Chirp evoked VEMPs: A test-retest reliability study

[1]Tarryn Marisca Reddy, [2]Barbara Heinze, [3]Leigh Biagio-de Jager and [4,5] Leen Maes

[1 3] Department of Speech-Language Pathology and Audiology, University of Pretoria, South Africa

[2] Ear Science Institute Australia, Western Australia

[4] Department of Rehabilitation Sciences, Ghent University, Belgium

[5] Department of Ear Nose Throat, Ghent University Hospital, Belgium

Financial disclosures/conflicts of interest: There are no conflicts of interest, financial, or otherwise.

All correspondence should be addressed to:

Tarryn M. Reddy, Department of Speech-Language Pathology and Audiology, University of Pretoria, South Africa, Gauteng, South Africa. Email: tarryn.reddy@up.ac.za

Keywords: cervical VEMP, ocular VEMP, test-retest reliability, CE-Chirp, saccule, utricle, reproducibility

Abstract

Objectives: To determine the test-retest reliability of cervical and ocular vestibular evoked myogenic potentials (c&oVEMP) evoked by 500 Hz narrowband (NB) CE-Chirp and broadband (BB) CE-Chirp stimuli.

Design: Twenty healthy participants (10 female) were tested twice on the same day to determine the within-session reliability and one week later to determine the between-session reliability. The latency, amplitude, and asymmetry ratio of c&oVEMPs elicited by 95 dB nHL air conducted (AC) 500 Hz NB CE-Chirp and BB CE-Chirp were recorded bilaterally.

Results: A moderate to good between-session reliability with ICC values ranging from 0.52 to 0.82 was observed for cVEMP latency, amplitude, and asymmetry ratio evoked by 500 Hz NB CE-Chirp, as well as for the BB CE-Chirp cVEMP amplitude (ICC of 0.70 and 0.84). In contrast, an overall poor reliability ICC values between 0.30 to 0.42 for latency and asymmetry ratio was observed for BB CE-Chirp. For the oVEMP, overall poor between-session reliability for all response parameters evoked by the 500 Hz NB CE-Chirp and the BB CE-Chirp was observed.

Conclusion: The 500 Hz NB CE-Chirp was more reliable than the BB CE-Chirp in terms of cVEMP latency, amplitude, and asymmetry ratio, and comparable to the 500 Hz TB stimulus. Further investigation utilizing the standard electrode montage is necessary to assess the test-retest reliability of the chirp evoked oVEMP.

INTRODUCTION

Vestibular evoked myogenic potentials (VEMPs) were first described by Colebatch, Halmagyi and Skuse (1994) as a component of the vestibular test battery. This test can assist in the differential diagnosis of neuro-otological diseases, including, but not limited to, Meniere's disease, superior semicircular canal dehiscence, vestibular neuronitis, vestibular schwannomas and perilymphatic fistulas (Cal & Bahmad 2009). VEMPs are considered widely as a useful tool to supplement caloric, rotational chair, and (video) head impulse testing, which are limited to the evaluation of the horizontal semicircular canal function and superior vestibular nerve integrity (Akin et al. 2003; Desmond 2011).

There are two types of VEMP responses, namely, cervical VEMP (cVEMP) and ocular VEMP (oVEMP). The cVEMP is an ipsilaterally inhibited potential, recorded by an electrode placed on the sternocleidomastoid muscle (SCMm) in response to a sound stimulus, resulting in the activation of the saccular macula, inferior vestibular nerve, vestibular nucleus, the medial vestibulo-spinal tract, accessory nucleus and nerve, and the ipsilateral SCMm (Akin et al. 2003; Rosengren et al. 2010). The oVEMP reflects predominantly utricular function and is a contralateral excitatory potential, recorded from the extraocular muscles due to activation of the utricle, superior vestibular nerve and nucleus, and the oculomotor nuclei (Rosengren et al. 2010). The cervical and ocular VEMP (c&oVEMP) responses are interpreted according to specific parameters, i.e., P1 latency, N1 latency, P1-N1 amplitude, VEMP threshold and asymmetry ratio (Isaradisaikul et al. 2012).

It has been clearly established that the VEMP response rate, latency and amplitude parameters are largely dependent on the stimulus design and frequency, as well as the intensity of the stimulus (Eleftheriadou et al. 2009). Initially, the click stimulus was thought to be the most precise, with the highest response rate, shorter latencies and, larger amplitudes compared to tone bursts (TB) (Cheng et al. 2003). However, it was later discovered that AC TB stimuli between 500 and 1000 Hz resulted in larger amplitudes (Murofushi et al. 1999; Akin et al. 2003; Singh et al. 2014), greater reliability, and smaller inter-laboratory variability than click-evoked VEMPs (Meyer et al. 2015), and the 500 Hz TB became the preferred stimulus to reliably perform cVEMPs. Regarding oVEMPs, electromechanical vibrators, such as the minishaker (type 4810, Bruel and Kjaer), haven proven to be the preferred stimulus due to the high response rates obtained (Cheng et al. 2009; Wang et al. 2010). However, due to the minishaker being costly and not certified for clinical use, the 500 Hz AC TB stimulus has been most frequently used (Rosengren et al. 2010). Several studies have, therefore, focused on the clinical applicability and validation of VEMPs evoked by a click and 500 Hz TB stimulus (Colebatch et al. 1994; Akin et al. 2003; Rosengren et al. 2010; Singh et al.

2014). Subsequently, there is a substantial amount of literature on the test-retest reliability of cVEMPs and oVEMPs evoked by click and TB stimuli (Isaradisaikul et al. 2008; Eleftheriadou et al. 2009; Maes et al. 2009; Nguyen et al. 2010; de Oliveira et al. 2014). Isaradisaikul et al. (2008) reported fair to good test-retest reliability for cVEMPs evoked by an air-conducted (AC) 500 Hz TB stimulus at 110 dB HL with and without the use of electromyography (EMG). The P1-N1 amplitude was found to be more reliable than latency, with the N1 latency proving more reliable that the P1 latency. Eleftheriadou et al. (2009) found that cVEMPs evoked by a click stimulus at 140 dB SPL (105 dB nHL) produced reliable P1 latency, N1 latency, and P1-N1 amplitudes, suggesting that this method is reproducible and feasible. Nguyen et al. (2010) investigated the test-retest reliability of c&oVEMPs evoked by click at 140 dB SPL (105 dB nHL), 500 Hz TB at 125 dB SPL (110 dB nHL), and two midline vibration stimuli delivered at the Fz position, i.e., hammer taps and mini-taps. The authors found better test-retest reliability for oVEMP parameters than cVEMP parameters.

