Within- and Across-Subject Variability of Repeated Measurements of Medial Olivocochlear-Induced Changes in Transient-Evoked Otoacoustic Emissions

Ian B. Mertes and Shawn S. Goodman

Objectives: Measurement of changes in transient-evoked otoacoustic emissions (TEOAEs) caused by activation of the medial olivocochlear reflex (MOCR) may have clinical applications, but the clinical utility is dependent in part on the amount of variability across repeated measurements. The purpose of this study was to investigate the within- and across-subject variability of these measurements in a research setting as a step toward determining the potential clinical feasibility of TEOAE-based MOCR measurements.

Design: In 24 normal-hearing young adults, TEOAEs were elicited with 35 dB SL clicks and the MOCR was activated by 35 dB SL broadband noise presented contralaterally. Across a 5-week span, changes in both TEOAE amplitude and phase evoked by MOCR activation (*MOC shifts*) were measured at four sessions, each consisting of four independent measurements. Efforts were undertaken to reduce the effect of potential confounds, including slow drifts in TEOAE amplitude across time, activation of the middle-ear muscle reflex, and changes in subjects' attentional states. MOC shifts were analyzed in seven 1/6-octave bands from 1 to 2 kHz. The variability of MOC shifts was analyzed at the frequency band yielding the largest and most stable MOC shift at the first session. Within-subject variability was quantified by the size of the standard deviations across all 16 measurements. Across-subject variability was quantified as the range of MOC shift values across subjects and was also described qualitatively through visual analyses of the data.

Results: A large majority of MOC shifts in subjects were statistically significant. Most subjects showed stable MOC shifts across time, as evidenced by small standard deviations and by visual clustering of their data. However, some subjects showed within- and across-session variability that could not be explained by changes in hearing status, middle ear status, or attentional state. Simulations indicated that four baseline measurements were sufficient to predict the expected variability of subsequent measurements. However, the measured variability of subsequent MOC shifts in subjects was often larger than expected (based on the variability present at baseline), indicating the presence of additional variability at subsequent sessions.

Conclusions: Results indicated that a wide range of within- and across-subject variability of MOC shifts was present in a group of young normal-hearing individuals. In some cases, very large changes in MOC shifts (e.g., 1.5 to 2 dB) would need to occur before one could attribute the change to either an intervention or pathology, rather than to measurement variability. It appears that MOC shifts, as analyzed in the present study, may be too variable for clinical use, at least in some individuals. Further study is needed to determine the extent to which changes in MOC shifts can be reliably measured across time for clinical purposes.

Key words: Auditory efferent system, Contralateral suppression, MOC, Reliability, Repeatability, TEOAE.

(Ear & Hearing 2016;37;e72–e84)

Department of Communication Sciences and Disorders, University of Iowa, Iowa City, Iowa, USA.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and text of this article on the journal's Web site (www.ear-hearing.com).

INTRODUCTION

The medial olivocochlear reflex (MOCR) reduces cochlear amplifier gain by decreasing the motility of the outer hair cells (Guinan 2006). MOCR activation can be investigated indirectly using otoacoustic emissions (OAEs), which are byproducts of outer hair cell motility (Brownell 1990). Experimental activation of the MOCR by presenting broadband noise to the contralateral ear (contralateral acoustic stimulation [CAS]) typically causes a magnitude decrease and/or a phase lead in the OAE (hereafter referred to as an *MOC shift*).

Measurements of MOCR function may be clinically useful for purposes such as predicting improvements in speech perception due to auditory training (de Boer & Thornton 2008), predicting the onset of presbycusis (Zhu et al. 2007), and detecting ototoxicity (Aran et al. 1994). However, such measurements have not yet been incorporated into clinical practice, which may be due in part to limited knowledge regarding the variability of MOC shifts. In determining the potential clinical utility of MOC shift measurements, the variability in MOC shifts across OAE frequencies must be examined because a dependence on frequency has been demonstrated (Marshall et al. 2014; Mishra & Abdala 2015). In addition, the variability within a subject is a determining factor for establishing a reliable baseline and for knowing what constitutes a significant change in MOC shift in an individual.

In describing measurements of MOC shift variability, the following definitions will be used in this article: a *buffer* is a single recorded waveform consisting of a stimulus and the associated cochlear response; a *set* is a collection of buffers averaged to obtain one estimate of an MOC shift; and a *session* is a single laboratory visit, during which multiple sets may be obtained. The variability of MOC shifts was quantified relative to three time intervals: *across buffer* (within a set), *across set* (within a session), and *across session* (within a subject). In addition, the *across frequency* and *across subject* variability was examined.

Each type of variability has been previously investigated for different purposes: across frequency for determining which OAE frequency regions yield the most stable MOC shifts (Marshall et al. 2014; Mishra & Abdala 2015), across buffer for assessing the statistical significance of a single MOC shift in an individual (Backus & Guinan 2007; Goodman et al. 2013), across set for determining short-term test–retest reliability (Berlin et al. 1993; De Ceulaer et al. 2001), across session for determining longer-term test–retest reliability (Graham & Hazell 1994; Kumar et al. 2013; Mishra & Lutman 2013; Mishra & Abdala 2015), and across subject for quantifying the distribution of the size of MOC shifts in a normative group (De Ceulaer et al. 2001; Backus & Guinan 2007; Marshall et al. 2014). No previous studies have investigated all types of variability in the same group of subjects; such an investigation is

<zdoi; 10.1097/AUD.0000000000000244>

needed for determining potential clinical utility. Also, potential limitations of previous studies include examining variability only for brief intervals (≤2 weeks; Kumar et al. 2013; Mishra & Lutman 2013; Marshall et al. 2014) and in small numbers of subjects (≤6; Graham & Hazell 1994; Mishra & Abdala 2015), possibly limiting the generalizability of results. The effects of attention, which can alter the stability of MOC shifts (de Boer & Thornton 2007), also have not been controlled for adequately.

In the present study, the variability in MOC shifts was quantified in normal-hearing young adults across buffers, sets, sessions, and subjects. Transient-evoked OAEs (TEOAEs) were used to examine MOC shifts over a relatively broad range of frequencies, thus allowing for the measurement of variability across frequencies. The variability in MOC shifts was assessed using a longer time span than most studies. Subjects performed a visual attention task to maintain a consistent attentional state. The question of how many initial measurements of MOC shift are needed to establish a reliable baseline function for a given subject was also examined.

MATERIALS AND METHODS

Subjects

Twenty-four normal-hearing adults participated (13 females, mean age = 23.3 years, SD = 3.3 years). All subjects had normal otoscopic findings, normal 226-Hz tympanograms, pure-tone air conduction thresholds ≤15 dB HL (re: ANSI 1996) at octave frequencies from 0.25 to 8 kHz, bilaterally, and no air-bone gaps >10 dB at octave frequencies from 0.5 to 4 kHz. In addition, no subjects reported significant difficulty communicating in noisy situations, outer and/or middle ear surgeries, significant noise exposure, tinnitus, vertigo, or use of ototoxic medications. All subjects had normal or corrected-to-normal vision. Tympanograms and audiometric thresholds were stable in all subjects across sessions, and no subjects reported noise exposure between sessions. The research protocol was approved by the University of Iowa Institutional Review Board and written informed consent was obtained from all subjects.

Equipment

Subjects were tested in a double-walled sound-treated booth while seated comfortably in a recliner. Measurements of MOC shifts were made using a PC running custom MATLAB software (The MathWorks, Inc., Natick, MA), the software utility Playrec (Humphrey 2008), and Windows Media Player (Microsoft, Redmond, WA). Stimuli were generated in MATLAB and presented using two sound cards: a 24-bit card (LynxTwo, Lynx Studio Technology, Inc., Costa Mesa, CA) for clicks and an onboard 16-bit card (Dell, Round Rock, TX) for noise. Stimuli were routed through power amplifiers (GFA-5002, ADCOM, Bangkok, Thailand; HB7, Tucker-Davis Technologies, Alachua, FL) and variable attenuators (PA5, Tucker-Davis Technologies; 350A, Hewlett-Packard, Palo Alto, CA) to insert earphones (ER-2, Etymōtic Research). Ear canal pressure recordings in both ears were made using two OAE probe microphone systems (ER-10B+, Etymōtic Research) connected to the 24-bit sound card.

Otoacoustic Emissions Testing

Stimuli consisted of clicks for eliciting TEOAEs and broadband acoustic noise for activating the MOCR. Both stimuli were generated using a sampling rate of 44.1 kHz. Clicks and noise had a nominal bandwidth of 0.6 to 10 kHz. The electrical drive for the click stimuli had a duration of 1.2 msec. Clicks were presented at a rate of 26/sec. Click and noise stimuli were created to have an acoustically flat magnitude (±1.5 dB) from 0.6 to 10 kHz (Goodman et al. 2009).

Click and noise stimuli were presented at 35 dB sensation level (SL) to maximize activation of the MOCR and minimize activation of the middle-ear muscle reflex (MEMR). The MEMR changes the impedance characteristics of the middle ear, which can alter OAE amplitudes in ways similar to the MOCR, thus complicating interpretation of the results (Guinan et al. 2003). In addition, a direct check of MEMR activation was implemented (described below).

At the first session, detection thresholds for the stimuli were measured. To ensure consistent presentation of stimulus levels across measurements, the SPLs corresponding to 35 dB SL for each subject were measured in subjects' ear canals and presented at these SPLs at all measurements. Prior to each set of TEOAE recordings, click and noise levels were calibrated in subjects' ear canals to be within ±0.3 and ±1.0 dB, respectively, of the target SPL values. The mean root mean square (RMS) click level across subjects was 55.0 dB SPL (SD = 3.4 dB). The mean RMS noise level was 58.1 dB SPL (SD = 3.5 dB). TEOAEs were measured in subjects' right ears, and CAS was presented to left ears because larger MOC shifts are typically demonstrated in this configuration (Khalfa et al. 1997).

