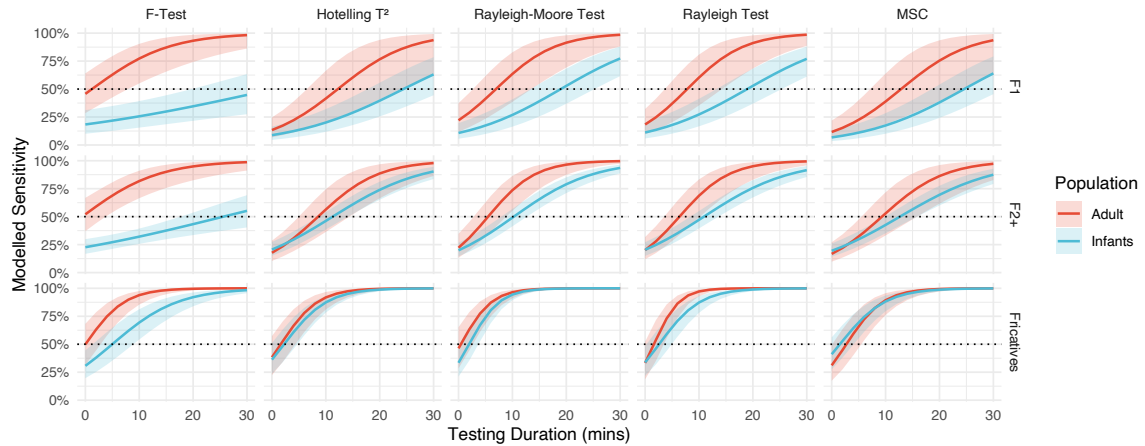


Figure 2-3: Comparison of modelled response detection sensitivity between populations. Columns are separated by the statistical indicator used, while the rows are split by stimulus type. Shaded area indicates the 95% confidence interval of the sensitivity estimate for the population.

2.3.2 Minimum testing duration for clinical benefit

Minimum testing time required for potential clinical benefit was determined by identifying sensitivities that were significantly greater than chance – those with 95% confidence intervals falling above 50%. If testing durations do not exceed these minimums, it cannot be determined that the presence of a response at a given time is greater than chance. The minimum testing time necessary for sensitivities to meet this criterion was dependent on the statistical indicator implemented (Table 2-3).

Furthermore, the minimum testing time decreased as the frequency of the stimulus increased in both populations. Minimum testing times in infants for F2+ and fricative elicited responses fell within 5 minutes of minimum testing times in adults for all indicators other than the F-test. Larger differences were observed for low-frequency F1 stimuli. Importantly, for the F-test in infants, the minimum testing time necessary for better-than-random response detection in vowel elicited stimuli exceeded maximum testing durations employed in this study. This indicates the F-test may not be clinically feasible as a statistical indicator in an infant population. Two other tests, the Hotelling's T² test and the MSC test, also exceeded maximum testing durations employed in this

study. However, they were only over the maximum by less than 5 minutes each, and thus may still prove clinically feasible.

Estimated test duration (minutes) for significantly higher than 50% sensitivity						
Function	F1		F2+		Fricative	
	adults	infants	adults	infants	adults	infants
F-Test	8	73*	4	45*	3	9
Hotelling T ²	25	33*	15	15	5	5
Rayleigh-Moore Test	14	26	9	13	3	4
Rayleigh Test	15	26	11	14	4	5
MSC	25	32*	16	16	6	5

* Extrapolated from model, exceeded maximum testing duration

Table 2-3: Minimum test duration necessary for better than chance sensitivity using each statistical indicator for both infants and adults across low- (F1), mid- (F2+), and high-frequency (fricative) EFR stimuli. Asterisks denote that the minimum testing duration exceeded the testing duration of the current experiment. Lower durations indicate greater efficiency.

2.4 Discussion

The current study had two main objectives. Firstly, the sensitivity of statistical indicators used for objective response detection in infants was assessed. Secondly, the sensitivity of EFR detection between infants and adults was compared. The principal findings in the study were: (1) the statistical indicator has a significant impact on EFR detection in both infants and adults; (2) infants required longer testing durations to achieve an acceptable level of sensitivity when compared to adults, especially for low-frequency stimuli and with statistics that do not explicitly incorporate response phase into detection.

Consequently, choosing one indicator over another can have a significant effect on the overall accuracy and duration of testing needed for response detection, both of which are important in a clinical setting. This effect was present in both infants and adults.

In infants, the F-test performed poorly, failing to detect responses with accuracy significantly greater than chance for all vowel stimuli. In comparison, the Rayleigh-Moore and Rayleigh tests, had the best performance across all stimuli, with the Hotelling's T^2 and MSC statistics performing better than the F-test, but worse than the Rayleigh and Rayleigh-Moore tests.

In adults, the F-test performed well across stimuli, as did the Rayleigh-Moore and Rayleigh tests when test duration exceeded about 10 minutes for vowels. In both populations, Hotelling's T^2 and MSC performed significantly worse than the best performing statistical indicators in most conditions, indicating that the use of either of these statistics in speech-evoked EFR detection may not be preferred. The decreased accuracy of the Hotelling's T^2 and MSC may be caused by a deviation from perfectly random background EEG amplitudes at the response frequency, an assumption of both statistical indicators. Deviations in this assumption may lead to decreased sensitivity when compared to tests that do not incorporate amplitude and those that do not rely on such assumptions (Dobie & Wilson, 1994; Lachaux et al., 1999.; Zhu et al., 2013).

Previous research on statistical indicator performance of tone-evoked EFR detection has shown that indicators using phase and amplitude have shown comparable performance to those that use only amplitude or only phase in an adult population (Picton et al., 1987; Valdes et al., 1997; Dobie, Wilson, 1994). For speech-evoked EFR, results on statistical indicators have been mixed, with some studies indicating an improvement in F-test performance over phase-incorporating statistics (Aiken and Picton, 2006) and others indicating the opposite (Vanheusdan et al., 2019, Easwar 2020). Differences between speech-evoked EFR studies, such as testing duration, stimuli, and EFR estimation technique (DFT vs. FA) makes direct comparison between studies difficult and may alter EFR estimates. Furthermore, work on statistical indicator accuracy has primarily been completed in adults rather than in infants. Previous results regarding statistical indicator accuracy in adults may not generalize accurately to an infant population. Based on the current findings, there is a significant and complex difference in response detection between populations.

The comparison of the sensitivity of EFR detection between infants and adults demonstrated that decreased sensitivity in infant populations was dependent on statistical indicator, the frequency of the stimulus, and the interaction between these factors.

Notably, the F-test exhibited a much larger difference in sensitivity over time between adult and infant population for all stimuli. However, regardless of the statistical indicator used, infants generally required more testing time to achieve the same sensitivity of detection as in adults. It is well known that characteristics and detection of the speech-evoked EFR differ between infants and adults, which is in agreement with the decreased sensitivity in infants (Easwar et al., 2020; Van Dyke et al., 2017, Savio et al., 2001). This decreased sensitivity has been mainly attributed to decreased amplitude rather than phase locking differences between populations (Easwar et al., in review; Lins et al., 1996; Van Dyke et al., 2017). The F-test uses only the response amplitude with an implicit dependence on intertrial reliability (John & Purcell, 2008; Lins et al., 1996; Picton et al., 2003), whereas other statistical indicators explicitly make use of phase information. Indicators that used phase showed much smaller effects of population on overall sensitivity, providing further evidence that amplitude primarily differed between populations. However, all differences cannot be attributed to the spectral magnitude. The Rayleigh test, which does not use the spectral magnitude information, also showed an effect of the population on overall sensitivity. Decreased sensitivity in the Rayleigh test suggests a difference in phase coherence values in infants, albeit a smaller effect than the observed decrease in other statistics likely caused by amplitude differences.

Additionally, there are other benefits to using the Rayleigh test. Statistics that incorporate amplitude information may introduce potential confounds caused by frequency-specific noise when response amplitude is used for response detection (Zhu et al., 2013). The Rayleigh test does not use amplitude information (Picton et al., 2003; Stapells et al., 1987), which may improve validity in comparisons of EFR responses between multiple frequency conditions (Zhu et al., 2013). Consequently, the current results support the use of the Rayleigh test for speech-evoked EFR response detection. However, a clinically important difference was noted between infants and adults regardless of statistical indicator used: minimum testing duration.

In the current work, infants required longer testing durations for comparable accuracy to adults. This effect is greater for low frequency stimuli, where even the best performing statistical indicators require almost twice as long to reach significantly non-random sensitivities (Table 2-3). Even with the comparatively long testing duration, meaningful response detection was observed in less than 30 minutes of testing across all speech-evoked EFR stimuli, including low-frequencies. For hearing losses requiring aided objective outcome validation in only mid- and/or higher frequencies, highly sensitive results can be obtained in less time when compared to those requiring validation in low frequencies, decreasing the necessary testing duration (Table 2-1). For example, in an infant population assessed by the Rayleigh-Moore statistic, 20 minutes of testing would likely be sufficient for the validation of a mid-to-high frequency hearing loss. This test duration would provide approximately 100% sensitivity in fricative elicited responses and 79% sensitivity in F2+ responses. If low-frequency, F1 stimuli need to be validated, a testing duration of 20 minutes would provide no significant clinical benefit, with a sensitivity of 53% (CI₉₅ = 37-68%). A minimum of 26 minutes is necessary for the /susɑ̃ʃi/ stimulus using the Rayleigh-Moore test to gain significant benefit in F1 stimulus testing. After 30 minutes of testing, the Rayleigh-Moore test yielded an F1 stimulus detection sensitivity of 77% (CI₉₅ = 62-88%). The minimum testing duration necessary accounts for the length of the /susɑ̃ʃi/ stimulus, however, the actual stimulation and measurement duration for each individual phoneme is significantly shorter. Although untested in the current study, adjustments to the stimulus to remove detected phonemes is expected to save measurement time for undetected phonemes, further decreasing the necessary testing duration.

2.5 Conclusion

For EFR to be incorporated into a clinical workflow as an objective hearing aid outcome measure, efficient and sensitive response detection is needed for both infants and adults. Speech-evoked EFR detection sensitivity is dependent on the statistical indicator used, the stimulus, and the age of the participant. There are significant differences in EFR detection sensitivity between infant and adult populations that are modulated by the statistical indicator used. Amplitude-based indicators show a larger effect of population

and perform relatively poorly in an infant population. By assessing the sensitivity of response detection in infants and adults given a variable testing time, the current study has shown that using the Rayleigh test or Rayleigh-Moore test for speech-evoked EFR detection provide highly sensitive speech-evoked EFR response detection in under 30 minutes of testing.

2.6 References

- Aiken, S. J., & Picton, T. W. (2006). Envelope Following Responses to Natural Vowels. *Audiology and Neurotology*, *11*(4), 213–232. <https://doi.org/10.1159/000092589>
- Aiken, S. J., & Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hearing Research*, *245*(1–2), 35–47. <https://doi.org/10.1016/j.heares.2008.08.004>
- Anderson, S., & Kraus, N. (2013). The Potential Role of the cABR in Assessment and Management of Hearing Impairment. *International Journal of Otolaryngology*, *2013*, 1–10. <https://doi.org/10.1155/2013/604729>
- Bagatto, M. P., Scollie, S., Seewald, R. C., Moodie, K. S., & Hoover, B. M. (2002). Real-Ear-to-Coupler Difference Predictions as a Function of Age for Two Coupling Procedures. *Journal of the American Academy of Audiology*, *13*(8), 9.
- Bagatto, Moodie, S., Scollie, S., Seewald, R., Moodie, S., Pumford, J., & Liu, K. P. R. (2005). Clinical Protocols for Hearing Instrument Fitting in the Desired Sensation Level Method. *Trends in Amplification*, *9*(4), 199–226. <https://doi.org/10.1177/108471380500900404>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bidelman, G. M. (2018). Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. *NeuroImage*, *175*, 56–69. <https://doi.org/10.1016/j.neuroimage.2018.03.060>
- Cebulla, M., Stürzebecher, E., & Elberling, C. (2006). Objective Detection of Auditory Steady-State Responses: Comparison of One-Sample and q-Sample Tests. *Journal of the American Academy of Audiology*, *17*(02), 093–103. <https://doi.org/10.3766/jaaa.17.2.3>
- Choi, J. M., Purcell, D. W., Coyne, J.-A. M., & Aiken, S. J. (2013). Envelope Following Responses Elicited by English Sentences: *Ear and Hearing*, *34*(5), 637–650. <https://doi.org/10.1097/AUD.0b013e31828e4dad>
- Dobie, R. A., & Wilson, M. J. (1993). Objective response detection in the frequency domain. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *88*(6), 516–524. [https://doi.org/10.1016/0168-5597\(93\)90040-V](https://doi.org/10.1016/0168-5597(93)90040-V)
- Dobie, R. A., & Wilson, M. J. (1994). Phase weighting: A method to improve objective detection of steady-state evoked potentials. *Hearing Research*, *79*(1–2), 94–98. [https://doi.org/10.1016/0378-5955\(94\)90130-9](https://doi.org/10.1016/0378-5955(94)90130-9)
- Dobie, R. A., & Wilson, M. J. (1996). A comparison of *t* test, *F* test, and coherence methods of detecting steady-state auditory-evoked potentials, distortion-product otoacoustic emissions, or other sinusoids. *The Journal of the Acoustical Society of America*, *100*(4), 2236–2246. <https://doi.org/10.1121/1.417933>

- Easwar, V., Birstler, J., Harrison, A., Scollie, S., & Purcell, D. (2020). The accuracy of envelope following responses in predicting speech audibility. *Ear & Hearing, Publish Ahead of Print*. <https://doi.org/10.1097/AUD.0000000000000892>
- Easwar, V., Purcell, D., Aiken, S. J., Parsa, V., & Scollie, S. (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*. 36(6), 18.
- Easwar, V., Scollie, S., Aiken, S., & Purcell, D. (2020). Test-Retest Variability in the Characteristics of Envelope Following Responses Evoked by Speech Stimuli: *Ear and Hearing*, 41(1), 150–164. <https://doi.org/10.1097/AUD.0000000000000739>
- Easwar, V., Scollie, S., Lasarev, M., Urichuk, M., Aiken, S., & Purcell, D. (2021). Characteristics of speech-evoked envelope following responses in infancy. *Trends in Hearing*, 25, 1-17. DOI: 10.1177/23312165211004331.
- Henning, R. W., & Bentler, R. (2005). Compression-Dependent Differences in Hearing Aid Gain Between Speech and Nonspeech Input Signals: *Ear and Hearing*, 26(4), 409–422. <https://doi.org/10.1097/00003446-200508000-00004>
- Herdman, A. T., & Stapells, D. R. (2001). Thresholds determined using the monotic and dichotic multiple auditory steady-state response technique in normal-hearing subjects. *Scandinavian Audiology*, 30(1), 41–49. <https://doi.org/10.1080/010503901750069563>
- Hotelling, H. (1931). The generalization of student's ratio. *Ann. Math. Statist.*, 2(3), 360–378. <https://doi.org/10.1214/aoms/1177732979>
- Jeng, F.-C., Hu, J., Dickman, B., Montgomery-Reagan, K., Tong, M., Wu, G., & Lin, C.-D. (2011). Cross-Linguistic Comparison of Frequency-Following Responses to Voice Pitch in American and Chinese Neonates and Adults: *Ear and Hearing*, 32(6), 699–707. <https://doi.org/10.1097/AUD.0b013e31821cc0df>
- Jenkins, K. A., Fodor, C., Presacco, A., & Anderson, S. (2018). Effects of Amplification on Neural Phase Locking, Amplitude, and Latency to a Speech Syllable: *Ear and Hearing*, 39(4), 810–824. <https://doi.org/10.1097/AUD.0000000000000538>
- John, S., & Purcell, D. (2008). Introduction to Technical Principles of Auditory Steady-State Response Testing. In *The Auditory Steady State Response*.
- Lachaux, J.-P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). *Measuring Phase Synchrony in Brain Signals*. 1.
- Lenth, R. (2020). *emmeans: Estimated marginal means, aka least-squares means* [Manual]. <https://CRAN.R-project.org/package=emmeans>
- Lins, O. G., Picton, T. W., Boucher, B. L., Durieux-Smith, A., Champagne, S. C., Moran, L. M., Perez-Abalo, M. C., Martin, V., & Savio, G. (1996). Frequency-Specific Audiometry Using Steady-State Responses. *Ear and Hearing*, 17(2). https://journals.lww.com/ear-hearing/Fulltext/1996/04000/Frequency_Specific_Audiometry_Using_Steady_State.1.aspx

- Mardia, K. V., & Jupp, P. E. (2009). *Directional statistics*. Wiley.
<https://books.google.ca/books?id=PTNiCm4Q-M0C>
- Mijares, E., Pérez Abalo, M. C., Herrera, D., Lage, A., & Vega, M. (2013). Comparing statistics for objective detection of transient and steady-state evoked responses in newborns. *International Journal of Audiology*, *52*(1), 44–49.
<https://doi.org/10.3109/14992027.2012.736030>
- Moodie, S., Seewald, R. C., & Sinclair, S. (1994). Procedure for Predicting Real-Ear Hearing Aid Performance in Young Children. *American Journal of Audiology*, *3*(1), 23–31. <https://doi.org/10.1044/1059-0889.0301.23>
- Moore, B. R. (1980). A modification of the Rayleigh test for vector data. *Biometrika*, *67*(1), 175–180. <https://doi.org/10.1093/biomet/67.1.175>
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133–142. <https://doi.org/10.1111/J.2041-210x.2012.00261.X>
- Picton, T. W., John, M. S., Purcell, D. W., & Plourde, G. (2003). Human Auditory Steady-State Responses: The Effects of Recording Technique and State of Arousal: *Anesthesia & Analgesia*, 1396–1402.
<https://doi.org/10.1213/01.ANE.0000082994.22466.DD>
- R Core Team. (2020). *R: A language and environment for statistical computing* [Manual]. <https://www.R-project.org/>
- Savio, G., Cárdenas, J., Pérez Abalo, M., González, A., & Valdés, J. (2001). The Low and High Frequency Auditory Steady State Responses Mature at Different Rates. *Audiology and Neurotology*, *6*(5), 279–287. <https://doi.org/10.1159/000046133>
- Scollie, S., & Seewald, R. C. (2002). Evaluation of Electroacoustic Test Signals I: Comparison with Amplified Speech: *Ear and Hearing*, *23*(5), 477–487.
<https://doi.org/10.1097/00003446-200210000-00009>
- Sininger, Y. S., Abdala, C., & Cone-Wesson, B. (1997). Auditory threshold sensitivity of the human neonate as measured by the auditory brainstem response. *Hearing Research*, *12*.
- Sininger, Y. S., Martinez, A., Eisenberg, L., Christensen, E., Grimes, A., & Hu, J. (2009). Newborn Hearing Screening Speeds Diagnosis and Access to Intervention by 20-25 Months. *Journal of the American Academy of Audiology*, *20*(1), 49–57.
<https://doi.org/10.3766/jaaa.20.1.5>
- Stapells, D. R., Makeig, S., & Galambos, R. (1987). Auditory steady-state responses: Threshold prediction using phase coherence. *Electroencephalography and Clinical Neurophysiology*, *67*(3), 260–270. [https://doi.org/10.1016/0013-4694\(87\)90024-1](https://doi.org/10.1016/0013-4694(87)90024-1)
- The Joint Committee on Infant Hearing. (2019). Year 2019 Position Statement: Principles and Guidelines for Early Hearing Detection and Intervention Programs. *The Journal of Early Hearing Detection and Intervention*, *4*(2), 1–44.

- Uus, K., & Bamford, J. (2006). Effectiveness of Population-Based Newborn Hearing Screening in England: Ages of Interventions and Profile of Cases. *PEDIATRICS*, *117*(5), e887–e893. <https://doi.org/10.1542/peds.2005-1064>
- Vaisberg, J. M., Folkeard, P., Pumford, J., Narten, P., & Scollie, S. (2018). Evaluation of the Repeatability and Accuracy of the Wideband Real-Ear-to-Coupler Difference. *Journal of the American Academy of Audiology*, *29*(06), 520–532. <https://doi.org/10.3766/jaaa.17007>
- Valdes, Jorge, Perez-Abalo, Maria, Martin, Vivian, Savio, Guillermo, Sierra, Carlos, Rodriguez, Ernesto, & Lins, Octavio. (1997). Comparison of Statistical Indicators for the Automatic Detection of 80 Hz Auditory Steady State Responses. *Ear and Hearing*, *18*, 420–429.
- Van Dyke, K. B., Lieberman, R., Presacco, A., & Anderson, S. (2017). Development of Phase Locking and Frequency Representation in the Infant Frequency-Following Response. *Journal of Speech, Language, and Hearing Research*, *60*(9), 2740–2751. https://doi.org/10.1044/2017_JSLHR-H-16-0263
- Vanheusden, F., Bell, S., Chesnaye, M., & Simpson, D. (2018). Improved detection of vowel envelope frequency following responses using hotelling's T2 analysis. *Ear and Hearing*, *40*. <https://doi.org/10.1097/AUD.0000000000000598>
- Victor, Jonathon, & Mast, Joelle. (1991). A new statistic for steady-state evoked potentials. *Electroencephalography and Clinical Neurophysiology*, *78*, 378–388.
- Zhu, L., Bharadwaj, H., Xia, J., & Shinn-Cunningham, B. (2013). A comparison of spectral magnitude and phase-locking value analyses of the frequency-following response to complex tones. *The Journal of the Acoustical Society of America*, *134*(1), 384–395. <https://doi.org/10.1121/1.4807498>

Chapter 3

3 Thévenin-equivalent source parameter reliability: intra- and inter-sessional effects on stimulus level calibration and wideband reflectance²

3.1 Introduction

Accurate sound level calibration of acoustic stimuli in the human ear is necessary for almost every measurement in audiology. From diagnostic tests to hearing aid fittings, determination of the sound level presented to the individual's auditory system is a component of test accuracy. Measurements calibrated *in-situ* using forward-pressure level (FPL) and integrated pressure level (IPL) provide better accuracy than SPL-calibration (Lewis et al., 2009; Souza et al., 2014) and provide a robust wideband acoustic immittance (WAI) measurement of middle ear-function. FPL, IPL and WAI measures all rely on accurate determination of the transducer's effects on the stimulus, which are determined prior to testing using a technique known as Thévenin-equivalent source-parameter calibration (Voss & Allen, 1994). Source parameter calibration has been completed at least once daily in WAI studies (Rosowski et al., 2012; Shahnaz et al., 2009; Voss & Allen, 1994) and in studies that analyze IPL or FPL calibrated stimuli (Boothalingam et al., 2018; Burke et al., 2010; Richmond et al., 2011; Rogers et al., 2010; Withnell et al., 2009). Daily source parameter calibration may present a barrier to the incorporation of WAI, IPL and/or FPL into a clinical workflow given the additional clinical time necessary for the calibration procedure. Ensuring reliability in WAI, IPL, and FPL clinical measurements while decreasing the need for frequent calibrations may improve clinical uptake of these tests. The reliability of source parameter calibration, and the impact on FPL, IPL, and WAI reliability is explored in the current study.

² A version of this chapter has been submitted: Urichuk, M., Purcell, D., Allen, P., Scollie, S. Thévenin-equivalent source parameter reliability: intra- and inter-sessional effects on stimulus level calibration and wideband reflectance. Submitted to *International Journal of Audiology*.

3.1.1 Calibration methods: ear canal sound pressure level

The stimulus level presented to the auditory system may be calibrated using coupler-referenced sound level. Calibrating the level of the output in a coupler does not account for the unique acoustic properties of an individual's peripheral auditory system (Chan & Geisler, 1990; Gilman & Dirks, 1986; Voss et al., 2000; Voss & Herrmann, 2005) which results in up to 40 dB of inter-subject variability in adult ear canals, measured in SPL (Saunders & Morgan, 2003; Valente et al., 1994, 1997). To account for this inter-subject variability, the sound level measurement can be performed in the individual's ear canal. Ear canal sound level measurement can be performed distal to the eardrum using a probe-housed microphone, such as for otoacoustic emissions (Siegel, 2007), or using a probe-tube microphone placed close to the eardrum, such as for hearing aid verification (Mueller, 2001).

The sound level at the microphone is used as an estimate of the sound level entering the auditory system. However, at high frequencies, the sound-field in the ear canal is non-uniform due to the interaction of forward and reflected waves. Interference from the reflected wave results in measurement errors up to 20 dB along the length of the ear canal (Siegel, 1994; Stinson & Lawton, 1989). The distance between the microphone and the eardrum is inversely related to the frequency of the standing wave. Measurements made by relatively distal microphones housed in the transducer result in measurement errors above 2 kHz (Lapsley Miller et al., 2018). Placing a probe-tube near the eardrum, as done in hearing aid verification, shifts the reflected wave interference to a frequency above 4 kHz (McCreery et al., 2009). Although the accuracy and reliability of probe-tube microphone measurements are excellent below 4 kHz, the accuracy and reliability of the measures decrease at higher frequencies (Vaisberg et al., 2016, 2018). Accuracy in probe-tube measurements requires accurate probe-tube placement (Vaisberg et al., 2016), so may be dependent upon clinician skill and training. Furthermore, while the probe-tube measurement mitigates standing wave errors at frequencies below 4 kHz, standing wave errors remain at higher frequencies, particularly at 8 kHz and higher. As a result, probe-tube measurements cannot fully eliminate standing wave interference effects.

3.1.2 Calibration methods: alternatives to ear canal SPL

Recent studies have shown that estimating sound levels in the ear from a remote pressure response measurement (i.e. from the sound-source itself) using a source parameter calibrated transducer is more accurate than using measurements of ear canal SPL. This is due to the elimination of standing wave errors from the measurement (Farmer-Fedor & Rabbitt, 2002; Keefe & Schairer, 2011; McCreery et al., 2009). Using the transducer's source parameters, it is possible to measure stimuli in terms of only the forward-moving pressure level (FPL), or the integrated pressure level (IPL). FPL is a measurement of only the forward-moving sound-wave component without any interaction of the reflected sound-wave (Scheperle et al., 2008). IPL describes the sound-pressure level present at the termination of a cylindrical cavity, where forward and reflected sound-waves interact perfectly in-phase. For measurements in the ear canal, this means the IPL measured by the probe will be an estimate of the SPL measured by a microphone at the eardrum (Lewis et al., 2009; Scheperle et al., 2011). Measuring ear canal stimulus level using either FPL or IPL has a number of advantages over SPL-based approaches. FPL and IPL measures eliminate standing wave interference, do not require highly skilled probe placements, account for transducer effects, and provide better accuracy even at extended high frequencies (Scheperle et al., 2008, 2011; Souza et al., 2014). Additionally, these in-situ measurements yield the wideband acoustic immittance (WAI), providing information about the individual's middle ear that shows promise as a means of identifying various middle ear pathologies, including otitis media, and conductive hearing losses (Beers et al., 2010; Feeney et al., 2009; Hunter et al., 2010; Keefe et al., 2000; Nakajima et al., 2013; Voss et al., 2012).