A newly developed CE-Chirp stimulus has produced larger amplitudes (Walther & Cebulla, 2016) and shorter latencies for cVEMP (Wang et al. 2014; Ozgur et al. 2015) and oVEMP (Bas et al. 2020; Karacayli, 2020) compared to TB and click stimuli, at lower intensity levels of 100 – 105 dB nHL. Reddy et al. (2022) found that the 500 Hz NB CE-chirp provides the highest response rates, shorter P1 and N1 latencies and overall, larger amplitudes, at a lower intensity level of 95 dB nHL, and therefore seems a promising stimulus for reliably measuring c&oVEMPs in clinical practice. The differences reported in latency is possibly related to the differences in stimulus shape and rise time of the CE-Chirp. The chirp is a frequency and time modulated stimulus which increases the temporal synchrony within the cochlea by compensating for the time delay (Bas et al. 2020) caused by the traveling wave theory described by Bèkèsy (1952). The chirp stimulus can be designed to include a wide frequency range (i.e. broad band (BB)) or as frequency specific (i.e. narrow-band (NB)), octave-band chirps (500, 1000, 2000 and 4000 Hz) (Wang et al. 2014). Recently, Bas et al. (2020) found that a BB chirp, which included frequencies between 10 and 10 000 Hz, produced the highest oVEMP response rate, largest N1-P1 amplitude, and shortest latency compared to 500 Hz TB and click stimuli. Karacayli et al. (2020) also found larger N1-P1 amplitudes and shorter latencies for oVEMPs evoked by a 500 Hz NB CE-Chirp as compared to that evoked with a 500 Hz TB stimulus. Significantly larger cVEMP P1-N1 amplitudes for 500 Hz NB chirp compared to 500 Hz TB have been reported (Moinudeen (2020). Shorter latencies and smaller amplitudes for cVEMPs evoked by BB chirp stimulus containing several frequencies (500-4000Hz) compared to a 500 Hz TB stimulus have been described (Ozgur et al. (2015). Walther and Cebulla (2016) designed a chirp stimulus (250-1000 Hz) specifically for c&oVEMP testing and reported significantly larger c&oVEMP amplitudes compared to the click and 500Hz tone burst stimuli. Therefore, the literature suggests that the CE-Chirp seems a promising stimulus to estimate saccular and utricular function in clinical practice.

There are, however, no studies, to our knowledge, that focus on the test-retest reliability of chirp evoked c&oVEMPs. In addition, it is apparent that the available literature on chirp evoked VEMPs is limited and lacks agreement on whether a BB or NB CE-Chirp stimulus should be utilized. For this reason, the aim of this study was to determine the test-retest reliability of c&oVEMPs evoked by 500 Hz NB CE-Chirp and BB CE-Chirp stimulus.

MATERIALS AND METHODS

Ethical clearance and informed consent

Prior to the commencement of the data collection, ethical approval was obtained from the University of Pretoria Research and Ethics Committee of the Faculty of Humanities (approval number GW20170407HS). Written informed consent was obtained from all participants who met the sample selection criteria and were included in the study.

Participants and study design

A quantitative, exploratory, and repeated measures within participant research design was utilized in the study. A purposive sampling method was used to recruit participants. Twenty participants (10 females) ranging from 19 to 39 years of age, with a mean age of 26 years (SD ± 1.188), were enrolled in the study. This age range was selected to exclude the effects of age-related hearing loss and/or vestibular dysfunction. Each participant underwent within and between-session testing. The data from each session was used for the test-retest reliability analysis of the 500 Hz NB CE-Chirp and the BB CE-Chirp.

Procedures

Audiological investigations

Prior to VEMP testing, an audiological assessment was conducted, including bilateral otoscopic examinations to exclude outer ear pathologies. Tympanometry (y-226 Hz), acoustic reflex testing (MT10, Interacoustics, Denmark), and pure tone air and bone conduction audiometry (Kuduwave, eMoyo, South Africa) were conducted to exclude conductive and sensorineural hearing loss. Pure tone air conduction audiometry thresholds within the normal limits of -10-25 dB HL across the frequency range of 250-8000 Hz were required for each participant (Gelfand, 2001). All equipment used in the study was calibrated prior to the commencement of data collection.

VEMP procedure

The Interacoustics Eclipse EP25 two-channel VEMP system (Interacoustics, AS, Assens, Denmark) was used to conduct cVEMPs and oVEMPs on all participants. The Interacoustics Eclipse was calibrated by a trained technician, prior to the commencement of data collection, according to output level, frequency, and time (Wilber, 2002), and certified under ISO 389-6: 2007, which specifies requirements for medical devices. Calibration of the short duration CE-Chirp stimulus level was done in peak-to-peak equivalent sound pressure level (peSPL). The EAR 3A insert earphones (EAR 3A, Etymotic research, USA) were connected to an occluded ear simulator, which was then connected to a sound level meter and oscilloscope. At each frequency the output was adjusted to correspond to 100 dB SPL on the sound level meter and the peak-to-peak voltage was recorded on the oscilloscope in millivoltage (mV). The short duration stimulus is then adjusted to match the mV as closely as possible. The sound level meter and microphone used to perform the calibration was calibrated prior to the calibration of the Interacoustics Eclipse.

The electrode montage proposed by Colebatch et al. (1994) was utilized in the recording of cVEMP responses. Non-disposable silver disc electrodes were used. The skin was cleaned at each electrode site with Nuprep abrasive paste. The non-inverting electrode was placed at the midpoint of the left and right sternocleidomastoid (SCM) muscle, the inverting electrode was placed on the upper sternum, and the ground electrode was placed on the forehead. Participants were seated in an upright position with their head rotated 45 degrees to the opposite side of the test ear, resulting in unilateral SCMm activation. An external electromyogram (EMG) monitor was self-monitored by participants to ensure equal and sustained contraction of the SCMm muscle. The external EMG monitor ensured that participants maintained their muscle contraction within a lower (50 μV) and upper (150 μV) limit. Thereafter, a mathematical correction for amplitude normalization was applied to determine the right and left cVEMP response. The current study utilized the Interacoustics Eclipse which determines the mean rectified EMG for each sweep of recording from a 100 ms pre-stimulus period and a root mean square (RMS) of the rectified EMG, i.e., the cVEMP response amplitude is the pre-stimulus EMG minus the EMG recorded when the stimulus is presented. This ensures that the cVEMP response, following the presentation of the stimulus, is not included in the average EMG contraction.

The oVEMP electrode montage comprised of the non-inverting electrode on the lateral canthus of the eye contralateral to the stimulated ear (Sandhu et al. 2013; Govender et al. 2016), the inverting electrode on the chin to limit reference contamination from surrounding facial muscles (Piker et al. 2011), and the ground electrode placed on the forehead (Todd et al. 2007). Participants were seated and asked to look up at a reference on the ceiling to maintain maximal upward gaze for the duration of the test. Electrode impedance was less than or equal to 5 kOhm. The participants rested for a minimum of 10 minutes between trials to prevent muscle fatigue.

<u>Stimulus parameters</u>

A rarefaction 500 Hz NB CE-Chirp stimulus (duration 9 ms, 95 dB nHL equivalent to 120.5 dB peSPL) and BB CE-Chirp stimulus (200 Hz-11 kHz; duration 8 ms, 95 dB nHL equivalent to 126.5 dB peSPL) were presented bilaterally via insert earphones (EAR 3A, Etymotic research, USA) and ER3-14B disposable foam eartips. Each stimulus was presented twice, with averaging of the response to 100-150 stimulus repetitions for cVEMPs and 500 repetitions for oVEMPs. Longer averaging is required for oVEMPs due to the reduced amplitude of the response compared the cVEMP, and an increased number of artefacts from periocular and surrounding facial muscles (Rosengren et al. 2019). A bandpass filter of 10-1000 Hz was utilized for both oVEMP and cVEMP recordings. An artefact rejection level of 800 μV was used for cVEMP and 400 μV was used for oVEMP. A VEMP wave reproducibility rate of >85% was accepted. The wave reproducibility score was calculated by the VEMP software, and established the quality and reliability of the VEMP response (Interacoustics A/S, 2020).

Response parameters

VEMP latency, amplitude, and asymmetry ratio response parameters were recorded in order to investigate the within-session and between-session reliability. All waveforms for each participant were recorded and marked by one trained audiologist. A cVEMP and oVEMP response was determined to be 'present' when a biphasic waveform within the specified time frame could be recorded. The first positive peak on the waveform was marked P1 and first negative deflection was marked N1 for cVEMPs. For oVEMPs the first negative peak in the waveform was marked N1 and the first positive deflection was marked P1. The P1 latency, N1 latency, P1-N1 amplitude, and asymmetry ratio were recorded for each stimulus. The asymmetry ratio calculation outlined by Akin and Murnane (2008) was utilized, i.e. [(AL – AR)/(AL+AR)] x 100. AL refers to the amplitude of the left ear and AR refers to the amplitude of the right ear.