TEOAEs were extracted using a linear paradigm to allow for measurement of both the linear and nonlinear portions of MOC shifts (Guinan 2006). All recorded buffers were written directly to computer hard disk, with synchronous averaging and artifact rejection performed post-hoc based on methods described in Goodman et al. (2009).

As described earlier, *buffer* refers to a recording of a single click stimulus presentation and the 38.46 msec immediately following, which contains the TEOAE. Each approximately 10-second grouping of buffers (38.46 msec × 260 buffers), obtained either with or without CAS, is referred to as a *block*. For each block with CAS, the CAS was first turned on for 500 msec before presenting clicks to allow for "fast" and "medium" MOCR effects to achieve steady state (Backus & Guinan 2006). Likewise, after a block with CAS was recorded, the CAS and clicks were turned off and there was a silent period of 500 msec to ensure that the MOCR was not activated during the subsequent block without CAS. A *set* consisted of 16 interleaved blocks: eight made with CAS and eight made without CAS (see Figure, Supplemental Digital Content 1, for a schematic of the recording paradigm; http://links.lww.com/EANDH/A222). Interleaving helps reduce the impact of small drifts in TEOAE amplitude on measurement of MOC shifts. In addition, TEOAE amplitudes across time were de-trended before further analysis using methods described by Goodman et al. (2013). After recording a set (nominally lasting 160 s), the blocks were deinterleaved, so that all of the buffers obtained without CAS were grouped together, and those obtained with CAS were grouped together, forming two matrices (*A* and *B*, respectively).

Because subject attention can modulate the MOCR (e.g., de Boer & Thornton 2007) and potentially increase MOC shift variability across time, subjects performed a visual attention task during recordings, based on methods described by Meric and Collet (1992). Subjects were instructed to press a button as

soon as they saw a white box randomly appear on a black computer screen. The OAE probe cables were situated away from the subjects' bodies to minimize recorded noise due to muscle movement. Feedback on response times was provided to subjects to engage them in the task. The experimenter also monitored performance to detect any substantial drifts in attention, signaling the need to provide the subject with a short break.

In addition to the TEOAE test condition described above, a control condition was recorded during each visit, which was identical except that the electrical cable attached to the loudspeaker delivering CAS was unplugged so that no acoustic noise was delivered to the contralateral ear. This condition served as a check to ensure that measured MOC shifts represented true effects resulting from the presence of CAS, rather than any unexpected artifacts because it was expected that no significant differences would be seen between the amplitudes of the waveforms in the *A* and *B* matrices.

Testing Schedule

Each subject participated in four sessions. The first three sessions occurred within a 7-day period, each separated by at least 1 day. The fourth session took place 4 weeks after the third session. All testing was conducted by the same researcher to help ensure consistency in experimental procedures across visits and across subjects. At each session, five TEOAE recording sets were measured: four in the test condition and one in the control condition. Subjects were given a 2-min break between sets. The OAE probes were reinserted before the next recording set. Care was taken to place the probe in the same location in the ear canal each time.

Analysis

Time- and Frequency-Domain Analysis Windows • The relative size of MOC shifts depends on OAE frequency (Goodman et al. 2013; Marshall et al. 2014; Mishra & Abdala 2015). The variability was examined at the frequency yielding the largest and most stable MOC shifts in a given subject. A restricted analysis bandwidth from 1 to 2 kHz was considered, based on previous studies showing that this band typically contains the largest MOC shifts (e.g., Hood et al. 1996; Goodman et al. 2013). The time window for analyzing MOC shifts was based on the 95% confidence intervals for expected OAE latencies reported by Shera et al. (2002). The full time window for measuring MOC shifts started at time = 5 msec and ended at time = 14.2 msec, including a 1 msec onset ramp and 2 msec offset ramp (see Fig. 1).

Due to their potential for affecting measurement of MOC shifts, the presence of synchronized spontaneous OAEs (SSOAEs) and MEMR activation were examined. SSOAEs are spontaneous cochlear oscillations that can become entrained to click stimuli and ring out in time, affecting measured TEOAE amplitudes and phases. Using the same waveforms recorded without CAS, SSOAEs were analyzed within the same frequency band as the TEOAEs using an analysis window that was the same length as the MOC window, but shifted closer to the end of the buffer (23.29 to 32.41 msec; see Fig. 1). The presence of energy in this time window was examined via fast Fourier transform (FFT); a signal to noise ratio (SNR) >6 dB was taken as evidence that SSOAEs were present. The noise floor was computed as the standard error of the mean computed across all recorded buffers (Goodman et al. 2009).

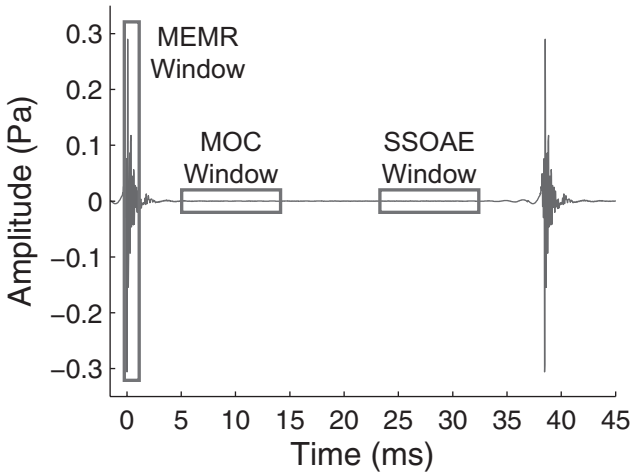

Fig. 1. The temporal locations of the three analysis windows. Tracing represents the recorded waveform for a single buffer that includes the stimulus and the TEOAE, as well as the adjacent stimulus. The MEMR window encompasses the stimulus while excluding any TEOAEs from 1 to 2 kHz. The MOC window encompasses TEOAEs from 1 to 2 kHz, while excluding stimulus ringing. The SSOAE window encompasses OAEs that persist beyond the expected latency for a TEOAE from 1 to 2 kHz. MEMR indicates middle-ear muscle reflex; MOC, medial olivocochlear; OAE, otoacoustic emission; SSOAE, synchronized spontaneous otoacoustic emission; TEOAE, transient-evoked otoacoustic emission.

Because MEMR activation complicates the interpretation of measured MOC shifts, MEMR activation was detected by comparing the amplitudes of the click stimuli obtained with and without CAS. Some previous studies have analyzed MEMR activation at a lower frequency than the frequencies at which MOC shifts were analyzed (e.g., Goodman et al. 2013; Kumar et al. 2013; Mishra & Lutman 2013); however, this may not provide an accurate indication of the influence of MEMR at higher frequencies. We ascertained the presence of MEMR after bandpass filtering the click stimuli in two frequency regions: from 600 to 800 Hz (lower than the frequency region at which MOC shifts were analyzed), and from 1 to 2 kHz (the same frequency region at which MOC shifts were analyzed). Although some differences in the presence of MEMR were found when filtering the click stimuli from 600 to 800 Hz versus from 1 to 2 kHz, these differences appeared to be minimal. Results are reported only for the 1 to 2 kHz region; therefore, the data reported in this article indicate presence or absence of MEMR in the same frequency region at which MOC shifts were analyzed. The filtered click stimulus amplitudes were examined in a time window extending from −0.2 to 1.16 msec (time zero being the peak of the stimulus waveform) (see Fig. 1). Waveforms in this window were multiplied by a raised cosine that extended across the full duration of the window. A difference between the *A* and *B* matrices in this window that exceeded a criterion amount was taken as evidence of MEMR activation (described in more detail below).

Quantification of Magnitude and Phase Changes • Although MOC shifts are typically quantified in terms of the change in magnitude, quantifying the simultaneous change in magnitude and phase may allow for more reliable detection of MOC shifts (Backus & Guinan 2007; Marshall et al. 2014). Therefore, MOC shifts were quantified as a complex vector that takes into account changes in both magnitude and phase (see Text, Supplemental

Digital Content 2, which provides MATLAB code for calculating MOC shifts; http://links.lww.com/EANDH/A223). Within the MOC window, the means across matrices *A* and *B* were computed to yield $\bar{a}$ and $\bar{b}$. The mean change in the TEOAEs due to the presence of CAS was found by computing the ratio of $\bar{b}$ to $\bar{a}$ in the frequency domain,

$$\Delta = F\bar{b}/F\bar{a}, \tag{1}$$

where Δ is a $n\times1$ vector of complex ratios obtained by complex pointwise division, and *F* is the $n\times n$ Fourier matrix having elements $f_{kj}=\omega^{jk}$, $\omega=e^{-2\pi i/n}$ (*j* and *k* designate rows and columns, and *i* is the imaginary operator, $\sqrt{-1}$). The quantities $F\bar{a}$ and $F\bar{b}$ were complex values composed of the real and imaginary output of discrete Fourier transforms. Because Δ was computed by complex pointwise division, this was equivalent to dividing the magnitudes and subtracting the phases expressed in polar form. The quantity Δ was reduced to a single complex value, δ, by taking the mean of the subset of elements in Δ corresponding to a band of desired frequencies:

$$\delta = \frac{1}{n}\sum D,\ D=\left\{D\in\Delta : i_{f_L}\le D\le i_{f_H}\right\} \tag{2}$$

where *D* is the set containing all elements of Δ with indices *i* which are greater than or equal to the index of the desired low cutoff frequency $\left(i_{f_L}\right)$ and less than or equal to the index of the high desired cutoff frequency $\left(i_{f_H}\right)$, the set *D* containing *n* elements. This averaging procedure was computed on the real and imaginary components separately. The magnitude of the complex ratio, δ, quantifies the amplitude of $\bar{b}$ relative to the amplitude of $\bar{a}$, while the phase angle of the ratio quantifies the phase difference of $\bar{b}$ relative to the phase of $\bar{a}$.

