Otology & Neurotology
xx:xx–xx © 2016, Otology & Neurotology, Inc.

Reliability and Normative Data for the Dynamic Visual Acuity Test for Vestibular Screening

*†Kristal M. Riska and *‡Courtney D. Hall

**Auditory and Vestibular Research Enhancement Award Program, James H. Quillen VA Medical Center, Mountain Home; †Department of Audiology and Speech-Language Pathology; and ‡Department of Physical Therapy, East Tennessee State University, Johnson City, Tennessee*

Hypothesis: The purpose of this study was to determine reliability of computerized dynamic visual acuity (DVA) testing and to determine reference values for younger and older adults.
Background: A primary function of the vestibular system is to maintain gaze stability during head motion. The DVA test quantifies gaze stabilization with the head moving versus stationary. Commercially available computerized systems allow clinicians to incorporate DVA into their assessment; however, information regarding reliability and normative values of these systems is sparse.
Methods: Forty-six healthy adults, grouped by age, with normal vestibular function were recruited. Each participant completed computerized DVA testing including static visual acuity, minimum perception time, and DVA using the NeuroCom inVision System. Testing was performed by two examiners in the same session and then repeated at a follow-up session 3 to 14 days later. Intraclass correlation coefficients (ICCs) were used to determine inter-rater and test–retest reliability.
Results: ICCs for inter-rater reliability ranged from 0.323 to 0.937 and from 0.434 to 0.909 for horizontal and vertical head movements, respectively. ICCs for test–retest reliability ranged from 0.154 to 0.856 and from 0.377 to 0.9062 for horizontal and vertical head movements, respectively. Overall, raw scores (left/right DVA and up/down DVA) were more reliable than DVA loss scores.
Conclusion: Reliability of a commercially available DVA system has poor-to-fair reliability for DVA loss scores. The use of a convergence paradigm and not incorporating the forced choice paradigm may contribute to poor reliability.
Key Words: Dynamic visual acuity—Reliability—Vestibular.

Otol Neurotol **37:**xxx–xxx, 2016.

Dizziness is among the most prevalent complaints for which people seek medical help and the incidence of dizziness increases with age (1,2). Although dizziness has many etiologies, it is often related to vestibular pathology which is treated effectively with vestibular rehabilitation (3). Uncompensated vestibular loss results in subjective complaints of imbalance, oscillopsia, and/or vertigo and postural and gaze instability. These impairments can result in decreased activity and avoidance of driving with resultant diminished independence and social isolation (4,5). Appropriate clinical management of patients with vestibular deficits necessitates that valid and reliable tools are available to clinicians.

Gaze stability during head movements is the primary function of the vestibulo-ocular reflex (VOR). When functioning normally, the VOR generates eye movements equal and opposite to head rotation, which enables images to remain stable on the fovea during head motion. The dynamic visual acuity test (DVA) is a functional measure of the VOR and quantifies the difference in visual acuity with the head still and then moving. Computerized DVA for horizontal head movement was introduced as a research measurement tool and demonstrated very good reliability (ICC = 0.83–0.87) and excellent sensitivity (94%) and specificity (95%) for identifying vestibular disorders (6). Computerized DVA for vertical head movements is also very reliable (ICC = 0.89 – 0.94), has excellent specificity (90%), but has modest sensitivity for identifying vestibular disorders (7). Recently, computerized DVA systems have become commercially available from Neurocom, Micromedical Technologies, and as part of the NIH toolbox making it feasible for clinicians to incorporate DVA into their vestibular assessment. However, information regarding reliability and normative values of these commercial

Address correspondence and reprint requests to Courtney D. Hall, Ph.D., Audiology (126), James H. Quillen VAMC, Mountain Home, TN 37684, U.S.A.; E-mail: hallcd1@etsu.edu
The views expressed here are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the U.S. government.The authors have no financial or personal relationships with other people or organizations that could inappropriately influence or bias their work. Funding for this study was provided by the Auditory and Vestibular Research Enhancement Award Program. C.D.H. was supported by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Rehabilitation Research and Development Service, grant no. E7613R.The authors disclose no conflicts of interest.
Supplemental digital content is available in the text.