Testing Sessions

To determine the internal consistency and reliability of c&oVEMP evoked by 500 Hz NB CE-Chirp and BB CE-Chirp, the variability between repeated trials was investigated. To record the within-session reliability, test 1 and test 2 were repeated approximately 15 minutes apart, without electrode replacement or insert earphone repositioning. To exclude the influence of systematic bias on the within-session reliability, test 3 was repeated one week later, to record the between-session reliability between test 1 and test 3.

Data Analysis

Descriptive and inferential statistical analyses were used in this study. All statistical analysis was performed using IBM SPSS (version 27) software for Windows. The distribution of the data was determined by the test of skewness, visual inspection of Q-Q plots, and by conducting the Shapiro-Wilk test across all variables of the study for each stimulus bilaterally. The data was found to be normally distributed for both cVEMP (W=0.489-0.983, $p > 0.05$) and oVEMP (W=0.718–0.9.77, $p > 0.05$) and parametric statistics were utilized to determine whether a statistically significant difference existed between the right and left ear results and between genders for each stimulus. A value of $p < 0.05$ was accepted as statistically significant.

An independent samples t-test indicated no statistically significant difference between gender across c&oVEMP parameters. In addition, a paired samples t-test indicated that a statistical difference was not observed between the right and left ear results for either c&oVEMP. Due to the statistical independence between the two ears, the data for the right and left ears were pooled for analysis (Coren & Hakstian, 1990). This pooled data was found to be normally distributed for cVEMP (W=0.905-0.985, $p > 0.05$) and not normally distributed for oVEMP (W=0.784–0.973, $p < 0.05$). Subsequently, data is presented for the right and left ear separately, and parametric statistics were utilized in the analysis of the data. The c&oVEMP data is presented as mean ± standard deviation (SD).

The intraclass correlation coefficient (ICC) estimates and their 95% confidence intervals were determined by an averaged-measures, absolute agreement, two-way mixed effects model (Koo & Li, 2016). The ICC is "a statistical estimate that measures the extent of agreement between at least two quantitative measurements" and determines the level of reliability, consistency, and stability (Bujang & Baharum 2017, pp 1). ICC values less than 0.5 were indicative of poor reliability, values between 0.5 and 0.75 indicated moderate reliability, values between 0.75 and 0.9 indicated good reliability, and values greater than 0.9 were indicative of excellent reliability (Koo & Li, 2016).

The Standard Error of Measurement (SEM) was used to determine the absolute reliability by providing the precision of individual measurements (Weir, 2005; Dontje et al. 2018). The SEM was calculated to provide an estimate of the band of confidence around each participant's raw score (Maes et al. 2009), using the formula: $\mathrm{SEM}=\mathrm{SD}\sqrt{1-\mathrm{ICC}}$ (Weir, 2005), where SD is the standard deviation and ICC is the intraclass correlation coefficient. The SEM provides a reference for assessing test scores over time by considering the differences in measurements. The SEM was then used to determine the minimal detectable difference (MDD) for each variable using the formula: $\mathrm{MDD}=\mathrm{SEM} \times 1.96 \times \sqrt{2}$ (Weir, 2005). The MDD provides the border between significant and non-significant by determining the effect size that relates to the critical value of the test statistic (Mair et al. 2020), thus, providing an indication of the minimal amount of change that can be interpreted as a real change in VEMP latency, amplitude, and asymmetry ratio for a participant over time.

A one-way repeated measures analysis of variance (ANOVA) was used to explore possible differences between the test sessions for each VEMP parameter. Mauchly's Test of Sphericity was used and when the assumption of sphericity was violated, the Greenhouse-Geisser correction was applied. A statistically significant difference ($p < 0.05$) obtained with ANOVA was further investigated using a post hoc pairwise significance test, also known as a pairwise comparison, to determine which groups were responsible for the significant difference obtained (Cohen & Lea, 2004).

RESULTS

The cVEMP and oVEMP results are presented separately.

cVEMP

Data for 20 right ears and 20 left ears were analyzed. The 500 Hz NB CE-Chirp P1 and N1 latencies were shorter than the BB CE-Chirp P1 and N1 latencies across all test sessions for both the right and left ear. In addition, the P1-N1 amplitude of the 500 Hz NB CE-Chirp was larger than that of the BB CE-Chirp for all test sessions bilaterally. Table 1 shows the mean ± SD of the P1 latency, N1 latency, P1-N1 amplitude, and asymmetry ratios for the right and left ear for each stimulus at 95 dB nHL for test 1, test 2 and test 3.

The repeated measures ANOVA revealed no significant differences for the 500 Hz NB CE-Chirp for the right ear P1 latency (f=1.213; p=0.474), N1 latency (f=1.361; p=0.613) and P1-N1 amplitude (f=1.744; p=0.132) and the left ear P1 latency (f=1.272; p=0.492), N1 latency

Table 1. Mean and standard deviation of right and left ear cVEMP P1 and N1 latencies, P1-N1 amplitude and asymmetry ratio for the 500 Hz NB CE-Chirp and BB CE-Chirp for each test session (*n*=20)

	Test 1				Test 2				Test 3			
	500 Hz NB CE-Chirp		BB CE-Chirp		500 Hz NB CE-Chirp		BB CE-Chirp		500 Hz NB CE-Chirp		BB CE-Chirp	
	Right ear	Left ear	Right ear	Left ear	Right ear	Left ear	Right ear	Left ear	Right ear	Left ear	Right ear	Left ear
P1 latency (ms)	10.60 ± 2.71	10.35 ± 3.12	14.02 ± 2.26	13.90 ± 3.58	11.27 ± 2.88	10.98 ± 2.33	14.82 ± 3.15	15.17 ± 2.81	10.68 ± 3.67	10.85 ± 3.06	14.62 ± 2.68	13.97 ± 2.91
N1 latency (ms)	17.60 ± 2.80	16.82 ± 2.95	19.62 ± 2.44	19.22 ± 4.23	18.05 ± 2.72	18.35 ± 1.82	22.18 ± 3.45	20.47 ± 3.54	17.45 ± 3.94	17.72 ± 2.87	20.65 ± 2.95	20.63 ± 3.06
P1-N1 amplitude (µV)	74.67 ± 30.58	75.01 ± 31.70	39.09 ± 17.78	40.38 ± 18.87	62.42 ± 29.09	64.66 ± 40.39	41.78 ± 14.69	43.01 ± 20.99	73.32 ± 38.96	80.11 ± 38.79	43.97 ± 22.29	48.58 ± 28.30
AR (%)	17.42 ± 13.52		12.64 ± 10.93		15.73 ± 14.99		18.18 ± 19.57		18.45 ± 8.24		15.39 ± 13.27	

Hz: Hertz; NB: Narrowband; BB: Broadband; ms: milliseconds; µV: microvolts; AR: asymmetry ratio (*f*=1.929; *p*=0.063) and P1-N1 amplitude (*f*=1.787; *p*=0.134) between two test moments, and this for the within as well as the between-session measurements. Also, no significant difference was obtained for cVEMP evoked by the BB CE-Chirp for within and between-session measurements for the right ear P1 latency (*f*=1.847; *p*=0.539), N1 latency (*f*=1.104; *p*=0.373) and P1-N1 amplitude (*f*=1.515; *p*=0.523), and the left ear P1 latency (*f*=2.000; *p*=0.316), N1 latency (*f*=1.882; *p*=0.348) and P1-N1 amplitude (*f*=1.734; *p*=0.215). A significant difference was not observed for 500 Hz NB CE-Chirp (*f*=3.980; *p*=0.061) and BB CE-Chirp (*f*=0.684; *p*=0.501) asymmetry ratio for within and between session measurements.