When plotted on the unit circle with real values on the *x* axis and imaginary values on the *y* axis, δ provides a clear visual interpretation of how the presence of CAS changes both TEOAE magnitude and phase: if δ falls on the unit circle $\left(|\delta|=1\right)$, CAS did not change TEOAE magnitude. If δ falls inside the unit circle $\left(|\delta|<1\right)$ or outside the unit circle $\left(|\delta|>1\right)$, CAS reduced or increased TEOAE magnitude, respectively. In terms of phase, if δ falls directly on the *x* axis $(\angle\delta=0)$, CAS did not cause a phase change. If δ falls above the *x* axis $(\pi/2>\angle\delta>0)$ or below the *x* axis $(-\pi/2<\angle\delta<0)$, CAS caused a phase lead or lag, respectively. It was expected that MOCR activation will usually cause a magnitude reduction and phase lead, so δ will most often fall inside the unit circle in quadrant I (see Figure, Supplemental Digital Content 3, which demonstrates three examples of δ values on the unit circle; http://links.lww.com/EANDH/A224).

Tolerance Regions • Since for real data, the ratio δ will almost never be exactly 1 (which would require $\bar{a}$ and $\bar{b}$ to be identical), it is desirable to determine what values of δ represent a statistically significant difference in $\bar{a}$ and $\bar{b}$ for any given recording set. While it is possible to compute statistical intervals for this purpose on magnitude and phase separately, in cases like MOCR activation where both kinds of shifts are expected, it may be preferable to compute a statistical region that simultaneously includes both magnitude and phase (see Text, Supplemental Digital Content 2, which provides MATLAB code for calculating statistical tolerance regions; http://links.lww.com/EANDH/A223).

To construct simultaneous intervals, the distribution of the sampling mean of δ must be estimated. A bootstrapping procedure for accomplishing this using magnitude only was described by Goodman et al. (2013). This same operation can be performed on complex values, such as δ. A statistical resampling with replacement (bootstrap) algorithm (Efron & Tibshirani 1993) for this purpose was implemented in MATLAB. Recall that the ratio δ was obtained from two matrices, *A* and *B*, each of size *n* samples by *m* buffers. The null hypothesis was that there was no difference between the populations from which the matrices were sampled, and under this hypothesis all the buffers were pooled into a single $n\times2m$ matrix. The distribution of ratios from this pooled matrix was estimated by randomly selecting (with replacement) two independently resampled matrices, each of size $n\times m$, and calculating the resulting complex ratio of resampled values, $\hat{\delta}$. This process was iterated K = 10,000 times (in practice, stable, repeatable results are usually obtained using as little as K = 1000 iterations). On each iteration, the complex ratio of the resampled pair of matrices was calculated using the same procedure described previously: Each resampled matrix was averaged, and the resulting mean vectors were Fourier transformed and divided pointwise, yielding a vector of complex ratios (see Eq. 1). The vector was reduced to a single complex number, $\hat{\delta}_k$ (see Eq. 2), where *k* indicates the *k*th of *K* iterations. On the completion of this process, $\hat{\delta}$ was a $K\times1$ vector of complex resampled estimates.

The distribution of $\hat{\delta}$ shows the expected ratios of $\bar{a}$ and $\bar{b}$ under the null hypothesis that there is no difference between the populations from which the matrices *A* and *B* were sampled. This distribution was used to calculate a statistical tolerance region as follows. First, the (complex) mean of the vector $\hat{\delta}$ was calculated, and $\hat{\delta}$ was centered about the origin on the complex plane by subtracting this mean. Next, for computational purposes, $\hat{\delta}$ was changed from a $K\times1$ complex vector into a $K\times2$ matrix of real values, by placing the real and imaginary parts of $\hat{\delta}$ in the first and second columns, respectively. The covariance matrix (unscaled) of the $\hat{\delta}$ matrix was then obtained by

$$S=\mathrm{cov}\left(\hat{\delta}\right)=\frac{1}{K-1}\hat{\delta}^T\hat{\delta}, \tag{3}$$

where the superscript uppercase *T* indicates the matrix transpose operator. The covariance matrix *S* is a 2×2 matrix, with the diagonal elements being the variances of the real and imaginary parts of $\hat{\delta}$ and the off-diagonal elements being the covariance between the real and imaginary parts of $\hat{\delta}$. The eigenvalues and eigenvectors of the covariance matrix *S* define the major and minor axes of an ellipse describing the bivariate distribution of the elements in $\hat{\delta}$. These were used to calculate the edges of the ellipse:

$$R=\left(\lambda\sqrt{\upsilon}C\right)^T, \tag{4}$$

where *R* is an $m\times2$ vector containing a set of coordinates on the Cartesian plane corresponding to the edges of the elliptical region, υ is the matrix of eigenvalues, and λ is the matrix of eigenvectors. The variable *C* designates the Cartesian coordinates of points on a unit circle described by a $2\times J$ matrix, where each element in the first row is defined by $\cos\left(\theta_j\right)$ and each element in the second row is defined by $\sin\left(\theta_j\right)$, θ a row vector of *J* equally spaced radian phase values from 0 to 2π. The choice of the number of columns, *J*, is somewhat arbitrary and is not critical, so long as it is relatively large, so as to allow reasonable estimation of the ellipse magnitude at any arbitrary phase value.

Finally, the elliptical region R was scaled to yield a tolerance region encompassing $100(1-\alpha)\%$ of the complex ratio values in $\hat{\delta}$. For a standard, bivariate normal vector, the squares of the major and minor axes can be estimated by the χ^2 distribution on 2 degrees of freedom (Chew 1966). A scaling factor was therefore found using the inverse χ^2 distribution, with degrees of freedom $1-\alpha$ and 2. R was then multiplied by the scaling factor, yielding

$$R_\alpha = R\sqrt{\mathrm{Inv}\chi^2(1-\alpha,2)}, \quad (5)$$

where R_α defines the edges of the elliptical tolerance region encompassing $100(1-\alpha)\%$ of the complex ratio values in $\hat{\delta}$. In most cases, α would be set to 0.05 or 0.01. The centering operation described toward the beginning of this section was undone (the original mean of the vector $\hat{\delta}$ was added to R_α), and R_α was changed from a $m \times 2$ matrix back into a $m \times 1$ complex vector.

The resampled tolerance region, R_α, defines which values of δ represent a statistically significant change: δ must fall outside of the ellipse described by R_α to be significant. In the context of MOCR studies using TEOAEs, a significant result for δ can be interpreted to mean that the presence of CAS changed the measured TEOAE, presumably as a result of MOCR activation (provided the influence of MEMR is ruled out). Significant changes could be due to change in TEOAE magnitude, phase, or both. When δ and R_α are plotted together on the unit circle, the relative contribution of magnitude and phase changes can be examined, as shown in Figure 2.

Total Quantity of Change • One method of computing an MOC shift is to compute the Fourier transform of the subtraction of waveforms with and without CAS, and then express the magnitude of this result relative to the magnitude of the Fourier transform of the emissions obtained without CAS (see, e.g., Marshall et al. 2014). This value may be thought of as representing the "total quantity" of change because it includes both magnitude and phase. Following the notation used in Equation 1, this quantification of change can be written as

$$\nabla = \frac{\left|F(\bar{a}-\bar{b})\right|}{\left|F\bar{a}\right|}. \quad (6)$$

Similarly to Δ, ∇ (nabla) can be reduced to a single, noncomplex value by taking the mean of the subset of elements in ∇ corresponding to a range of frequencies (Eq. 2). This reduced value, ϙ (turned delta), represents the "total quantity" of change, including both magnitude and phase, as a real, relative value (see Figure, Supplemental Digital Content 4, which illustrates the relationship between δ and ϙ; http://links.lww.com/EANDH/A225). The statistical significance of ϙ can be found using the same resampled tolerance region, R_α, that was used to determine the significance of δ.

Establishing Baseline • The area of the elliptical tolerance region, R_α, represents the pooled variance of the emissions recorded with and without CAS. Within a single recording set, the variance (and covariance) of the buffers are described by eigenvectors and eigenvalues, which are then scaled to yield the tolerance region. The tolerance region is, therefore, expected to account for all of the sources of variability that are present within that particular recording set. Hence, it is also related to the expected variability of δ across repeated measures for a given subject: if there are no unexpected changes in variability, $100(1-\alpha)\%$ of repeated measurements of δ would be expected

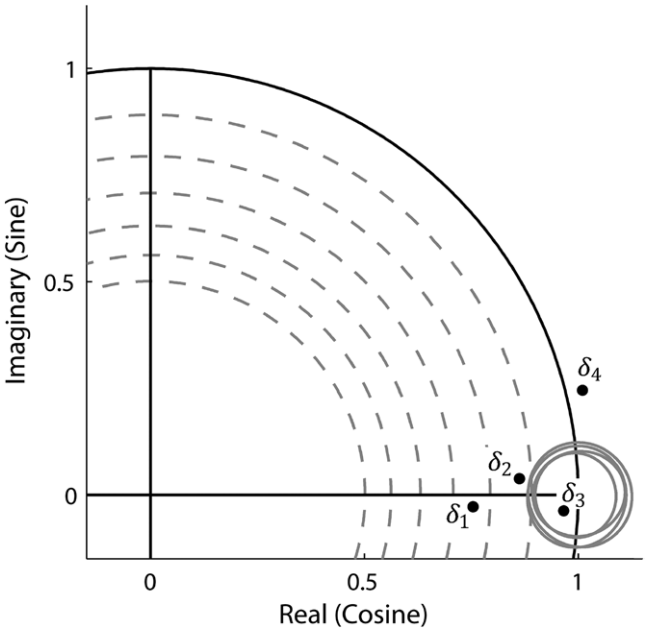

Fig. 2. Four examples of the complex ratio, δ, and their associated tolerance regions, R_α. Tolerance regions are plotted as gray ellipses centered around (1,0) on the unit circle. Each measurement of δ has its own associated tolerance region. For sake of visual clarity, the specific ellipse associated with each value of δ is not indicated in this figure; however, the significance of each δ was determined by its own tolerance region. Values of δ that fall outside of their associated tolerance region (δ_1, δ_2, δ_4) represent a statistically significant change. Values of δ that fall inside their associated tolerance region (δ_3) do not represent a statistically significant change.