Copyright © 2016 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.

systems is sparse (8–11). Two of these previous studies reported on a previous version of NeuroCom inVision System that used a series of mirrors to reflect the optotype onto the screen (9,11). Thus, the purpose of this study was twofold: 1) to determine test–retest and inter-rater reliability of the NeuroCom computerized DVA test; and 2) to determine reference values for healthy younger and older adults with the system.

METHODS

Participants. Forty-six participants were enrolled from the community and local university. Written informed consent was obtained from each participant before testing. The protocol was approved by VA/East Tennessee State University IRB Committee.

Inclusion criteria for the study were age and no history of vertigo or dizziness. Participants were grouped according to age: 1) younger healthy adults (YA) between the ages of 18 and 30 years and 2) older healthy adults (OA) at least 60 years of age. Participants were excluded if they had a history of neurological disease, head injury with loss of consciousness, or perforated tympanic membrane. Older participants completed the functional comorbidity index question to quantify comorbidities (diseases such as arthritis and visual impairment) that impact physical function (12). Each participant underwent an ocular motor screening, including search for spontaneous nystagmus, saccadic, and smooth pursuit eye movements using infrared goggles (ICS Chartr 200 system, Schaumburg, IL, U.S.A.) to rule out gross central nervous system abnormalities, and caloric testing to confirm normal horizontal semicircular canal function. Either monothermal warm calorics (44°C) or bithermal calorics (30 and 44°C) were performed (ICS Chartr Water Caloric Stimulator NCI-480). When warm caloric irrigations produced a monothermal asymmetry less than 10%, no additional irrigations were performed. When the monothermal asymmetry was 10% or greater, all four bithermal irrigations were performed and unilateral weakness was calculated using Jonkgees formula. Subjects with abnormal responses (unilateral weakness >25%) were excluded.

Equipment. The Neurocom SMART EquiTest with the InVision software package (version 8.4.0) (NeuroCom, a Division of Natus, Clackamas, OR, U.S.A.) was used for testing minimum perception time (MPT), static visual acuity (SVA), and DVA. A head-mounted sensor (InertiaCube2+, a three-axis integrating gyro) was used to measure the speed and range of head movements and was secured using an adjustable strap. The InVision system displays random sequences of orientation (up, down, left, or right) of the optotype ''E'' on a computer monitor. The size of the optotype was reported in units of logarithm of the minimum angle of resolution (logMAR) for all measures. Testing was performed in a well-lit and quiet room. Participants were seated in a chair and positioned 2 m from the monitor.

Using the manufacturer's convergence algorithm, visual acuity was determined for each measure in the testing protocol. This algorithm begins with an ''easier'' target presentation and then adjusts according to the participant's response. If the participant responds correctly, the next presentation is made more difficult (e.g., smaller optotype). If the participant responds incorrectly, the next target presentation is made easier (e.g., larger optotype). Testing continues until three out of five correct answers at the same level of difficulty are in agreement.

Testing Protocol. Participants underwent testing by two examiners on two separate days to evaluate inter-rater and test–retest reliability of each of the measures (described later). Participants who wore single-distance corrective lenses or contacts (n = 6) were tested with the corrective lenses. Participant who wore multifocal lenses (bifocals, trifocals, or progressive lenses; n = 14) were tested without corrective lenses. For all testing, participants were asked to identify the direction (up, down, left, or right) that the optotype ''E'' faced. Participants were instructed to say ''pass,'' when they were unsure of the direction of the optotype to discourage guessing. A ''pass'' was considered an incorrect response by the software and entered using the spacebar to advance the test. For each of the test measures, practice with the task was performed to ensure that the individual understood the task and instructions. For each individual, SVA, MPT, and DVA was completed in this order of testing, by one examiner and then repeated by a second examiner 5 to 10 minutes later. Examiner order was determined by a random number generator. Each participant returned 3 to 14 days after the initial test session and repeated the protocol with the same testers in the opposite order.