The 500 Hz NB CE-Chirp provided higher reliability scores and *p*-values < 0.05 for all cVEMP test parameters, i.e., P1 latency, N1 latency, P1-N1 amplitude, and asymmetry ratio, for both within and between-session reliability for the right and left ear. The smallest MDD and SEM values were obtained for the 500 Hz NB CE-Chirp P1 latency and N1 latency for within-session reliability. The ICC values with their 95% confidence intervals and significance levels, together with the SEM and MDD values for within and between-session reliability for cVEMPs evoked by 500 Hz NB CE-Chirp and BB CE-Chirp for the right and left ear are presented in Table 2 and Table 3, respectively.

Table 2. Right ear: The intraclass correlation coefficient (ICC) and their 95% confidence intervals (CI), standard error of measurement (SEM) and minimal detectable difference (MDD) as within-session and between-session reliability parameters for the cVEMP P1 latency, N1 latency, P1-N1 Amplitude and asymmetry ratio (AR) (*n*=20)

	500 Hz NB CE-Chirp								BB CE-Chirp							
Reliability parameters	**Within-session reliability**				**Between-session reliability**				**Within-session reliability**				**Between-session reliability**			
	P1 Latency (ms)	N1 Latency (ms)	P1-N1 Amp (μV)	AR	P1 Latency (ms)	N1 Latency (ms)	P1-N1 Amp (μV)	AR	P1 Latency (ms)	N1 Latency (ms)	P1-N1 Amp (μV)	AR	P1 Latency (ms)	N1 Latency (ms)	P1-N1 Amp (μV)	AR
ICC	0.93	0.87	0.78	0.75	0.64	0.64	0.82	0.71	0.62	0.60	0.76	0.31	0.34	0.30	0.84	0.33
ICC 95% CI	0.80 to 0.97	0.68 to 0.95	0.43 to 0.91	0.36 to 0.90	0.46 to 0.86	0.34 to 0.86	0.53 to 0.93	0.41 to 0.90	0.14 to 0.85	0.44 to 0.84	0.36 to 0.87	0.20 to 0.52	0.26 to 0.74	0.17 to 0.72	0.61 to 0.94	0.21 to 0.48
ICC Sig	0.000	0.000	0.000	0.002	0.019	0.018	0.000	0.002	0.006	0.021	0.001	0.660	0.198	0.221	0.000	0.726
SEM	0.72	1.01	14.34	13.52	2.20	2.36	16.53	8.24	1.40	1.54	10.37	10.93	2.17	2.47	8.92	13.27
MDD	1.99	2.80	39.76	37.48	6.11	6.55	45.82	22.84	3.87	4.27	28.74	30.30	6.03	6.85	24.72	36.78

Hz: Hertz; NB: Narrowband; BB: Broadband; ms: milliseconds; μV: microvolts; ICC: intraclass correlation coefficient; Sig: significance; SEM: standard error of measurement; MDD: minimal detectable differences

Table 3. Left ear: The intraclass correlation coefficient (ICC) and their 95% confidence intervals (CI), standard error of measurement (SEM) and minimal detectable difference (MDD) as within-session and between-session reliability parameters for the cVEMP P1 latency, N1 latency, P1-N1 Amplitude and asymmetry ratio (AR) (*n*=20)

	500 Hz NB CE-Chirp								BB CE-Chirp							
Reliability parameters	**Within-session reliability**				**Between-session reliability**				**Within-session reliability**				**Between-session reliability**			
	P1 Latency (ms)	N1 Latency (ms)	P1-N1 Amp (µV)	AR	P1 Latency (ms)	N1 Latency (ms)	P1-N1 Amp (µV)	AR	P1 Latency (ms)	N1 Latency (ms)	P1-N1 Amp (µV)	AR	P1 Latency (ms)	N1 Latency (ms)	P1-N1 Amp (µV)	AR
ICC	0.93	0.86	0.80	0.75	0.61	0.52	0.73	0.71	0.48	0.38	0.81	0.31	0.42	0.38	0.70	0.33
ICC 95% CI	0.80 to 0.97	0.63 to 0.92	0.51 to 0.92	0.36 to 0.90	0.47 to 0.85	0.31 to 0.80	0.42 to 0.85	0.41 to 0.90	0.23 to 0.79	0.15 to 0.75	0.51 to 0.92	0.20 to 0.52	0.26 to 0.72	0.11 to 0.76	0.27 to 0.88	0.21 to 0.48
ICC Sig	0.000	0.007	0.000	0.002	0.002	0.033	0.003	0.002	0.072	0.152	0.000	0.660	0.761	0.165	0.004	0.726
SEM	0.83	1.67	14.18	13.52	1.91	1.99	20.16	8.24	2.58	3.33	8.22	10.93	2.22	2.40	12.50	13.27
MDD	2.29	4.62	39.30	37.48	5.30	5.51	55.87	22.84	5.15	6.24	29.27	22.80	6.14	6.65	42.97	36.78

Hz: Hertz; NB: Narrowband; BB: Broadband; ms: milliseconds; µV: microvolts; ICC: intraclass correlation coefficient; Sig: significance; SEM: standard error of measurement; MDD: minimal detectable differences

The ICC for within-session measurements for cVEMPs evoked by 500 Hz NB CE-Chirp revealed excellent reliability for the P1 latency, good reliability for N1 latency, P1-N1 amplitude, and asymmetry ratio for both the right and left ear. Moderate reliability was observed for between-session measurements for P1 latency, N1 latency, asymmetry ratio and P1-N1 amplitude. Good reliability was observed for P1-N1 amplitude in the right ear (0.82), slightly higher than that observed in the left ear (0.73). Evidently, the within-session reliability was higher than the between session reliability with P1 latency providing the highest reliability. P1-N1 amplitude resulted in the highest between-session reliability for both the right and the left ear. Statistically significant ICC ($p < 0.05$) was evident for within and between-session measurements for amplitude, latency, and asymmetry ratio response parameters for cVEMP evoked by the 500 Hz NB CE-Chirp stimulus bilaterally.

The BB CE-Chirp ICC values for within-session measurements revealed moderate reliability in the right ear and poor reliability in the left ear for P1 latency and N1 latency. Good reliability was, however, observed for the within-session BB CE-Chirp P1-N1 amplitude, and moderate to good reliability was observed for the between session P1-N1 amplitude. Poor between-session reliability was obtained for P1 latency and N1 latency for both the right and left ear. In addition, poor within and between-session asymmetry ratio was recorded. Statistically significant ICC ($p < 0.05$) was evident for within and between-session P1-N1 amplitude bilaterally. However, a statistically significant ICC was not observed for between-session P1 latency and N1 latency, or within and between-session asymmetry ratio. SEM and MDD values were calculated and compared for within and between-session measurements with smaller SEM and MDD values suggesting higher reliability.