Although source parameter calibration is a relatively quick procedure (< 5 minutes), there are currently no guidelines for how frequently calibration needs to be completed or the effect that source parameter reliability may have on clinical measurements of FPL, IPL and WAI. In the present study, we assess the intra-session and inter-session reliability of source parameter calibration and the effect of variation on stimulus calibration and WAI measurements in an adult human ear and a reference coupler, by repeating calibration and WAI measures every day for 30 consecutive days.

3.2 Thevenin-equivalent source parameter calibration

3.2.1 Source parameter derivation

Source parameter calibration requires multiple (generally four) calibration cavities with known dimensions (Allen, 1986; Keefe et al., 1992). These cavities are generally tubes made of metal, such as stainless steel or copper, and are closed at one end with a right angle. A specific measurement is made in each cavity, and the resulting data are used to obtain the transducer's source parameters: the source pressure (P_{src}) and source impedance (Z_{src}). The source parameters estimate the transducer's effect on the stimulus *independent of the cavity it is measuring*, which allows for the conversion of subsequent pressure response measurements into individualized in-situ FPL and/or IPL measures for unknown cavities including ear canals. While probe-tube measurements quantify the acoustic effects of an individual's ear canal independent of the stimulus, the transducer's source parameters quantify the transducer's acoustic effects independent of the ear or cavity it will be stimulating.

To begin source parameter calibration, each cavity's length and radius must be known to calculate the characteristic impedance and the ideal acoustic impedance of each cavity (Nørgaard et al., 2017). The characteristic impedance (Z_0) of the cavity is the resistance the introduced sound would encounter if it were to pass through an infinitely long tube with the same radius as the cavity being measured. The Z_0 is a frequency-independent property of each cavity. Mathematically, this is calculated as (Keefe, 1984):

$$Z_0 = \frac{\rho c}{\pi r^2} \quad (3.1)$$

where cavity radius (r ; 0.2 cm for cavities used in this study), a constant for the density of air (ρ ; approximately 1.225 kg/m³), and speed of sound (c ; approximately 344 m/s) are known. In contrast, the ideal acoustic impedance (Z_{cav}) is the resistance the introduced sound is expected to encounter in the acoustic cavity. The acoustic impedance and each subsequent measure introduced in this section are functions of frequency and are calculated separately for each analyzed frequency, although not explicitly denoted. Using

the characteristic impedance, the ideal acoustic impedance (Z_{cav}) of each calibration cavity can be calculated (Scheperle et al., 2008):

$$Z_{cav} = -jZ_0 \cot(\Gamma l) \quad (3.2)$$

where j is the imaginary unit vector (used to represent a 90° phase shift between Z_0 and Z_{cav}), l is the length of the calibration cavity (1.2-2.0cm in this study), Γ is the wave-propagation constant (the frequency dependent phase change per unit length along the tube; calculated in Keefe 1984), and \cot is the trigonometric function cotangent. Pressure measurements are then made in each cavity (P_{cav}) using a probe containing the stimulus transducer and a microphone. The calculated Z_{cav} relates the pressure response measured in each cavity (P_{cav}) to the transducer's source pressure (P_{src}) and source impedance (Z_{src}) by (Scheperle et al., 2008):

$$\frac{P_{src}}{Z_{src} + Z_{cav}} = \frac{P_{cav}}{Z_{cav}} \quad (3.3)$$

With the analytically determined Z_{cav} and the corresponding measurement of P_{cav} for each of the four calibration cavities, we can isolate P_{src} and Z_{src} to solve for the transducer's source parameters by minimizing error between the theoretical pressure calculated using coupler dimensions and the pressure measurement made by the probe (Allen, 1986; Neely, Gorga 1998; Scheperle et al., 2008).

Traditionally, error minimization is the final step of the calibration procedure; the values of P_{src} and Z_{src} are used for IPL calibration and WAI measurement. However, the present study uses an updated calibration procedure introduced by Nørgaard and colleagues (2017) that applies further analytical corrections for evanescent (non-propagating) sound waves caused by the geometrical discontinuity at the transducer-cavity coupling location (i.e. the change in tube diameter at the probe tip; Huang et al., 2000; Keefe et al., 1993). These evanescent waves are present at the transducer, but decay exponentially and do not contribute to the sound level at the termination of the cavity (or at the eardrum). Evanescent wave effects are minimized and reliable transducer placement in the calibration cavities is ensured by using small diameter cavities that couple directly to the

transducer without a plastic (or foam) tip. A direct coupling to smaller calibration cavities minimizes effects of evanescent waves and improves accuracy in high frequency calibration (Nørgaard et al., 2017).

3.2.2 Ear canal stimulus level measurement

Once source parameter calibration is complete, the transducer can be used to measure the IPL and WAI of acoustic loads (i.e. ear canals, other couplers) with unknown dimensions. Both IPL and WAI measurements are derived from the individual's acoustic impedance (Z_{ec} ; Souza et al., 2014). Z_{ec} is calculated using:

$$Z_{ec} = \frac{Z_{src}P_{meas}}{P_{src} - P_{meas}} \quad (3.5)$$

where P_{meas} is the wideband pressure response measurement made in the ear canal or unknown acoustic cavity of interest. Using the individual's acoustic impedance Z_{ec} and an estimate of the ear's characteristic impedance (Z_0 ; where $Z_0 = \frac{\rho c}{\pi r^2}$), the pressure reflectance of the acoustic load can be calculated (Souza et al., 2014):

$$R_{ec} = \frac{Z_{ec} - Z_0}{Z_{ec} + Z_0} \quad (3.6)$$

The value of the Z_0 can be either estimated using the average adult ear, as in the current study, or can be determined in the individual by iteratively adjusting the Z_0 value used for Z_{ec} calculation. Ear canal sizes differ significantly between individuals; the smaller the ear canal, the larger the characteristic impedance. For an individual measurement, the Z_0 iteration that minimizes the calculated value of Z_{ec} at time $t = 0$, referred to as the surge impedance, will accurately estimate Z_0 of the ear canal at the location of the probe (Rasetshwane & Neely, 2011). While individualized Z_0 measurement improves FPL/IPL accuracy (Scheperle et al., 2011; Souza et al., 2014), any error introduced by the current study's approach using the average adult canal radius will be constant across all calibration trials. Because this will be equally present across all calibration trials, source calibration variability on WAI, IPL and FPL measurement are not expected to be significantly impacted.

With an estimate of R_{ec} , we can determine the proportion of the measured sound level (P_{meas}) that is moving toward the eardrum/cavity termination without reflected wave interference – the forward pressure level (P_{FPL} ; Scheperle et al., 2011):

$$P_{FPL} = \frac{P_{meas}}{1 + R_{ec}} \quad (3.7)$$

The P_{FPL} is measured in micro-Pascals and is convertible to dB SPL re 20 μ Pa (Lapsley Miller et al., 2018). When assessing P_{FPL} using decibels, we refer to the measurement as FPL. We can also determine the proportion of the total sound level reflected back towards the transducer (P_{RPL} ; Scheperle et al., 2011):

$$P_{RPL} = P_{meas} - P_{FPL} = R_{ec}P_{FPL} \quad (3.8)$$

The forward-pressure level has been shown to be more accurate than eardrum SPL estimated by conventional probe-tube measurements in real ears (McCreery et al., 2009; Souza et al., 2014). At low frequencies, FPL is approximately 6 dB less than the SPL, but as wavelength decreases at higher frequencies, the conversion between FPL and SPL differs based on the individual's middle-ear status, residual ear canal space, and the probe placement. Integrated pressure level (IPL) measurement has the same benefits as FPL: no standing wave contamination and more accurate than probe-tube microphone measures (Lewis et al., 2009). IPL also benefits from its conceptual equivalence to eardrum SPL (Lewis et al., 2009; Scheperle et al., 2011),

$$|P_{IPL}| = |P_{FPL}| + |P_{RPL}| \quad (3.9)$$

P_{IPL} amplitudes can also be converted to dB SPL. These are referred to as IPL, which are the same as eardrum SPL.

3.2.3 Wideband acoustic immittance measurement

The pressure reflectance of the ear canal R_{ec} as calculated by equation 6 above is directly used in the derivation of the wideband acoustic immittance measurement of power absorbance (A) and power reflectance for middle ear diagnostics (\mathcal{R}). The power reflectance is the square of the pressure reflectance R_{ec} and has no phase angle. Power

reflectance is constant along the ear canal and is of magnitude between 0 and 1. The power reflectance therefore has a simple interpretation as the percentage of power reflected by the eardrum of the individual and provides an assessment of middle ear function (Rosowski & Wilber, 2015). The same measurement can be considered as the power absorbed using the term power absorbance (A). There is a direct, inverse relationship between power reflectance and power absorbance, with both values calculated from a measurement of the pressure reflectance (Neely et al., 2013):

$$A = 1 - |R_{ec}|^2 \text{ or } \mathcal{R} = |R_{ec}|^2 \quad (3.10)$$

The current study considers the effects of the source parameter calibration on WAI measurements through the effects on power reflectance, however, these could also be described using the power absorbance (A).

3.3 Methods

3.3.1 Calibration

Source parameter calibration was completed daily 30 consecutive days for the inter-session calibration condition. Each measurement was completed in a quiet room. The testing room was not climate controlled, and small variations in temperature and humidity were expected as the weather fluctuated. On day 31, an additional 31 calibration measurements were made for the intra-session calibration condition. During each calibration, the frequency-specific source parameter calculations (P_{src} and Z_{src}) were recorded. All measurements were made using a commercially available Interacoustics Titan probe (Interacoustics A/S, Middelfart, Denmark). Reported measurements were made using a custom-written MATLAB (version 2020a, MathWorks) script to control the Interacoustics Research Platform.

The current study used four commercially available calibration cavities (Interacoustics A/S, Middelfart, Denmark) with a constant radius of 0.2 cm and lengths of 1.2, 1.45, 1.75, and 2.0 cm. The radius was chosen to minimize the effect of non-propagating waves on the calibration procedure (Nørgaard et al., 2017) and cavity lengths were chosen to minimize the measurement error caused by overlapping impedance

extrema (Scheperle et al., 2011). The calibration procedure used in the current study was introduced by Nørgaard and colleagues in 2017, as reviewed above.

Calibration cavity pressure responses were measured using a wideband click (226Hz-8000Hz) stimulus with rate of 21 clicks per second calibrated to 96 dB peak-to-peak equivalent SPL as measured by a microphone (Bruel and Kjaer Type 4192) in an ear simulator (B&K Type 4157). For the calibration of the click, the Titan transducer was coupled flush to the ear simulator opening using an ER38-14A small foam insert. The B&K type 4192 microphone was connected to a Brüel and Kjær Nexus conditioning amplifier set to 316 milliVolts/Pascal. The amplifier was connected to a USBPre 2 external soundcard (Sound Devices, WI, United States) set to full-scale connected to SpectraPLUS software (Pioneer Hill Software, WA, United States). All measurements are referenced to a SpectraPLUS calibration file recorded using a Brüel and Kjær type 4231 calibrator which output a 94 dB SPL tone at 1 kHz. Source parameter calibration cavity measurements were completed using thirty-two 1024-sample blocks, which were averaged together to reduce measurement noise. The average waveforms were then analyzed using a Fourier transform with sampling frequency of 22050 Hz to determine the pressure response of each cavity.

3.3.2 Experimental measurement

Using the calibrated transducer, the pressure response in one adult male ear canal was completed. The participant had no occluding earwax and had normal middle-ear function, determined by falling within the normative wideband acoustic immittance region measured by an Interacoustics Titan transducer (Kenny, 2011). Experimental measurements were completed with the same click stimulus as described in the calibration measurements and waveforms were filtered into 1/3rd octave bands from 250 through 8000 Hz.

3.3.3 Analysis

Analysis was completed offline using a custom-written MATLAB script. All 61 (30 intersession condition; 31 intrasession condition) source parameter values were paired with a single pressure response of the ear to determine the sound level referenced in FPL,

IPL and the individual's power reflectance (equations 5-10). The characteristic impedance of the ear was calculated using an estimated radius of 4 mm. This radius was used to calculate the characteristic impedance for all subjects in many early FPL studies (Scheperle et al., 2011). The current study did not account for the individual's deviation from the average characteristic impedance, which has been shown to reduce error in FPL, IPL, and WAI measurement (Rasetshwane & Neely 2011; Scheperle et al., 2011). Although the individual's deviation from average characteristic impedance will change the absolute magnitude of IPL, FPL, and WAI, it will not significantly alter the relative changes in the measurements caused by source parameter differences. By using only one pressure response measurement in the ear, we isolated any changes in FPL, IPL, and WAI to differences in source parameter calibration rather than to the in-ear measurement itself.

Using R (R core team, 2020) and the lme4 package (Bates et al., 2015), we created multiple fixed-effect linear models to assess the relationship between the calibration and the resulting IPL, FPL, and WAI in a male adult ear. Frequency and day of calibration are entered into the model as fixed effects (with interaction). Inspection of residual plots did not indicate obvious deviations from normality or homoskedasticity. All reported p-values were obtained by likelihood ratio (LRT) tests of the full model with the effect of interest against the model without the effect included.

Intra-session reliability was determined by differences in pairs of consecutive calibrations done on the final day. Each of the 31 calibrations were compared to the calibration completed immediately prior and the test-retest difference was calculated across all frequencies. A comparison to a pre-determined cut-off of clinical significance was also completed.

3.3.4 Clinical significance of variation

Different tests have different requirements for calibration precision and measurement reliability. Across audiometry, hearing aid verification, or otoacoustic emission testing, the acceptable test-retest reliability differs across both test and frequency of interest. For example, in hearing aid verification, the probe-tube microphone needs to be placed near

the eardrum to prevent changes in measurement exceeding 2 dB and requires probe-tube microphone calibration accuracy within 3 dB (ANSI, S3.46-2013). Other stimulus level measurements also require high precision, with Class 1 sound level meters requiring precision within ± 1 dB. To avoid affecting any clinical decisions, the variation caused by source-parameter calibration should be less than the minimum acceptable variation of the tests that depend on it. Stimulus level measurement can require accuracy as precise as ± 1 dB, as outlined above. As a result, we applied a cut-off for clinically significant changes caused by source parameter calibration differences as any change greater than 1 dB

Similarly, changes in WAI measurements caused by test-retest differences in source parameter calibration should fall below the minimum intra-subject reliability of the measurement. Work on intra-subject reliability has shown standard deviations of over 2.5% across all frequencies, with decreasing reliability in the high frequency range (Abur et al., 2014; Rosowski et al., 2012). A conservative cut-off of 2.5% across all frequencies was used in this study. Any change greater than 2.5% at any frequency was considered to be a clinically significant difference and would indicate a meaningful change in value between measurements.

3.4 Results

3.4.1 Inter-session source parameter calibration reliability

Source parameter calibration resulted in a systematic, though not clinically relevant, deviation from baseline over a one-month period for IPL and FPL measurements (Figure 3-1, top and middle rows). Generally, the absolute magnitude of change increased with frequency. Frequencies at or above 4 kHz showed the largest change. The effect of source parameter deviation was statistically significant in IPL determination ($\chi^2_{16} = 51.65$, $p = 0.001$) and FPL determination ($\chi^2_{16} = 248.73$, $p < 0.001$). There was also a significant interaction of source parameter change with frequency for both IPL ($\chi^2_{15} = 33.51$, $p = 0.004$) and FPL ($\chi^2_{15} = 141.53$, $p < 0.001$), indicating that the effect depends on frequency. For IPL measurement, the largest changes occurred in the 5000 Hz frequency band, with a weekly change in IPL of 0.048 dB/week ($CI_{95} = 0.021-0.075$; Figure 3-2A). For FPL measurement, the largest change occurred in the 6300 Hz frequency band, with a

weekly change in FPL of 0.075 dB/week ($CI_{95} = 0.040-0.109$; Figure 3-2A). The deviation caused by source parameter calibration for ear canal stimulus level measurements was not clinically significant, falling far below the set cut-off of 1 dB at all tested frequencies.

When analyzing the power reflectance for wideband acoustic immittance (WAI) testing, a systematic source parameter change was also observed (Figure 3-2B, right panel). Given that power reflectance is measured on a linear, percentage-based scale rather than logarithmically using dB, the effect of the inter-session source parameter differences was relatively large when compared to the effect that the same difference had on the ear canal stimulus level measurement. The power reflectance measurement was significantly altered by the day that the calibration took place ($\chi^2_{16} = 383.00$, $p < 0.001$). As expected, the effect of inter-session source parameter differences also had a significant interaction with frequency ($\chi^2_{15} = 291.15$, $p < 0.001$). The largest effect was seen in the high frequencies, where power reflectance changed by as much as 2.0% ($CI_{95} = 1.54-2.46\%$) per week (Figure 3-2, rightmost panel). This effect of weekly source parameter differences did not exceed the cut-off for a clinically meaningful change at any frequency. However, calibrations completed more than a week apart introduced the potential for clinically significant systematic error, especially in the high frequency range.

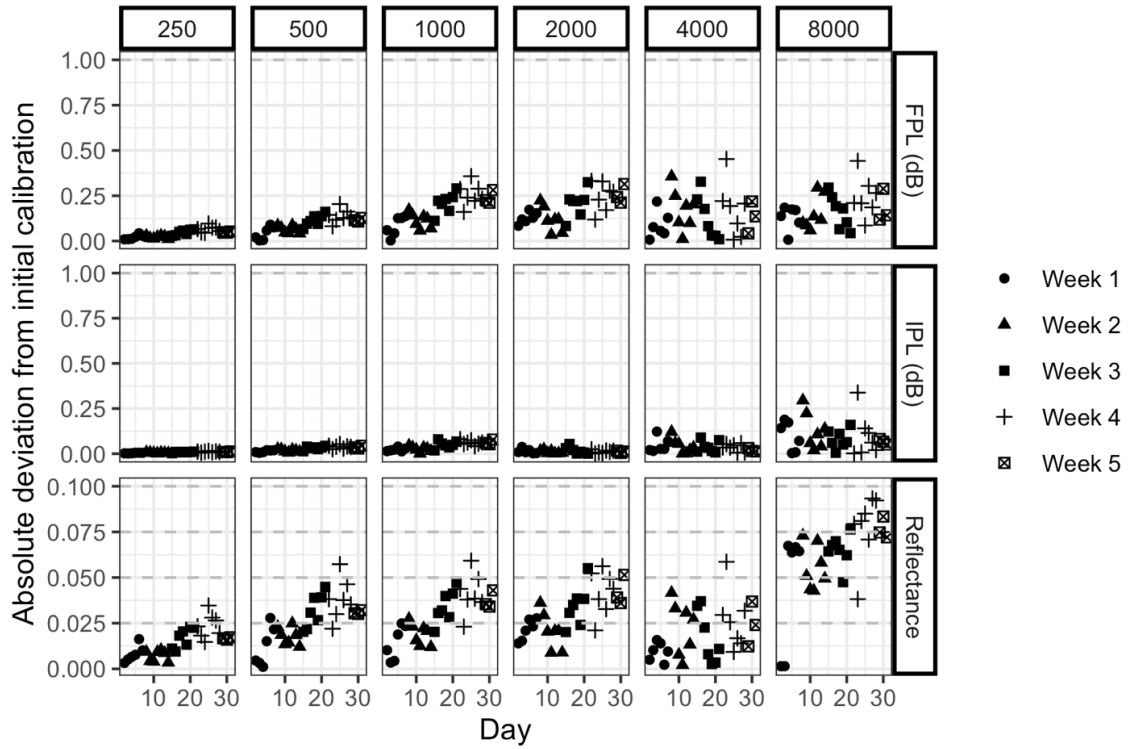


Figure 3-1: Absolute deviation in FPL, IPL, and power reflectance from first day results at all audiometric frequencies. A clinically meaningful difference is indicated by a dotted line at every 1 dB for FPL and IPL and every 0.025 unit (2.5%) change for power reflectance.

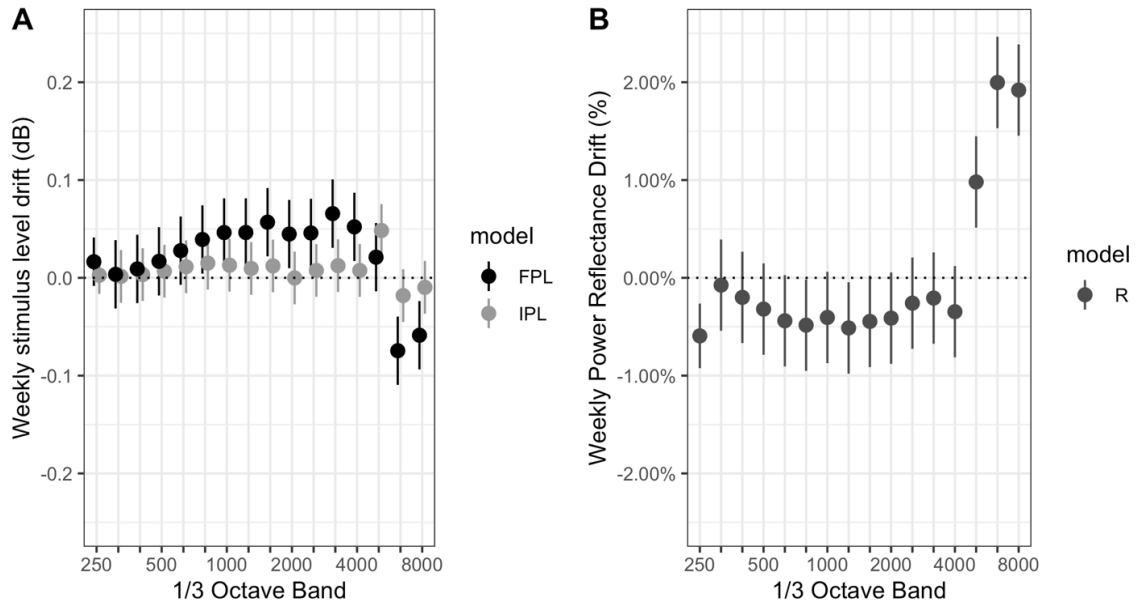


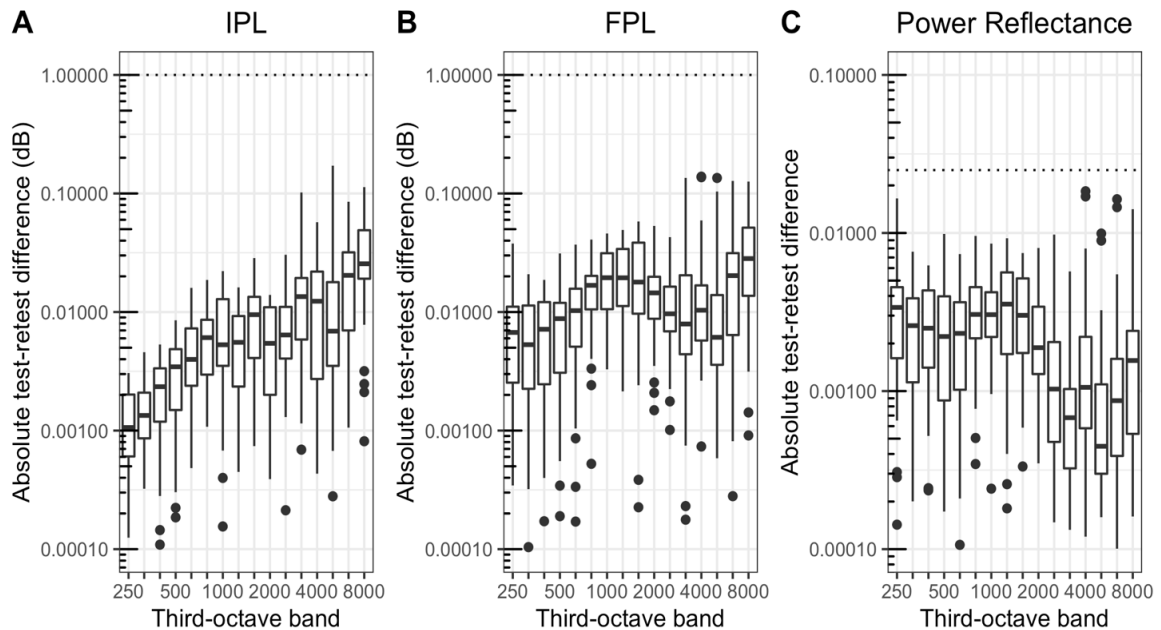
Figure 3-2: Weekly change for in-ear stimulus measurement (A) and power reflectance (B) caused by source parameter calibration differences. Points indicate the regression coefficient for the weekly change, with 95% confidence intervals ($\pm 2 \cdot SE$) visualized.

3.4.2 Cross-sectional source parameter calibration reliability

Intra-session test-retest reliability of source parameter calibration yielded no clinically significant differences for any calibration trial at any frequency across all ear canal stimulus measurements (Figure 3-3, Panel A and B). A general trend of decreasing reliability with increasing frequency occurred for both IPL and FPL measurements, however, even the most extreme outlier across both FPL and IPL measurements had a test-retest difference magnitude of less than 0.2 dB.

Similarly, the high source parameter calibration reliability resulted in no clinically significant variation in power reflectance measurements (Panel C of Figure 3-3). Unexpectedly, there was no clear decrease in test-retest reliability with increasing frequency. The median test-retest difference was numerically lower in many of the high-frequency bands.

With no individual test-retest difference falling above the clinically significant cut-off, it is unlikely that intra-session source parameter recalibration will alter wideband acoustic immittance or ear canal stimulus-level measurements in any meaningful way.