to fall within a region having the area of R_α, but shifted so as to be centered around μ_δ, the population mean of δ (Fig. 3). The value of μ_δ is unknown, but can be estimated from one or more measured samples of δ. The tolerance region is therefore defined as

$$\overleftarrow{R_\alpha} = \left(\frac{1}{n}\sum_{i=1}^{n} R_{\alpha_i}\right) + \left(\frac{1}{n}\sum_{i=1}^{n}\delta_i\right), \quad (7)$$

where $\overleftarrow{R_\alpha}$ is the tolerance region expected to encompass $100(1-\alpha)\%$ of subsequent repeated measurements of δ. $\overleftarrow{R_\alpha}$ becomes a more accurate estimate of the true distribution of δ as n, the number of estimates of δ and R_α, increases.

An important question then becomes how many estimates of δ are needed to obtain a reasonable estimate of $\overleftarrow{R_\alpha}$. While a large n is desirable, there are practical time and attentional constraints involved in taking repeated baseline measures from human subjects. To estimate the number of baseline measurements needed to account for 95% of repeated measures of δ, a computer simulation was implemented based on 1000 independently generated values of δ and their associated tolerance regions, R_α. For several different numbers of possible baseline measurements (1, 2, 4, or 8), 1000 iterations were computed. On each iteration, the desired number of baseline measures was taken by random sampling from the 1000 values of δ. From these randomly chosen baseline measures, $\overleftarrow{R_\alpha}$ was computed, and the percent of the remaining 1000 values of δ falling inside of $\overleftarrow{R_\alpha}$ was found. The results of these simulated data (not shown) indicated that adequate accuracy can be obtained using $n = 4$ baseline measurements of δ, where $\overleftarrow{R_\alpha}$ encompasses 95% of repeated measures of δ, 95% of the time.

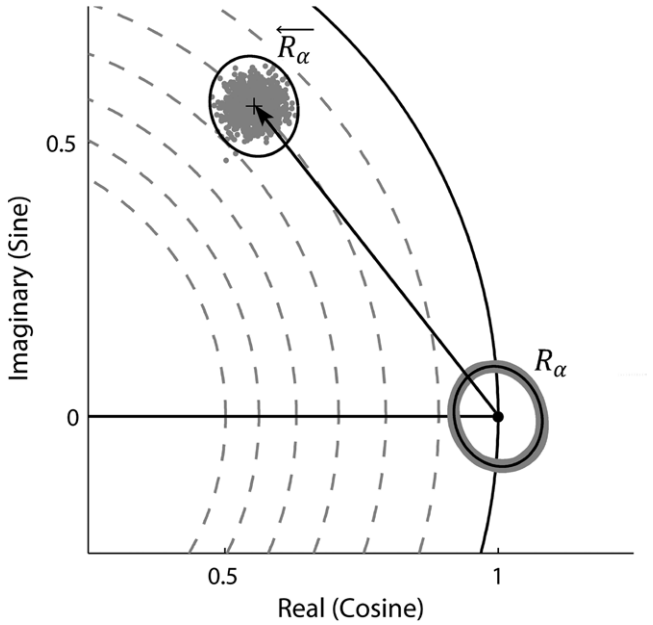

Fig. 3. Relationship between the tolerance region R_α and the shifted tolerance region, $\bar{R}_\alpha$. Data for this figure were generated by a computer simulation using 1000 independent estimates. The 1000 tolerance regions are shown by gray ellipses centered around (1,0). Because they overlap so closely, they appear as a single thick gray ellipse. The mean of the 1000 ellipses is shown as a thinner black ellipse. The 1000 computed values of δ are shown as *gray dots* clustered in the *upper left* of the figure. The mean value of delta is shown as a *black plus* (+) symbol. When the mean value of R_α is shifted by the mean value of delta (shown by the *black arrow*; see Eq. 7), it encompasses $100(1-\alpha)$% of the repeated measurements of δ.

Selection of Analysis Frequencies • The analysis bandwidth of 1 to 2 kHz was subdivided into seven 1/6th-octave wide analysis bands. Initial analyses revealed that the frequencies showing the largest MOC shifts were not necessarily the most stable across repeated measures. Because the frequencies that yielded the most stable MOC shifts were not known a priori, the across-set variability from the first session (i.e., the variability computed across the four sets obtained during the first laboratory visit) for each subject served as the baseline to which results from later sessions were compared. Baseline MOC shifts were computed in all seven frequency bands (referred to by the center frequency).

Baseline variability was quantified as the sum of squared distances of the four baseline δ values relative to their centroid. First, the centroid was computed as

$$\bar{\delta} = \frac{1}{n}\sum_{i=1}^{n}\delta_i. \tag{8}$$

The square root of the sum of squared distances $\left(\sqrt{\text{SSD}}\right)$ from the centroid was computed as

$$\sqrt{\text{SSD}} = \sqrt{\sum_{i=1}^{n}\text{Re}(\delta_i-\bar{\delta})^2 + \text{Im}(\delta_i-\bar{\delta})^2}. \tag{9}$$

Center frequencies with δ values that were more tightly clustered (i.e., lower across-set variability) had lower values of $\sqrt{\text{SSD}}$. The $\sqrt{\text{SSD}}$ values were computed for each of the seven frequency bands. In selecting the frequency band to analyze, not only the $\sqrt{\text{SSD}}$ was considered but also the magnitude and significance of the mean MOC shift at baseline (i.e., the centroid $\bar{\delta}$) because a small, nonsignificant MOC shift would not be useful for measuring longitudinal changes. Therefore, the magnitude of the mean MOC shift normalized by its $\sqrt{\text{SSD}}$ was computed as

$$\delta_{\text{norm}} = \frac{|\bar{\delta}|}{\sqrt{\text{SSD}}}. \tag{10}$$

A larger δ_{norm} indicates a larger MOC shift magnitude with smaller variability. The frequency band yielding the largest δ_{norm}, which also had a magnitude of $\bar{\delta}$ that exceeded the magnitude of R_α, was selected for the remaining analyses.

RESULTS

MEMR Activation

Before analyzing MOC shifts, the presence of MEMR was determined. Because the *A* and *B* matrices contain 2080 buffers, a determination of the statistical significance of MEMR shifts would require a Bonferroni correction that would be too stringent to be of practical value. Therefore, MEMR shifts were instead classified as *possible* and *probable*, where possible and probable MEMR shifts were complex differences that exceeded the 95% and 99% bootstrapped tolerance regions, respectively.

The validity of the MEMR check was assessed by examining the results obtained in the control condition (where the cable delivering CAS was unplugged so that no acoustic noise stimulus was presented to the contralateral ear). Because the "CAS" and "no CAS" conditions were presumably identical, differences between the *A* and *B* matrices, in addition to being unexpected, could not be attributable to MEMR activation. Five of 96 total measurements (24 subjects × 4 sessions × 1 control measurement) were found to have possible MEMR, yielding an acceptably low false positive rate of 5.2%.

In the test condition, there were a total of 384 measurements (24 subjects × 4 sessions × 4 measurements). Of these, 51 measurements (13.3%) were identified as having possible MEMR and 40 measurements (10.4%) were identified as having probable MEMR. Twelve of 24 subjects had at least 1 case of possible MEMR. Within these 12 subjects, there was a median of 4 cases and a range of 1 to 9 cases (out of 16 measurements). Seven of 24 subjects had at least 1 case of probable MEMR. Within these 7 subjects, there was a median of 5 cases and a range of 1 to 16 cases. MOC shift data from all subjects will be reported, with the presence of possible or probable MEMR noted when relevant.

SSOAEs

The presence of SSOAEs was examined at each of the 16 measurements in the test condition. Out of a total of 384 measurements, 216 measurements (56.3%) were identified as having SSOAEs. Fourteen subjects (58.3%) had SSOAEs at half or more of all measurements, with 8 subjects (33.3%) exhibiting SSOAEs at all measurements. The large prevalence of SSOAEs is likely due to several factors, including the frequency regions that were analyzed (1 to 2 kHz, where SSOAEs are common), the click rate of 26/sec which did not allow for SSOAEs to decay into the noise floor before the next presentation of the click, and the characteristics of our subject population (young, normal hearing individuals).