Table 4. Mean and standard deviation of right and left ear oVEMP N1 and P1 latencies, N1-P1 amplitude and asymmetry ratio for the 500 Hz NB CE-Chirp and BB CE-Chirp for each test session (*n*=20)

	Test 1				Test 2				Test 3			
	500 Hz NB CE-Chirp		BB CE-Chirp		500 Hz NB CE-Chirp		BB CE-Chirp		500 Hz NB CE-Chirp		BB CE-Chirp	
	Right ear	Left ear	Right ear	Left ear	Right ear	Left ear	Right ear	Left ear	Right ear	Left ear	Right ear	Left ear
N1 latency (ms)	7.67 ± 3.69	6.30 ± 4.09	10.75 ± 2.15	11.12 ± 1.88	7.83 ± 3.46	6.65 ± 4.37	10.33 ± 1.91	10.28 ± 2.37	8.73 ± 4.10	8.00 ± 4.19	11.30 ± 1.76	9.95 ± 2.05
P1 latency (ms)	11.60 ± 3.68	10.30 ± 3.87	14.18 ± 2.23	15.10 ± 2.06	11.37 ± 3.63	10.88 ± 4.37	13.80 ± 2.00	14.22 ± 2.55	12.95 ± 4.93	12.40 ± 4.38	15.37 ± 2.41	13.80 ± 2.20
N1-P1 amplitude (µV)	4.33 ± 1.94	5.11 ± 2.59	3.42 ± 1.44	3.66 ± 1.71	4.24 ± 1.98	4.95 ± 2.37	3.44 ± 1.73	2.81 ± 0.98	4.18 ± 1.83	4.50 ± 2.06	4.01 ± 1.70	4.26 ± 1.62
AR (%)	17.27 ± 12.34		16.90 ± 13.63		13.77 ± 15.49		10.19 ± 10.43		10.88 ± 7.79		11.16 ± 9.97	

Hz: Hertz; NB: Narrowband; BB: Broadband; ms: milliseconds; µV: microvolts; AR: asymmetry ratio

oVEMP

Data for 20 right ears and 20 left ears were analyzed. The 500 Hz NB CE-Chirp N1 and P1 latencies were shorter than the BB CE-Chirp N1 and P1 latencies across all test sessions for both the right and left ear. In addition, the N1-P1 amplitude of the 500 Hz NB CE-Chirp was larger than that of the BB CE-Chirp for all test sessions bilaterally. Table 4 shows the mean ± SD of the N1 latency, P1 latency, N1-P1 amplitude, and asymmetry ratios for the right and left ear for each stimulus at 95 dB nHL for test 1, test 2 and test 3.

The repeated measures ANOVA revealed no significant differences for the 500 Hz NB CE-Chirp for the right ear N1 latency (f=1.511; p=0.561), P1 latency (f=1.335; p=0.377) and N1-P1 amplitude (f=1.793; p=0.943) and the left ear N1 latency (f=1.364; p=0.237), P1 latency (f=1.421; p=0.174) and N1-P1 amplitude (f=1.779; p=0.522) between two test moments, and this for the within as well as the between-session measurements. Also, no significant difference was obtained for oVEMP evoked by the BB CE-Chirp for within and between-session measurements for the right ear N1 latency (f=1.978; p=0.154), P1 latency (f=1.513; p=0.065) and N1-P1 amplitude (f=1.725; p=0.461), and the left ear N1 latency (f=1.868; p=0.161), P1 latency (f=1.923; p=0.091) and N1-P1 amplitude (f=1.816; p=0.099). A significant difference was not observed for 500 Hz NB CE-Chirp (f=1.975; p=0.200) and BB CE-Chirp (f=0.501; p=0.115) asymmetry ratio for within and between session measurements.

The 500 Hz NB CE-Chirp provided higher reliability scores and p-values < 0.05 for all oVEMP test parameters, i.e., N1 latency, P1 latency, N1-P1 amplitude, and asymmetry ratio, for within-session reliability for the right and left ear. SEM values were comparable across all VEMP parameters for the 500 Hz NB CE-Chirp and the BB CE-Chirp. Slightly higher MDD values were observed for the within and between-session 500 Hz NB CE-Chirp N1 and P1 latency compared to that of the BB CE-Chirp. The ICC values with their 95% confidence intervals and significance levels, together with the SEM and MDD values for within and between-session reliability for oVEMPs evoked by 500 Hz NB CE-Chirp and BB CE-Chirp for the right and left ear are presented in Table 5 and Table 6, respectively.

The ICC for within-session measurements for oVEMPs evoked by 500 Hz NB CE-Chirp revealed moderate to good reliability for all VEMP parameters, i.e., N1 latency, P1 latency, N1-P1 amplitude, and asymmetry ratio bilaterally. Poor reliability was, however, observed for between-session measurements across all VEMP parameters. The within-session reliability was higher than the between-session reliability. Statistically significant ICC ($p < 0.05$) was evident for within-session measurements, and not for between-session measurements for amplitude, latency, and asymmetry ratio response parameters for oVEMP evoked by the 500 Hz NB CE-Chirp stimulus bilaterally.

The BB CE-Chirp ICC values for within-session measurements revealed moderate reliability for N1 and P1 latency in the right ear and poor reliability for N1 and P1 latency in the left ear. Poor reliability for N1-P1 amplitude, and moderate reliability for asymmetry ratio was obtained for within-in session measurements for oVEMP evoked by BB CE-Chirp. Furthermore, statistically significant ICC ($p < 0.05$) was evident for within-session measurements, except for N1 latency in the left ear. Statistically significant ICC ($p < 0.05$) was not evident and poor reliability was observed across all VEMP parameters for BB CE-Chirp for between-session measurements.

Table 5. Right ear: The intraclass correlation coefficient (ICC) and their 95% confidence intervals (CI), standard error of measurement (SEM) and minimal detectable difference (MDD) as within-session and between-session reliability parameters for the oVEMP P1 latency, N1 latency, N1-P1 Amplitude and asymmetry ratio (AR) (*n*=20)

	500 Hz NB CE-Chirp								BB CE-Chirp							
Reliability parameters	**Within-session reliability**				**Between-session reliability**				**Within-session reliability**				**Between-session reliability**			
	N1 Latency (ms)	P1 Latency (ms)	N1-P1 Amp (μV)	AR	N1 Latency (ms)	P1 Latency (ms)	N1-P1 Amp (μV)	AR	N1 Latency (ms)	P1 Latency (ms)	N1-P1 Amp (μV)	AR	N1 Latency (ms)	P1 Latency (ms)	N1-P1 Amp (μV)	AR
ICC	0.72	0.79	0.73	0.76	0.45	0.43	0.32	0.22	0.65	0.60	0.48	0.66	0.50	0.31	0.40	0.27
ICC 95% CI	0.27 to 0.89	0.47 to 0.92	0.31 to 0.90	0.37 to 0.90	0.28 to 0.44	0.32 to 0.76	0.18 to 0.74	0.10 to 0.70	0.13 to 0.87	0.40 to 0.91	0.37 to 0.80	0.16 to 0.87	0.22 to 0.80	0.05 to 0.71	0.12 to 0.66	0.11 to 0.71
ICC Sig	0.005	0.001	0.004	0.002	0.783	0.105	0.216	0.302	0.014	0.001	0.009	0.004	0.067	0.189	0.968	0.258
SEM	1.95	1.69	1.01	6.05	3.04	3.72	1.51	6.88	1.27	1.00	1.04	7.95	1.25	2.00	1.32	8.52
MDD	5.41	4.68	2.80	16.76	8.43	10.32	4.19	19.06	3.53	2.77	2.88	22.03	3.45	5.55	3.66	23.61

Hz: Hertz; NB: Narrowband; BB: Broadband; ms: milliseconds; μV: microvolts; ICC: intraclass correlation coefficient; Sig: significance; SEM: standard error of measurement; MDD: minimal detectable differences

Table 6. Left ear: The intraclass correlation coefficient (ICC) and their 95% confidence intervals (CI), standard error of measurement (SEM) and minimal detectable difference (MDD) as within-session and between-session reliability parameters for the oVEMP P1 latency, N1 latency, N1-P1 Amplitude and asymmetry ratio (AR) (*n*=20)