*Figure 3-3: Distribution of absolute values for intra-session test-retest differences in IPL (A), FPL (B), and power reflectance (C) calculations caused by source parameter calibration differences. Values were averaged across 1/3 octave bands. The transducer was re-inserted between each calibration procedure. The clinically significant cut-off is denoted by a dotted line at 1 dB (A, B) and 2.5% (C). The boxes indicate interquartile range which is the range between 25-75th percentile values with the median marked by a horizontal line. The error-bars extend to $1.5 * IQR$ and individual outliers are represented by single points.*

3.5 Discussion/Conclusion

This work provides preliminary information about the impact of time on the reliability of IPL and/or FPL calibration. Overall, the results suggest that weekly calibration provides sufficient reliability across all three clinical measurements analyzed while minimizing the need for repeated calibration of source parameters. Although there is a statistically significant change in source parameters over time, the recommended time interval

necessary for recalibration may vary with the purpose and type of measurement. If only stimulus level determination using FPL or IPL is required, the inter-session differences are small enough that calibrating monthly may not produce clinically meaningful differences in measured ear canal stimulus levels. A significant benefit of using a source parameter-based ear canal stimulus level measurement (i.e. IPL, FPL) is the simultaneous middle-ear assessment. Given the comparatively large, albeit clinically small, effect of source parameter changes in high frequencies of WAI measurement, the results indicate that if WAI measurement is expected to be completed, weekly calibration may be sufficient to ensure consistent results over time.

A limitation of the current study is that it was completed within a single ear using the same equipment over time and completed with the same operator. The generalizability of these results to other equipment, to multiple instances of the same model of equipment, and across operators and patients could be investigated in future.

3.6 References

- Abur, D., Horton, N. J., & Voss, S. E. (2014). Intrasubject Variability in Power Reflectance. *Journal of the American Academy of Audiology*, 25(5), 441–448. <https://doi.org/10.3766/jaaa.25.5.3>
- Allen, J. (1986). Measurement of Eardrum Acoustic Impedance. *Peripheral Auditory Mechanisms*, 44–51.
- American National Standards Institute (ANSI). (2013). Methods of Measurement of Real-Ear Performance Characteristics of Hearing Aids. ANSI S3.46-2013(R2018). New York, NY: Acoustical Society of America.
- Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01.
- Beers, A. N., Shahnaz, N., Westerberg, B. D., & Kozak, F. K. (2010). Wideband reflectance in normal school-aged children and in children with otitis media. *Ear & Hearing*, 31, 221–233.
- Boothalingam, S., Klyn, N. A. M., Stiepan, S. M., Wilson, U. S., Lee, J., Siegel, J. H., & Dhar, S. (2018). Revisiting gender, race, and ear differences in peripheral auditory function. *AIP Conference Proceedings*, 1965. <https://doi.org/10.1063/1.5038496>
- Burke, S. R., Rogers, A. R., Neely, S. T., Kopun, J. G., Tan, H., & Gorga, M. P. (2010). Influence of calibration method on distortion-product otoacoustic emission measurements: I. Test performance. *Ear and Hearing*, 31(4), 533–545. <https://doi.org/10.1097/AUD.0b013e3181d86b3d> [doi]
- Chan, J. C. K., & Geisler, C. D. (1990). Estimation of eardrum acoustic pressure and of ear canal length from remote points in the canal. *The Journal of the Acoustical Society of America*, 87(3), 1237–1247. <https://doi.org/10.1121/1.398799>
- Farmer-Fedor, B. L., & Rabbitt, R. D. (2002). Acoustic intensity, impedance and reflection coefficient in the human ear canal. *The Journal of the Acoustical Society of America*, 112(2), 600–620. <https://doi.org/10.1121/1.1494445>
- Feeney, M. P., Grant, I. L., & Mills, D. M. (2009). Wideband energy reflectance measurements of ossicular chain discontinuity and repair in human temporal bone. *Ear and Hearing*, 30(4), 391–400. <https://doi.org/10.1097/AUD.0b013e3181a283ed>
- Gilman, S., & Dirks, D. D. (1986). Acoustics of ear canal measurement of eardrum SPL in simulators. *The Journal of the Acoustical Society of America*, 80(3), 783–793. <https://doi.org/10.1121/1.393953>
- Huang, G. T., Rosowski, J. J., Puria, S., & Peake, W. T. (2000). A noninvasive method for estimating acoustic admittance at the tympanic membrane. *The Journal of the Acoustical Society of America*, 108(3 Pt 1), 1128–1146. <https://doi.org/10.1121/1.1287024>

- Hunter, L. L., Feeney, M. P., Lapsley Miller, J. A., Jeng, P. S., & Bohning, S. (2010). Wideband reflectance in newborns: Normative regions and relationship to hearing-screening results. *Ear and Hearing, 31*(5), 599–610. <https://doi.org/10.1097/AUD.0b013e3181e40ca7>
- Keefe, D. H. (1984). Acoustical wave propagation in cylindrical ducts: Transmission line parameter approximations for isothermal and nonisothermal boundary conditions. *The Journal of the Acoustical Society of America, 75*(1), 58–62. <https://doi.org/10.1121/1.390300>
- Keefe, D. H., Bulen, J. C., Arehart, K. H., & Burns, E. M. (1993). Ear-canal impedance and reflection coefficient in human infants and adults. *The Journal of the Acoustical Society of America, 94*(5), 2617–2638. <https://doi.org/10.1121/1.407347>
- Keefe, D. H., Folsom, R. C., Gorga, M. P., Vohr, B. R., Bulen, J. C., & Norton, S. J. (2000). Identification of Neonatal Hearing Impairment: Ear canal measurements of acoustic admittance and reflectance in neonates. *Ear and Hearing, 21*(5), 443–461. <https://doi.org/10.1097/00003446-200010000-00009>
- Keefe, D. H., Ling, R., & Bulen, J. C. (1992). Method to measure acoustic impedance and reflection coefficient. *The Journal of the Acoustical Society of America, 91*(1), 470. <https://doi.org/10.1121/1.402733>
- Keefe, D. H., & Schairer, K. S. (2011). Specification of absorbed-sound power in the ear canal: Application to suppression of stimulus frequency otoacoustic emissions. *The Journal of the Acoustical Society of America, 129*(2), 779–791. <https://doi.org/10.1121/1.3531796>
- Kenny, S. (2011). Clinical application of the Interacoustics REFLWIN system wideband reflectance machine in the assessment of the eustachian tube. (Master's Thesis). University of British Columbia.
- Lapsley Miller, J. A., Reed, C. M., Robinson, S. R., & Perez, Z. D. (2018). Pure-Tone Audiometry With Forward Pressure Level Calibration Leads to Clinically-Relevant Improvements in Test-Retest Reliability. *Ear and Hearing, 39*, 946-957. <https://doi.org/10.1097/AUD.0000000000000555>
- Lewis, J. D., McCreery, R. W., Neely, S. T., & Stelmachowicz, P. G. (2009). Comparison of in-situ calibration methods for quantifying input to the middle ear. *The Journal of the Acoustical Society of America, 126*(June 2009), 3114–3124. <https://doi.org/10.1121/1.3243310>
- MATLAB 2020a, The MathWorks, Inc., Natick, Massachusetts, United States.
- McCreery, R. W., Pittman, A., Lewis, J., Neely, S. T., & Stelmachowicz, P. G. (2009). Use of forward pressure level to minimize the influence of acoustic standing waves during probe-microphone hearing-aid verification. *The Journal of the Acoustical Society of America, 126*(1), 15–24. <https://doi.org/10.1121/1.3143142>
- Mueller, H. G. (2001). Probe microphone measurements: 20 Years of progress. *Trends in Amplification, 5*(2), 35–68. <https://doi.org/10.1177/108471380100500202>

- Nakajima, H. H., Rosowski, J. J., Shahnaz, N., & Voss, S. E. (2013). Assessment of Ear Disorders Using Power Reflectance. *Ear and Hearing, 34*(SUPPL. 1), 48S–53S.
- Neely, S. T., Stenfelt, S., & Schairer, K. S. (2013). Alternative ear canal measures related to absorbance. *Ear and Hearing, 34*(SUPPL. 1), 1–15.
<https://doi.org/10.1097/AUD.0b013e31829c7229>
- Nørgaard, K. R., Fernandez-Grande, E., & Laugesen, S. (2017). Incorporating evanescent modes and flow losses into reference impedances in acoustic Thévenin calibration. *The Journal of the Acoustical Society of America, 142*(5), 3013–3024.
<https://doi.org/10.1121/1.5010891>
- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Rasetshwane, D. M., & Neely, S. T. (2011). Inverse solution of ear canal area function from reflectance. *The Journal of the Acoustical Society of America, 130*(6), 3873–3881. <https://doi.org/10.1121/1.3654019>
- Richmond, S. A., Kopun, J. G., Neely, S. T., Tan, H., & Gorga, M. P. (2011). Distribution of standing-wave errors in real-ear sound level measurements. *The Journal of the Acoustical Society of America, 129*(5), 3134–3140.
<https://doi.org/10.1121/1.3569726>
- Rogers, A. R., Burke, S. R., Kopun, J. G., Tan, H., Neely, S. T., & Gorga, M. P. (2010). Influence of calibration method on distortion-product otoacoustic emission measurements: II. Threshold prediction. *Ear and Hearing, 31*(4), 546–554.
<https://doi.org/10.1097/AUD.0b013e3181d86b59>
- Rosowski, J. J., & Wilber, L. A. (2015). Acoustic immittance, absorbance, and reflectance in the human ear canal. *Seminars in Hearing, 36*(1), 11–28.
<https://doi.org/10.1055/s-0034-1396924>
- Rosowski, J. J., Nakajima, H. H., Hamade, M. A., Mahfoud, L., Merchant, G. R., Halpin, C. F., & Merchant, S. N. (2012). Ear canal reflectance, umbo velocity, and tympanometry in normal-hearing adults. *Ear and Hearing, 33*(1), 19–34.
<https://doi.org/10.1097/AUD.0b013e31822ccb76>
- Saunders, G. H., & Morgan, D. E. (2003). Impact on hearing aid targets of measuring thresholds in dB HL versus dB SPL. *International Journal of Audiology, 42*(6), 319–326. <https://doi.org/10.3109/14992020309101324>
- Scheperle, R. A., Goodman, S. S., & Neely, S. T. (2011). Further assessment of forward pressure level for in situ calibration. *The Journal of the Acoustical Society of America, 130*(6), 3882–3892. <https://doi.org/10.1121/1.3655878>
- Scheperle, R. A., Neely, S. T., Kopun, J. G., & Gorga, M. P. (2008). Influence of in situ , sound level calibration on distortion-product otoacoustic emission variability. *The Journal of the Acoustical Society of America, 124*(1), 288–300.
<https://doi.org/10.1121/1.2931953>

- Shahnaz, N., Bork, K., Polka, L., Longridge, N., Bell, D., & Westerberg, B. D. (2009). Energy reflectance and tympanometry in normal and otosclerotic ears. *Ear and Hearing, 30*(2), 219–233. <https://doi.org/10.1097/AUD.0b013e3181976a14>
- Siegel, J. H. (1994). Ear canal standing waves and high frequency sound calibration using otoacoustic emission probes. *The Journal of the Acoustical Society of America, 95*(5), 2589–2597. <https://doi.org/10.1121/1.409829>
- Siegel, J. H. (2007). Calibration of Otoacoustic Emission Probes: Clinical Applications. In M. S. Robinete, & T. J. Glatcke (Eds.), *Otoacoustic Emissions: Clinical Applications* (3rd ed.). Thieme.
- Souza, N. N., Dhar, S., Neely, S. T., & Siegel, J. H. (2014). Comparison of nine methods to estimate ear canal stimulus levels. *The Journal of the Acoustical Society of America, 136*(4), 1768–1787. <https://doi.org/10.1121/1.4894787>
- Stinson, M. R., & Lawton, B. W. (1989). Specification of the geometry of the human ear canal for the prediction of sound-pressure level distribution. *The Journal of the Acoustical Society of America, 85*(6), 2492–2503. <https://doi.org/10.1121/1.397744>
- Vaisberg, J. M., Folkeard, P., Pumford, J., Narten, P., & Scollie, S. (2018). Evaluation of the Repeatability and Accuracy of the Wideband Real-Ear-to-Coupler Difference. *Journal of the American Academy of Audiology, 29*(6), 520–532. <https://doi.org/10.3766/jaaa.17007>
- Vaisberg, J. M., Macpherson, E. A., & Scollie, S. D. (2016). Extended bandwidth real-ear measurement accuracy and repeatability to 10 kHz. *International Journal of Audiology, 55*(10), 580–586. <https://doi.org/10.1080/14992027.2016.1197427>
- Valente, M., Potts, L. G., Valente, L. M. (1997). Differences and intersubject variability of loudness discomfort levels measured in sound-pressure level and hearing level for TDH-50P and ER-3A earphones. *Journal of the American Academy of Audiology, 8*(1), 59-67.
- Valente, M., Potts, L. G., Vass, W., & Goebel, J. (1994). Intersubject variability of real-ear sound pressure level: conventional and insert earphones. *Journal of the American Academy of Audiology, 5*(6), 390–398. <http://www.ncbi.nlm.nih.gov/pubmed/7858300>
- Voss, S. E., & Allen, J. B. (1994). Measurement of acoustic impedance and reflectance in the human ear canal. *The Journal of the Acoustical Society of America, 95*(1), 372–384. <https://doi.org/10.1121/1.408329>
- Voss, S. E., & Herrmann, B. S. (2005). How does the sound pressure generated by circumaural, supra-aural, and insert earphones differ for adult and infant ears? *Ear and Hearing, 26*(6), 636–650. <https://doi.org/10.1097/01.aud.0000189717.83661.57>
- Voss, S. E., Merchant, G. R., & Horton, N. J. (2012). Effects of Middle-Ear Disorders on Power Reflectance Measured in Cadaveric Ear Canals. *Ear and Hearing, 33*, 195–208.

- Voss, S. E., Rosowski, J. J., Shera, C. A., & Peake, W. T. (2000). Acoustic mechanisms that determine the ear canal sound pressures generated by earphones. *The Journal of the Acoustical Society of America*, *107*(3), 1548–1565. <https://doi.org/10.1121/1.428440>
- Voss, S. E., Stenfelt, S., Neely, S. T., & Rosowski, J. J. (2013). Factors that introduce intrasubject variability into ear canal absorbance measurements. *Ear and Hearing*, *34*(SUPPL. 1), 60S-64S.
- Withnell, R. H., Jeng, P. S., Waldvogel, K., Morgenstein, K., & Allen, J. B. (2009). An in situ calibration for hearing thresholds. *The Journal of the Acoustical Society of America*, *125*(3), 1605–1611. <https://doi.org/10.1121/1.3075551>

Chapter 4

4 Measurement of hearing aid tubing length using a wideband stimulus³

4.1 Introduction

Measurement of the sound level produced in an individual's ear is a crucial step in hearing aid verification. The acoustic variability of ear canals can transform a signal such that the overall sound-level near the eardrum varies by 20 dB between individuals and occasionally introducing differences as large as 40 dB (Saunders & Morgan, 2003). To account for this variability, probe-tube microphones are placed near the eardrum (within 4 mm; Vaisberg et al., 2016) to measure the sound-pressure level (SPL) in the ear canal during hearing aid fittings. The specific test signals and measurement formats vary (American National Standards Institute [ANSI], ANSI S3.46-2013). In pediatric hearing aid fittings, one common measurement is the real-ear to coupler difference (RECD; Moodie et al., 2016b). Briefly stated, the RECD is the difference between the acoustic response of the individual's ear and a reference HA-1 2.0 cc coupler when both are measured with the same stimulus, transducer, and coupling method (ANSI, S3.46-2013). The RECD is used as an acoustic transform, both for the audiometric data referenced to dB hearing level (HL) and for transforming hearing aid responses to the SPL in the ear canal (Bagatto et al., 2005). The RECD measurement procedure has been described in detail in elsewhere (American Academy of Audiology, 2013; Bagatto et al., 2005; Moodie et al., 1994; Munro & Lazenby, 2001; Scollie et al., 1998). Following RECD measurement, the rest of the fitting process can be completed in a hearing aid coupler without patient cooperation, which is especially advantageous when working with infants.

³ A version of this chapter has been submitted: Urichuk, M., Purcell, D., Scollie, S. Measurement of hearing aid tubing length using a wideband stimulus. Submitted to *The Journal of the Acoustical Society of America*.

In practice, however, the real-ear portion of the RECD is often measured with the acoustic transducer coupled to a segment of hearing aid tubing attached to the individual's custom earmold placed in the ear canal. The length of the tubing segment is typically 30-50 mm, chosen to suit the individual's anatomy. Such a measurement is formally defined as an ear canal level difference (ECLD; ANSI, S3.46-2013), although this terminology is rarely used in clinical equipment. The current paper will refer to ECLD using the term more common in audiology, "earmold RECD". Similarly, "foam-tip RECD" refers to RECD measurements made with a standard disposable Etymotic foam tip, which attaches to the acoustic transducer using the foam-tip's own integrated 25 mm long segment of tubing. The choice of coupling method directly impacts the SPL produced by the transducer (Egolf et al., 1985; Gilman et al., 1981; Munro & Davis, 2003; Munro & Salisbury, 2002; Sanborn, 1998). If different coupling methods are used for hearing threshold measurement (e.g., transducer to foam-tip) and RECD measurement (e.g., transducer to custom hearing aid tubing and earmold), the coupling mismatch introduces errors exceeding 10 dB (American Academy of Audiology, 2013; Bagatto et al., 2005; Moodie et al., 2016a; Munro & Hatton, 2000; Munro & Salisbury, 2002). The coupling-mismatch is caused by differences in impedance between coupling methods, which is mainly due to differences in tubing lengths: 25 mm integrated into the foam-tip versus a custom length for hearing aid tubing to an individual's earmold (Bagatto et al., 2005; Gustafson et al., 2013; Moodie et al., 2016a; Munro & Salisbury, 2002). More specifically, the difference between the earmold RECD and foam-tip RECD is directly impacted by the interaction between the impedance of the tubing segment, which varies depending on the coupling method used, and the impedance of the transducer, which is constant. The impedance of the transducer (i.e., RE-770, ER-3A) itself is consistent regardless of the whether the earmold RECD or foam-tip RECD is being measured. While the RECD has shown some dependence on sound-source impedance differences between transducers, such effects are small in comparison to the dependence on the tubing length of the coupling method (Munro & Salisbury 2002). The portion of the coupling mismatch discussed in the present study is caused by the custom length of hearing aid tubing associated with an individual's earmold. The approach proposed here improves upon the use of an assumed average hearing aid tubing length. Prediction and

correction of the error caused by this coupling-mismatch is of particular interest in clinical cases for which the foam-tip RECD is measured, but an earmold RECD is necessary or vice versa. Such situations can arise when the earmold is not immediately available during the hearing aid fitting or the foam-tip RECD cannot be directly measured in an individual due to discomfort with the procedure.

One method proposed to minimize the error introduced by a coupling mismatch is a foam-tip-to-earmold transform (Moodie et al., 2016a). The foam-tip-to-earmold transform is available in some hearing aid analyzers and is added to the foam-tip RECD values in a frequency-specific manner. The frequency-specific corrections were determined using average differences between earmold and foam-tip RECDs of an HA-1 coupler and two simulated ears created to approximate the volume and impedance of an adult and infant, respectively (Moodie et al., 2016a). The corrections were validated in a sample of infants and adults with tubing varying between 22 and 59 mm (Moodie et al., 2016a). This correction has been shown to more accurately predict earmold RECD values, especially above 1 kHz, when compared to age specific RECD averages (Moodie et al., 2016a). However, the foam-tip-to-earmold transform is the average across hearing aid tubing lengths between 30 and 50 mm. As a result, this transform does not account for individual variation in hearing aid tubing length. The current study proposes and validates a rapid method for estimating the hearing-aid tubing length from the half-wavelength resonance associated with this tubing for more accurate and individual foam-tip-to-earmold transforms.

4.1.1 Sources of variability

The variability in sound pressure level associated with coupling to the ear is largely caused by impedance differences due to the length of tubing between the transducer and earmold (Gustafson et al., 2013; Moodie et al., 2016a; Munro & Salisbury, 2002). The choice of hearing aid tubing length has a frequency-specific impact on the sound-level measured in the individual's ear (Moodie et al., 2016a). The largest effects of tubing length are seen in frequencies near the first impedance minima and maxima, which both are typically present in the mid-frequencies (approximately 1-4 kHz; Moodie et al., 2016). Infant earmolds typically have shorter tubing lengths than adult earmolds due to

the smaller size of their external ears. Consequently, an average foam-tip-to-earmold transform that is reasonably accurate for the average adult may overcorrect for the acoustic differences between earmold and foam tip coupling methods if used on an infant with short tubing lengths. Similarly, the average foam-tip-to-earmold transform also systematically under-corrects the acoustic differences when the hearing aid tubing segment is longer than the tubing length of an average adult ear. Accurate physical measurement of the individual's tubing length is challenging in a clinical setting due to the curvature of the tubing, infection control considerations, and embedding of the the tubing within the individual's earmold. If individualized hearing-aid tubing length measurements were implemented into a clinical instrument, the validity of the foam-tip-to-earmold correction may improve and increase accuracy of hearing aid fittings to prescriptive targets.

4.1.2 Potential solution: reflectance-based tubing length estimation

One potential solution is the estimation of hearing-aid tubing length with a Thevenin-equivalent source-parameter calibrated transducer. Source-parameter calibrated transducers can accurately assess the impedance of an acoustic load, which can include the source transducer, coupling method, and the middle-ear system using a wideband acoustic immittance measurement (Allen, 1986; Hunter et al., 2010; Keefe et al., 1993; Voss & Allen, 1994). The transducer houses a microphone flush with the sound-source to accurately determine the effects that the transducer has on the stimulus in acoustic loads of differing immittance (Huang et al., 2000; Souza et al., 2014; Voss & Allen, 1994; Withnell et al., 2009). Source-parameter calibration enables accurate in-situ determination of the pressure reflectance coefficient of the acoustic load (Keefe et al., 1993). Using the pressure reflectance coefficient, the total sound-level can be decomposed into the forward-moving and reflected components (Keefe et al., 1993; Lewis et al., 2009; Scheperle et al., 2008, 2011; Souza et al., 2014). Implementation of in-situ calibration with forward-pressure level (FPL) has been of interest as an alternative calibration approach for hearing aid verification (Lewis et al., 2009; McCreery et al., 2009), otoacoustic emissions (Charaziak & Shera, 2017; Scheperle et al., 2008, 2011), and other audiological measurements (Lapsley Miller et al., 2018; Souza et al., 2014;

Withnell et al., 2014). However, such a calibration procedure has not been applied to estimate the length of an unknown open tube previously. The current study attempts to apply in-situ calibration with FPL for tube length estimation.

4.1.3 Reflectance-based methods to estimate tubing length

The pressure reflectance coefficient can be used to acoustically measure the length of an unknown segment of hearing aid tubing using an arbitrary wideband stimulus. The first step to determining the length of a hearing aid (open) tube segment is measuring the frequency-location of the half-wavelength resonance (F_{res}). The F_{res} can be determined by the frequency location where the pressure reflectance coefficient is closest to -1, which corresponds to complete negative reflection (with inverted phase relative to the incident wave). Alternatively, and perhaps more intuitively, the F_{res} can also be determined by minimizing the difference between the SPL at the entrance of the tube (SPL measured by the transducer) and the SPL at the termination of the tube-segment. The SPL at the termination of a closed tube is theoretically equivalent to the integrated pressure level (IPL), which is the summed magnitude of the forward-moving and reflected sound-waves (Lewis et al., 2009; Scheperle et al., 2011; Withnell et al., 2009, 2014). Mathematically, the frequency location is described as:

$$F_{res} = \arg \min(SPL - IPL) \quad (4.1)$$

Once the frequency of the half-wavelength resonance is determined, the length of an ideal cylindrical tube can be determined:

$$L_{ideal} = \frac{c}{2F_{res}} \quad (4.2)$$

Where c is the speed of sound ($\sim 343\text{m/s}$) and L_{ideal} is the length of an ideal cylindrical tube. For physical tubes of non-zero diameter, the pressure null occurs just outside the mouth of the tube, requiring an open-end physical length correction of (Levine & Schwinger, 1948):

$$\Delta L_{OE} = 1.2266 * r_{tube} \quad (4.3)$$

Where r_{tube} is the internal radius of the hearing aid tube segment. For this study, the radius is 0.96 mm using #13 hard walled hearing aid tubing, yielding the ΔL value used here of 1.184 mm. With the open-end correction adjustments, the physical length equation is:

$$\hat{L} = \frac{c}{2F_{res}} - \Delta L_{OE} = L_{ideal} - \Delta L_{OE} \quad (4.4)$$

Where \hat{L} is the estimated length of the tubing being measured.

4.1.4 Applications of an estimate of the tubing length of earmolds

Determining the length of the individual's earmold tubing using an accurate, rapid acoustic measurement could allow for personal foam-tip-to-earmold corrections based on the actual length of hearing aid tubing used by the individual. This individualized measurement will improve the accuracy of fitting whenever a foam-tip-to-earmold correction is used. A method to accomplish this using a source-parameter calibrated transducer is proposed and validated in the current study. Clinical implications for individualized foam-tip-to-earmold RECD corrections are discussed.

4.2 Methods

4.2.1 Instrumentation

All acoustic length measurements were made using a commercially available Interacoustics Titan probe (Interacoustics A/S, Middelfart, Denmark). Measurements were made using a custom-written Matlab script and the Interacoustics Research Platform (Matlab, 2020a). The individualized transforms were determined using an RE-770 transducer inserted into couplers manufactured by Audioscan (a division of Etymotic; Dorchester, Ontario). Couplers included an HA-1 coupler, a simulated adult ear canal, and a simulated infant ear canal. Simulated ear canals, designed to approximate the volume and impedance of an adult and infant ear, respectively, were created and used to determine the average foam-tip-to-earmold transform (Moodie et al., 2016). Details of simulated ear canal construction can be found in the study completed by Moodie and colleagues (2016).

4.2.2 Calibration

Source parameter calibration was completed on the same day as the hearing aid tubing length measurements in a quiet, climate-controlled room. Four calibration cavities of constant radius (0.2 cm) and length (1.2, 1.45, 1.75, 2.0 cm) were used for source parameter determination. The radii and lengths of these tubes were chosen to minimize the effects of non-propagating waves and overlapping impedance extrema on measurement accuracy (Nørgaard et al., 2017; Scheperle et al., 2011). Calibration cavity pressure responses were measured using a transient wideband stimulus (226-8000 Hz; repetition rate 21 Hz) calibrated to 96 dB peak-to-peak equivalent SPL as measured by a microphone (Bruel and Kjaer Type 4192) in an ear simulator (B&K Type 4157). The response in each calibration cavity was computed using thirty-two 1024-sample blocks, averaged together to reduce measurement noise. The average waveforms were analyzed using a Fourier transform with a sampling frequency of 22050 Hz.