Across all 216 cases of SSOAEs, the median SNR was 14.2 dB and the range was 6.0 to 34.3 dB. The median amplitude

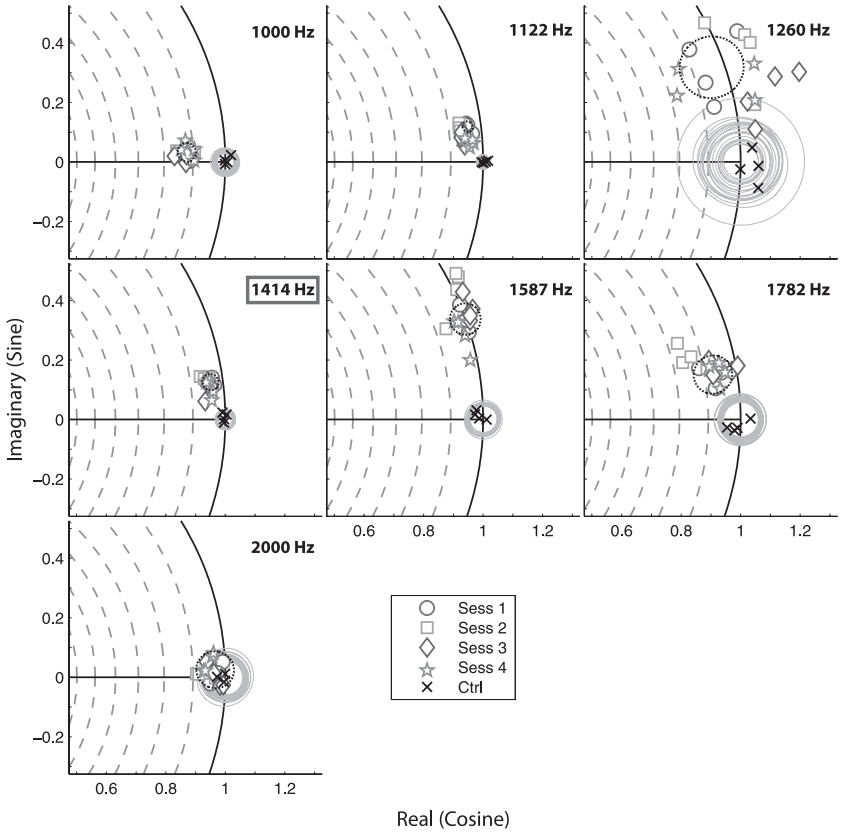

Fig. 4. MOC shifts at all seven frequency bands for one representative subject (05). Each panel shows data for a different frequency indicated in the *upper right-hand corners*. *Markers* represent δ values obtained at sessions 1 to 4 in the test condition, differentiated by marker style as indicated in the figure legend. *Cross markers* represent δ values obtained in the control condition, not differentiated by session in the figure. *Thin gray ellipses* centered about the point (1,0) represent the R_α obtained at each visit, not differentiated by session in the figure. The *black dotted ellipse* represents $\overline{R_\alpha}$, which was derived from the data obtained at the first session (note that in some instances, $\overline{R_\alpha}$ has a very small area). 1414 Hz (highlighted using a *rectangle*) had the largest δ_{norm} value, and was therefore the frequency selected for subsequent analyses for this subject. MOC indicates medial olivocochlear.

was −21.4 dB SPL and the range was −34.2 to 0.2 dB SPL. Despite this large range, the 90th percentile was −9.4 dB SPL, indicating that most SSOAEs were relatively low in amplitude and were thus unlikely to impact the click response levels.

To compare SSOAE levels to TEOAE levels, the ratio of energy (squared amplitude) of SSOAEs to TEOAEs was also computed, where a smaller ratio indicates lower SSOAE energy relative to the TEOAE energy. The median ratio was 1.56%. There were three outliers at 8.3%, 27.6%, and 75.8%, showing that the SSOAE energy constituted >50% of the TEOAE energy in only one subject.

MOC Shifts

Within-Subject Variability • Figure 4 shows a representative example of the across-frequency variability for one subject. All δ values from all sets and sessions are plotted for each frequency. The size of MOC shifts varied across frequencies, consistent with previous reports (e.g., Goodman et al. 2013; Marshall et al. 2014; Mishra & Abdala 2015). Importantly, the across-set and across-session variability also differed across frequency. Relatively low variability can be seen at some frequencies (e.g., 1000, 1122, and 1414 Hz), whereas high across-set and across-session variability can be seen at others (e.g., 1260 and 1587 Hz). The frequency band with the largest δ_{norm} was 1414 Hz. Although other frequencies exhibited larger magnitude and/or phase changes (e.g., 1000, 1587, and 1782 Hz), their across-set variability at baseline was greater.

Qualitatively similar results across frequency were seen for the other subjects. Figure 5 shows the across-set variability $\left(\sqrt{\text{SSD}}\right)$ obtained at all seven frequencies for each subject. It was expected that frequencies with better TEOAE SNRs would be less variable. Therefore, subjects were ordered by their median TEOAE SNRs in descending order from left to right. Contrary to this expectation, a visual inspection of the data suggested that SNR was not related to $\sqrt{\text{SSD}}$. This observation was confirmed by a lack of significant statistical correlation between SNR and $\sqrt{\text{SSD}}$ ($p > 0.05$). Because MOC shifts can demonstrate high variability across frequency in some subjects (e.g., subject 06), an individualized frequency analysis may be important to minimize measurement variability across time.

Across-Subject Variability • Examples of MOC shifts from six individual subjects are shown in Figure 6 to highlight the variety of results seen across subjects. The results are shown for the frequency band having the largest δ_{norm}. Panel A shows a representative example of MOC shifts occurring in the expected directions (i.e., a magnitude decrease and a phase lead). This subject had 6 instances of possible MEMR (i.e., 6 out of 16 recorded sets), and no instances of probable MEMR or SSOAEs. MOC shifts were relatively small, but all the δ values were clustered together, indicating low across-set and across-session

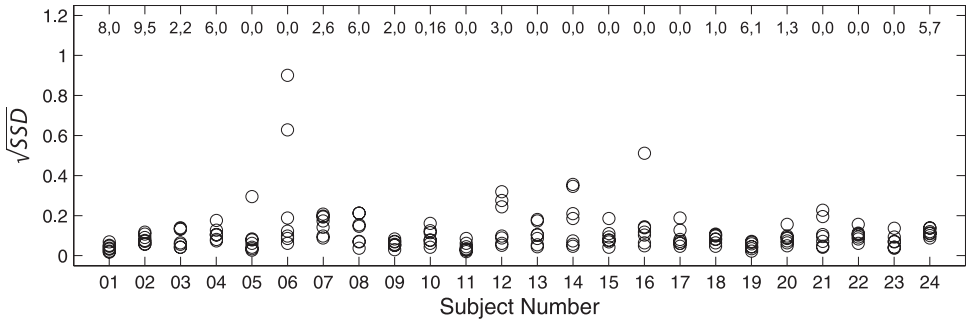

Fig. 5. Across-set variability $\left(\sqrt{\text{SSD}}\right)$ obtained at all frequency bands for each subject. Each $\sqrt{\text{SSD}}$ value (one for each frequency band) is plotted as an *unfilled circle*. Subjects are ordered by their median TEOAE SNR (ranging from 41.1 to 21.9 dB, in descending order from left to right). *Numbers* at the *top* of the panel indicate the number of measurements (out of 16 total) that contained possible and probable MEMR (first and second numbers, respectively) for each subject. MEMR indicates middle-ear muscle reflex; SNR, signal to noise ratio; SSD, sum of squared distances; TEOAE, transient-evoked otoacoustic emission.

variability. Most subjects (20 of 24) also showed median MOC shifts in the expected directions.

In contrast to the data shown in panel A, the data in panel B show low across-set variability at session 1 but high within-set and across-set variability at the remaining sessions. This subject had no instances of possible MEMR, probable MEMR, or SSOAEs. However, this subject had higher measurement variability than the subject in panel A, as evidenced by the larger areas of R_α. It is unclear what contributed to the change in MOC shift variability between session 1 and the remaining sessions. However, these results demonstrated that MOC shift variability at baseline can sometimes fail to account for the variability seen at subsequent sessions.

Panels C and D show data from two subjects that differed in MEMR results: the subject in panel C had no possible or probable MEMR, whereas the subject in panel D had no possible MEMR but 16 instances of probable MEMR. Both subjects had 15 to 16 instances of SSOAEs. The subject in panel C demonstrated apparent MOC shifts that were in the expected directions at baseline, with low within-set variability. However, high within-set and across-set variability was demonstrated at additional sessions. The data shown in panel D demonstrate low within-set variability at baseline and session 2, with more variability seen at sessions 3 and 4. Interestingly, the MOC shifts were much larger in phase than in magnitude. No other subjects demonstrated such a large number of probable MEMR cases, so it is not known the extent to which these large phase shifts indicate MEMR activity. Example C suggests that MOC shift variability may not necessarily be related to the presence of MEMR.

Panels E and F show data from two subjects that differed in SSOAE results: the subject in panel E had no instances of SSOAEs, whereas the subject in panel F had SSOAEs at all measurements. The subject in panel E had small magnitude and phase shifts, but with low within-set and across-set variability.

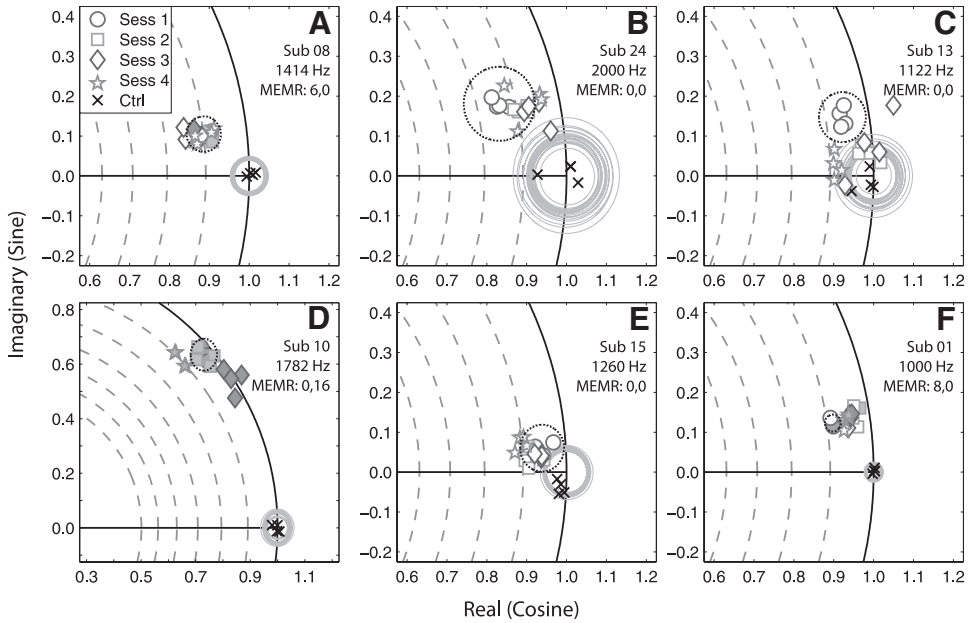

Fig. 6. Six examples of MOC shift results. Each panel represents data obtained from a different subject. The subject numbers, analysis frequency bands, and number of measurements containing possible and probable MEMR (first and second numbers, respectively) are displayed in the *upper right-hand corners*. Figure format is identical to that in Figure 4, with the addition of filled markers representing instances of possible or probable MEMR. Note that the results shown in (D) are plotted on a different scale from the other panels. See text for discussion of each panel. MEMR indicates middle-ear muscle reflex; MOC, medial olivocochlear.