	500 Hz NB CE-Chirp								**BB CE-Chirp**							
Reliability parameters	**Within-session reliability**				**Between-session reliability**				**Within-session reliability**				**Between-session reliability**			
	N1 Latency (ms)	P1 Latency (ms)	N1-P1 Amp (μV)	AR	N1 Latency (ms)	P1 Latency (ms)	N1-P1 Amp (μV)	AR	N1 Latency (ms)	P1 Latency (ms)	N1-P1 Amp (μV)	AR	N1 Latency (ms)	P1 Latency (ms)	N1-P1 Amp (μV)	AR
ICC	0.87	0.81	0.76	0.76	0.50	0.41	0.37	0.22	0.44	0.51	0.41	0.66	0.25	0.43	0.24	0.27
ICC 95% CI	0.68 to 0.95	0.53 to 0.93	0.37 to 0.90	0.37 to 0.90	0.17 to 0.80	0.33 to 0.76	0.16 to 0.75	0.10 to 0.70	0.32 to 0.78	0.11 to 0.80	0.29 to 0.75	0.16 to 0.87	0.06 to 0.68	0.36 to 0.77	0.11 to 0.70	0.11 to 0.71
ICC Sig	0.000	0.000	0.002	0.002	0.005	0.105	0.167	0.302	0.096	0.004	0.009	0.004	0.238	0.106	0.272	0.258
SEM	1.48	1.69	1.27	6.05	2.96	3.36	1.51	6.88	1.41	1.44	1.31	7.95	1.78	1.66	1.41	8.52
MDD	4.09	4.67	3.52	16.76	8.21	9.32	4.19	19.06	3.90	4.00	3.63	22.03	4.93	4.60	3.92	23.61

Hz: Hertz; NB: Narrowband; BB: Broadband; ms: milliseconds; μV: microvolts; ICC: intraclass correlation coefficient; Sig: significance; SEM: standard error of measurement; MDD: minimal detectable differences

DISCUSSION

The aim of this study was to determine the test-retest reliability of c&oVEMP evoked by 500 Hz NB CE-Chirp and BB CE-Chirp. The literature suggests that the chirp is a promising stimulus to determine saccular and utricular function, however there is a lack of consensus on which type of chirp stimulus should be used. A recent study by Reddy et al. (2022) reported that cVEMPs evoked by the 500 Hz NB CE-Chirp provided the highest response rates, shorter P1 and N1 latencies, and overall, larger VEMP amplitudes when compared to BB CE-Chirp, 500 Hz TB, and click stimulus. However, there are no studies to date that investigate the reliability of the c&oVEMP parameters evoked by the 500 Hz NB CE-Chirp or the BB CE-Chirp. Reliability is defined as the extent to which test measurements can be replicated (Koo & Li, 2016), or the consistency of measurements in the absence of measurement error (Atkinson & Nevill, 1998). The main considerations of measurement error are systematic bias, possibly attributed to inadequate recovery between each test session to a general learning effect, or to a random error due to biological or equipment variability (Atkinson & Nevill, 1998).

The first measure of reliability in this study was an investigation of changes in mean values for c&oVEMP between test trials at an intensity level of 95 dB nHL. A significant difference was not observed for the c&oVEMP latency, amplitude, and asymmetry ratio parameters for both within and between-session measurements. This suggests that the effects of random or systematic error did not influence the c&oVEMP responses between test sessions.

cVEMP

The second measure of reliability in this study was the intraclass correlation coefficient (ICC) for cVEMPs. The results indicate a good to excellent within-session reliability, and moderate to good between-session reliability for all cVEMP response parameters evoked by a 500 Hz NB CE-Chirp stimulus. cVEMPs evoked by BB CE-Chirp, however, resulted in poor to good within-session reliability and, essentially, poor between-session reliability for latency and asymmetry ratio, and moderate to good reliability for the amplitude parameter.

Previous studies have focused on the test-retest reliability of c&oVEMPs evoked by click and TB stimuli using a variation of methodologies and subsequently varying results (Versino et al. 2001; Maes et al. 2009; Vanspauwen et al. 2009; Eleftheriadou et al. 2009; Nguyen et al. 2010; Venhovens et al. 2015). Some authors utilized unilateral SCMm activation (Maes et al. 2009; Vanspauwen et al. 2009; Eleftheriadou et al. 2009; Behtani et al. 2018), as was done in the current study, whereas others employed bilateral SCMm activation (Nguyen et al. 2010; Venhovens et al. 2015; Versino et al. 2001). Some studies were conducted with participants seated in the upright position (Versino et al. 2001; Maes et al. 2009; Vanspauwen et al. 2009; Venhovens et al. 2015), as in the present study, while other studies required that participants were supine during assessment (Isaradisaikul et al. 2008; Nguyen et al. 2010; Behtani et al. 2018). Lastly, some studies included the use of EMG monitoring (Behtani et al. 2018), while others relied on feedback mechanisms instead (Maes et al. 2009). Variations in each of these parameters are likely to have contributed to the discrepancies in the literature regarding the test-retest reliability of the P1 latency, N1 latency, P1-N1 amplitude, and asymmetry ratio. There is, however, consensus that the cVEMP P1-N1 amplitude delivers the highest reliability across several studies, as was the case in the present study for between-session measurements. This led to Isaradisaikul et al. (2008) suggesting that the cVEMP P1-N1 amplitude may be a good parameter with which to monitor pathologic findings over time.

The current study measured good between-session test-retest reliability for cVEMP P1-N1 amplitude evoked by the 500 Hz NB CE-Chirp, and moderate to good between-session test-retest reliability for BB CE-Chirp. Several studies have reported moderate to excellent test-retest reliability of the P1-N1 amplitude utilizing a click or 500 Hz TB stimulus (Versino et al. 2001; Isaradisaikul et al. 2008; Maes et al. 2009; Nguyen at al. 2010; Behtani et al. 2018). Historically, a large intrasubject variation in VEMP amplitude has been observed due to varying SCMm contraction levels, the type and frequency of the stimulus, and the mode of stimulation used (Isaradisaikul et al. 2008; Anupriya & Kumar 2019). Nevertheless, the P1-N1 amplitude is still mostly relied upon in the interpretation of VEMP responses. The results of the current study suggest that both the 500 Hz NB CE-Chirp and BB CE-Chirp amplitude parameter is as reliable as previously reported literature on the 500 Hz TB stimulus, with the 500 Hz NB CE-Chirp proving slightly more reliable than the BB CE-Chirp. The present study also reported better between-session reliability for cVEMP asymmetry ratio (0.71) with the 500 Hz NB CE-Chirp compared to that reported by previous studies, and demonstrated poor between-session reliability for BB CE-Chirp (0.33). For clinical use of cVEMPs, therefore, the most reliable stimulus would be the 500 Hz NB CE-Chirp, with response amplitudes and asymmetry ratio being the most reliable parameters.