4.2.3 Experimental measurement

The calibrated transducer was coupled to hearing aid tubing segments comprised of #13-gauge hard walled tubing of various lengths. Hearing aid tubing lengths were created for every 5 mm interval between 25-55 mm (inclusive). These values were chosen because most clinical hearing aid tubing lengths, as well as standard Etymotic foam tip tubing, are expected to fall within this range (Moodie et al., 2016a). The transducer and tested hearing aid tubing segments have similar outer diameters of 0.130 inches (3.3 mm). The transducer was coupled to each tube segment using a soft, silicone coupling sleeve with length 17.5 mm, inner diameter 3.1 mm, and outer diameter of 15.8 mm. The coupling sleeve inner diameter size was chosen to allow for a tight mechanical coupling that minimizes leakage during the measurement. The inner diameter of the hearing aid tubing is 0.076 inches (1.93 mm), however, the exact inner diameter of the transducer could not be determined. All acoustic measurements of the hearing aid tubing segment were completed with the same transient wideband stimulus used to calibrate the transducer's source-parameters. The Fourier transform had a frequency resolution of 21.5 Hz per bin. Fourier transform output was converted to dB SPL re 20 μ Pa root-mean-square (RMS) values.

Measurements were completed in three sessions, each in a quiet room. The transducer and hearing aid tubing were re-inserted into the coupler for each measurement to determine acoustic measurement test-retest reliability. To minimize the potential impact of low-frequency environmental noise contaminating the measurement, only frequencies above 500 Hz were analyzed for the presence of a resonance. Acoustic measurements were made for tubing lengths between 30 and 50 mm in increments of 5 mm.

The sound levels referenced in FPL, SPL, and IPL were calculated using a MATLAB script. The characteristic impedance of the hearing-aid tubing was calculated using an inner diameter constant of 1.93 mm. All visualizations and comparisons to a pre-determined 2 dB cut-off of clinical significance (ANSI, S3.46-2013) were completed using the R Statistical Programming Environment (R Core Team, 2020).

4.2.4 Individualized transform

An RE-770 transducer was attached to a foam-tip and inserted in three different cavities: (1) an HA-1 coupler, (2) simulated adult ear canal, and (3) simulated infant ear canal. The sound level output was measured, filtered into 1/3 octave bands, and the average response produced in each cavity was recorded. The same cavities used for the foam-tip were also employed with an earmold and tubing of lengths 30, 35, 40, 45, and 50 mm (Moodie et al., 2016a). The average sound level was measured in each cavity the same way as with the foam-tip. The measurements in the three cavities were averaged together for each coupling condition to attempt to minimize effects of cavity volume, cavity volume, and measurement error. Individualized transforms were computed as the difference between the average sound level measured for a particular length of hearing aid tubing and the foam-tip measurement. For lengths between measured values, transforms were linearly interpolated to provide a specific foam-tip-to-earmold correction value at all frequencies and for all lengths between 30 and 50 mm, referenced to the average foam-tip response across cavities.

4.3 Results

4.3.1 Acoustic tubing length measurement

The physically measured tubing length was compared to the estimated tubing length from the acoustic measurements (Figure 4-1A). There was excellent agreement between the physical hearing aid tubing length and the estimated acoustic length. Observed errors were less than 1 mm in all conditions and trials with an intraclass correlation coefficient of 0.99 (Figure 4-1B). Maximum test-retest variation between trials, including hearing aid tubing re-insertion and additional source-parameter calibration, is denoted by differences in bar height in Figure 4-1B and did not exceed 1 mm. These lengths were used to derive individualized foam-tip-to-earmold RECD corrections, below.

The variability of the acoustic estimation of tubing length was also evaluated by the average error in estimating the frequency of the half-wavelength resonance (Table 4-1). The estimated frequency of the resonant wave (calculated by equation 4, solving for F_{res}) fell within the 250 to 5500 Hz range for all tubing lengths analyzed in this study, as noted in Table 4-1. Generally, the average frequency error was inversely proportional to the length of the tubing. In other words, the shorter the tubing segment, the larger the error in Hertz. As previously noted, all length estimates were within 1 mm of the physical tubing length, even with the differences observed between measured and analytic resonant frequency for shorter tubing (Table 4-1). The differences in the acoustic transform can be visualized using both the difference between IPL and SPL across frequencies (Figure 4-2A) and the magnitude of the pressure reflectance (Figure 4-2B).

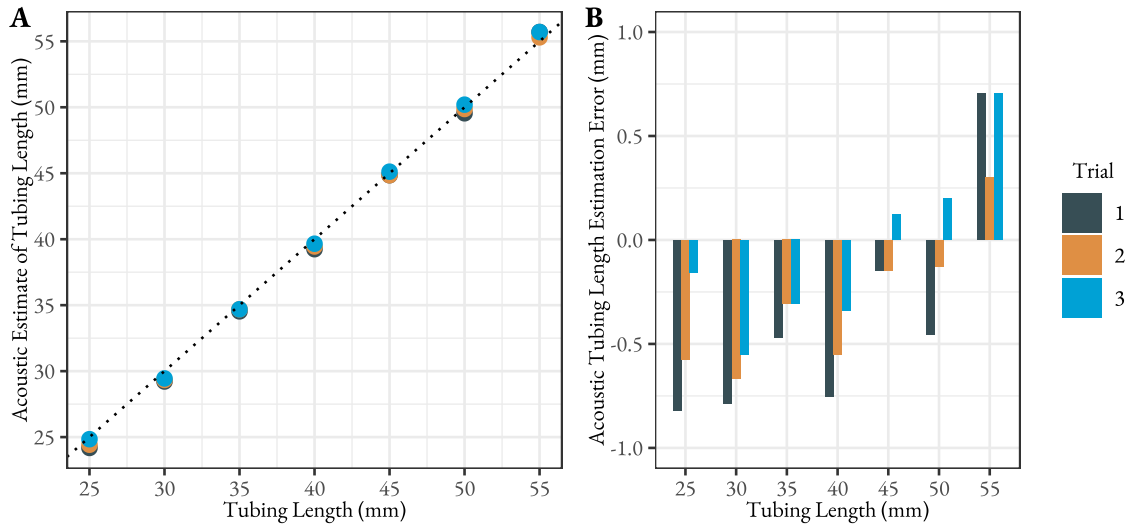


Figure 4-1: (A) Comparison of estimated tubing lengths (y-axis) and the actual physically measured tubing lengths. Points falling on the dotted line indicate perfect agreement. Points above the line indicate overestimation of tubing lengths while points falling below the dotted line indicate underestimation of the tubing length. (B) Difference between estimated tubing length and the physical tubing lengths. Values below 0 indicate underestimation of tubing length, whereas positive values indicate overestimation. All three trials are shown.

Tubing Length (mm)	Resonant Frequency (Hz)	Acoustic Resonance Estimate (Hz)	Average Error (Hz)
25	6550	6682	-132
30	5500	5620	-120
35	4740	4788	-48
40	4164	4221	-57
45	3713	3718	-5
50	3351	3359	-8
55	3052	3022	30

Table 4-1: Resonance measurement accuracy across different tubing length conditions.

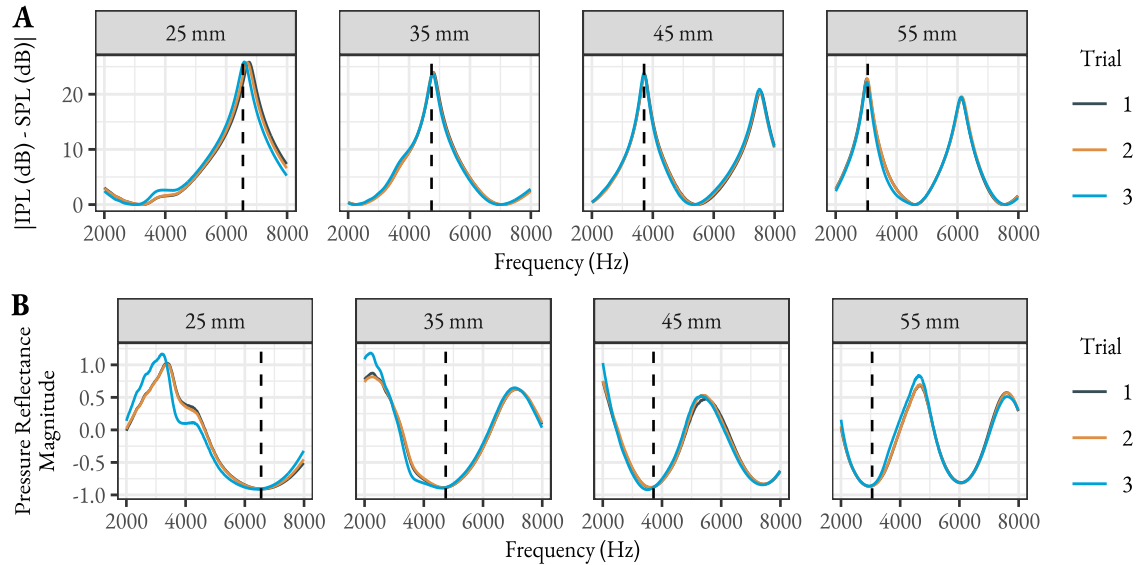


Figure 4-2: (A) Absolute differences (dB) between IPL and SPL measurements and (B) Measured pressure reflectance magnitude for frequencies between 2000-8000 Hz for four hearing aid tubing length (25, 35, 45, 55 mm). Vertical dashed lines are the analytic frequency at which the maximum difference and minimum reflectance values are expected; this is frequency corresponding to the half-wave resonance of the tube.

4.3.2 Individualized foam-tip-to-earmold correction using tubing length estimation

To evaluate whether individualized estimation of tubing length would improve the accuracy of the foam-tip-to-earmold correction, individual tubing length corrections were derived as follows. Recall that individualized transforms were computed as the difference in sound-level produced at the termination of an HA-1 coupler by the acoustic transducer attached using a foam-tip and the same transducer attached using a specified length of hearing aid tubing. The individualized transforms were compared to the fixed, average, and frequency-specific foam-tip-to-earmold RECD correction (from Moodie et al., 2016a), which is implemented in some clinical hearing aid analyzers. This was completed to determine whether the individually derived earmold tubing length measurements would improve the transform in a clinically meaningful manner. Error was quantified as the difference between the actual response measured for each tubing length and the averaged transform, which indicates the improvement of the length-specific correction.

These results show reliable acoustic length estimation across testing sessions and between reinsertions with an intraclass correlation (ICC) value of 0.99, indicating good reliability (Portney & Watkins, 2015). It follows that a single measurement of hearing aid tubing length is sufficient for the determination of an individualized foam-tip-to-earmold transform for the individual within ~1 dB unless any physical changes to the hearing aid tubing occur.

The effect of the hearing aid tubing on the RECD measurement across a range of realistic hearing aid tubing lengths can be seen in Figure 4-3A. The currently implemented average transform (from Moodie et al., 2016a) is indicated by black asterisks in the plot. On the right side of the figure, deviation from the averaged transform as a function of hearing aid tubing length can be observed, with dotted lines at ± 2 dB; differences exceeding these values are clinically significant (ANSI, S3.46-2013). This deviation represents the signed error introduced to the RECD measurement when using the average foam-tip-to-earmold correction, versus measurement of the RECD with the earmold itself. Compared to the averaged transform, the differences reach clinical significance for at least one experimental tubing length tested at all frequencies for which the averaged transform is used except for 250 and 2000 Hz. The largest deviation from the average transform was observed at 4000 Hz with the shortest hearing aid tubing segments, where errors exceeded 6 dB for a segment length of 30 mm.

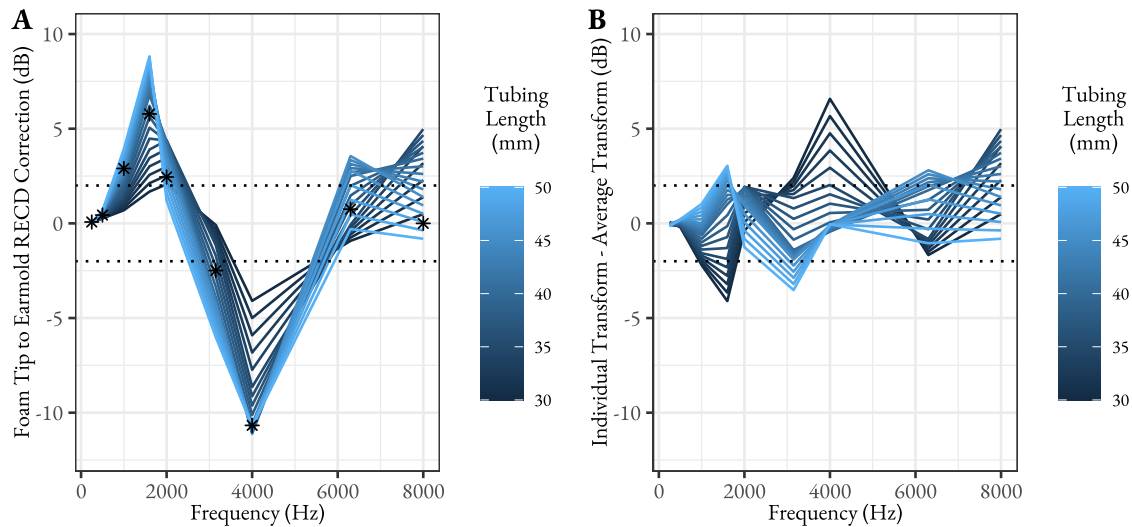


Figure 4-3: Plot (A) indicates the individualized foam-tip to earmold RECD corrections across frequencies for 1 mm increments for hearing aid tubing measurements. Averaged transforms calculated by Moodie et al., (2016) are indicated by black asterisks (*). A value of 0 indicates no difference between the earmold RECD and the foam-tip RECD. Plot (B) indicates error caused by using the average foam-tip to earmold RECD rather than the individualized transform. Clinical significance for real-ear measurements is denoted by horizontal dotted lines at ± 2 dB. Averaged foam-tip-to-earmold RECD transforms do not exist for 8 kHz, so a transform value of 0 is used. Positive values indicate that the average transform underestimates the correction factor necessary for a given length of tubing whereas negative values indicate overestimation.

4.4 Discussion

The current study has proposed a rapid method for acoustic estimation of hearing aid tubing length using a source-parameter calibrated transducer. Based on the current results, hearing aid tubing length measurements will significantly improve the validity of foam-tip-to-earmold RECD transforms, particularly for infants or those whose ears are significantly smaller or larger than the average adult with hearing aid tubing of ~ 40 mm. Furthermore, these acoustic length measurements using a source-parameter calibrated transducer are reliable across trials. In sum, the acoustic length measurement is expected to improve individually predicted earmold RECDs.

The average foam-tip-to-earmold correction used in current clinical practice yields significantly improved results over age-predicted norms (Moodie et al., 2016a). However, for newborns and others with short hearing aid tubing (i.e., approximately 30 mm), use of the average foam-tip-to-earmold correction implemented in clinical systems can result in the underestimation of an individual's RECD by as much as 6 dB near 4 kHz. Previous studies have noted that the acoustic effects of hearing aid tubing shorter than 25 mm are minor, when compared to responses from standard foam-tips (Bagatto et al., 2005; Gustafson et al., 2013; Munro & Hatton, 2000). However, if the average foam-tip-to-earmold correction is used, there will be an error in the RECD, resulting in an overestimation near 1500 Hz and an under-estimation near 4 kHz. The individualized measurement of the hearing aid tubing removes this source of error and better estimates the difference between coupling methods for the individual.

Good reliability of acoustic measurement was observed between trials (Portney & Watkins, 2015), with the errors in acoustic length estimation falling well below the ± 2 dB clinically significant criterion established for real-ear measurement previously (ANSI, S3.46-2013; Vaisberg et al., 2016). In a companion study, daily calibration results were also reliable across multiple source parameter estimates. Re-calibration is not necessary prior to every measurement, which removes a potential barrier to clinical implementation (Urichuk et al., 2021). The measurement of acoustic length takes seconds, and the proposed individualized transforms could be used directly in clinical settings with the commercially available wideband transducer that was used in this study. Across other transducers, however, the effect of the tubing length and additional series impedance may differ, as the pressure response of an acoustic load is dependent on not only the impedance of the load, but also on the impedance of the sound-source and coupling system (Egolf et al., 1985; Gilman et al., 1981; Sanborn, 1998). In comparisons between earmold RECD measurements using the RE-770 transducer and ER-3A transducer, the effect of a length of hearing aid tubing on sound-level output is partially dependent on the transducer itself (Bagatto et al., 2005; Munro & Salisbury, 2002). This is in part due to the nature of the ER-3A transducer design, which uses 250 mm of tubing to couple with any load, resulting in a higher overall sound-source impedance which minimizes the

effect of additional load impedances (Voss & Herrmann, 2005). As such, the individualized foam-tip-to-earmold transforms introduced in the present study are limited to the RE-770 transducer, although the same method can be applied if further experimental measurements are made with other transducers of interest.

4.4.1 Limitations

In the current study, the precise inner diameter of the transducer was not known. It is possible that slight impedance mismatches between the inner diameter of the transducer and the hearing aid tubing introduced minor errors into the measurement. Furthermore, it is possible that the coupling of the transducer and the hearing aid tubing was imperfect due to the inability to visually inspect inside the coupling sleeve. However, even with these possible sources of clinically realistic error in the measurement, the measurement of hearing aid tubing fell within a millimeter of the physically measured tubing length. In the current study, we assessed a range of hearing aid tubing lengths between 25 mm and 55 mm, which is applicable for most clinical hearing aid tubing segments (Moodie et al., 2016a). These segments result in resonant frequencies within the bandwidth tested, however, for tubing lengths shorter than 25 mm, it is possible that the expected frequency will be above 8000 Hz, which will require a stimulus with a wider bandwidth. Due to the upper frequency limit of the instrument used in this study, these outlier segment lengths were not assessed.

4.5 Conclusion

Thevenin-equivalent source parameter calibrated transducers can be configured to measure earmold tubing length accurately and quickly and provide clinically significant improvements over currently implemented average foam-tip-to-earmold RECD corrections. These improvements may contribute to improved validity of certain clinical measurements of hearing aid frequency response that use either couplers or RECD corrections as part of clinical processes. The benefits of individualized transforms would be most evident for those with hearing aid tubing lengths that significantly differ from the average adult, such as infants.

4.6 References

- Allen, J. B. (1986). Measurement of Eardrum Acoustic Impedance. In J. B. Allen, J. L. Hall, A. E. Hubbard, S. T. Neely, & A. Tubis (Eds.), *Peripheral Auditory Mechanisms* (Vol. 64, pp. 44–51). Springer Berlin Heidelberg.
https://doi.org/10.1007/978-3-642-50038-1_6
- American Academy of Audiology. (2013). American Academy of Audiology Clinical Practice Guidelines Pediatric Amplification.
- American National Standards Institute (ANSI). (2013). *Methods of Measurement of Real-Ear Performance Characteristics of Hearing Aids*. Acoustical Society of America.
- Bagatto, M., Moodie, S., Scollie, S., Seewald, R., Moodie, S., Pumford, J., & Liu, K. P. R. (2005). Clinical Protocols for Hearing Instrument Fitting in the Desired Sensation Level Method. *Trends in Amplification*, 9(4), 199–226.
<https://doi.org/10.1177/108471380500900404>
- Charaziak, K. K., & Shera, C. A. (2017). Compensating for ear-canal acoustics when measuring otoacoustic emissions. *The Journal of the Acoustical Society of America*, 141(1), 515–531. <https://doi.org/10.1121/1.4973618>
- Egolf, D. P., Feth, L. L., Cooper, W. A., & Franks, J. R. (1985). Effects of normal and pathologic eardrum impedance on sound pressure in the aided ear canal: A computer simulation. *The Journal of the Acoustical Society of America*, 78(4), 1281–1285. <https://doi.org/10.1121/1.392897>
- Gilman, S., Dirks, D. D., & Stern, R. (1981). The effect of occluded ear impedances on the eardrum SPL produced by hearing aids. *The Journal of the Acoustical Society of America*, 70(2), 370–386. <https://doi.org/10.1121/1.386772>
- Gustafson, S., Pittman, A., & Fanning, R. (2013). Effects of Tubing Length and Coupling Method on Hearing Threshold and Real-Ear to Coupler Difference Measures. *American Journal of Audiology*, 22(1), 190–199. [https://doi.org/10.1044/1059-0889\(2012/12-0046\)](https://doi.org/10.1044/1059-0889(2012/12-0046))
- Huang, G. T., Rosowski, J. J., Puria, S., & Peake, W. T. (2000). A noninvasive method for estimating acoustic admittance at the tympanic membrane. *The Journal of the Acoustical Society of America*, 108(3), 1128. <https://doi.org/10.1121/1.1287024>
- Hunter, L. L., Feeney, M. P., Lapsley Miller, J. A., Jeng, P. S., & Bohning, S. (2010). Wideband Reflectance in Newborns: Normative Regions and Relationship to Hearing-Screening Results: *Ear and Hearing*, 31(5), 599–610.
<https://doi.org/10.1097/AUD.0b013e3181e40ca7>
- Keefe, D. H., Bulen, J. C., Arehart, K. H., & Burns, E. M. (1993). Ear-canal impedance and reflection coefficient in human infants and adults. *The Journal of the Acoustical Society of America*, 94(5), 2617–2638.
<https://doi.org/10.1121/1.407347>
- Lapsley Miller, J. A., Reed, C. M., Robinson, S. R., & Perez, Z. D. (2018). Pure-Tone Audiometry With Forward Pressure Level Calibration Leads to Clinically-

- Relevant Improvements in Test–Retest Reliability: *Ear and Hearing*, 39(5), 946–957. <https://doi.org/10.1097/AUD.0000000000000555>
- Levine, H., & Schwinger, J. (1948). On the Radiation of Sound from an Unflanged Circular Pipe. *Physical Review*, 73(4), 383–406. <https://doi.org/10.1103/PhysRev.73.383>
- Lewis, J. D., McCreery, R. W., Neely, S. T., & Stelmachowicz, P. G. (2009). Comparison of *in-situ* calibration methods for quantifying input to the middle ear. *The Journal of the Acoustical Society of America*, 126(6), 3114–3124. <https://doi.org/10.1121/1.3243310>
- McCreery, R. W., Pittman, A., Lewis, J., Neely, S. T., & Stelmachowicz, P. G. (2009). Use of forward pressure level to minimize the influence of acoustic standing waves during probe-microphone hearing-aid verification. *The Journal of the Acoustical Society of America*, 126(1), 15–24. <https://doi.org/10.1121/1.3143142>
- Moodie, Shane, Seewald, R. C., & Sinclair, S. (1994). Procedure for Predicting Real-Ear Hearing Aid Performance in Young Children. *American Journal of Audiology*, 3(1), 23–31. <https://doi.org/10.1044/1059-0889.0301.23>
- Moodie, Sheila, Pietrobon, J., Rall, E., Lindley, G., Eiten, L., Gordey, D., Davidson, L., Moodie, K. S., Bagatto, M., Haluschak, M. M., Folkeard, P., & Scollie, S. (2016a). Using the Real-Ear-to-Coupler Difference within the American Academy of Audiology Pediatric Amplification Guideline: Protocols for Applying and Predicting Earmold RECDs. *Journal of the American Academy of Audiology*, 27(03), 264–275. <https://doi.org/10.3766/jaaa.15086>
- Moodie, S., Rall, E., Eiten, L., Lindley, G., Gordey, D., Davidson, L., Bagatto, M., & Scollie, S. (2016b). Pediatric Audiology in North America: Current Clinical Practice and How It Relates to the American Academy of Audiology Pediatric Amplification Guideline. *Journal of the American Academy of Audiology*, 27(03), 166–187. <https://doi.org/10.3766/jaaa.15064>
- Munro, Kevin J., & 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 and Hearing*, 24(2), 100–110. <https://doi.org/10.1097/01.AUD.0000058114.20741.4D>
- Munro, Kevin J., & Hatton, N. (2000). Customized Acoustic Transform Functions and Their Accuracy at Predicting Real-Ear Hearing Aid Performance. *Ear and Hearing*, 21(1), 59–69.
- Munro, Kevin J., & Lazenby, A. (2001). Use of the ‘real-ear to dial difference’ to derive real-ear SPL from hearing level obtained with insert earphones. *British Journal of Audiology*, 35(5), 297–306. <https://doi.org/10.1080/00305364.2001.11745248>
- Munro, Kevin J., & Salisbury, V. (2002). Is the real-ear to coupler difference independent of the measurement earphone. *International Journal of Audiology*, 41(7), 408–413.