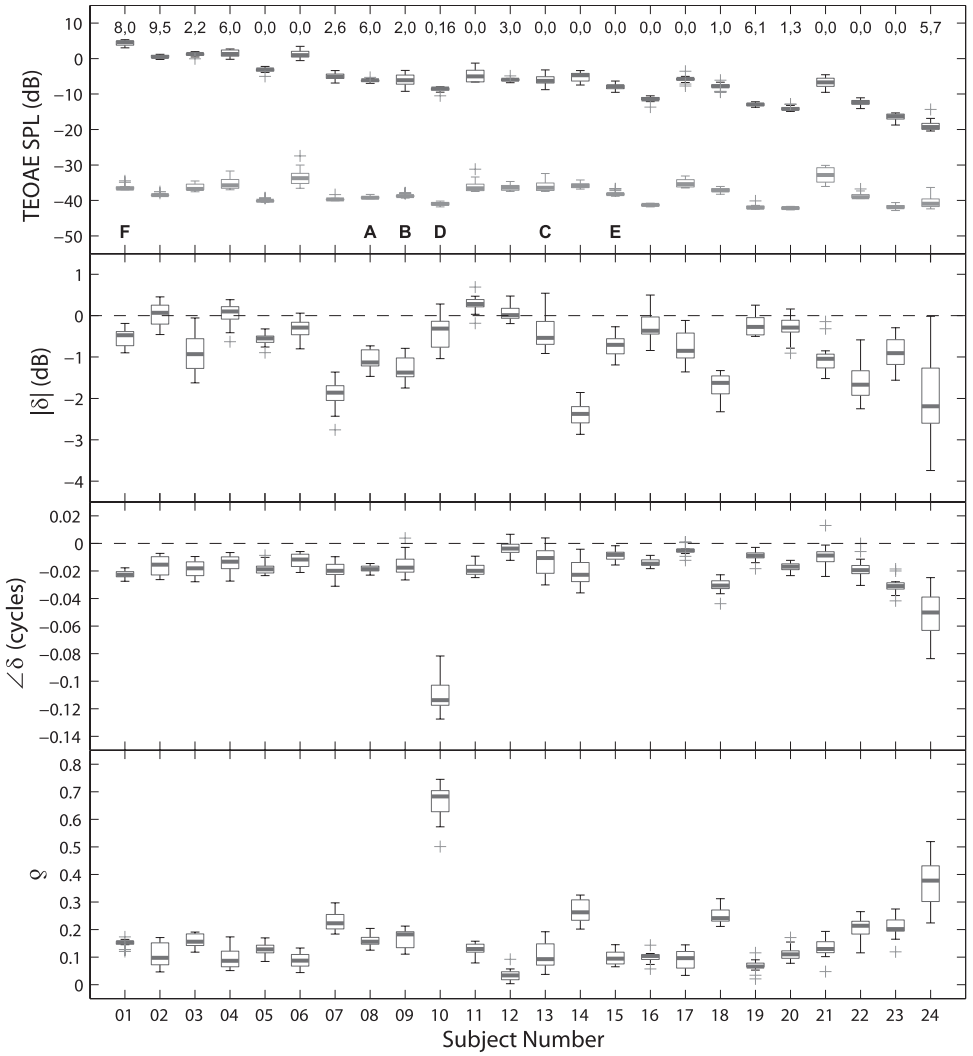

Fig. 7. TEOAE data (panel 1) and MOC shift data (panels 2–4) for all individual subjects. MOC shift data were obtained at the frequency with the largest δ_{norm}. Each box plot represents the distribution of values obtained in a subject at all 16 measurements in the test condition. Boxes represent the first and third quartiles. *Horizontal lines* within each box represent the medians. *Gray plus* symbols represent outliers, defined as any values falling above or below 1.5 × the IQR. Whiskers represent the smallest and largest values not considered outliers. For all panels, subjects are ordered by their median TEOAE SNR values, in descending order from left to right. *Dashed horizontal lines* on panels 2 and 3 represent no change. *Panel 1*: Distributions of RMS amplitudes of TEOAE signals (box plots at *top* of panel) and TEOAE noise floors (gray box plots at *bottom* of panel) obtained from the *A* matrices (i.e., no CAS). *Numbers* at the *top* of the panel indicate the number of measurements (out of 16 total) that contained possible and probable MEMR (first and second numbers, respectively). *Bold letters* (A–F) at the *bottom* of the panel refer to the subjects shown in the respective panels in Figure 6. *Panel 2*: Distribution of magnitude changes ($|\delta|$). *Panel 3*: Distribution of phase changes ($\angle\delta$). *Panel 4*: Distributions of total quantities of change (ϙ). CAS indicates contralateral acoustic stimulation; IQR, interquartile range; MEMR, middle-ear muscle reflex; MOC, medial olivocochlear; RMS, root mean square; SNR, signal to noise ratio; TEOAE, transient-evoked otoacoustic emission.

The subject in panel F also had small magnitude shifts but with larger phase shifts than the subject in panel E. There was also very low within-set variability but somewhat higher across-set variability than the subject in panel E (but lower than most of the subjects presented in Fig. 6). These results qualitatively suggest that SSOAEs may exhibit minimal or no impact on MOC shift variability (quantitative examinations are described below).

Across all subjects, 351 MOC shifts (91.4%) were statistically significant. It should be noted that 11.4 to 13.4% of these MOC shifts may have included contributions from MEMR (using $\alpha = 0.01$ and $\alpha = 0.05$, respectively). Taking into account repeated measures, 19 out of 24 subjects (79.2%) showed significant MOC shifts at 15 or 16 (out of 16) measurements. Four subjects showed significant MOC shifts at 10 to 12 measurements. Finally, 1 subject showed significant MOC shifts at only 3 measurements. These results are consistent with Goodman et al. (2013), who reported significant shifts for at least 1 frequency from 1 to 2 kHz in 14 of 16 subjects (87.5%).

To allow for visual comparison of MOC shifts across subjects, data for all individuals are plotted in Figure 7 as box and whisker plots. TEOAE signal and noise floor amplitudes are plotted in the top panel, in order of descending SNR from left to right (as in Fig. 5). Results for three measures of MOC shift ($|\delta|$, $\angle\delta$, and ϙ) are plotted in the second, third, and fourth panels, respectively. Median TEOAE amplitudes (calculated across the 16 measurements without CAS for each subject) ranged from −19.3 to 4.6

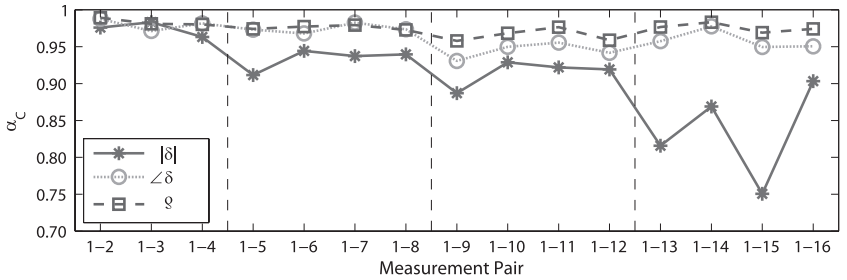

Fig. 8. Cronbach's alpha (α_C) for the three quantifications of MOC shifts, computed between measurement 1 and all subsequent measurements. *Vertical dashed* lines are used to separate each of the subsequent measurements in terms of the session (1–4) in which it occurred. MOC indicates medial olivocochlear.

dB SPL. Median noise floor amplitudes ranged from −42.2 to −32.8 dB SPL. Median SNRs ranged from 21.9 to 41.1 dB.

There was considerable variability in the size of MOC shifts exhibited across subjects. Median $|\delta|$ values ranged from −2.37 to 0.28 dB. A majority of magnitude shifts (327 of 384, or 85.2%) were in the expected (negative) direction. There was no significant correlation between $|\delta|$ and TEOAE noise floor ($p > 0.05$). Median $\angle\delta$ values ranged from −0.11 to −0.004 cycles. A majority of the phase shifts (374 of 384, or 97.4%) were in the expected direction (phase lead). Median ϙ values across subjects ranged from 0.03 to 0.68.

Across-Set and Across-Session Variability • Across-set and across-session variability was quantified for each subject by computing the standard deviations of MOC shifts ($|\delta|$, $\angle\delta$, and ϙ) across all 16 measurements in the test condition. For $|\delta|$, across all subjects the average SD was 0.33 dB (10th percentile = 0.20 dB, 90th percentile = 0.44 dB, range = 0.14 to 0.92 dB). For $\angle\delta$, across all subjects the average SD was 0.0062 cycles (10th percentile = 0.0027 cycles, 90th percentile = 0.0106 cycles, range = 0.0024 to 0.0158 cycles). For ϙ, across all subjects the average SD was 0.034 (10th percentile = 0.020, 90th percentile = 0.050, range = 0.015 to 0.081).