The cVEMP latency is considered a clinically relevant diagnostic parameter in the interpretation of the VEMP response. However, there are conflicting reports throughout the literature on the test-retest reliability of cVEMP latencies evoked by 500 Hz TB and click stimuli, making it a poor choice for the purpose of within-subject monitoring over time (Venhovens et al. 2015), despite the diagnostic value of prolonged VEMP latencies (Murofushi, 2001). It has been suggested that stimulus duration, type of stimulus used, i.e., air conduction, bone conduction, forehead taps or vibration, seated or supine testing position and the placement of electrodes between sessions, may account for and contribute to, differences in reliabilities for latencies (Kumar et al. 2018). The current study reported between-session ICC values with moderate reliability for P1 and N1 latency evoked by a 500 Hz NB CE-Chirp stimulus bilaterally. The N1 latency is regarded as the most robust and reliable cVEMP latency parameter (Isaradisaikul et al. 2008; Nguyen at al. 2010; Behtani et al. 2018), but diagnostically, the P1 latency is more sensitive to lesions in the vestibulospinal tract and is the better parameter to evaluate the latency of a cVEMP response (Murofushi et al. 2001; Maes 2009). In the current study it was observed that the P1 latency evoked by the 500 Hz NB CE-Chirp had excellent within-session reliability (0.93), and both the P1 and N1 latency were more reliable than that of the BB CE-Chirp. In fact, the BB CE-Chirp in the current study showed poor reliability for both P1 latency (0.33) and N1 latency (0.28). The 500 Hz NB CE-Chirp produced test-retest reliability scores that were comparable to the 500 Hz TB stimulus for N1 latency, regardless of unilateral or bilateral SCMm activation, or whether the participants were seated or in the supine position. However, the P1 latency evoked by the 500 Hz NB CE-Chirp proved to be more reliable than that of the BB CE-Chirp in the current study, and the previously reported reliability of the 500 Hz TB P1 latency. In addition, this was achieved at a lower intensity level of 95 dB nHL than the higher intensity levels used in previous studies, i.e., 105 – 110 dB nHL (Eleftheriadou et al. 2009; Nguyen et al. 2010). This suggests that, with regard to the evaluation of cVEMP latency, the 500 Hz NB CE-Chirp is more reliable than the BB CE-Chirp stimulus.

The second measure of reliability in this study was the Standard Error of Measurement (SEM) to calculate the spread of errors as they relate to the true cVEMP response compared to the observed cVEMP response parameters. The SEM increases as the standard deviation increases, and decreases as the test reliability increases (Musselwhite & Wesolowski, 2018). Therefore, a good test reliability together with a small SEM will result in an observed response being similar to a true response. In the current study, P1 latency and N1 latency SEM values for within and between-session cVEMP evoked by the 500 Hz NB CE-Chirp were smaller than BB CE-Chirp SEM values for the right and left ears, suggesting a smaller range for the 500 Hz NB CE-Chirp between which the true latencies are expected. A greater variation in the standard deviation of the cVEMP P1-N1 amplitude, compared to the latency parameter, was observed for both 500 Hz NB CE-Chirp and BB CE-Chirp for the right and left ears. Smaller SEM values were observed for the P1-N1 amplitude evoked by the BB CE-Chirp for within and between-session reliability for both the right and left ears. This suggests that although the 500 Hz NB CE-Chirp resulted in larger P1-N1 amplitudes for within and between-sessions, the range in which the true P1-N1 amplitude can be expected was greater with the 500 Hz NB CE-Chirp compared to the BB CE-Chirp.

The Minimal Detectable Difference (MDD) was the final method used to evaluate the reliability of the cVEMP. The MDD determines the amount of change that must be achieved for each variable to reflect an actual test-retest difference (Mair, 2020), with a deviation greater than the MDD value for any test parameter indicating an abnormal response. The results of the current study demonstrate stricter MDD values for the 500 Hz NB CE-Chirp compared to BB CE-Chirp for both within and between-session variables. This suggests that the 500 Hz NB CE-Chirp is the more reliable stimulus of the two with which to interpret cVEMP responses in a clinical setting.

oVEMP

The ICC reliability measure for oVEMPs evoked by the 500 Hz NB CE-Chirp in the current study revealed good reliability for all VEMP parameters for within-session measurements, compared to the poor to moderate reliability obtained for BB CE-Chirp. With regards to oVEMP, both Nguyen et al. (2010) and Venhovens et al. (2015) found that the N1 and P1 latencies result in poor reliability with good to excellent reliability for N1-P1 amplitude. It has been postulated that the reliability of the oVEMP amplitude is excellent as it is an excitatory potential recorded in the presence of limited background noise from the extra-ocular muscles (Anupriya & Kumar, 2019). This contrasts with the present study which found poor between-session test-retest reliability for oVEMP latency, amplitude, and asymmetry ratio parameters for both 500 Hz NB CE-Chirp and BB CE-chirp. It is well known that the mode of stimulation and differences in test protocols could affect the reliability of oVEMP responses (Isaradisaikul et al. 2008). In the current study the active electrode was placed on the lateral canthus of the eye, which has shown to provide variable oVEMP amplitudes and response rates (Sandhu et al. 2013; Govender et al. 2016). Furthermore, in the current study non-disposable silver disc electrodes were used, which are smaller than the traditional disposable silver disc electrodes utilized for VEMP testing. The smaller electrode surface area may have resulted in greater variation in the placement of the electrodes between sessions, which may have contributed to the poor test-retest reliability. For this reason, it is imperative that oVEMPs evoked by the CE-Chirp stimulus are investigated under different electrode montage configurations to confirm whether the poor test-retest reliability persists across each method.

With regards to oVEMP asymmetry ratio, good within-session (0.76) and poor between-session (0.22) reliability was observed for 500 Hz NB CE-Chirp, and poor within-session (0.40) and between session (0.27) reliability for BB CE-Chirp was found. Previous studies have reported moderate to good reliability for oVEMP asymmetry ratio (Nguyen et al. 2010; Piker et al. 2011; Kumar et al. 2018). These studies utilized the standard electrode montage, which involves placing the active electrode below the midpoint of the eye and a reference electrode approximately 1-2 cm below the active electrode. In the current study, the active electrode was placed on the lateral canthus of the eye which has been reported to produce smaller oVEMP amplitudes (Sandhu et al. 2013; Govender et al. 2016). This may have negatively affected the reliability obtained for the asymmetry ratio in the current study. Historically, asymmetry ratio is not relied upon in isolation as it cannot be utilized when only a unilateral response can be obtained.

Conclusion

The available literature on chirp evoked c&oVEMPs currently lacks consensus on which type of chirp stimulus should be used. Therefore, this study aimed to investigate the test-retest reliability of the latency, amplitude, and asymmetry ratio parameters for c&oVEMP evoked by 500 Hz NB CE-Chirp and BB CE-Chirp. The 500 Hz NB CE-Chirp proved to be more reliable than the BB CE-Chirp across all VEMP parameters, with the cVEMP amplitude being the most reliable parameter. Therefore, the NB CE-Chirp is a reliable tool to estimate saccular function in clinical practice. Further evaluation of oVEMP reliability using a variety of electrode montage configurations is recommended.

Author Contribution

All authors contributed equally to this manuscript. All authors discussed the results and implications and commented on the manuscript at all stages. T.R. wrote the main paper. B.H. provided the concept and contributed to the integration of the results. L.B. and L.M. provided interpretive analysis.

There are no competing interests or funding sources to disclose.

References

Akin, F.W., Murnane, O.D., Proffitt, T.M. (2003). The effects of click and tone-burst stimulus parameters on the vestibular evoked myogenic potential (VEMP). *AM. J. Audiol*, 14(9), 500-509

Anupriya, E., Kumar, K. (2019). Test-retest reliability of cervical and ocular vestibular evoked myogenic potential with simultaneous and sequential recording. *AM. J. Audiol*, 28, 414-421

Atkinson, G., Nevill, A.M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med, 26(4), 217-238

Behtani, L., Maheu, M., Delcenserie, A., Nooristani, M. (2018). State-of-the-art assessment allows for improved vestibular evoked myogenic potential test-retest reliability. *Audiol. Res*, 8, 34-36

Bèkèsy, G. (1952). Direct observations of the vibrations of the cochlear partition under a microscope. *Acta Oto-Laryngol* 42:3, 197-201, DOI: 10.3109/00016485209120346

Bujang, M.A., Baharum, N. (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. *Arch Orofac*, 12(1), 1-11

Cheng, P.W., Huang, T.W., Young, Y. (2003). The influence of clicks versus short tone bursts on the vestibular evoked myogenic potentials. *Ear Hear*, 24, 195-197

Cheng, P.W., Chen, C.C., Wang, S.J., Young, Y.H. (2009). Acoustic, mechanical and galvanic stimulation modes elicit ocular vestibular-evoked myogenic potentials. *Clin Neurophysiol*, 120(2009), p.1841-1844.