- Munro, K. J., & Toal, S. (2005). Measuring the Real-Ear to Coupler Difference Transfer Function With an Insert Earphone and a Hearing Instrument: Are They the Same? *26*(1), 8.
- Nørgaard, K. R., Fernandez-Grande, E., & Laugesen, S. (2017). Incorporating evanescent modes and flow losses into reference impedances in acoustic Thévenin calibration. *The Journal of the Acoustical Society of America*, *142*(5), 3013–3024. <https://doi.org/10.1121/1.5010891>
- Portney, L. G., & Watkins, M. P. (2015). *Foundations of Clinical Research: Applications to Practice* (3rd Edition). F. A. Davis Company.
- R Core Team. (2020). *R: A language and environment for statistical computing* [Manual]. <https://www.R-project.org/>
- Sanborn, P.-E. (1998). Predicting hearing aid response in real ears. *The Journal of the Acoustical Society of America*, *103*(6), 3407–3417. <https://doi.org/10.1121/1.423082>
- Saunders, G. H., & Morgan, D. E. (2003). Impact on hearing aid targets of measuring thresholds in dB HL versus dB SPL: El impacto en la medición de los umbrales en dB HL o en dB SPL, en las metas de un auxiliar auditivo. *International Journal of Audiology*, *42*(6), 319–326.
- Scheperle, R. A., Goodman, S. S., & Neely, S. T. (2011). Further assessment of forward pressure level for *in situ* calibration. *The Journal of the Acoustical Society of America*, *130*(6), 3882–3892. <https://doi.org/10.1121/1.3655878>
- Scheperle, R. A., Neely, S. T., Kopun, J. G., & Gorga, M. P. (2008). Influence of *in situ* , sound-level calibration on distortion-product otoacoustic emission variability. *The Journal of the Acoustical Society of America*, *124*(1), 288–300. <https://doi.org/10.1121/1.2931953>
- Scollie, S., Seewald, Richard, Cornelisse, Leonard, & Jenstad, Lorianne. (1998). Validity and Repeatability of Level-Independent HL to SPL Transforms. *Ear and Hearing*, *19*, 407–413.
- Souza, N. N., Dhar, S., Neely, S. T., & Siegel, J. H. (2014). Comparison of nine methods to estimate ear-canal stimulus levels. *The Journal of the Acoustical Society of America*, *136*(4), 1768–1787. <https://doi.org/10.1121/1.4894787>
- Urichuk, M., Purcell, D., Allen, P., Scollie, S. (2021). *Thevenin-equivalent source parameter reliability: intra- and inter-session effects on stimulus level calibration and wideband reflectance*. Manuscript submitted for publication.
- Vaisberg, J. M., Macpherson, E. A., & Scollie, S. (2016). Extended bandwidth real-ear measurement accuracy and repeatability to 10 kHz. *International Journal of Audiology*, *55*(10), 580–586. <https://doi.org/10.1080/14992027.2016.1197427>
- Voss, S. E., & Allen, J. B. (1994). Measurement of acoustic impedance and reflectance in the human ear canal. *The Journal of the Acoustical Society of America*, *95*(1), 372–384. <https://doi.org/10.1121/1.408329>

- Voss, S. E., & Herrmann, B. S. (2005). How Does the Sound Pressure Generated by Circumaural, Supra-aural, and Insert Earphones Differ for Adult and Infant Ears?: *Ear and Hearing*, 26(6), 636–650. <https://doi.org/10.1097/01.aud.0000189717.83661.57>
- Withnell, R. H., Jeng, P. S., Parent, P., & Levitt, H. (2014). The clinical utility of expressing hearing thresholds in terms of the forward-going sound pressure wave. *International Journal of Audiology*, 53(8), 522–530. <https://doi.org/10.3109/14992027.2014.898122>
- Withnell, R. H., Jeng, P. S., Waldvogel, K., Morgenstein, K., & Allen, J. B. (2009). An *in situ* calibration for hearing thresholds. *The Journal of the Acoustical Society of America*, 125(3), 1605–1611. <https://doi.org/10.1121/1.3075551>

Chapter 5

5 Validity and reliability of integrated pressure level Real-Ear-to-Coupler Difference measurements

5.1 Introduction

Clinical fitting procedures for hearing aids include routine verification and fine-tuning of ear-canal sound levels to ensure safe and beneficial aided responses (AAA, 2013; Valente et al., 2006). The ear-canal and middle-ear properties have a significant impact on the total sound-level presented to the eardrum, frequently causing 20 dB of variability between individuals and occasionally introducing differences as large as 40 dB (McCreery et al., 2015; Saunders & Morgan, 2003; Watts et al., 2020). Hearing aid verification systems account for this variability to accurately fit hearing aids to the combined effects of the individual's hearing loss and unique ear canal acoustics (see reviews by Fabry, 2003; Mueller, 2001; and Seewald et al., 2005). Some transform procedures are used to estimate ear canal sound levels for audiometric data (Gustafson et al., 2013; Scollie et al., 1998) while others focus on hearing aid responses, which are discussed in more detail below.

Verification systems measure an individual's hearing aid response in a variety of formats. One such format is the real-ear aided response (REAR; American National Standards Institute [ANSI], S3.46-2013). The individual's REAR can be measured directly using a probe-tube microphone system or estimated from coupler-based measures of hearing aid output. These estimated, or coupler-based, approaches use a transform, the Transform for Estimating Real Ear Output (TEREO; Mueller & Hall, 1998; Seewald et al., 1997), that accounts for the microphone location effect and Real-Ear to Coupler Difference (RECD; ANSI, S3.46-2013; Bagatto et al., 2005). Direct measurement of the REAR requires sustained cooperation from the individual, which is not always possible in all clinical populations, such as with infants and young children (Bagatto et al., 2005; Moodie et al., 1994). Alternatively, the RECD measurement can be made in seconds with minimal participation from the patient and used within the TEREO transform. This makes RECD measurement a recommended and common choice for verification in infants and other

clinical populations for whom direct REAR measurements are challenging (AAA, 2013; Moodie et al., 2016). RECDs are also an option for pre-fitting or remote fitting.

Measurement of the RECD can be done using either a foam-tip transducer or the individual's personalized earmold. The RECD measured with a personal earmold does not meet the standard definition of the RECD (ANSI, S3.46-2013) and is instead considered an Ear to Coupler Level Difference (ECLD). Earmold "RECDs" are necessary, however, to predict the on-ear responses for behind-the-ear (BTE) hearing aids due to the effects of earmold acoustics. This can help to ensure accurate hearing aid fitting at frequencies up to 4 kHz (Munro & Davis, 2003; Munro & Hatton, 2000; Vaisberg et al., 2018).

The reduced accuracy of the real-ear probe-tube microphone measurement at high frequencies is well understood in direct REAR and RECD measurement. It is attributable to interactions between the incident pressure wave and acoustic reflections from the eardrum (Chan & Geisler, 1990; McCreery et al., 2009). This interference results in standing wave errors which introduce variations of up to 20 dB in the amplitude measurements near frequencies where the 3/4-wavelength is similar to the distance between the location of measurement and the eardrum (Chan & Geisler, 1990; Dirks & Kincaid, 1987; Stinson & Lawton, 1989). By placing the probe-tube microphone near the eardrum, the frequency of the standing wave is increased but the error is not eliminated. The standing wave is merely shifted to higher frequencies, which impacts both REAR and RECD measurements above 4 kHz (Bagatto et al., 2005; Feigin et al., 1989; Munro & Hatton, 2000; Tharpe et al., 2001; Vaisberg et al., 2018; Valente et al., 1994). The coupler used in the RECD is traditionally a 2.0 cc coupler, however, recently a modified procedure using a 0.4 cc coupler was developed to minimize standing-wave errors (International Electrotechnical Commission, 2016). This procedure, named the wideband RECD (wRECD), improves the validity of the coupler measurement at high frequencies. However, it does not mitigate the standing-wave errors in the ear-canal.

The purpose of the current study is to evaluate a new wRECD measurement paradigm for use in clinical hearing aid fitting that quantifies individual ear-canal acoustics

independent of the standing wave error. The proposed wRECD paradigm is measured *in-situ* with the use of a reflectance-based probe rather than with a probe-tube microphone. Such a measurement is expected to improve high-frequency hearing aid fitting accuracy and validity without the need to place a probe-tube microphone near the eardrum.

Historically, hearing aids have been limited in providing high-frequency functional gain (Moore et al., 2001), reducing concern about standing wave error in REAR determination. With modern advancements in hearing aid technology, the bandwidth of functional gain has been extended as high as 10 kHz (Moore et al., 2008). Extended bandwidths improve speech recognition, sound quality, and improve subjective preference for some listeners. Some prescriptive methods provide audibility at 8 kHz and above, increasing the clinical relevance of high-frequency hearing aid output (Alexander & Rallapalli, 2017; Brennan et al., 2014; Folkeard et al., 2021; Füllgrabe et al., 2010; Hornsby et al., 2011; Levy et al., 2015; Moore et al., 2010; Ricketts et al., 2008; Vaisberg et al., 2021a,b; Van Eeckhoutte et al., 2020). Extending the bandwidth of a hearing aid fitting requires prescriptive targets and verification methods that are valid across the entire bandwidth. For such verification to be accurate, the acoustic input into the auditory system needs to be measured in a valid and reliable manner. Given the limited reliability and validity of probe-tube microphone measures in the high frequencies, an alternative measurement procedure may assist in ensuring accuracy in hearing aid fitting with extended bandwidth devices.

5.1.1 Accounting for individualized ear-canal acoustics using a source parameter calibrated transducer

In-ear sound-level calibration using a Thevenin-equivalent source parameter calibrated transducer has been shown to more accurately quantify input into the eardrum without reflected-wave interference. Pressure measurements can be made using a microphone housed in the transducer. Accurate calibration of the transducer's Thevenin-equivalent source parameters, referred to in this study as "source parameters", can be completed through a reliable weekly calibration procedure on a set of calibration cavities in less than five minutes (Urichuk et al., 2021a). Source parameter-based calibration approaches accurately and reliably measure the sound input to the auditory system up to 16 kHz with

less than 3 dB of error (Souza et al., 2014). During the in-ear sound-level calibration measurement, source-parameter calibrated transducers also simultaneously provide an assessment of wideband acoustic immittance. To emphasize this point: one pressure response measured by the transducer in the ear-canal can produce accurate sound-level calibration and accurate middle-ear assessment.

Source parameter calibration, composed of the measurement of multiple calibration cavities, determines the transducer's source pressure (P_{src}) and source impedance (Z_{src}). Obtaining these values through the calibration procedure is explained in-depth elsewhere (Nørgaard et al., 2017; Rosowski et al., 2013; Urichuk et al., 2021a; Voss & Allen, 1994). After the calibration procedure, the transducer can be coupled to the ear-canal using an acoustic immittance tip and the pressure response (P_{ec}) of the ear-canal to a wideband stimulus can be made in seconds. Using the known source parameters and the wideband P_{ec} measurement, the wideband power absorbance, and sound-pressure level (SPL) present at the eardrum (independent of the standing wave) can be determined through a series of steps, which can be integrated into the measurement instrument. The first step in this process is the determination of the ear-canal acoustic impedance (Z_{ec}):

$$Z_{ec} = \frac{Z_{src}P_{ec}}{P_{src} - P_{ec}} \quad (5.1)$$

Once Z_{ec} is calculated, an estimate of the ear-canal characteristic impedance (Z_0) needs to be determined. The characteristic impedance is dependent on the diameter of the ear-canal where the transducer is located. Previously, estimates of Z_0 using average adult ear-canal dimensions (radius = 3.75 mm) have been employed (e.g. McCreery et al., 2009; Scheperle et al., 2008). However, a recent analytic procedure introduced by Rasetshwane and Neely (2011) can accurately estimate Z_0 for an individual measurement by iteratively adjusting Z_0 values used in the ear-canal reflectance calculation below. The Z_0 value that minimizes the calculated ear-canal reflectance at a time of $t = 0$ provides an accurate estimate of the Z_0 , and therefore cross-sectional area, of the individual's ear canal at the measurement location. This approach to individualized Z_0 estimates was implemented in the current study. Using the calculated Z_{ec} and Z_0 values provides us with the information

necessary to accurately calculate ear-canal pressure reflectance (R_{ec} ; Keefe et al., 1993; Voss & Allen, 1994):

$$R_{ec} = \frac{Z_{ec} - Z_0}{Z_{ec} + Z_0} \quad (5.2)$$

The pressure reflectance is then used for two calculations – one relating to power absorbance and the other relating to sound-level estimation. Power absorbance, the final value used for wideband acoustic immittance, is obtained using the pressure reflectance by:

$$A = 1 - |R_{ec}|^2 \quad (5.3)$$

Pressure reflectance is also used to quantify the sound pressure moving towards the eardrum (forward-pressure level; FPL) and the pressure reflected by the eardrum (reverse-pressure level; RPL):

$$FPL = \frac{P_{ec}}{1 + R_{ec}} \quad (5.4)$$

$$RPL = R_{ec}FPL \quad (5.5)$$

Forward-pressure level has been validated as a stimulus calibration method that maintains phase information. Phase information is important for some calibration uses (such as otoacoustic emissions calibration), however, phase information is not explicitly necessary for accurate hearing aid verification.

Unfortunately, direct comparison between stimuli referenced in FPL versus SPL at high frequencies is complex. Hearing aid prescriptive targets are currently calculated in SPL, which doesn't have a simple one-to-one conversion with FPL, making the incorporation of FPL based measurements difficult. Fortunately, at the eardrum, the FPL waves and reflected sound-waves interact in-phase for closed tube (Lewis et al., 2009). We can therefore find the magnitude of the forward-moving wave and the magnitude of the reflected wave and sum them together in-phase to determine the SPL at the eardrum, referred to as the integrated pressure level (IPL; Lewis et al., 2009):

$$\text{IPL} = |\text{FPL}| + |\text{RPL}| \quad (5.6)$$

Integrated pressure level has been validated with termination SPL in cylinders (Scheperle et al., 2011) and is not impacted by standing-wave interference (i.e. the quarter-wavelength pressure null present at the measurement location of the transducer), yielding an accurate and reliable method of calibration (Lewis et al., 2009; Souza et al., 2014; Withnell et al., 2009).

5.1.2 IPL wRECD: Method for foam-tip and earmold wRECD determination

By obtaining the IPL response, which is theoretically equivalent to the SPL at the termination of the cavity being measured, we can obtain the wRECD transform with the source-parameter calibrated transducer. The IPL wRECD can therefore be defined as:

$$\text{IPL wRECD} = [\text{Real ear IPL response}] - [0.4 \text{ cc coupler IPL response}] \quad (5.7)$$

Given the theoretical equivalence to eardrum SPL, we hypothesize that IPL wRECD can be used in the hearing aid fitting workflow as a direct replacement to traditional probe-tube microphone wRECD. Furthermore, we hypothesize that the acoustic determination of an individual ear's characteristic impedance will improve the agreement between IPL wRECD and traditional probe-tube microphone wRECD measurements. However, some consideration of wRECD subtypes is necessary to determine appropriate comparison conditions, as discussed below.

Current clinical probe-tube microphone wRECD values can be completed with either a foam-tip coupling, or with the individual's ear-mold attached to a variable length of hearing aid tubing. The choice of coupling, as well as the choice of the transducer, have been shown to significantly alter wRECD values (Bagatto et al., 2005; Gustafson et al., 2013; Moodie et al., 2016; Munro & Salisbury, 2002; Munro & Toal, 2005). These differences are largely due to the interaction between the impedances of the hearing aid tubing and the transducer, with increased tubing length leading to larger errors in REAR prediction (Bagatto et al., 2005; Gustafson et al., 2013; Moodie et al., 2016). To mitigate

these errors, an averaged correction between earmold and foam-tip transducer has been implemented in at least one hearing aid verification system (Moodie et al., 2016). However, it is still recommended to directly measure the RECD using the individual's ear-mold when possible to minimize the transducer mismatch errors (Moodie et al., 2016).

Recently, a method for acoustically determining hearing aid tubing length with the source parameter calibrated transducer has been developed and validated (Urichuk et al., 2021b). Using an individualized tubing length correction, accurate estimation of an individual's earmold wRECD using the wRECD measured using a generic insert may be possible. At time of the current study, it is not possible to calibrate the source parameters of an individual's hearing aid and individualized earmold in a clinically feasible manner. Consequently, IPL wRECDs cannot be measured directly using an individual's earmold. Incorporating an accurate, individualized generic-tip-to-earmold transform for IPL wRECD measurement may allow accurate conversion between IPL wRECD and estimated earmold wRECD independent of the individual's hearing aid tubing length.

The purpose of the current study can be summarized as follows:

- (1) The primary purpose was to compare IPL wRECD measurement to probe-tube microphone wRECD measurements completed using either a generic insert-tip or the individual's ear-mold.
- (2) The secondary purpose was to evaluate individualized foam-tip-to-earmold correction using in earmold IPL wRECD determination.
- (3) The third purpose was to examine the test-retest reliability of the IPL wRECD measurement.
- (4) Finally, the current study aimed to assess the impact of individualized characteristic impedance estimation at the transducer's measurement position on agreement between IPL wRECD and probe-tube microphone wRECD measurement.

We hypothesized that no clinically significant differences will be observed between IPL wRECD measurements and traditional probe-tube microphone wRECD measurements

for frequencies below 4 kHz. Above 4kHz, we hypothesize the IPL wRECD values to be larger than probe-tube microphone measured wRECD measurements due to standing-wave errors being present in the probe-tube microphone wRECD measurements (specifically, the effects of the $\frac{3}{4}$ wavelength pressure null approaching the eardrum and the probe-tube microphone). A similar relationship is expected between the probe-tube microphone measured earmold wRECD and the IPL estimated earmold wRECD using individualized hearing aid tubing length measurement to account for variations in tubing length. Finally, an improved agreement between IPL wRECD and probe-tube microphone measured wRECD in the mid-frequency range was expected, consistent with other studies validating the use of acoustic characteristic impedance estimation for in-ear stimulus level calibration (Scheperle et al., 2011; Souza et al., 2014).

5.2 Methods

5.2.1 Participants

Participants were recruited from the Western University School of Communication Sciences and Disorders. A total of 22 adults (4 male; 18 female) between 21-30 years of age participated in this study. All participants presented with normal middle ear status, as determined by type A tympanograms. All participants passed a basic audiometric screening at all octave frequencies between 250-8000 Hz with hearing thresholds at or below 25 dB HL. Otoscopy was performed prior to testing. All subjects had typical ear-canal anatomy and minimal wax present.

5.2.2 Procedure

Testing was completed in a quiet laboratory at the National Centre for Audiology at the Western University. Probe-tube microphone wRECD measurements were completed with an Audioscan VeriFit 2 using the RE-770 transducer coupled to a generic foam-tip insert or to the individual's earmold with a variable length of #13 hard wall hearing aid tubing. IPL wRECD measurements were completed using an Interacoustics Titan transducer using a modified MATLAB (Mathworks, Natick, MA, United States) script built on the Interacoustics Research Platform. Source parameter calibration was completed weekly (Urichuk et al., 2021a).

The testing session consisted of nine measurements from three testing conditions: (1) bilateral foam-tip wRECD measurement (two measurements total), (2) unilateral earmold wRECD measurement with either short, medium, or long tubing lengths (three measurements total), (3) bilateral IPL wRECD measurement (two measures with reinsertion; four measurements in total). Measurement condition order was randomized across participants. Measurements were used to determine the five measurement types: (1) probe-tube microphone measured foam-tip wRECD, (2) estimated earmold wRECD calculated using the foam-tip wRECD with an average foam-tip-to-earmold correction (3) probe-tube microphone measured earmold wRECD, (4) IPL wRECD, and (5) earmold IPL wRECD using the procedure outlined in Figure 5-1. The earmold IPL wRECD is calculated using two measurements: (1) a foam-tip IPL wRECD measurement in the individual and (2) an acoustic measurement of the hearing aid tubing length for a given test condition (Figure 5-1B). Using the tubing length measurement, an individualized foam-tip-to-earmold transform was determined for the individual (Urichuk et al., 2021b). The earmold IPL wRECD is defined as the sum of the IPL wRECD and the individualized foam-tip-to-earmold transform at each measured 1/3 octave band.

To measure the real-ear response in each probe-tube microphone condition, the probe-tube was inserted into the ear-canal and placed within a maximum of 5 mm from the individual's tympanic membrane. Initial probe-tube placement was completed at a depth 30 mm from the intratragal notch for male participants and 28 mm from the intratragal notch for female participants. Probe-tube microphone placement was verified to be within 5 mm of the eardrum using a combination of a validated probe-tube insertion guide (Folkeard et al., 2019) and visualization using otoscopy. Probe-tube placement was reassessed if movement occurred with transducer placement, as observed by the tube position marker. Earmold tubing was coupled to the IPL wRECD transducer using a silicone coupling sleeve with inner diameter of 3.125 mm to complete an acoustic length measurement (Urichuk et al., 2021b) prior to probe-tube microphone earmold wRECD measurement. For IPL wRECD measurement, a plastic acoustic immittance tip was coupled with the source-parameter calibrated transducer. A broadband transient stimulus, described below, was presented to the ear-canal and the pressure response was measured

by the transducer-housed microphone. The transducer was fully removed and re-inserted between all measurements.

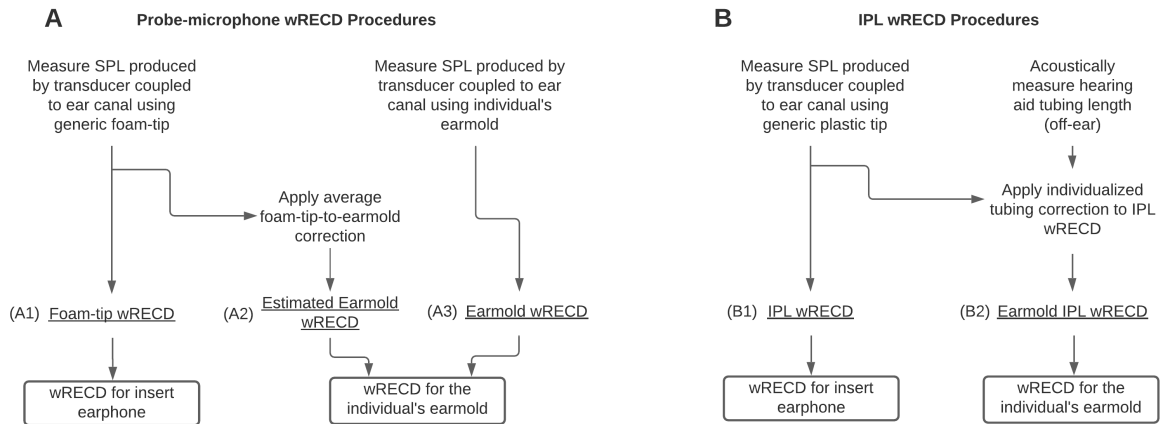


Figure 5-1: Measurement procedure for determining the insert earphone wRECD and earmold wRECD using (A) probe microphone wRECD measurement procedures and (B) IPL wRECD measurement procedures. Acoustic measurements are taken at the top of each pathway. Each measurement type is referred to in the current study using the underlined term associated with each pathway. The resulting value for use in hearing aid fitting procedures (i.e. the insert-earphone wRECD or earmold wRECD) is indicated in the bottom row.

5.2.3 Stimulus calibration

A wideband transient click-like stimulus (226 Hz-8000 Hz; 21 Hz presentation rate) was presented by the Interacoustics Titan transducer at 96 dB peak-to-peak equivalent SPL (peSPL). The stimulus was calibrated in a Brüel and Kjær type 4157 ear simulator (Nærum, Denmark) using a Brüel and Kjær type 4192 microphone. The Titan transducer was coupled flush to the ear simulator opening using an ER38-14A small foam insert. The Brüel and Kjær type 4192 microphone was connected to a Brüel and Kjær conditioning amplifier set to 316 milliVolts/Pascal. The amplifier was connected to a USBPre 2 external soundcard (Sound Devices, WI, United States) set to full-scale which sent output to SpectraPLUS software (Pioneer Hill Software, WA, United States). All

measurements are referenced to a SpectraPLUS calibration file measuring a Brüel and Kjær type 4231 calibrator output a 94 dB SPL tone at 1 kHz. All measurements were completed in a quiet room.

5.2.4 Analysis

All analysis was completed using 1/3 octave bands between 250 and 8000 Hz. Using R (R Core Team, 2020) and lme4 (Bates et al., 2015), we constructed mixed-effect linear models to assess the relationships between predictor variables and outcome variables of interest. Participants were coded as random effects across all models. All reported p-values were obtained using likelihood ratio tests of the full model compared against the model without the effect of interest included. Testing was completed by comparing estimated marginal means of specified groups with Tukey honest significant difference (HSD) multiple comparison corrected p-values as necessary.

The clinically significant difference between measurements is determined using estimated 95% confidence intervals for the difference between real-ear measurement and RECD measurement as reported by Munro and Davis (2003). At each analyzed frequency, two standard deviations were added to each mean difference and then results were averaged across all frequencies. This method of determining clinical significance has been used previously to validate the probe-tube microphone measured wRECD procedure with a cut-off of 3 dB (Vaisberg et al., 2018). Mean absolute differences between IPL wRECD and probe-tube microphone measured wRECD measurements that fell below 3 dB were not deemed to be clinically significant.

5.2.5 Data exclusion

Offline analysis of all measurements was completed. Real-ear IPL measurements obtained with foam tips were excluded after all testing was completed when a leak was found to be present in the measurement. A leak was determined by a low-frequency power absorbance magnitude >0.29 and low-frequency admittance phase <44 degrees (Groon et al., 2015). Two real-ear IPL measurements and one earmold wRECD were found to include significant acoustic leaks and excluded from analysis.

5.3 Results

5.3.1 Comparison between IPL wRECD and probe-tube microphone measured foam-tip wRECD

There were no clinically significant differences (greater than 3 dB) observed between probe-tube microphone measured foam-tip wRECD and IPL wRECD measurement below 5 kHz (Figure 5-2). For 1/3 octave bands at or above 5 kHz, a significant difference was observed, with IPL wRECD measurements yielding higher wRECD values. A linear mixed model with fixed effects of measurement methodology and frequency of measurement, with interactions was constructed. The model included random intercepts for participant and random slopes for measurement methodology. Group average IPL wRECD values exceeded probe-tube microphone foam-tip wRECD values by 5.72 dB ($CI_{95} = 4.85-6.60$) and 8.52 dB ($CI_{95} = 7.65-9.40$) at 6300 and 8000 Hz, respectively. These large differences are due to the varying effect of standing wave interference observed in the foam-tip wRECD probe-tube microphone measurements in individuals (Figure 5-3). In general, larger deviations are noted for the individuals that show a standing wave in the high-frequencies, or in individuals where a resonance of the ear-canal space is not seen in the probe-tube measurement, resulting in flat foam-tip wRECD probe-tube microphone values. This can happen when the probe-tube microphone is placed approximately half-way between the eardrum and the location of the $\frac{3}{4}$ -wavelength pressure node nearest the eardrum. When the probe-tube microphone is placed at this location, reflected sound-waves do not sum completely in phase as they would at the eardrum, and the probe-tube microphone doesn't observe evidence of the standing wave as it would if it were placed shallower in the canal. There was also a significant main effect of the measurement methodology on wRECD values ($\chi^2_1 = 6.87$, $p = 0.009$), as well as an interaction between measurement methodology and 1/3-octave band ($\chi^2_{15} = 1328$, $p < 0.001$).