The repeatability of MOC shifts across set and session was also computed using Cronbach's alpha (Cronbach 1951). Cronbach's alpha is an index of the reliability of results assessed by different raters. In this study, the "raters" were considered the different test measurements. Cronbach's alpha was computed as

$$\alpha_C = \frac{k\bar{c}}{\bar{v} + (k-1)\bar{c}}, \tag{11}$$

where k is the number of measurements, $\bar{c}$ is the mean correlation between all measurements, and $\bar{v}$ is the mean variance of all measurements. The possible values of α_C range from 0 to 1, with 1 indicating perfect reliability.

Based on previous results (Mishra & Lutman 2013), it was expected that all α_C values would be at least 0.7. α_C values were computed separately for $|\delta|$, $\angle\delta$, and ϙ. To examine the repeatability across time, α_C was computed for each pairwise comparison between the first measurement and each subsequent measurement. Each quantification of MOC shift showed high repeatability ($\alpha_C > 0.75$) across all measurement pairs (Fig. 8), with $|\delta|$ showing reduced repeatability across measurements relative to $\angle\delta$ and ϙ.

Impact of SSOAEs on MOC Shift Variability • Separate correlations were computed between each quantification of MOC shift variability ($|\delta|$, $\angle\delta$, and ϙ) and the following: count of SSOAEs, SSOAE amplitude, SSOAE SNR, and ratio of SSOAE energy to TEOAE energy. All correlations were nonsignificant ($p > 0.05$). In addition, there were no significant correlations between the magnitude of the tolerance regions $\left(\left|R_\alpha\right|\right)$ and SSOAEs (quantified by amplitude, SNR, and energy ratio; $p > 0.05$ in all cases).

Accuracy of Baseline Variability for Explaining Subsequent Variability • Computer simulations (not shown) suggested that using four baseline measurements to estimate $\overline{R_\alpha}$ was sufficient for encompassing 95% of subsequent values of δ. However, $\overline{R_\alpha}$ only encompassed 95% of subsequent values for 1 subject, suggesting that additional variability was present at subsequent sessions that was not present at baseline in most subjects. It was assumed that the variability in δ values obtained in all our subjects primarily represented measurement variability rather than changes in MOC function that represented improvements or decrements. Therefore, determining whether a significant change in MOC function occurs based on where a given δ value falls relative to $\overline{R_\alpha}$ would result in an unacceptable number of false positive responses. The amount of scaling of $\overline{R_\alpha}$ needed to encompass 95% of subsequent δ values was examined. Across all subjects, the median scaling factor was 1.61 and the range was 1 to 5.6, meaning that the amount that $\overline{R_\alpha}$ had to be expanded before it encompassed subsequent MOC shifts varied across subjects. Because these scaled $\overline{R_\alpha}$ ellipses encompassed 95% of repeated measures, the magnitude of scaled $\overline{R_\alpha}$ can be used to determine the size (magnitude) of the minimum detectable change that falls outside of measurement variability. The median magnitude of the minimum detectable change was 0.77 dB, the 10th and 90th percentiles were 0.54 and 1.43 dB, and the range was 0.42 to 2.02 dB, respectively.

DISCUSSION

The primary purpose of the present study was to assess the variability across repeated measures of MOCR-induced changes in TEOAEs. The variability of the MOCR across time was of interest because measurements of MOC shifts may have clinical applications, such as predicting benefit from auditory training (de Boer & Thornton 2008), detecting the onset of presbycusis (Zhu et al. 2007), and detecting ototoxicity (Aran et al. 1994). The variability of MOC shift measurements across time is important to examine because high variability may preclude the ability to reliably interpret small changes (e.g., 0.5 dB) as being caused by treatments or pathologies. The remainder of this section will discuss issues important to repeated measurements of MOC shifts.

Impact of MEMR and SSOAEs on MOC Shifts

The presence of MEMR found in the present study was similar to the prevalence reported in previous studies utilizing

similar CAS levels (Goodman et al. 2013; Marshall et al. 2014). However, it is unclear the extent to which MEMR may have influenced the detection of MOC shifts as well as the variability. Future studies need to consider the tradeoff between using CAS levels to maximize MOCR activation and minimize MEMR activation.

SSOAEs were present in over half of the measurements for almost 2/3 of our subjects, consistent with previous reports (Sisto et al. 2001; Keefe 2012). Marshall et al. (2014) reported that the presence of SSOAEs did not interfere with the detection of MOC shifts. In our results, SSOAEs did not appear to impact the variability of MOC shifts in an observable way, given the lack of significant correlations seen between MOC shift variability and the amplitude, SNR, and count of SSOAEs. Despite this observed lack of association in the present study, a similar lack of impact in other measurement paradigms cannot be ruled out. A click rate of 50/sec is common in clinical and research protocols (the present click rate was 26/sec); such a fast rate may not allow for substantial decay of SSOAE amplitudes that could thus affect the measured MOC shift variability more than what was seen in the present study. Higher click levels will also tend to elicit larger SSOAEs. Therefore, assessing the potential effect of SSOAEs on MOCR measurements in any given study warrants the careful consideration of the measurement protocol.

Assessment of Magnitude and Phase Changes

A majority of MOC shifts exhibited the expected changes in both magnitude and phase. A small number of subjects exhibited unexpected increases in magnitude and/or phase lags. These "enhancements" in TEOAE magnitude have been reported previously for a minority of subjects (Hood et al. 1996; De Ceulaer et al. 2001; Garinis et al. 2008; Goodman et al. 2013). Enhancements were smaller changes overall in both magnitude and phase, relative to changes in the expected directions (i.e., magnitude decrease and a phase lead). The origin of such enhancements is not known. Garinis et al. saw enhancements in adults with learning disabilities and speculated that the enhancements may have resulted from abnormalities in brainstem function or neurotransmitter release. An alternative explanation may be a differential MOC effect on distortion and reflection TEOAE components; such a differential effect has been established for distortion-product OAEs (e.g., Abdala et al. 2009). Although some evidence suggests that TEOAEs elicited by high-level clicks may contain distortion and reflection components (e.g., Moleti et al. 2012), the TEOAEs measured in the present study were likely composed of primarily reflection components due to the use of relatively low stimulus levels (Lewis & Goodman 2015).

Repeatability was higher when MOC shifts were quantified as phase shifts and total quantities of change than for magnitude changes alone (Fig. 8). It may therefore be preferable to quantify MOC shifts in terms of phase changes or total quantities of change, at least for the purposes of repeated measures.

The results were analyzed at the frequency band exhibiting the least variability at baseline, coupled with the largest mean MOC shift magnitude. A sum of squares approach was used to reduce the likelihood of selecting a frequency band that had an outlier, even if the remaining δ values were tightly clustered. It is possible that alternative approaches to quantifying MOC shifts and their variability might yield more stable results. TEOAEs are useful because responses can be elicited relatively quickly and the most stable response frequencies can be chosen post-hoc.

Establishing a Reliable Baseline

A large and/or unreliable baseline may render it difficult or impossible to determine whether subsequent MOC shift measurements represent a true change in MOCR function, or whether the MOC shifts fall within measurement variability. Marshall et al. (2014) reported that at a single session, the number of measurements needed to obtain a stable MOC shift varied across individual subjects, with the median number of measurements being 3. For repeated measures of MOC shifts, the present results indicated that four baseline measurements were inadequate for all but one subject. This suggests that in some cases, very large changes in MOC shifts (e.g., 1.5 to 2 dB changes) may be required before the change could be considered significant. More study is needed to determine optimum methods of quantifying baseline MOC function to which repeated measures can be compared.

MOC Shift Variability Within and Across Subjects

A range of within-subject variability in MOC shifts was exhibited, as demonstrated quantitatively by the standard deviations and also visually by the sizes of interquartile ranges (Fig. 7, panels 2–4). These results are consistent with Marshall et al. (2014). However, no significant correlations were found between TEOAE SNR and MOC shift variability in the present study, unlike Marshall et al. (see their Fig. 3),which may be explained by differences in TEOAE SNR between the two studies. It is possible that variability becomes less reliant on TEOAE SNRs exceeding 22 dB or so. The variability in MOC shifts was consistent with previous reports. Marshall et al. reported standard deviations that ranged from approximately 1 to 10% (see their Fig. 3). For the present dataset, the standard deviations of ϙ ranged from 0.01 to 0.07, equivalent to a range of 1 to 7%. Mishra and Lutman (2013) reported test–retest differences that rarely exceeded 5%.

The high values of α_C in the present study were quantitative evidence of the high repeatability of MOC shifts across time, on average. Mishra and Lutman (2013) reported smaller α_C values than in the present study. They did not control for slow drifts in TEOAE amplitude or for changes in the subjects' attentional states, and their TEOAE SNRs were lower than in the present study. One or more of these factors may explain their lower, but still acceptable, α_C values.

Potential Limitations and Future Directions

MEMR was only examined in terms of the likelihood of its activation, and responses with possible or probable MEMR were flagged. For clinical applications, it may be preferable to either avoid elicitation of MEMR or to determine the relative contribution of MEMR shifts to measured MOC shifts. Avoidance of MEMR elicitation would require either presenting CAS stimuli at a level that does not activate the MEMR in any subject (e.g., 50 dB SPL; Guinan et al. 2003) or presenting below each subject's threshold for MEMR activation (although this would result in different stimulus levels for different individuals). The relative contributions of MEMR versus MOCR could

be determined through an analysis of group delay (Guinan et al. 2003). Further refinements to the detection and quantification of MEMR shifts may be needed in order for MOC shift measurements to be clinically useful.