Cohen, B.H., Lea, B. (2004). *Essentials of statistics for the social and behavioural sciences*. Canada: Wiley & Sons.

Colebatch, J.G., Halmagyi, G.M., Skuse, N.F. (1994). Myogenic potentials generated by click-evoked vestibulocollic reflex. *J. Neurol. Neurosurg. Psychiatry*, 57(2), 190-197

Coren, S., Hakstian, A.R. (1990). Methodological implications of interaural correlation: Count heads not ears. *Percept. Physcophys*. 48(3), p.291-294.

Desmond, A.L. (2011). *Vestibular function: Clinical and practice management.* New York: Thieme

de Oliveira, A.C., de Lemos Menezes, P., Pereira, L.D. (2014). Reproducibility (test-retest) of vestibular evoked myogenic potential. *Braz J Otorhinolaryngol*, 81, 264-269

Dontje, M.L., Dall, P.M., Skelton, D.A., Gill, J.M.R., & Chastin, S.F.M. (2018). Reliability, minimal detectable change and responsiveness to change: Indicators to select the best method to measure sedentary behaviours in older adults in different study designs. PloS One, 13(4), doi.org/10.1371/journal.pone.0195424

Eleftheriadou, A., Deftereos, S., Zarikas, V., Panagopoulos, G., Sfetsos, S., Karageorgiou, K., Ferekidou, E., Korres, S., Kandiloros, D. (2009). Test-retest reliability of p13n23 and n34p44 components of vestibular evoked myogenic potentials in a large healthy population. *J Otolaryngol-Head N*, 38(4), 462-467

Gelfand, S. A. (2001). *Essentials of audiology.* New York: Thieme

Govender, S., Cheng, P.Y., Dennis, D.L., Colebatch, J.G. (2016). Electrode montage and gaze effects on ocular vestibular evoked myogenic potentials (oVEMP). *J. Clin. Neurophysiol.*, 127, 2846-2854

Interacoustics A/S. (2020). What is wave reproducibility? Retrieved, July 31, 2020, from https://www.interacoustics.com/academy/faq/abr/what-is-wave-reproducibility

Isaradisaikul, S., Navacharoen, N., Hanprasertpong, C., Kangsanarak, J. (2012). Cervical vestibular-evoked myogenic potentials: Norms and protocols. *Int. J. Otolaryngol*. 2012;2012. doi: 10.1155/2012/913515

Isaradisaikul, S., Strong, D.A., Moushey, J.M., Gabbard, S.A., Ackley,S.R., Jenkins, H.A. (2008). Reliability of vestibular evoked myogenic potentials in healthy subjects. *Otol. Neurotol.*, 29, 542-544

Kumar, K., Bhat, J.S., Varghese, A.L. (2018). Test-retest reliability of vestibular evoked myogenic potential across different age groups. *Int. Tinnitus J.*, 22(2), 103-108. DOI: 10.5935/0946-5448.20180018

Koo, T.K. & Li, M.Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.*, 15, 155-163

Maes, L., Vink, B.M., De Vel, E., D'haenens, W., Bockstael, A., Keppler, H., Philips, B., Swinnen, F., Dhooge, I. (2009). The vestibular evoked myogenic potential: A test-retest reliability study. *J Clin Neurophysiol*, 120 (2009), 594-600

Mair, M.M., Kattwinkel, M., Jakoby, O., Hartig, F. (2020). The minimum detectable difference (MDD) concept for establishing trust in nonsignificant results: A critical review. *Environ. Toxicol. Chem.*, 39(11), 2109-2123. doi.org/10.1002/etc.4847

Meyer, N., Vinck, B., Heinze, B. (2015). cVemps: A systematic review and meta-analysis. *Int. J. Audiol*, 54, 143-151

Moinudeen, K., Varshini, A., Wesley, J., 2020. Comparison of 500Hz toneburst and 500Hz octave chirps for cervical vestibular evoked potentials. *Int. J. Sci. Res.*, 10(3), p.332-335.

Murofushi T., Matsuzaki M., Wu C.H. (1999). Short tone-burst-evoked myogenic potentials on the sternocleidomastoid muscle. *Arch. Otolaryngol. Head Neck Surg.*, 125, 660 – 664

Murofushi, T., Shimizu, K., Takegoshi, H., Cheng, W. (2001). Diagnostic value of prolonged latencies in the vestibular evoked myogenic potential. *Arch. Otolaryngol. Head Neck Surg.*, 127, 1609-1072

Musselwhite, D.J., Wesolowski, B.C. (2018). Standard error of measurement. The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation. Thousand Oaks: SAGE. DOI: http://dx.doi.org/10.4135/9781506326139.n658

Nguyen, K.D., Welgampola, M.S., Carey, J.P. (2010). Test-retest reliability and age-related characteristics of the ocular and cervical vestibular evoked myogenic potential test. *Otol. Neurotol.*, 31, 793-802

Reddy, T.M., Heinze, B., Biagio-de Jager, L., Maes, L. (2022). Cervical and ocular vestibular evoked myogenic potential: A comparison of narrowband chirp, broadband chirp, tone burst and click stimulation. *Int. J Audiol*. Advance online publication. https://doi.org/10.1080/14992027.2022.2064924

Rosengren, S.M., Welgampola, M.S., Colebatch, J.G. (2010). Vestibular evoked myogenic potentials: Past, present and future. *J. Clin. Neurophysiol.*, 121, 636-651

Rosengren, S.M., Colebatch, J.G., Young, A.S., Govender, S., Welgampola, M.S. (2019). Vestibular evoked myogenic potentials in practice: Methods, pitfalls and clinical applications. *Clin. Neurophysiol. Pract.*, 4(2019), 47-68

Sandhu, J.S., George, S.R., Rea, P.A. (2013). The effect of electrode positioning on the ocular evoked myogenic potential to air-conducted sound. *J. Clin. Neurophysiol.*, 124(6), 1232-1236

Singh, N.K., Kumar, P., Aparna, T.H., Barman, A. (2014). Rise/fall and plateau time optimization for cervical-evoked myogenic potential elicited by short tone bursts of 500 Hz. *Int. J. Audiol.*, 53, 490-496

Todd, N.P., Rosengren, S.M., Colebatch, J.G. (2007). Ocular vestibular evoked myogenic potentials (oVEMPs) produced by air and bone conducted sound. *J. Clin. Neurophysiol.*, 118(2), 381-390

Venhovens, J., Meulstee, J., Verhagen, W.I.M. (2015). Ocular and cervical vestibular evoked myogenic potentials (VEMPs) in healthy volunteers: the intra-, interobserver, and the test re-test reliability. *J Vestib Res*, 25, 161-167

Versino, M., Colnaghi, S., Callieco, R., Cosi, V. (2001). Vestibular evoked myogenic potentials: Test-retest reliability. *Funct. Neurol.*, 16(4), 299

Wang, S.J., Weng, W.J., Jaw, F.S., Young, Y.H., 2010. Ocular and cervical vestibular-evoked myogenic potentials: A study to determine whether air- or bone-conducted stimuli are optimal. *Ear Hear*, p.283-288

Weir, J.P. (2005). Quantyifing test-retest reliability using the intraclass correlation coefficient and the SEM. *J. Strength Cond. Res.*, 19(1), 231-240