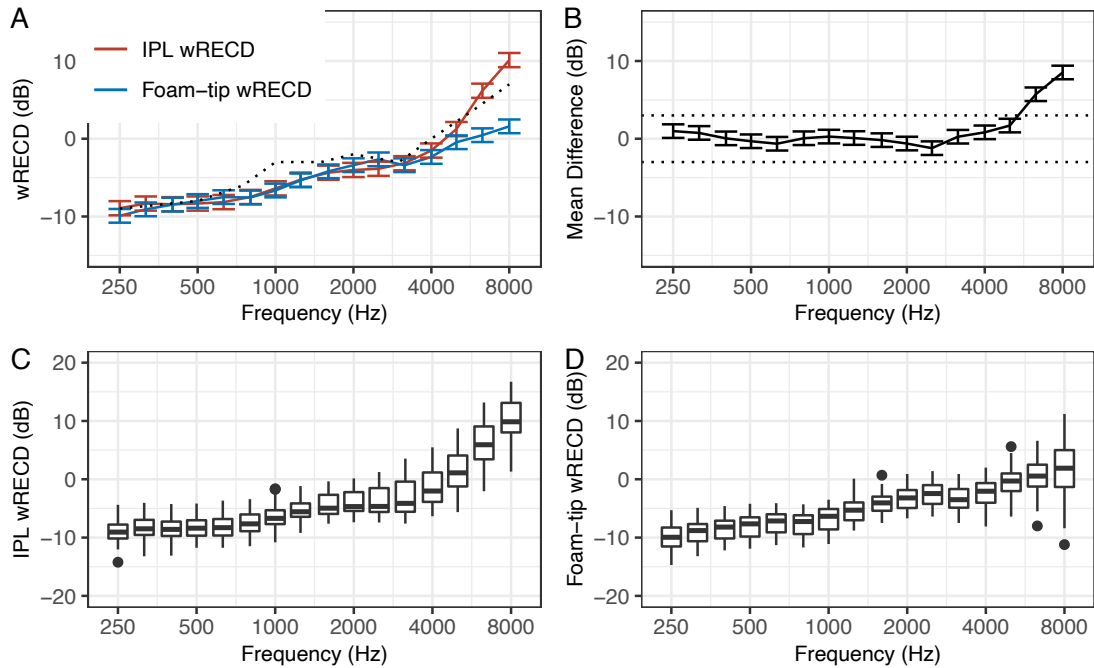


Figure 5-2: (A) Average measurement of wRECD values as a function of frequency using either the IPL wRECD (red) or probe-tube microphone SPL wRECD (blue). Average wRECD values for the normative adult population are overlaid with a dotted line (as implemented in the Audioscan Verifit 2). (b) signed difference in wRECD average between IPL wRECD minus probe-tube microphone foam-tip wRECD. The cut-off for clinical significance is denoted by dotted lines. Positive values indicate increased average wRECD values at a specific frequency for IPL wRECD. Error bars indicate the 95% confidence interval for each estimate. Subplots C and D show the distribution of all wRECD measured in the current study. Boxplots represent 25-75th percentile of the distribution, with interior horizontal line at the median and whiskers extending to 1.5 * interquartile range (IQR). Outliers are indicated by black dots.

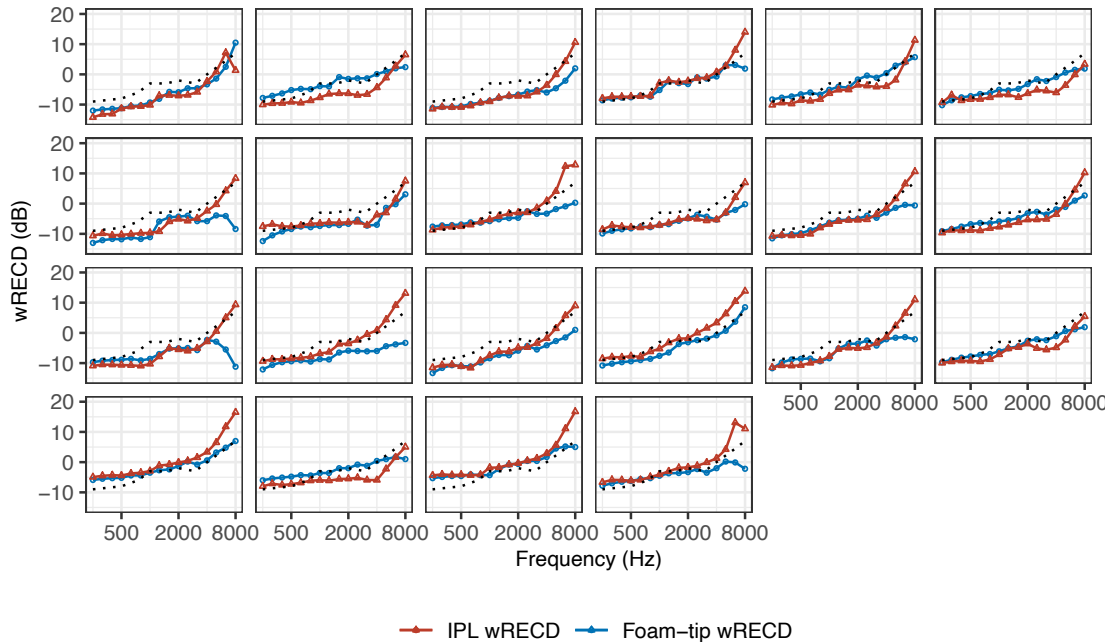


Figure 5-3: Individual measurements of wRECD values in all participants' right ears as a function of frequency using IPL wRECD measurement (red; triangles) and probe-tube microphone measured foam-tip wRECD measurement (blue; circles). Average wRECD values for normative adult population (as implemented in the Audioscan Verifit 2) are overlaid with dotted black lines.

5.3.2 Comparison between earmold IPL wRECD and probe-tube microphone measured earmold wRECD

Earmold IPL wRECD measurements yielded similar values to the probe-tube measured earmold wRECD in the mid-frequency region (Figure 5-4). In low frequencies, however, the range of measured earmold IPL wRECD measurements was much smaller, due in part to less slit-leakage. Earmold IPL wRECD values were also greater in the high-frequencies, likely due to the lack of standing-wave interference. To evaluate whether prediction of the earmold wRECD using the earmold IPL wRECD (IPL wRECD with individualized hearing aid tubing correction) estimated the probe-tube earmold wRECD, differences between the two were assessed as the outcome of a linear mixed model. Frequency and hearing aid tubing length (with interactions) were set as fixed effects with the random intercept of participant. A significant main effect of frequency was observed

($\chi^2_{15} = 4957$, $p < 0.001$), with no significant main effect of hearing aid tubing length ($\chi^2_1 = 0.10$, $p = 0.75$). A significant interaction between hearing aid tubing length and frequency was also observed ($\chi^2_{15} = 118.47$, $p < 0.001$), indicating that frequency specific differences between estimated and measured probe-tube microphone earmold wRECD caused by tubing length remained present. Differences between average earmold IPL wRECD and the probe-tube microphone measured earmold wRECD were less than 3 dB for all frequencies 500-2500 Hz (Figure 5-5). At frequencies lower than 500 Hz, the IPL wRECD was larger than the probe-tube microphone wRECD by more than the clinically significant cut-off of 3 dB. This was caused by slit-leakage present when measuring the probe-tube microphone earmold wRECD. The difference was largest at 250 Hz by almost 10 dB on average (Figure 5-5). This slit-leakage is anticipated to be absent when the probe-tube microphone isn't creating gaps between earmold and the ear-canal, or at the very least minimized, meaning that this difference is mainly due to errors present in the probe-tube microphone measured earmold wRECD rather than errors in estimation.

At all 1/3 octave bands above 2500 Hz, average earmold IPL wRECDs exceeded probe-tube microphone wRECDs by more than 3 dB. Differences at specific frequencies were 4.13 dB ($CI_{95} = 3.41-4.85$) at 3150 Hz, 5.4 dB at 4000 Hz ($CI_{95} = 4.68-6.12$), 6.3 dB at 5000 Hz ($CI_{95} = 5.82-6.83$), 9.01 dB at 6300 Hz ($CI_{95} = 8.29-9.73$), and 11.03 dB at 8 kHz ($CI_{95} = 10.31-11.75$). At these frequencies, standing-wave error is assumed to be present in probe-tube measured values resulting in smaller values than earmold IPL wRECD. Errors at these frequencies were largely independent of tubing length, indicating that there was an unaccounted-for systematic acoustic difference between the earmold IPL wRECD and the probe-tube microphone measured earmold wRECD.

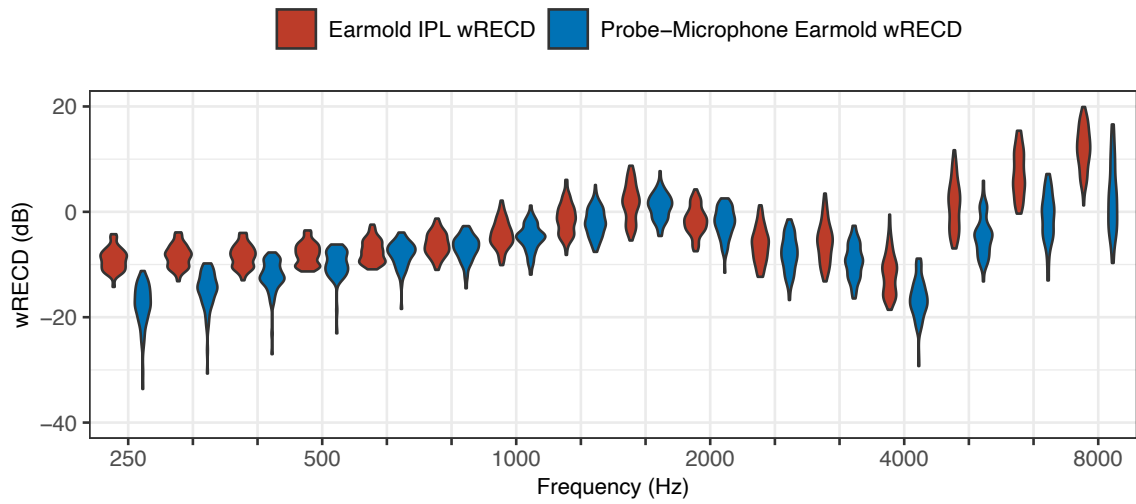


Figure 5-4: Distribution of all earmold wRECD measurements across all tubing length conditions using both earmold IPL wRECD and probe-tube measured earmold wRECD. Width of distribution indicates density of measurements yielding a specific wRECD for a given 1/3 octave band.

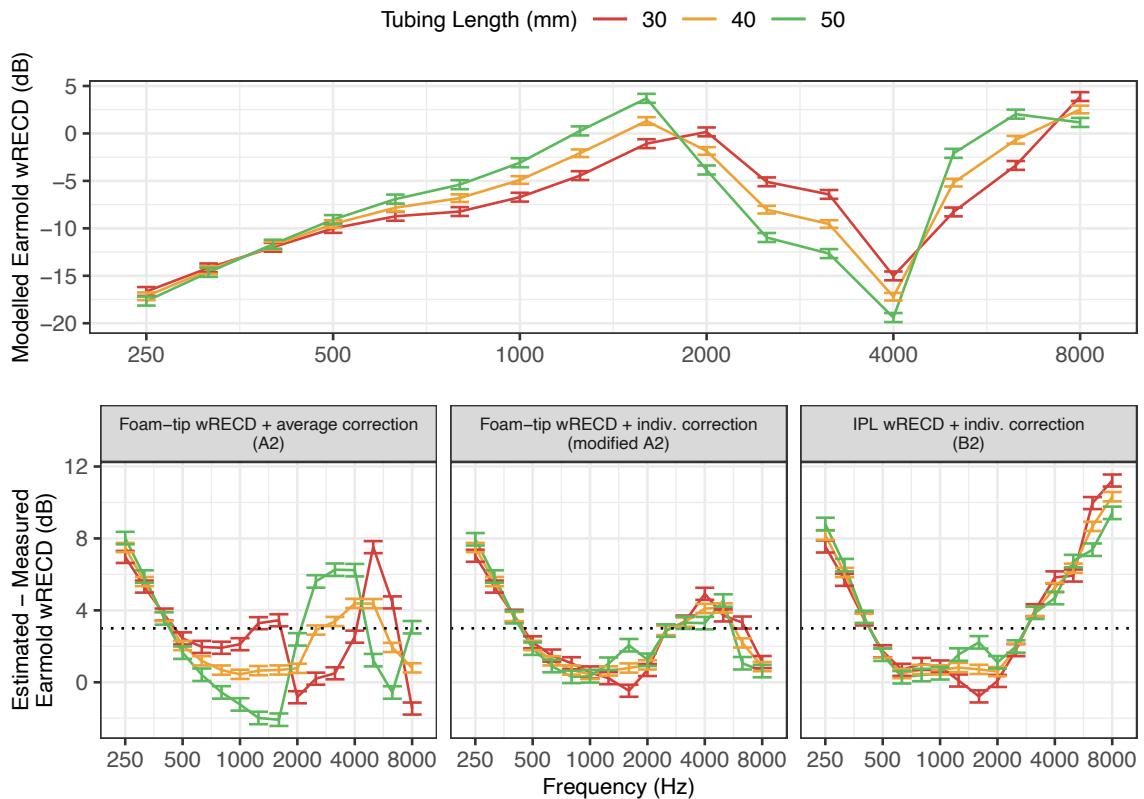


Figure 5-5: (Top row) Average modelled earmold wRECD for earmolds attached to 30, 40, and 50 mm of hearing aid tubing from probe-tube microphone measures. Error bars indicate ± 1 standard error. (Bottom row) Difference between measured and estimated earmold wRECD values for the three methods of earmold wRECD estimation shown in Figure 1. Clinically significant error is indicated by the dotted line. Positive values indicate larger estimated wRECD values than measured by probe-tube microphone.

5.3.3 Validation of individualized hearing tubing length correction

Implementation of the individualized hearing-aid tubing length correction minimized systematic deviations in earmold wRECD estimation when compared to averaged transforms (Figure 5-5). As seen in the top row of Figure 5-5, there is a systematic effect of hearing aid tubing length on the measured earmold wRECD. All else being equal, shorter tubing lengths result in smaller earmold wRECD values for frequencies below 2 kHz when compared to the same measurement completed with a longer segment of tubing. Above 2 kHz, the shorter tubing length has an opposite effect, producing larger

earmold wRECD values than are found when longer tubing is used. When individualized hearing aid transforms are incorporated into estimated earmold wRECDs, whether using probe-tube microphone measured foam-tip wRECD estimations (Figure 5-5; bottom center) or for earmold IPL wRECD measurement (Figure 5-5; bottom right), systematic variation caused by tubing length is minimized. In contrast, when average transforms are used, systematic error caused by tubing length is apparent (Figure 5-5; bottom left). These corrected plots indicate, at least descriptively, that the acoustic effects of hearing aid tubing length are being accounted for in the measurement. The individualized transform with probe-tube microphone did not account for all differences observed between estimated and measured earmold wRECD. An unexpected overestimation by probe-tube microphone measured foam-tip wRECD (with individualized hearing aid tubing correction) was also observed in the mid-to-high frequencies (Figure 5-5B; right panel). This unexpected difference exceeded average measured wRECD values by 3.21 dB ($CI_{95} = 2.67-3.75$) at 3150 Hz, 3.92 dB at 4000 Hz ($CI_{95} = 3.38-4.45$), and 3.94 dB ($CI_{95} = 3.40-4.47$) at 5000 Hz, independent of the tubing length attached to the earmold.

Errors associated with estimating tubing segment length with an individualized tubing length correction are shown in Figure 5-6, which shows the systematic deviation in “estimate error” modelled for earmold wRECD completed with 50 mm of tubing and 30 mm of tubing. If the individualized tubing length correction fully accounted for all the acoustic differences caused by changes in tubing length, systematic deviations should be 0 dB. Any frequency where the 95% confidence interval, indicated by error bars, does not cross the 0 threshold indicates a statistically significant effect of hearing aid tubing on estimated earmold wRECD error. This effect is most pronounced near 1600 Hz and above 5 kHz. These results indicate that the individualized tubing length corrections overcorrected at 1600 Hz and 8000 Hz, while under-correcting at 6300 Hz. However, these differences are much smaller than those seen in the currently implemented average foam-tip-to-earmold corrections, which can result in as much as ± 6 dB error above 1000 Hz due to individual variation in tubing segment length. Both the average and individualized methods of accounting for tubing lengths provide a better estimate of the

earmold wRECD than a foam-tip wRECD without any corrections for the additional tubing length segment associated with the individualized earmold.

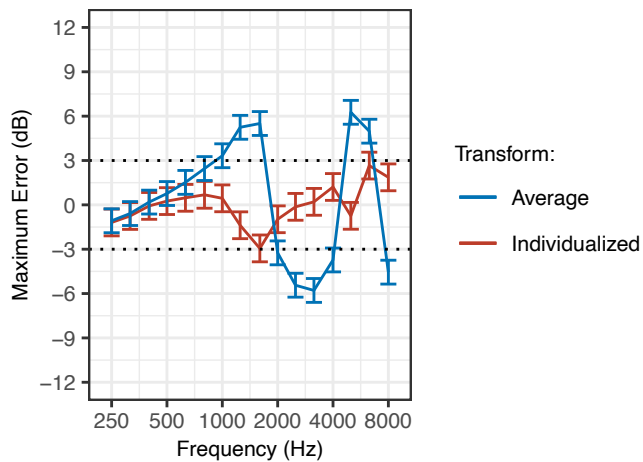


Figure 5-6: Differences in estimation of earmold wRECD for two extreme tubing lengths (average 50 mm earmold wRECD minus average 30 mm earmold wRECD) using two tubing-length compensation methods: individualized IPL (red) and average (blue). Clinically significant cut-offs are shown by the dotted lines. Error bars indicate 95% confidence interval for each estimate ($2*SE$).

5.3.4 Test-retest reliability of IPL wRECD measurement

IPL wRECD was highly reliable in all test subjects, with median differences of less than 1.5 dB across all 1/3 octave bands (Figure 5-7B). All individual test-retest differences fell below the 3 dB clinically significant criterion at all 1/3 octave bands for over 95% of the participants. In comparison, intrasession test-retest reliability of wRECD measurement using re-analysis of data collected by Vaisberg and colleagues (2018) found that absolute test-retest differences exceeded 3 dB in more than 25% of all measurements at 250, 6300, and 8000 Hz. The test-retest of both probe-tube microphone measured foam-tip wRECD and IPL wRECD of the current study was analyzed at each frequency-band using intraclass correlation coefficients for test-retest reliability, based on absolute agreement between two measurements. Reliability was classified as either excellent ($ICC > 0.9$), good ($0.9 > ICC > 0.75$), moderate ($0.75 > ICC > 0.5$), or poor ($0.5 > ICC$; Koo & Li, 2016). ICC values across 1/3 octave bands for each method of wRECD measurement are

found in Table 5-1. IPL wRECD ICC values come from the data collected for the present study, while probe-tube microphone measured foam-tip and earmold wRECD ICC values are obtained from a reanalysis of data from Vaisberg et al., (2018). IPL wRECD was found to be more reliable than probe-tube microphone wRECD measurement across the entire frequency range. At frequencies of 4000 Hz or greater, IPL wRECD reliability was good-to-excellent while probe-tube microphone foam-tip wRECD values produced poor-to-moderate reliability. In the mid-frequencies between 1000-4000 Hz, IPL wRECD had excellent reliability whereas probe-tube microphone measured foam-tip wRECD yielded moderate-to-good reliability. Finally, low-frequency reliability in IPL wRECD was found to be good, whereas probe-tube microphone foam-tip wRECD reliability was found to be moderate, likely due to the presence of slit-leakage. The current results suggest that IPL wRECD is a valid and reliable alternative to probe-tube microphone measured foam-tip wRECD measurement. Given that no clinically significant differences are present below 5 kHz, it may be possible to substitute IPL wRECD measurement for probe-tube microphone measured foam-tip wRECD measurement without the need for further corrections.

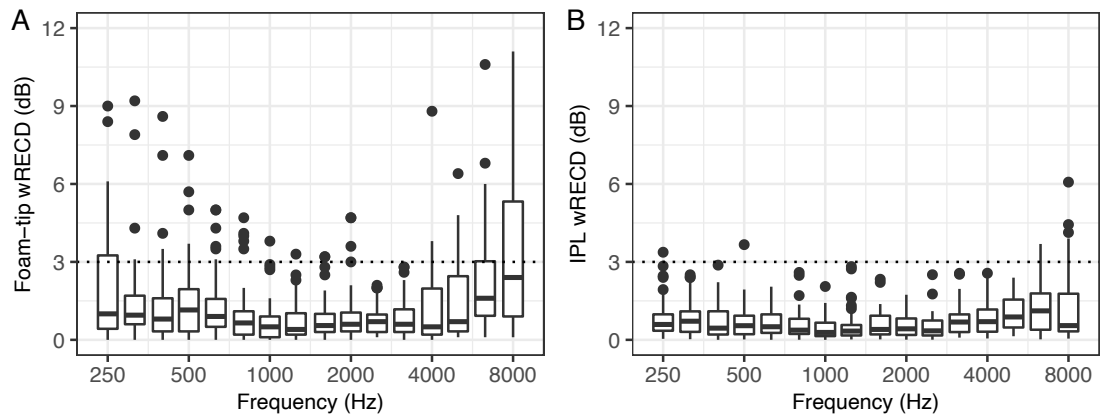


Figure 5-7: Absolute test-retest differences in decibels for each wRECD testing methodology across 1/3 octave bands. Boxes indicate 25-75th percentile; horizontal line indicates median test-retest difference; whiskers represent $1.5 * IQR$. Individual outlier points are denoted using a single black dot. (a) Test-retest data taken from 22 individuals, as reported in Vaisberg (2016) for probe-tube microphone wRECD measurements using a generic foam insert. (b) Test-retest data obtained in the current cohort of 21 individuals for IPL wRECD determination.

Frequency (Hz)	IPL wRECD	Foam-tip wRECD	Earmold wRECD
250	0.84 [0.72-0.91]	0.69 [0.50-0.82]	0.64 [0.42-0.79]
500	0.88 [0.79-0.94]	0.61 [0.39-0.77]	0.74 [0.57-0.85]
1,000	0.96 [0.93-0.98]	0.73 [0.55-0.85]	0.86 [0.76-0.92]
2,000	0.95 [0.90-0.97]	0.69 [0.49-0.82]	0.89 [0.81-0.94]
2,500	0.96 [0.92-0.98]	0.85 [0.73-0.91]	0.82 [0.69-0.90]
3,150	0.94 [0.90-0.97]	0.74 [0.57-0.85]	0.76 [0.60-0.87]
4,000	0.93 [0.88-0.96]	0.50 [0.23-0.70]	0.72 [0.53-0.84]
5,000	0.93 [0.87-0.96]	0.49 [0.23-0.69]	0.84 [0.73-0.91]
6,300	0.93 [0.88-0.96]	0.44 [0.16-0.65]	0.55 [0.30-0.73]
8,000	0.89 [0.80-0.94]	0.70 [0.50-0.82]	0.70 [0.51-0.83]

Table 5-1: Intraclass correlation coefficients for select 1/3 octave bands. Values within square brackets indicate 95% confidence intervals.

5.3.5 Performance of individualized ear canal cross-sectional area estimation

Individualized characteristic impedance measurement significantly improved agreement between IPL wRECD measurement and probe-tube microphone measured foam-tip wRECD values when compared to average characteristic impedance estimates (Figure 5-8). The accuracy of IPL wRECD determination using either measured or averaged characteristic impedance was quantified at and below 4 kHz by determining the absolute error between the IPL wRECD and the probe-tube microphone foam-tip wRECD. The absolute error between the IPL wRECD and probe-tube microphone measured wRECD was coded as the dependent variable in a mixed linear model with fixed effects (with interaction) of characteristic impedance method and the measurement frequency. Random intercepts of participant and random slope of characteristic impedance method were

included. There was a significant main effect of frequency band on overall error ($\chi^2_{11} = 566.4$, $p < 0.001$), with expected error generally increasing with frequency regardless of Z_0 determination method used. There was also a significant main effect of using individualized characteristic impedance measurement on overall error ($\chi^2_1 = 10.05$, $p = 0.002$). Individualized characteristic impedance measurement decreased absolute error by 0.49 dB when averaged across frequency ($CI_{95} = 0.39-0.59$). The interaction between characteristic impedance determination method and frequency was significant ($\chi^2_{12} = 236.39$, $p < 0.001$). Consistent with previous studies, individualized characteristic impedance measurement improved IPL determination in frequencies near the quarter-wavelength nulls during in-situ IPL measurement (approximately 2-3kHz; Figure 5-8). More specifically, individualized characteristic impedance measurement improved IPL wRECD agreement with probe-tube microphone measurements at all 1/3 octave bands between (and including) 1600Hz – 3100Hz. At 1600 Hz, 2000 Hz, 2500 Hz, and 3100 Hz, individualized Z_0 estimation improved agreement with wRECD values by 0.74 dB ($CI_{95} = 0.31-1.17$), 1.6 dB ($CI_{95} = 1.28 - 2.00$), 2.44 dB ($CI_{95} = 2.01-2.86$), and 1.00 dB ($CI_{95} = 0.57-1.43$) respectively. At 4000 Hz and above, agreement with probe-tube microphone foam-tip wRECD does not necessarily indicate increased or decreased improvement, due to the standing-wave errors present in the probe-tube microphone measurement.

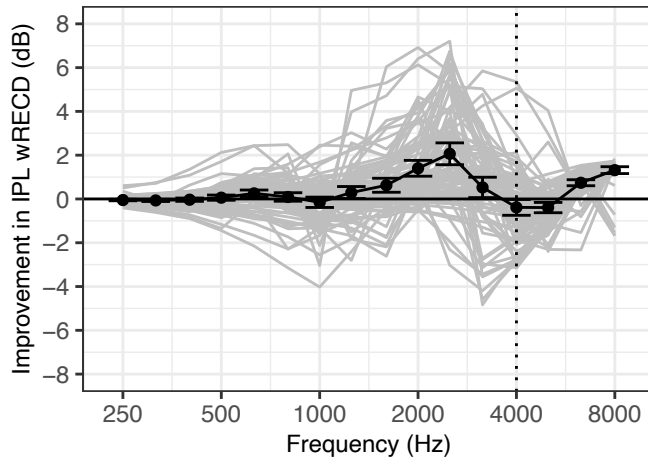


Figure 5-8: Improvement in agreement between the IPL wRECD and probe-tube microphone foam-tip wRECD when an individualized measurement of characteristic impedance (obtained with individualized cross-sectional area) is used compared to average characteristic impedance. Gray lines show individuals, and the black line is the group average. Error bars indicate 95% confidence interval for each frequency. A dotted line indicates 4000 Hz (see text for discussion).

5.4 Discussion

5.4.1 Main findings

We compared IPL wRECD values to SPL-based probe-tube microphone measured foam-tip and earmold wRECD values. Although statistically significant differences were noted between IPL wRECD and probe-tube microphone measured foam-tip wRECD values, these differences were only clinically significant above 4 kHz, where standing wave errors are expected to be present in probe-tube microphone measures but not IPL response measurement. At all frequencies up to 4 kHz, where probe-tube microphone measurements are expected to approximate the true SPL present at the eardrum, no clinically significant differences were found. Therefore, it appears that the advantages of IPL for reducing error associated with standing waves, as reported in previous studies of stimulus level calibration in the ear canal (Souza et al., 2014; Withnell et al., 2009), also apply to the measurement of the wRECD.