It is possible that different analysis methods may reveal MOC shifts that are larger and/or more stable than what was reported in the present study. MOC shifts were quantified across different frequency bands by averaging the complex ratio of Fourier transforms of TEOAE waveforms obtained with versus without CAS (see Eqs. 1 and 2). It is possible that averaging in this way could have in some instances canceled out MOC effects occurring in opposite directions across frequencies, if significant phase differences were present across the FFT frequency bins included in the average. Backus and Guinan (2007) showed variability in closely spaced stimulus frequency OAEs (within 40 Hz). However, it is not known whether these changes were caused by magnitude differences or phase differences. More study is needed to quantify the optimal analysis bandwidth. Using the present methodology, the number of FFT bins in the average can be reduced by modifying the cutoff frequencies in Equation 2 (in the limit of a single FFT bin). Either as part of this approach or separately, the number of FFT bins can also be reduced by first downsampling the time domain waveform. Alternatively, amplitude and phase changes can be examined in the time domain using bandpass filtered waveforms rather than an FFT analysis. Given the potential clinical utility of measurements of MOC shifts, further investigation into the optimal analysis bandwidth to limit MOC shift variability appears warranted.

CONCLUSIONS

MOC shifts in young normal-hearing subjects demonstrated low variability across a 5-week time span, on average. Repeatability was higher when quantifying the shifts as changes in phase and as total quantities of change, relative to changes in magnitude alone. Despite the low variability found in group data, some individuals demonstrated large variability within and across test sessions, despite implementation of methods to reduce variability. Simulations suggested that four baseline measurements of MOC shifts would adequately predict variability at subsequent measurements, but human data showed that four measurements underestimated the variability at subsequent measurements. SSOAEs did not appear to show consistent effects on MOC shift variability, but it should be remembered that this finding may not extend to other stimulus rate/level paradigms. The impact of MEMR activation, when it did occur, on MOC shift variability was unclear and will require further investigation to delineate its effects on measurements of MOC shifts. For some subjects, a change in MOC shift on the order of 1.5 to 2 dB would be required before any change could be attributed to something other than measurement variability. It therefore appears that MOC shifts, as analyzed in the present study, may be too variable to reliably detect small changes across time, at least in some individuals.

ACKNOWLEDGEMENTS

This investigation was supported by grants from the University of Iowa Graduate College and the University of Iowa Executive Council of Graduate and Professional Students (both to I.B.M.).

Portions of this article were presented at the 2014 MidWinter Meeting of the Association for Research in Otolaryngology (ARO), February 22–26, 2014, and at the 2014 SoCal Hearing Conference, September 13, 2014. The first author received a travel award from the American Academy of Audiology Foundation to present the results at the ARO Meeting. This research served as part of the first author's doctoral dissertation work.

All authors contributed equally to this work. I.B.M. designed and performed experiments, analyzed data, and wrote the paper and Supplemental Digital Content. S.S.G. designed experiments, analyzed data, and wrote the paper and Supplemental Digital Content.

The authors have no conflicts of interest to disclose.

Address for correspondence: Ian B. Mertes, Research Service (151), VA Loma Linda Healthcare System, 11201 Benton Street, Loma Linda, CA 92357, USA; and Loma Linda Veterans Association for Research and Education, Redlands, CA, USA. E-mail: ian.mertes@va.gov

Received June 12, 2015; accepted September 24, 2015.

REFERENCES

Abdala, C., Mishra, S. K., Williams, T. L. (2009). Considering distortion product otoacoustic emission fine structure in measurements of the medial olivocochlear reflex. *J Acoust Soc Am*, *125*, 1584–1594.

American National Standards Institute (ANSI) (1996). *Specification for audiometers*. *ANSI Report No. S3.6-1996*. New York, NY: American National Standards Institute.

Aran, J. M., Erre, J. P., Avan, P. (1994). Contralateral suppression of transient evoked otoacoustic emissions in guinea-pigs: Effects of gentamicin. *Br J Audiol*, *28*, 267–271.

Backus, B. C., & Guinan, J. J., Jr. (2006). Time-course of the human medial olivocochlear reflex. *J Acoust Soc Am*, *119*(5 Pt 1), 2889–2904.

Backus, B. C., & Guinan, J. J., Jr. (2007). Measurement of the distribution of medial olivocochlear acoustic reflex strengths across normal-hearing individuals via otoacoustic emissions. *J Assoc Res Otolaryngol*, *8*, 484–496.

Berlin, C. I., Hood, L. J., Wen, H., et al. (1993). Contralateral suppression of non-linear click-evoked otoacoustic emissions. *Hear Res*, *71*, 1–11.

Brownell, W. E. (1990). Outer hair cell electromotility and otoacoustic emissions. *Ear Hear*, *11*, 82–92.

Chew, V. (1966). Confidence, prediction, and tolerance regions for the multivariate normal distribution. *J Am Statist Assoc*, *61*, 605–617.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.

de Boer, J., & Thornton, A. R. (2007). Effect of subject task on contralateral suppression of click evoked otoacoustic emissions. *Hear Res*, *233*, 117–123.

de Boer, J., & Thornton, A. R. (2008). Neural correlates of perceptual learning in the auditory brainstem: Efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *J Neurosci*, *28*, 4929–4937.

De Ceulaer, G., Yperman, M., Daemers, K., et al. (2001). Contralateral suppression of transient evoked otoacoustic emissions: Normative data for a clinical test set-up. *Otol Neurotol*, *22*, 350–355.

Efron, B., & Tibshirani, R. J. (1993). *An Introduction to the Bootstrap* (1st ed.). New York, NY: Chapman and Hall.

Garinis, A. C., Glattke, T., Cone-Wesson, B. K. (2008). TEOAE suppression in adults with learning disabilities. *Int J Audiol*, *47*, 607–614.

Goodman, S. S., Fitzpatrick, D. F., Ellison, J. C., et al. (2009). High-frequency click-evoked otoacoustic emissions and behavioral thresholds in humans. *J Acoust Soc Am*, *125*, 1014–1032.

Goodman, S. S., Mertes, I. B., Lewis, J. D., et al. (2013). Medial olivocochlear-induced transient-evoked otoacoustic emission amplitude shifts in individual subjects. *J Assoc Res Otolaryngol*, *14*, 829–842.

Graham, R. L., & Hazell, J. W. (1994). Contralateral suppression of transient evoked otoacoustic emissions: Intra-individual variability in tinnitus and normal subjects. *Br J Audiol*, *28*, 235–245.

Guinan, J. J., Jr. (2006). Olivocochlear efferents: Anatomy, physiology, function, and the measurement of efferent effects in humans. *Ear Hear*, *27*, 589–607.

Guinan, J. J., Jr, Backus, B. C., Lilaonitkul, W., et al. (2003). Medial olivocochlear efferent reflex in humans: Otoacoustic emission (OAE) measurement issues and the advantages of stimulus frequency OAEs. *J Assoc Res Otolaryngol*, *4*, 521–540.

Hood, L. J., Berlin, C. I., Hurley, A., et al. (1996). Contralateral suppression of transient-evoked otoacoustic emissions in humans: Intensity effects. *Hear Res*, *101*, 113–118.

Humphrey, R. (2008). *Playrec Version 2.1.0*. Retrieved August 11, 2015, from http://www.playrec.co.uk.

Keefe, D. H. (2012). Moments of click-evoked otoacoustic emissions in human ears: Group delay and spread, instantaneous frequency and bandwidth. *J Acoust Soc Am*, *132*, 3319–3350.

Khalfa, S., Morlet, T., Micheyl, C., et al. (1997). Evidence of peripheral hearing asymmetry in humans: Clinical implications. *Acta Otolaryngol*, *117*, 192–196.

Kumar, U. A., Methi, R., Avinash, M. C. (2013). Test/retest repeatability of effect contralateral acoustic stimulation on the magnitudes of distortion product ototacoustic emissions. *Laryngoscope*, *123*, 463–471.

Lewis, J. D., & Goodman, S. S. (2015). Basal contributions to short-latency transient-evoked otoacoustic emission components. *J Assoc Res Otolaryngol*, *16*, 29–45.

Marshall, L., Lapsley Miller, J. A., Guinan, J. J., et al. (2014). Otoacoustic-emission-based medial-olivocochlear reflex assays for humans. *J Acoust Soc Am*, *136*, 2697–2713.

Meric, C., & Collet, L. (1992). Visual attention and evoked otoacoustic emissions: A slight but real effect. *Int J Psychophysiol*, *12*, 233–235.

Mishra, S. K., & Abdala, C. (2015). Stability of the medial olivocochlear reflex as measured by distortion product otoacoustic emissions. *J Speech Lang Hear Res*, *58*, 122–134.

Mishra, S. K., & Lutman, M. E. (2013). Repeatability of click-evoked otoacoustic emission-based medial olivocochlear efferent assay. *Ear Hear*, *34*, 789–798.

Moleti, A., Botti, T., Sisto, R. (2012). Transient-evoked otoacoustic emission generators in a nonlinear cochlea. *J Acoust Soc Am*, *131*, 2891–2903.

Shera, C. A., Guinan, J. J., Jr, Oxenham, A. J. (2002). Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proc Natl Acad Sci USA*, *99*, 3318–3323.

Sisto, R., Moleti, A., Lucertini, M. (2001). Spontaneous otoacoustic emissions and relaxation dynamics of long decay time OAEs in audiometrically normal and impaired subjects. *J Acoust Soc Am*, *109*, 638–647.

Zhu, X., Vasilyeva, O. N., Kim, S., et al. (2007). Auditory efferent feedback system deficits precede age-related hearing loss: Contralateral suppression of otoacoustic emissions in mice. *J Comp Neurol*, *503*, 593–604.