Static Visual Acuity (SVA). First, SVA testing was completed to determine the smallest optotype for which a participant could correctly identify the orientation whereas the head remained stationary. An optotype was presented on the screen for 2 seconds. The size of the optotype varied based on the response of the participant. If the orientation was correctly identified, the size was reduced. Static visual acuity was defined as the smallest optotype in which the orientation could be determined in 60% of the trials (i.e., three out of five) and was based on the manufacturer's convergence algorithm.

Minimum Perception Time (MPT). Next MPT testing was completed according to manufacturer protocol and is the minimum duration for the optotype to appear on the screen and be correctly identified. Specifically, the optotype is set to 0.2 logMAR above the SVA and the presentation time varies until the participant is unable to accurately identify the optotype direction. The shortest presentation duration is 20 milliseconds. On the basis of manufacturer recommendations, MPTs equal to or greater than 70 milliseconds were considered prolonged and could result in inaccurate DVA scores. In patients in whom MPTs were 70 milliseconds or longer, the task was repeated. If performance did not improve to 60 milliseconds or less, the participant was disqualified.

Dynamic Visual Acuity (DVA). Horizontal and Vertical DVA were tested separately for each participant. The direction of head rotation for each examiner and session were randomized. For horizontal testing, subjects rotated their head left and right (approximately 20 degrees in each direction) with a head velocity of at least 120 degrees/s for the optotype to appear. For vertical DVA, subjects rotated their head up and down (approximately 20 degrees in each direction) with a head velocity at least 100 degrees/s for the optotype to appear. The displacement and velocity of head rotations were controlled using a feedback bar on the screen that provides visual feedback to the participant. The following variables were recorded in LogMAR: DVA left, DVA right, DVA up, DVA down. DVA loss scores were calculated as the difference between the DVA score and SVA score and reported for each direction.

Data Analysis. Descriptive statistics were performed for the entire sample and by age group. To determine inter-rater and test–retest reliability intraclass correlation coefficients (ICC) were calculated and the 95% confidence intervals reported. For inter-rater reliability, consistency, rather than exact agreement was tested with a two-way random effects model (model

Copyright © 2016 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.

TABLE 1. *Participant characteristics*

Measure	Young Adults (n = 23)	Older Adults (n = 23)
Age (yr)		
Mean ± SD	22.4 ± 1.9	72.1 ± 6.4
Range	19–26	63–86
Sex (n)		
Male/female	7/16	14/9
Assistive device (n)		
None	23	22
Cane	0	1
Walker	0	0
Functional Comorbidity Index		
Mean ± SD	ND	2.2 ± 1.8
Range		0–7
Falls in past year (n)		
Mean ± SD	ND	0.3 ± 0.9
Range		0–4
No/yes		20/3
Eye glasses (n)		
None	19	2
Distance	4	2
Reading	0	5
Multifocal	0	14

ND indicates no data; younger adults did not complete the Functional Comorbidity Index or questions regarding falls in the past year.

2,1) (13). For inter-rater reliability ICCs were calculated for Session 1 and Session 2 separately (13). For test–retest reliability, consistency, rather than exact agreement, was tested using a two-way mixed model (model 3,1). To determine whether test–retest and inter-rater reliability differed by participant age, separate ICCs were performed for the younger and older groups. ICC values were interpreted as follows: excellent (0.75–1.0), fair to good (0.40–0.74), and poor (<0.39) (14). To determine age differences, independent samples *t* tests were performed for age groups (YA versus OA; 60-year olds versus 70-year olds) on DVA scores and head velocity collapsed across tester and time. Level of significance was set at alpha <0.05.

RESULTS

Participants

A description of the general characteristics of the YA (n = 23) and OA (n = 23) who completed at least one session can be observed in Table 1. Overall, the OA were healthy with a mean functional comorbidity index of 2.1. One OA did not complete all trials in the first session and two YA and one OA did not return for the second testing session.

Inter-rater Reliability

Inter-rater reliability for SVA, horizontal, and vertical DVA was assessed using ICCs (Table 2). Inter-rater reliability for SVA was excellent (ICCs = 0.957 and 0.981 for Sessions 1 and 2, respectively). Inter-rater reliability for horizontal head movements (left DVA, right DVA, left DVA loss, right DVA loss) ranged from 0.323 to 0.937 and from 0.434 to 0.909 for vertical head movements (up DVA, down DVA, up DVA loss, down DVA loss). Overall, raw scores (left/right DVA and up/down DVA) were more reliable than DVA loss scores.