Similar to comparisons with foam-tip wRECD determination, earmold IPL wRECD measurements produced larger high frequency values than probe-tube microphone earmold wRECD measurements. Such a deviation was anticipated, due to the reduction of standing wave error. Specifically, earmold IPL wRECD, calculated from IPL wRECD and the individualized foam-tip-to-earmold correction, fell within 3 dB of the probe-tube microphone measured earmold wRECD for all frequencies between 500-2500 Hz regardless of the earmold's tubing length. At frequencies below 500 Hz, both earmold IPL wRECD and estimated earmold wRECD produced larger values than the probe-tube microphone measured earmold wRECD due to probe-tube-induced slit-leak error being present in the probe-tube microphone earmold wRECD measurement. Unexpectedly, differences between the probe-tube microphone measured earmold wRECD and earmold wRECD estimated using probe-tube microphone measured foam-tip wRECD were observed near 4 kHz. This systematic deviation was independent of the tubing length, indicating potential systematic differences in acoustic responses between coupling methods unrelated to the tubing length. Individualized foam-tip-to-earmold correction transforms minimized error in earmold IPL wRECD and estimated earmold wRECD caused by tubing length variability.

Test-retest reliability of IPL wRECD was also assessed. IPL wRECD reliability was excellent, with almost all absolute errors between measurements falling below a clinical significance criterion of 3 dB across all analyzed frequencies. The test-retest reliability of IPL wRECD was found to be better than the reliability of probe-tube microphone measurements, as measured by each method's ICC values. Reliability improvements were especially marked in the low frequencies and above 4 kHz, due to reduced slit-leak error and standing-wave error, respectively. The current results support the replacement of probe-tube microphone measured foam-tip wRECD measurement with the IPL wRECD measurement paradigm, at least in an adult population. IPL wRECD improves test-retest reliability, high-frequency accuracy, and does not significantly change wRECD determination in frequencies where probe-tube microphone measurements are known to be valid. Further evaluation of this measurement in children is indicated.

The individualized determination of characteristic impedance (Rasetshwane & Neely, 2011) improved IPL wRECD agreement with probe-tube microphone measured wRECD values in the mid-frequencies by as much as 2.5 dB at 2500 Hz. The characteristic impedance measurement was obtained with the acoustic impedance during the same measurement as the IPL wRECD, requiring no additional time or changes in testing.

5.4.2 Comparison between IPL wRECD and foam-tip wRECD

Clinical recommendations for probe-tube insertion depth are largely based on minimizing the influence of the reflected-wave interference on the probe-tube microphone measurement below 4 kHz. Even with gold-standard probe-tube placement, standing waves still influence probe-tube microphone measurements as low as 4 kHz, as shown by decreased reliability in the measurement (McCreery et al., 2009; Vaisberg et al., 2016). Based on the findings in the current study, hearing aid fittings that incorporate probe-tube microphone measurements result in underestimation of the individual's wRECD above 4 kHz. The magnitude of the underestimation corresponds to the depth of the standing-wave in the individual. Standing wave error depends on probe-tube microphone placement and can be significantly shifted by even small (less than or equal to 2 mm) changes in probe-tube location (Vaisberg et al., 2016). Visual detection of the $\frac{3}{4}$ -wavelength standing wave in the frequency band of interest can help minimize error by allowing exclusion of that measurement from datasets, or in the case of clinical measurement, repositioning of the probe-tube microphone. However, the depth and width of the $\frac{3}{4}$ -wavelength standing wave depends on numerous factors, such as ear canal geometry, which makes accurate and reliable detection difficult (Chan & Geisler, 1990; Stinson & Lawton, 1989). Such challenges might explain why the average IPL wRECD value was greater than current adult norms even when standing waves were visually inspected and excluded when normative curves were developed (Bagatto et al., 2005). Regardless, source-parameter calibrated approaches such as IPL have been shown to be independent of the quarter-wavelength standing wave, and accurately measure behavioural thresholds, showcasing significantly less error in the higher frequencies when compared to SPL calibrated stimuli (Lapsley Miller et al., 2018; McCreery et al., 2009; Souza et al., 2014; Withnell et al., 2014). As such, the high-frequency difference

between the measured IPL wRECD values and probe-tube microphone measured foam-tip wRECD values in the present study should be interpreted primarily as an underestimation of wRECD by probe-tube microphone measurements rather than an overestimation of IPL wRECD measurements.

5.4.3 Comparison between earmold IPL wRECD and probe-tube microphone measured earmold wRECD and validation of individualized tubing length corrections

Averaged across all measurement conditions, earmold IPL wRECD and probe-tube measured earmold wRECD produced similar distributions in the mid-frequency region (Figure 5-4). However, earmold wRECD values were comparatively small in both the low- (≤ 500 Hz) and high- (≥ 4000 Hz) frequency regions, due to slit-leakage and standing-wave interference, respectively. Recall that the acoustic response of each individual tubing length was estimated using the acoustic length estimation measurement by the source parameter calibrated transducer, using methods described by Urichuk et al., (2021b). This allows a two-stage estimation of the earmold wRECD. In the first stage, the IPL wRECD (with generic insert) is measured in the ear. In the second stage, the length of the tubing is estimated using the response of the earmold tubing measured on the desk, which takes seconds to complete. The length of the hearing aid tubing is used to create an individualized foam-tip-to-earmold transform for that hearing aid, which is added to the IPL wRECD values to produce an estimate of the earmold IPL wRECD. This is similar in concept to the foam-tip-to-earmold transform developed by Moodie and colleagues (2016), which uses the average acoustic effect of hearing aid tubing lengths between 30 and 50 mm in length. By acoustically determining the length of the tubing, the impact of the tubing length can be determined more accurately, allowing for different foam-tip-to-earmold transforms to be used based on tubing length. Individualized foam-tip-to-earmold transforms largely eliminate error in earmold wRECD estimation caused by variation in tubing length that have been observed in previous studies (Gustafson et al., 2013; Moodie et al., 2016; Munro & Davis, 2003), as well as in the current study (Figure 5-5; top row). The acoustic length measurement did not require additional cooperation from the patient as it is completed off-ear and took only seconds to complete. Mid-frequency earmold wRECD estimation using the IPL wRECD measurement was not

significantly different than direct earmold wRECD measurement, validating the estimation procedure. Furthermore, estimation using IPL wRECD measurement minimized slit leak errors that may be caused by gaps between the probe-tube microphone and ear-mold that are not present when the individual is wearing the hearing aid outside of this verification measurement. Slit leak errors are further minimized by robust leak-detection criteria resulting from the simultaneous wideband acoustic immittance measurement, adapted for use in the current study (Groon et al., 2015). The larger values obtained in the high-frequencies for earmold IPL wRECD relative to the probe-tube microphone approach were similar in both shape and magnitude to the same comparison with probe-tube microphone measured foam-tip wRECD. The known presence of standing-wave error in both foam-tip and ear-mold wRECD is the most likely explanation for the observed effect.

Tubing-length independent differences were observed between directly measured earmold wRECD and indirect earmold wRECD measurement approaches, namely the earmold IPL wRECD and estimated earmold wRECD where the individualized foam-tip-to-earmold transform was used. If the tubing length was the only source of difference, we would expect to see no differences between the directly measured and the estimated earmold wRECD that uses the individualized foam-tip-to-earmold transform. It is possible that this clinically significant error between 3-6 kHz may be caused by either unintentional probe-tube depth differences, overestimation of the foam-tip-to-earmold correction, or an unaccounted-for acoustic difference between coupling methods and requires further investigation. Nonetheless, incorporation of the individualized tubing correction minimized systematic ± 6 dB variations in both earmold IPL wRECD and estimated earmold wRECD measurement caused by tubing length differences. Estimating the acoustic effects of individual hearing aid tubing length eliminated the statistically significant impacts of the tubing length on the accuracy of foam-tip-to-earmold wRECD estimation.

While IPL calibrated measurements accurately quantify the sound-input for a particular transducer placement, systematic errors in difference measurements may be present near the half-wavelength resonance of the residual ear canal length. Transducer placement

depth alters the frequency of this half-wavelength resonance (Souza et al., 2014). While these half-wavelength resonances can exceed 10 dB, there will be significant overlap between resonances in two adequate transducer placements, leading to minimal error introduced. This is supported by the relatively minimal test-retest variability noted in IPL wRECD measurement. However, if two transducer placements differ substantially from one another, such as a relatively shallow immittance tip placement compared with a deep earmold placement, errors due to half-wavelength resonance differences may be introduced at frequencies as low as 8 kHz in adult ear canals. These errors are of most concern in longer, adult ear canals due to the frequency location of the resonance being inversely proportional to the residual ear canal length. In shorter ear canals and, perhaps more importantly, for infants and young children, the shift will be well beyond the bandwidth of interest and will not be a significant source of error for wRECD measurement. In the current study, the deep probe-placements, relatively shallow earmolds, and small ears in the sample population make it unlikely that the half-wave resonance was a significant source of error. The potential for such errors still highlights the importance of deep transducer placements – regardless of the acoustic measurement being completed. While transducer-depth mismatches have the potential to introduce errors in high frequencies, it should be noted that such errors will be significantly smaller in magnitude, and at significantly higher frequencies than the quarter-wavelength standing-wave errors currently present in clinical measurements.

The current results suggest that measurement of an earmold IPL wRECD does not showcase clinically meaningful errors due to hearing aid tubing length and that earmold IPL wRECD values fall within 3 dB of probe-tube microphone measured earmold wRECD between 500-2500 Hz, regardless of the tubing length. As a result, the earmold IPL wRECD builds upon the individual results obtained by Moodie and colleagues (2016) and is recommended as a more accurate alternative to estimating the earmold wRECD using a foam-tip wRECD measurement and the average foam-tip-to-earmold transform. Results between direct probe-tube microphone measured earmold wRECD and the indirect earmold IPL wRECD differed in high and low frequency regions. While it is expected that this is mainly caused by slit leakage and standing wave errors, it is not possible to determine which measurement would be more accurate in a clinical setting.

Measurements using a wider set of earmold materials and styles, across a wider range of ages, are necessary to further generalize systematic deviations between wRECD measurements made with generic inserts and personalized earmolds. Earmold audiograms could also be completed using stimuli calibrated using each earmold wRECD method to isolate which method provides more accurate results in the extreme frequency ranges, similar to audiometric comparisons completed in previous studies (McCreery et al., 2009; Souza et al., 2014).

5.4.4 Performance of individualized ear canal cross-sectional area estimation

Measuring the individual's characteristic impedance significantly improved agreement of the IPL wRECD with probe-tube microphone foam-tip wRECD measurement up to 4 kHz when compared to the average characteristic impedance. Improvements of individualized characteristic impedance were largest in the mid-frequencies, consistent with previous studies (Scheperle et al., 2011; Souza et al., 2014). The largest improvement was observed near the quarter-wavelength resonance with mean improvement within 3 dB at all frequencies. The magnitude and frequency of this improvement is relatively consistent with improvements seen by Souza and colleagues (2014), who found minor (~1 dB) improvements on group-level threshold determination reliability when using the individualized characteristic impedance measurement. In the present study, improvement in real-ear IPL determination beyond 4 kHz could not be assessed using probe-tube microphone measurements due to standing-wave interference present in probe-tube microphone measurements. Results in the current study add to the literature by providing further validation of the Z_0 measurement procedure proposed by Rasetshwane and Neely (2011).

Overall improvements in agreement due to individual estimation of Z_0 are modest, however, the improvements are expected to be greatest in ears that significantly deviate from the average Z_0 . Although the statistical impact of radius was not formally assessed in the current study due to sample size limitations and deviating from the study's main aim, visual inspection of the data found a pronounced improvement caused by individualized Z_0 measurement in ear canals significantly smaller than the reference

coupler. The smallest ears showcased some of the largest improvement of Z_0 measurement over averaged methods. In smaller ears, benefit of Z_0 measurement routinely exceeded 5 dB at 2500 Hz. This is of particular interest for the accurate calibration of stimuli in small ear canals, such as those similar to infant ears, and validates the implementation of Z_0 measurement.

5.5 Conclusion

IPL wRECD, although theoretically intimidating, is a simple clinical measurement, requiring similar precision and less time than a typical 226-Hz tympanometric test and yields the individual's wRECD and a robust middle ear analysis using wideband acoustic immittance. The speed of the measurement, lack of sustained cooperation necessary, and extended valid bandwidth make IPL wRECD an enticing clinical measurement for implementation into hearing instrument fitting workflows. No probe-tube placement is necessary, and such a measurement can be completed with minimal training. This introduces exciting potential and implications for tele-audiological and screening environments where a highly trained practitioner may not be physically present.

5.6 Acknowledgements

This work was supported by Ontario Research Fund (RE-08-072). The authors are grateful to Bernafon Canada for providing all earmolds used in the current study and to Jonathan Vaisberg for sharing raw data from previous work. The authors are also grateful to Audioscan, a Division of Etymotic Design Incorporated, and Interacoustics A/S for their technical support throughout the study.

5.8 References

- Alexander, J. M., & Rallapalli, V. (2017). “Acoustic and Perceptual Effects of Amplitude and Frequency Compression on High-Frequency Speech.” *The Journal of the Acoustical Society of America*, 142(2): 908–923.
<https://doi.org/10.1121/1.4997938>.
- 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.
- Bagatto, Moodie, S., Scollie, S., Seewald, R., Moodie, S., Pumford, J., & Liu, K. P. R. (2005). Clinical Protocols for Hearing Instrument Fitting in the Desired Sensation Level Method. *Trends in Amplification*, 9(4), 199–226.
<https://doi.org/10.1177/108471380500900404>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
<https://doi.org/10.18637/jss.v067.i01>
- Brennan, M. A., McCreery, R., Kopun, J., Hoover, B., Alexander, J., Lewis, D., & Stelmachowicz, P. G. (2014). Paired Comparisons of Nonlinear Frequency Compression, Extended Bandwidth, and Restricted Bandwidth Hearing Aid Processing for Children and Adults with Hearing Loss. *Journal of the American Academy of Audiology*, 25(10), 983–998. <https://doi.org/10.3766/jaaa.25.10.7>
- Chan, J. C. K., & Geisler, C. D. (1990). Estimation of eardrum acoustic pressure and of ear canal length from remote points in the canal. *The Journal of the Acoustical Society of America*, 87(3), 1237–1247. <https://doi.org/10.1121/1.398799>
- Dirks, Donald & Kincaid, Gail. (1987). Basic Acoustic Considerations of Ear Canal Probe Measurements. *Ear and Hearing*, 8(5), 60–67.
- Fabry, David. (2003). Nonlinear Hearing Aids and Verification of Fitting Targets. *Trends in Amplification*, 7(3). 99-115.
- Feigin, J. A., Kopun, J., Stelmachowicz, P., & Gorga, M. (1989). Probe-tube microphone measures of ear-canal sound pressure levels in infants and children. *Ear and Hearing*, 10(4), 254–258.
- Folkeard P, Pumford J, Pietrobon J, Scollie S. (2019). Evaluation of Probe Guide: Software-assisted probe tube placement in hearing aid fittings. *Hearing Review*. 26(11).
- Folkeard, P., Van Eeckhoutte, M., Levy, S., Dundas, D., Abbasalipour, P., Glista, D., Agrawal, S., Scollie, S. (2021). Detection, Speech Recognition, Loudness, and Preference Outcomes With a Direct Drive Hearing Aid: Effects of Bandwidth. *Trends in Hearing* 25, 1-17.
- Füllgrabe, C., Baer, T., Stone, M. A., & Moore, B. C. J. (2010). Preliminary evaluation of a method for fitting hearing aids with extended bandwidth. *International Journal of Audiology*, 49(10), 741–753. <https://doi.org/10.3109/14992027.2010.495084>

- Groon, K. A., Rasetshwane, D. M., Kopun, J. G., Gorga, M. P., & Neely, S. T. (2015). Air-Leak Effects on Ear-Canal Acoustic Absorbance: *Ear and Hearing*, 36(1), 155–163. <https://doi.org/10.1097/AUD.0000000000000077>
- Gustafson, S., Pittman, A., & Fanning, R. (2013). Effects of Tubing Length and Coupling Method on Hearing Threshold and Real-Ear to Coupler Difference Measures. *American Journal of Audiology*, 22(1), 190–199. [https://doi.org/10.1044/1059-0889\(2012/12-0046\)](https://doi.org/10.1044/1059-0889(2012/12-0046))
- Hornsby, B. W. Y., Johnson, E. E., & Picou, E. (2011). Effects of Degree and Configuration of Hearing Loss on the Contribution of High- and Low-Frequency Speech Information to Bilateral Speech Understanding: *Ear and Hearing*, 32(5), 543–555. <https://doi.org/10.1097/AUD.0b013e31820e5028>
- International Electrotechnical Commission (IEC). (2016). *Electroacoustics – Hearing aids – Method for measuring electroacoustic performance up to 16 kHz*. IEC TS 62886:2016.
- Keefe, D. H., Bulen, J. C., Arehart, K. H., & Burns, E. M. (1993). Ear-canal impedance and reflection coefficient in human infants and adults. *The Journal of the Acoustical Society of America*, 94(5), 2617–2638. <https://doi.org/10.1121/1.407347>
- Kenny, S. (2011). Clinical application of the interacoustics REFLWIN system wideband reflectance machine in the assessment of the eustachian tube. *Unpublished master's thesis*, University of British Columbia, Vancouver, Canada.
- Lapsley Miller, J. A., Reed, C. M., Robinson, S. R., & Perez, Z. D. (2018). Pure-Tone Audiometry With Forward Pressure Level Calibration Leads to Clinically-Relevant Improvements in Test–Retest Reliability: *Ear and Hearing*, 39(5), 946–957. <https://doi.org/10.1097/AUD.0000000000000555>
- Levy, S. C., Freed, D. J., Nilsson, M., Moore, B. C. J., & Puria, S. (2015). Extended High-Frequency Bandwidth Improves Speech Reception in the Presence of Spatially Separated Masking Speech: *Ear and Hearing*, 36(5), e214–e224. <https://doi.org/10.1097/AUD.0000000000000161>
- Lewis, J. D., McCreery, R. W., Neely, S. T., & Stelmachowicz, P. G. (2009). Comparison of *in-situ* calibration methods for quantifying input to the middle ear. *The Journal of the Acoustical Society of America*, 126(6), 3114–3124. <https://doi.org/10.1121/1.3243310>
- McCreery, R. W., Pittman, A., Lewis, J., Neely, S. T., & Stelmachowicz, P. G. (2009). Use of forward pressure level to minimize the influence of acoustic standing waves during probe-microphone hearing-aid verification. *The Journal of the Acoustical Society of America*, 126(1), 15–24. <https://doi.org/10.1121/1.3143142>
- McCreery, R. W., Walker, E. A., Spratford, M., Bentler, R., Holte, L., Roush, P., Oleson, J., Van Buren, J., & Moeller, M. P. (2015). Longitudinal Predictors of Aided Speech Audibility in Infants and Children. *Ear and hearing*, 36 Suppl 1(0 1), 24S–37S. <https://doi.org/10.1097/AUD.0000000000000211>

- Moodie, Shane, Seewald, R. C., & Sinclair, S. (1994). Procedure for Predicting Real-Ear Hearing Aid Performance in Young Children. *American Journal of Audiology*, 3(1), 23–31. <https://doi.org/10.1044/1059-0889.0301.23>
- Moodie, S., Pietrobon, J., Rall, E., Lindley, G., Eiten, L., Gordey, D., Davidson, L., Moodie, K. S., Bagatto, M., Haluschak, M. M., 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. *Journal of the American Academy of Audiology*, 27(03), 264–275. <https://doi.org/10.3766/jaaa.15086>
- Moodie, S. Rall, E., Eiten, L., Lindley, G., Gordey, D., Davidson, L., Bagatto, M., & Scollie, S. (2016). Pediatric Audiology in North America: Current Clinical Practice and How It Relates to the American Academy of Audiology Pediatric Amplification Guideline. *Journal of the American Academy of Audiology*, 27(03), 166–187. <https://doi.org/10.3766/jaaa.15064>
- Moore, B. C. J., Glasberg, B. R., and Stone, M. A. (2010). Development of a new method for deriving initial fittings for hearing aids with multi-channel compression: CAMEQ2-HF *Int. J. Audiol.* 49, 216-227.
- Moore, B. C. J., M. A. Stone, and J. I. Alcantara. (2001). Comparison of the Electroacoustic Characteristics of Five Hearing Aids. *British Journal of Audiology* 35(5): 307–325. doi:[10.1080/00305364.2001.11745249](https://doi.org/10.1080/00305364.2001.11745249).
- Moore, B. C. J., Stone, M. A., Füllgrabe, C., Glasberg, B. R., & Puria, S. (2008). Spectro-Temporal Characteristics of Speech at High Frequencies, and the Potential for Restoration of Audibility to People with Mild-to-Moderate Hearing Loss: *Ear and Hearing*, 29(6), 907–922. <https://doi.org/10.1097/AUD.0b013e31818246f6>
- Mueller HG, Hall JW. (1998) Audiologists' Desk Reference. San Diego: Singular Publishing Group.
- Munro, Kevin J., & 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 and Hearing*, 24(2), 100–110. <https://doi.org/10.1097/01.AUD.0000058114.20741.4D>
- Munro, K. J., & Hatton, N. (2000). Customized Acoustic Transform Functions and Their Accuracy at Predicting Real-Ear Hearing Aid Performance. *Ear and Hearing*, 21(1), 59–69.
- Munro, K. J., & Salisbury, V. (2002). Is the real-ear to coupler difference independent of the measurement earphone. *International Journal of Audiology*, 41(7), 408–413.
- Munro, K. J., & Toal, S. (2005). *Measuring the Real-Ear to Coupler Difference Transfer Function With an Insert Earphone and a Hearing Instrument: Are They the Same?* 26(1), 8.
- Nørgaard, K. R., Fernandez-Grande, E., & Laugesen, S. (2017). Incorporating evanescent modes and flow losses into reference impedances in acoustic Thévenin calibration. *The Journal of the Acoustical Society of America*, 142(5), 3013–3024. <https://doi.org/10.1121/1.5010891>

- R Core Team. (2020). *R: A language and environment for statistical computing* [Manual]. <https://www.R-project.org/>
- Rasetshwane, D. M., & Neely, S. T. (2011). Inverse solution of ear-canal area function from reflectance. *The Journal of the Acoustical Society of America*, *130*(6), 3873–3881. <https://doi.org/10.1121/1.3654019>
- Ricketts, T. A., Dittberner, A. B., & Johnson, E. E. (2008). High-Frequency Amplification and Sound Quality in Listeners With Normal Through Moderate Hearing Loss. *Journal of Speech, Language, and Hearing Research*, *51*(1), 160–172. [https://doi.org/10.1044/1092-4388\(2008/012\)](https://doi.org/10.1044/1092-4388(2008/012))
- Rosowski, J. J., Stenfelt, S., & Lilly, D. (2013). An Overview of Wideband Immittance Measurements Techniques and Terminology: You Say Absorbance, I Say Reflectance. *Ear and Hearing*, *34*, 9s–16s. <https://doi.org/10.1097/AUD.0b013e31829d5a14>
- Saunders, G. H., & Morgan, D. E. (2003). Impact on hearing aid targets of measuring thresholds in dB HL versus dB SPL: El impacto en la medición de los umbrales en dB HL o en dB SPL, en las metas de un auxiliar auditivo. *International Journal of Audiology*, *42*(6), 319–326.
- Scheperle, R. A., Goodman, S. S., & Neely, S. T. (2011). Further assessment of forward pressure level for *in situ* calibration. *The Journal of the Acoustical Society of America*, *130*(6), 3882–3892. <https://doi.org/10.1121/1.3655878>
- Scheperle, R. A., Neely, S. T., Kopun, J. G., & Gorga, M. P. (2008). Influence of *in situ* , sound-level calibration on distortion-product otoacoustic emission variability. *The Journal of the Acoustical Society of America*, *124*(1), 288–300. <https://doi.org/10.1121/1.2931953>
- Scollie, S., Seewald, Richard, Cornelisse, Leonard, & Jenstad, Lorianne. (1998). Validity and Repeatability of Level-Independent HL to SPL Transforms. *Ear and Hearing*, *19*, 407–413.
- Seewald, R. C., Cornelisse L. E., Richert F. M., Block M.G. (1997). Acoustic transforms for fitting CIC hearing instruments. In Chasin M. (ed.) *CIC Handbook*. San Diego: Singular Publishing Group; 83–100.
- Souza, N. N., Dhar, S., Neely, S. T., & Siegel, J. H. (2014). Comparison of nine methods to estimate ear-canal stimulus levels. *The Journal of the Acoustical Society of America*, *136*(4), 1768–1787. <https://doi.org/10.1121/1.4894787>
- Stinson, M. (1990). Revision of estimates of acoustic energy reflectance at the human eardrum. *The Journal of the Acoustical Society of America*, *88*(4), 1773–1778. <https://doi.org/10.1121/1.400198>
- Stinson, M., & Lawton, B. W. (1989). Specification of the geometry of the human ear canal for the prediction of sound-pressure level distribution. *The Journal of the Acoustical Society of America*, *85*(6), 2492–2503. <https://doi.org/10.1121/1.397744>

- Tharpe, A. M., Sladen, D., Huta, H. M., & Rothpletz, A. M. (2001). Practical Considerations of Real-Ear-to-Coupler Difference Measures in Infants. *American Journal of Audiology*, 10(1), 41–49. [https://doi.org/10.1044/1059-0889\(2000/006\)](https://doi.org/10.1044/1059-0889(2000/006))
- Urichuk, M., Purcell, D., Allen, P., Scollie, S. (2021a). *Thevenin-equivalent source parameter reliability: intra- and inter-session effects on stimulus level calibration and wideband reflectance*. Manuscript submitted for publication.
- Urichuk, M., Purcell, D., Scollie, S. (2021b). *Measurement of hearing aid tubing segments using a wideband stimulus*. Manuscript submitted for publication.
- Vaisberg, J. M., Beaulac, S., Glista, D., Macpherson, E., Scollie, S. (2021). Perceived Sound Quality Dimensions Influencing Frequency-Gain Shaping Preferences for Hearing Aid-Amplified Speech and Music. *Trends in Hearing*, 25, 1-17
- Vaisberg, J. M., Folkeard, P., Levy, S., Dundas, D., Agrawal, S., Scollie, S. (2021b). Sound Quality Ratings of Amplified Speech and Music Using a Direct Drive Hearing Aid: Effects of Bandwidth. *Otology & Neurotology*, 42(2), 227-234
- Vaisberg, J. M., Folkeard, P., Pumford, J., Narten, P., & Scollie, S. (2018). Evaluation of the Repeatability and Accuracy of the Wideband Real-Ear-to-Coupler Difference. *Journal of the American Academy of Audiology*, 29(06), 520–532. <https://doi.org/10.3766/jaaa.17007>
- Vaisberg, J. M., Macpherson, E. A., & Scollie, S. (2016). Extended bandwidth real-ear measurement accuracy and repeatability to 10 kHz. *International Journal of Audiology*, 55(10), 580–586. <https://doi.org/10.1080/14992027.2016.1197427>
- Valente, M. (2006) Guideline for audiologic management of the adult patient. Audiology Online. <https://www.audiologyonline.com/articles/guideline-for-audiologic-management-adult-966>. Accessed May 3, 2021.
- Valente, M., Potts, L. G., Valente, M., Vass, W., & Goebel, J. (1994). Intersubject Variability of Real-Ear Sound Pressure Level: Conventional and Insert Earphones. *Journal of the American Academy of Audiology*, 5(6), 9.
- Van Eeckhoutte, M., Folkeard, P., Glista, D., & Scollie, S. (2020). Speech recognition, loudness, and preference with extended bandwidth hearing aids for adult hearing aid users. *International Journal of Audiology*, 59(10), 780–791. <https://doi.org/10.1080/14992027.2020.1750718>
- Voss, S. E., & Allen, J. B. (1994). Measurement of acoustic impedance and reflectance in the human ear canal. *The Journal of the Acoustical Society of America*, 95(1), 372–384. <https://doi.org/10.1121/1.408329>
- Voss, S. E., Merchant, G. R., & Horton, N. J. (2012). Effects of Middle-Ear Disorders on Power Reflectance Measured in Cadaveric Ear Canals: *Ear and Hearing*, 33(2), 195–208. <https://doi.org/10.1097/AUD.0b013e31823235b5>
- Watts, K. M., Bagatto, M., Clark-Lewis, S., Henderson, S., Scollie, S., Blumsack, J. (2020). Relationship of Head Circumference and Age in the Prediction of the Real-Ear-to-Coupler Difference (RECD). *Journal of the American Academy of Audiology* . 31(7), 496-505.

- Withnell, R. H., Jeng, P. S., Parent, P., & Levitt, H. (2014). The clinical utility of expressing hearing thresholds in terms of the forward-going sound pressure wave. *International Journal of Audiology*, *53*(8), 522–530. <https://doi.org/10.3109/14992027.2014.898122>
- Withnell, R. H., Jeng, P. S., Waldvogel, K., Morgenstein, K., & Allen, J. B. (2009). An *in situ* calibration for hearing thresholds. *The Journal of the Acoustical Society of America*, *125*(3), 1605–1611. <https://doi.org/10.1121/1.3075551>

Chapter 6

6 Contributions, limitations, and conclusions

6.1 Research aims and summary of findings

The aim of this thesis was to improve the detection and accuracy of objective measurements for procedures used in audiology, with a focus on pediatric procedures related to detection of auditory evoked potentials, and for in-ear stimulus calibration. Chapter 2 assessed the clinical efficacy of the speech-evoked envelope following response (EFR) as an objective outcome measurement across different testing durations and using an array of different statistical indicators in both infants and adults. Chapters 3-5 focused on the development of an individualized in-ear calibration using a Thevenin-equivalent source parameter calibrated transducer. More specifically, Chapter 3 assessed the test-retest reliability of the Thevenin-equivalent source calibration procedure and the impact of reliability on subsequent in-ear measurements. Chapter 4 proposed and validated an acoustic method of estimating an individual's hearing aid tubing length for use in individualized foam-tip-to-earmold wideband real-ear-to-coupler difference (wRECD) transforms. Chapter 5 incorporated the calibration approach discussed in Chapter 3 and the acoustic length estimation in Chapter 4 and validated a novel method of measuring wRECD using integrated pressure level (IPL) without the use of probe-tube microphones. Validity and reliability of the IPL wRECD procedure was assessed against current foam-tip and earmold wRECD measurement approaches.

Chapter 2 found that speech-evoked envelope-following response (EFR) detection is dependent on the statistical indicator used. The results in this chapter indicated that statistical indicator accuracy is not fixed across the lifespan and the accuracy differs depending on the population of interest (children vs. adults). This indicates a likely maturational difference between populations. Statistical indicators that used the phase of the speech-evoked EFR were the most accurate and efficient in infant response detection, which has also been observed previously in adult populations (Easwar et al. 2020, Zhu et al., 2013). Phase-incorporating statistical indicators provided clinical utility as an objective outcome measurement tool with less than 30 minutes of testing duration. This

chapter contributed by providing a comparison of statistical indicators as a function of testing duration and by highlighting differences that exist between infant and adult populations. A limitation of this chapter was the use of a broadband real-ear-to-coupler-difference rather than the frequency-specific RECD that is used frequently in clinical practice. Accounting for the individual's ear-canal acoustics improves the accuracy and the validity of EFR detection (Easwar et al., 2021). However, the placement of a probe-tube microphone is not always possible during testing and may be difficult to complete prior to EFR measurement which requires the child to sleep/be relaxed during the duration of testing. The second half of this thesis (Chapter 3-5) proposed and validated a method to measure the wRECD simultaneously with a robust analysis of the middle ear to alleviate this clinical barrier. In this work, the integrated pressure level wRECD measurement was found to provide gold-standard calibration, improve high-frequency wRECD validity, and require no probe-tube microphone placement or sustained cooperation from the individual. The implementation of the IPL procedure would allow for the incorporation of wRECD measurement into speech-evoked EFR measurement protocols, as well as other hearing aid verification procedures.

Chapters 3 and 4 also addressed several barriers that could be introduced using source-parameter calibrated wRECD measurement in clinical practice. More specifically, Chapter 3 analyzed the reliability of the Thevenin-equivalent source-parameter calibration procedure over multiple sessions. If source-parameter calibrated stimuli were implemented in a clinical setting, knowledge of the within-session and between-session reliability, and how frequently a calibration needs to be completed is required. Previous work has varied in how often calibration occurs, ranging from before any given measurement, to daily, to weekly. The current work confirmed that weekly calibration is sufficient for both accurate stimulus calibration *and* middle-ear assessment, minimizing concerns that a clinician may have with such a calibration procedure. Furthermore, the duration of calibration was relatively short (less than 5 minutes) and is completed on calibration cavities without a patient present. Taken together, this attribute of weekly, brief calibration provide further information about the clinical feasibility of source-parameter techniques.

Source-parameter calibrated stimuli require measurement using a transducer with a microphone flush with the speaker. At time of publication, there is no clinically available method of determining the source-parameters of a hearing aid with a variable length of hearing aid tubing with the individual's personalized earmold. The coupling method, especially the length of tubing, has a significant impact on the acoustics of the signal that cannot be easily determined given an acoustic measurement using a generic acoustic immittance or foam tip (Gustafson et al., 2013; Moodie et al., 2016; Munro & Davis, 2003). Chapter 4 proposed a method of acoustically determining the length of hearing aid tubing coupled in an individual's hearing aid using the source-parameter calibrated transducer. With the acoustic length known, individualized foam-tip-to-earmold wRECD transforms were developed based on measurements made in an HA-1 2.0cc coupler and two simulated ear-canals with similar volume and impedance to an adult and infant ear canal (Moodie et al., 2016). By incorporating an individualized foam-tip-to-earmold wRECD transform, an estimated "earmold IPL wRECD" could be determined from the real-ear IPL response measured using a generic coupling method, without systematic error caused by variability in tubing length. Acoustic length measurement was validated and found to be precise and reliable, and was applied to create individualized transforms that were validated in Chapter 5.

The primary contribution of Chapter 5 was the proposal and validation of a novel method of measuring the wideband real-ear-to-coupler difference using integrated pressure level stimuli. Previous studies have found that in-ear calibration approaches using Thevenin-equivalent source parameter calibrated transducers yield more accurate and reliable measurements of stimulus input to the auditory system (Farmer-Fedor & Rabbitt, 2002; Keefe & Schairer, 2011; Lapsley Miller et al., 2018; Lewis et al., 2009; McCreery et al., 2009; Scheperle et al., 2008, 2011; Souza et al., 2014; Withnell et al., 2009, 2014). Previously, such calibration procedures have been mainly focused on diagnostic testing and have not been implemented into a hearing aid fitting and verification workflow using ear-canal SPL based prescriptive targets. The measurement proposed herein assesses the reliability and validity of a source-parameter calibrated measurement (IPL wRECD), and directly compares IPL wRECD values to traditional foam-tip and earmold wRECD measurements. IPL wRECD provided results that were clinically similar to foam-tip

wRECD measurements, with improved performance in high-frequencies. The results indicate that IPL wRECD can be used interchangeably with foam-tip wRECD measurements without clinically significant differences below 5 kHz, and increased validity in the high frequencies.

Collectively, results in Chapters 4 and 5 indicate that earmold IPL wRECD is a viable alternative to probe-tube microphone measured earmold wRECD. In combination with independence from standing-wave error and thus improving high-frequency reliability, earmold IPL wRECD measurement resulted in significantly decreased slit leakage when compared to probe-tube microphone measured earmold wRECD. An additional, often overlooked, clinical disadvantage of probe-tube microphone measured earmold wRECD is the infection control concerns present with direct coupling and handling of the individual's earmold to the reference coupler (Ahmad et al., 2007; Sturgulewski et al., 2006), frequently attached using putty. Cleaning with an alcohol cleaning wipe is insufficient for complete earmold sanitization; more robust sanitization strategies are required that may require substantial cleaning time sufficient to disrupt normal scheduling. Insufficient sanitization can increase risk of chronic otitis externa, and transfer microflora between patients and between ears (Ahmad et al., 2007). For these reasons, current procedures allow hearing aid verification without puttying to the coupler, and offer corrections based on earmold wRECDs in the ear or earmold corrections. Similarly, earmold IPL wRECD does not require direct coupling to the individual's earmold or between the earmold and coupler. Instead, the earmold IPL wRECD uses a disposable generic plastic tip for in-ear measurement and a measurement in the coupler directly by the bare transducer. Consequently, earmold IPL wRECD measurement overcomes infection control barriers present with probe-tube microphone earmold wRECD measurements, combined with improved measurement validity compared to the current standard of practice.

In summary, the current work strengthens the justification of speech-evoked EFR as an aided outcome measurement and provides a new method of accounting for ear-canal acoustics in hearing-aid verification procedures. IPL wRECD measurement provides an improved, clinically viable method of measuring an individual's ear canal acoustics

without interference of the standing-wave or requiring precise probe-tube microphone placement. Furthermore, IPL wRECD measurement provides a simultaneous assessment of the individual's middle ear using wideband acoustic immittance. Because both values are obtained from one measurement, it is possible to obtain **both** a wRECD measurement and gold-standard assessment of the middle-ear in approximately the same, or less, time required by a typical tympanometric measurement.

6.2 Limitations

There are several limitations in the current work that need to be acknowledged:

- First and foremost, the determination and validation of the IPL wRECD measurement was only completed in adults with normal external auditory system physiology. Although smaller ear-canals are hypothesized to further reduce variability in IPL wRECD determination, further work in infants is necessary. Furthermore, the impact of middle-ear disorders and abnormal ear-canal physiology should be investigated. Middle ear disorders are known to alter power absorbance measurements (Merchant et al., 2009, 2014; Nakajima et al., 2013; Shahnaz et al., 2013; Voss et al., 2001, 2012), but the effect on IPL wRECD determination remains uninvestigated.
- Chapters 3 and 5 are limited by the frequencies tested. That is, only frequencies up to 8 kHz were assessed. There is evidence to show that the benefit of the source-parameter based measurements such as integrated pressure level and forward pressure level over SPL measurement will increase with bandwidth (Souza et al., 2014), however, in the current work, this was untested. Given that current probe-tube microphone wRECD bandwidth provides measures up to 12.5 kHz, assessment in the extended bandwidth is warranted, as it is likely that accuracy of current hearing aid fittings in this extended frequency range is subpar due to physical limitations of the probe-tube microphone measurement.

- Only normal hearing infants were assessed in Chapter 2. It is possible that hearing loss impacts statistical detection of speech-evoked EFR and/or the duration of testing necessary.
- In Chapter 3, only the calibration method validated by Nørgaard and colleagues (2017) was completed. This source-parameter calibration approach doesn't incorporate the rubber acoustic immittance tip and differs from the calibration approach utilized by some source-parameter calibration transducers (like the Mimoso Acoustics probe). While extremely unlikely, longitudinal effects unobserved in this chapter may be present in other calibration procedures.

6.3 Future work

For implementation of IPL wRECD in speech-evoked EFR testing and in a broad clinical setting, more work and further advancement may be necessary. Future directions of research may include:

- Assessment of the impact of various middle ear disorders (i.e. otosclerosis, cholesteatoma, eardrum perforation, etc) on the IPL wRECD measurement and how that may alter conductive disorder hearing thresholds is necessary. IPL wRECD also needs to be validated in a wider population of ears, including infants and children, to ensure no unexpected systematic difference with probe-tube microphone measurements arise. It is possible that the forward pressure level (FPL) measurement strategy, which is related to the IPL strategy, could have advantages with this population. A disordered middle ear will have slightly reduced impact on the realized sound-level than when IPL is used, due to the decreased effect of the reflected wave amplitude on overall stimulus level when using FPL rather than IPL. While the total realized difference is expected to be small, if wRECD measurement in a disordered population is completed, this has the potential to negatively impact the validity of such measurements and will need to be assessed formally to quantify the impact on wRECD determination.

- A standardization of source-parameter calibration procedures and an assessment of the effect that varying source parameter calibration procedures have on the measurement should be completed.
- Assessment of estimated earmold wRECD using IPL wRECD measurement should be completed in a wider population with a variety of earmold materials and tubing types, and in a clinical population to ensure that tubing independent differences between estimates and probe-tube measured wRECD do not exist. Procedures for use with vented earmolds should also be considered.
- Simultaneous IPL wRECD measurement coupled directly to the hearing aid tubing/individualized earmold may be possible and should be explored further.
- Measurement of the IPL wRECD should extend to the extended high frequency region and normative extended high frequency values should be produced.
- Assessment of slit-leakage and whether it is truly present in earmold couplings when the probe-tube microphone is not interfering with placement is necessary.

6.4 Conclusion

The procedures investigated for objective hearing aid verification and validation are expected to improve evaluation of pediatric hearing aid fittings. This preliminary study evaluated accuracy of speech-evoked EFR and identified barriers to IPL wRECD measurement. The main contribution of this thesis is the development of a novel wRECD measurement paradigm using a source-parameter calibrated transducer. The measurement of the IPL wRECD improves reliability of wRECD measurement. Values do not significantly differ from probe-tube microphone measurements below 4 kHz when measured with a generic coupling method. IPL wRECD is independent of standing wave errors, and more accurately represents the sound-level presented to the individual's auditory system at high frequencies when compared to probe-tube microphone measurement. Prior to clinical implementation, validation in a set of infants and those with middle-ear disorders is necessary to ensure that the presence of a middle-ear disorder is not a contraindication of IPL wRECD measurement. IPL wRECD shows great

promise as a complete replacement of probe-tube wRECD measurements and implementation is anticipated to improve pediatric hearing aid fitting accuracy. Increased ease of measurement and simultaneous middle ear analysis are also anticipated to improve routine clinical measurement with the IPL wRECD.

6.5 References

- Ahmad, N., Etheridge, C., Farrington, M., Baguley, DM. (2007). Prospective study of the microbiological flora of hearing aid moulds and the efficacy of current cleaning techniques. *J Laryngol Otol* 121(2):110-113
- Easwar, V. Scollie, S., Lasarev, Urichuk, M., Aiken, S., Purcell, D. (2021). Characteristic of speech-evoked envelope following responses in infancy. *Trends in Hearing*, 25, 1-17. DOI: 10.1177/23312165211004331.
- Easwar, V., Birstler, J., Harrison, A., Scollie, S., & Purcell, D. (2020). The accuracy of envelope following responses in predicting speech audibility. *Ear & Hearing, Publish Ahead of Print*. <https://doi.org/10.1097/AUD.0000000000000892>
- Farmer-Fedor, B. L., & Rabbitt, R. D. (2002). Acoustic intensity, impedance and reflection coefficient in the human ear canal. *The Journal of the Acoustical Society of America*, 112(2), 600–620. <https://doi.org/10.1121/1.1494445>
- Gustafson, S., Pittman, A., & Fanning, R. (2013). Effects of Tubing Length and Coupling Method on Hearing Threshold and Real-Ear to Coupler Difference Measures. *American Journal of Audiology*, 22(1), 190–199. [https://doi.org/10.1044/1059-0889\(2012/12-0046\)](https://doi.org/10.1044/1059-0889(2012/12-0046))
- Keefe, D. H., & Schairer, K. S. (2011). Specification of absorbed-sound power in the ear canal: Application to suppression of stimulus frequency otoacoustic emissions. *Journal of Acoustical Society of America*. 129(2). 779-791
- Lapsley Miller, J. A., Reed, C. M., Robinson, S. R., & Perez, Z. D. (2018). Pure-Tone Audiometry With Forward Pressure Level Calibration Leads to Clinically-Relevant Improvements in Test–Retest Reliability: *Ear and Hearing*, 39(5), 946–957. <https://doi.org/10.1097/AUD.0000000000000555>
- Lewis, J. D., McCreery, R. W., Neely, S. T., & Stelmachowicz, P. G. (2009). Comparison of in-situ calibration methods for quantifying input to the middle ear. *The Journal of the Acoustical Society of America*, 126(6), 3114–3124. <https://doi.org/10.1121/1.3243310>
- McCreery, R. W., Pittman, A., Lewis, J., Neely, S. T., & Stelmachowicz, P. G. (2009). Use of forward pressure level to minimize the influence of acoustic standing waves during probe-microphone hearing-aid verification. *The Journal of the Acoustical Society of America*, 126(1), 15–24. <https://doi.org/10.1121/1.3143142>
- Merchant, G. R., Röösl, C., Niesten, M. E. F., Hamade, M. A., Lee, D. J., McKinnon, M. L., Ulku, C. H., Rosowski, J. J., Merchant, S. N., & Nakajima, H. H. (2014). Power Reflectance as a Screening Tool for the Diagnosis of Superior Semicircular Canal Dehiscence: *Otology & Neurotology*, 1. <https://doi.org/10.1097/MAO.0000000000000294>
- Merchant, S. N., Rosowski, J. J., & McKenna, M. J. (2009). Superior Semicircular Canal Dehiscence Mimicking Otosclerotic Hearing Loss. *Adv Otorhinolaryngol.*, 65. 137-145.

- Moodie, S., Pietrobon, J., Rall, E., Lindley, G., Eiten, L., Gordey, D., Davidson, L., Moodie, K. S., Bagatto, M., Haluschak, M. M., 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. *Journal of the American Academy of Audiology*, 27(03), 264–275. <https://doi.org/10.3766/jaaa.15086>
- Munro, Kevin J., & 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 and Hearing*, 24(2), 100–110. <https://doi.org/10.1097/01.AUD.0000058114.20741.4D>
- Nakajima, H. H., Rosowski, J. J., Shahnaz, N., & Voss, S. E. (2013). Assessment of Ear Disorders Using Power Reflectance: *Ear and Hearing*, 34, 48s–53s. <https://doi.org/10.1097/AUD.0b013e31829c964d>
- Nørgaard, K. R., Fernandez-Grande, E., & Laugesen, S. (2017). Incorporating evanescent modes and flow losses into reference impedances in acoustic Thévenin calibration. *The Journal of the Acoustical Society of America*, 142(5), 3013–3024. <https://doi.org/10.1121/1.5010891>
- Scheperle, R. A., Goodman, S. S., & Neely, S. T. (2011). Further assessment of forward pressure level for in situ calibration. *J. Acoust. Soc. Am.*, 130(6), 12.
- Scheperle, R. A., Neely, S. T., Kopun, J. G., & Gorga, M. P. (2008). Influence of in situ , sound-level calibration on distortion-product otoacoustic emission variability. *The Journal of the Acoustical Society of America*, 124(1), 288–300. <https://doi.org/10.1121/1.2931953>
- Shahnaz, N., Feeney, M. P., & Schairer, K. S. (2013). Wideband Acoustic Immittance Normative Data: Ethnicity, Gender, Aging, and Instrumentation. *Ear and Hearing*, 34(9), 27S-35S.
- Souza, N. N., Dhar, S., Neely, S. T., & Siegel, J. H. (2014). Comparison of nine methods to estimate ear-canal stimulus levels. *The Journal of the Acoustical Society of America*, 136(4), 1768–1787. <https://doi.org/10.1121/1.4894787>
- Sturgulewski, SK., Bankaitis, AU, Klodd, DA, Haberkamp, T. (2006). What’s still growing on your patients’ hearing aids? *Hear J* 59(9):45-48
- Voss, S. E., Merchant, G. R., & Horton, N. J. (2012). Effects of Middle-Ear Disorders on Power Reflectance Measured in Cadaveric Ear Canals: *Ear and Hearing*, 33(2), 195–208. <https://doi.org/10.1097/AUD.0b013e31823235b5>
- Voss, S. E., Rosowski, J. J., Merchant, S. N., & Peake, W. T. (2001). Middle-ear function with tympanic-membrane perforations. I. Measurements and mechanisms. *The Journal of the Acoustical Society of America*, 110(3), 1432–1444. <https://doi.org/10.1121/1.1394195>
- Withnell, R. H., Jeng, P. S., Parent, P., & Levitt, H. (2014). The clinical utility of expressing hearing thresholds in terms of the forward-going sound pressure wave. *International Journal of Audiology*, 53(8), 522–530. <https://doi.org/10.3109/14992027.2014.898122>

- Withnell, R. H., Jeng, P. S., Waldvogel, K., Morgenstein, K., & Allen, J. B. (2009). An in situ calibration for hearing thresholds. *The Journal of the Acoustical Society of America*, *125*(3), 1605–1611. <https://doi.org/10.1121/1.3075551>
- Zhu, L., Bharadwaj, H., Xia, J., & Shinn-Cunningham, B. (2013). A comparison of spectral magnitude and phase-locking value analyses of the frequency-following response to complex tones. *The Journal of the Acoustical Society of America*, *134*(1), 384–395. <https://doi.org/10.1121/1.4807498>

Appendices

Appendix A: HSREB Approval



Date: 14 December 2020
To: Dr. David Purcell
Project ID: 116805
Study Title: Ear Canal Sound Level Calibration: Measurement with a Calibrated Sound Source
Application Type: HSREB Initial Application
Review Type: Delegated
Meeting Date / Full Board Reporting Date: 12/Jan/2021
Date Approval Issued: 14/Dec/2020
REB Approval Expiry Date: 14/Dec/2021

Dear Dr. David Purcell

The Western University Health Science Research Ethics Board (HSREB) has reviewed and approved the above mentioned study as described in the WREM application form, as of the HSREB Initial Approval Date noted above. This research study is to be conducted by the investigator noted above. All other required institutional approvals must also be obtained prior to the conduct of the study.

Documents Approved:

Document Name	Document Type	Document Date	Document Version
Protocol References v1.0	Protocol	30/Oct/2020	1.0
Data Collection Sheet v1.0	Other Data Collection Instruments	30/Oct/2020	1.0
Recruitment Poster v1.0	Recruitment Materials	23/Nov/2020	1.0
REB 116805 Protocol v2.0	Protocol	07/Dec/2020	2.0
Protocol_116805 Questionnaire Paper Copy v2.0	Paper Survey	07/Dec/2020	2.0
Qualtrics Survey Online Copy v2.0	Online Survey	07/Dec/2020	2.0
Social Media Poster	Recruitment Materials	02/Dec/2020	1.0
Section 12.13_LOI_adult_online consent v2.0	Written Consent/Assent	07/Dec/2020	2.0
Section 12.13_LOI_adult consent v2.0	Written Consent/Assent	07/Dec/2020	2.0
Recruitment Email Script v3.0	Email Script	11/Dec/2020	3.0
Protocol_116805 Recruitment Telephone Script	Telephone Script	14/Dec/2020	3.0

Documents Acknowledged:

Document Name	Document Type	Document Date	Document Version
Qualtrics security pdf	Technology Review document	30/Oct/2020	1.0
rationale references v2.0	References	07/Dec/2020	2.0
Budget for Protocol 116805	Study budget	11/Dec/2020	3.0

No deviations from, or changes to, the protocol or WREM application should be initiated without prior written approval of an appropriate amendment from Western HSREB, except when necessary to eliminate immediate hazard(s) to study participants or when the change(s) involves only administrative or logistical aspects of the trial.

REB members involved in the research project do not participate in the review, discussion or decision.

Curriculum Vitae

Name: Matthew Urichuk

Post-secondary Education and Degrees: University of Manitoba
Winnipeg, Manitoba, Canada
2013-2017 B.A. Adv.

The University of Western Ontario
London, Ontario, Canada
2017-2021 M.Cl.Sc./PhD

Related Work Experience Teaching Assistant
The University of Western Ontario
2019-2021

Infant Hearing Program CQI Reporting Consultant
Ontario Ministry of Children, Community and Social Services
2017-2021

Research and Development Intern
Audioscan (a division of Etymotic Inc.)
2019

Publications:

- Urichuk, M., Purcell, D., Allen, P., Scollie, S.** Submitted. Thevenin-equivalent source parameter reliability: intra- and inter-session effects on stimulus level calibration and wideband reflectance. Submitted to International Journal of Audiology.
- Urichuk, M., Purcell, D., Scollie, S.** Submitted. Measurement of hearing aid tubing segments using an arbitrary wideband stimulus. Submitted to Journal of Acoustic Society of America.
- Urichuk, M., Easwar, V., Scollie, S., Purcell, D.** Submitted. Performance of statistical indicators in the objective detection of speech-evoked envelope following responses. Submitted to Ear and Hearing
- Easwar, V. Scollie, S., Lasarev, **Urichuk, M., Aiken, S., Purcell, D.** (2021). Characteristic of speech-evoked envelope following responses in infancy. Trends in Hearing. DOI: 10.1177/23312165211004331.
- Urichuk, M., Easwar, V., Purcell, D.** (2021). Montage-related Variability in the Characteristics of Envelope Following Responses. Ear and Hearing. DOI: 10.1097/aud.0000000000001018.
- Urichuk, M., Loureiro-Rodriguez, V.** (2019). Brocatives: Self-reported use of masculine nominal vocatives in Manitoba (Canada). It's not all about you: New perspectives on address research.