Reliability was also examined according to age groups. In the YA, the ICC for SVA was 0.573 suggesting fair-to-good reliability for Session 1 and 0.902 for Session 2 suggesting excellent reliability (Table 3). In the YA, ICCs for horizontal DVA (left/right DVA, left/right DVA loss) ranged from 0.112 to 0.815. In the YA, ICCs for vertical head movements (up/down DVA, up/down DVA loss) ranged from 0.146 to 0.772. Overall, ICCs for DVA raw scores were fair to excellent and DVA loss scores were poor to fair. In the OA, SVA reliability was excellent for both sessions (0.968 and 0.967). In the OA, ICCs for horizontal ranged from 0.265 to 0.933 and for vertical DVA from 0.413 to 0.913. Similar to findings in the YA, ICCs for DVA loss scores were poor to fair.

Test–Retest Reliability

Test–retest reliability for SVA was excellent (ICCs = 0.934 and 0.955 for Testers 1 and 2, respectively). Reliability ranged from 0.154 to 0.856 for horizontal head movements and from 0.377 to 0.906 for vertical head movements (Table 4). Raw scores were more reliable than DVA loss scores.

Test–retest reliability was evaluated by age group. In YA, test–retest reliability for SVA was 0.434 and 0.499 for Testers 1 and 2, respectively. In YA, ICCs were −0.025–0.417 for horizontal DVA suggesting poor-

TABLE 2. *Inter-rater reliability for DVA variables for all participants (n = 45 unless otherwise noted)*

	Session 1		Session 2	
Variable	ICC	95% CI	ICC	95% CI
SVA	0.957	0.924–0.976	0.981 (n = 43)	0.966–0.990
Left DVA	0.919	0.857–0.955	0.882 (n = 43)	0.793–0.935
Right DVA	0.937	0.888–0.965	0.884 (n = 42)	0.795–0.936
Up DVA	0.882 (n = 43)	0.793–0.934	0.909 (n = 43)	0.838–0.950
Down DVA	0.886 (n = 44)	0.800–0.936	0.895 (n = 43)	0.814–0.942
Left DVA loss	0.323	0.036–0.561	0.468 (n = 43)	0.199–0.672
Right DVA loss	0.548	0.30–0.723	0.404 (n = 42)	0.118–0.629
Up DVA loss	0.447 (n = 43)	0.173–0.657	0.458 (n = 43)	0.186–0.665
Down DVA loss	0.466 (n = 44)	0.199–0.668	0.434 (n = 43)	0.157–0.647

CI indicates confidence interval; DVA, dynamic visual acuity; ICC, intraclass correlation coefficients; SVA, static visual acuity.

Copyright © 2016 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.

TABLE 3. *Inter-rater reliability for DVA variables at Session 1 and Session 2 by age groups*

	Younger Adults				Older Adults			
	Session 1 (n = 23)		Session 2 (n = 21)		Session 1 (n = 22)		Session 2 (n = 22)	
Variable	ICC	95% CI	ICC	95% CI	ICC	95% CI	ICC	95% CI
SVA	0.573	0.219–0.794	0.902	0.775–0.959	0.968	0.925–0.987	0.967	0.923–0.986
Left DVA	0.815	0.613–0.917	0.660	0.330–0.847	0.811	0.598–0.917	0.762	0.508–0.894
Right DVA	0.477	0.090–0.739	0.628	0.280–0.831	0.933	0.845–0.971	0.776 (n = 21)	0.526–0.903
Up DVA	0.447	0.052–0.721	0.772	0.518–0.901	0.912 (n = 20)	0.793–0.964	0.823	0.622–0.923
Down DVA	0.466	0.076–0.732	0.442	0.023–0.728	0.852 (n = 21)	0.670–0.937	0.913	0.802–0.963
Left DVA loss	0.401	−0.004–0.693	0.223	−0.220–0.590	0.265	−0.166–0.611	0.544	0.169–0.782
Right DVA loss	0.112	−0.306–0.494	0.130	−0.309–0.524	0.776	0.534–0.900	0.516 (n = 21)	0.119–0.770
Up DVA loss	0.212	−0.211–0.567	0.564	0.185–0.797	0.749 (n = 20)	0.467–0.892	0.413	0.000–0.706
Down DVA loss	0.201	−0.222–0.560	0.146	−0.295–0.535	0.642 (n = 21)	0.301–0.837	0.632	0.296–0.829

CI indicates confidence interval; DVA, dynamic visual acuity; ICC, intraclass correlation coefficients; SVA, static visual acuity.

to-fair reliability (Table 5). In YA, ICCs were 0.175 to 0.690 for vertical DVA suggesting poor-to-good reliability. In OA, ICCs were 0.157 to 0.856 for horizontal DVA suggesting poor-to-excellent reliability. In OA, ICCs were 0.485 to 0.895 for vertical DVA suggesting good-to-excellent reliability. ICCs for DVA loss scores were less reliable than the raw values for both age groups.

Normative Values and Age Differences

There were significant differences between YA and OA for SVA as well as raw horizontal and vertical DVA scores ($p < 0.001$). There were no significant age group differences ($p > 0.05$) between YA and OA for horizontal or vertical DVA loss scores (Table 6). Additionally, there were no significant age group differences ($p > 0.05$) between 60- and 70-year olds for SVA or horizontal or vertical DVA (see table, supplemental digital content 1, http://links.lww.com/MAO/A375).

To understand potential factors contributing to DVA performance, age group differences in head velocity were examined (see table, supplemental digital content 2, http://links.lww.com/MAO/A376). There were no significant age group differences ($p > 0.05$) for average head velocity achieved for horizontal or vertical head movements. Head velocity for error trials in which the optotype was incorrectly identified was significantly different between older and younger adults for Right DVA and Down DVA ($p = 0.023$ and 0.009, respectively). Furthermore, the percentage of error trials in which head velocity exceeded 180 degrees/s was low (1.9% for YA and 3.5% for OA).

On the basis of an estimate of measurement error for DVA (www.rehabmeasures.org accessed 9/17/15) that included ICCs for test–retest reliability and 95% confidence intervals, the minimal detectable change (MDC_{95}) would vary from 0.247 to 0.326 LogMARs for YA and from 0.251 to 0.383 LogMARs for OA (the equivalent of two to four lines difference on the Early Treatment of Diabetic Retinopathy Study chart).

DISCUSSION

Reliability

Multiple studies have examined the reliability of computerized versions of DVA testing. Reliability results vary widely across the literature depending on study

TABLE 4. *Test–retest reliability for DVA variables for all participants (n = 43 for Tester 1 and n = 42 for Tester 2, unless otherwise noted)*

	Tester 1 (n = 43)		Tester 2 (n = 42)	
Variable	ICC	95% CI	ICC	95% CI
SVA	0.934	0.881–0.964	0.955	0.919–0.976
Left DVA	0.846	0.733–0.914	0.844	0.729–0.913
Right DVA	0.832 (n = 42)	0.709–0.906	0.856	0.747–0.920
Up DVA	0.906 (n = 42)	0.831–0.948	0.897	0.817–0.944
Down DVA	0.848	0.737–0.915	0.897 (n = 41)	0.814–0.943
Left DVA loss	0.154	−0.151–0.431	0.295	-0.006–0.547
Right DVA loss	0.371 (n = 42)	0.079–0.605	0.366	0.073–0.601
Up DVA loss	0.443 (n = 42)	0.164–0.656	0.459	0.184–0.668
Down DVA loss	0.442	0.167–0.654	0.377 (n = 41)	0.082–0.611

CI indicates confidence interval; DVA, dynamic visual acuity; ICC, intraclass correlation coefficients; SVA, static visual acuity.

Copyright © 2016 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.

TABLE 5. *Test–retest ICCs for DVA variables by age groups*

	Tester 1		Tester 2	
Variable	Younger (n = 21)	Older (n = 22)	Younger (n = 21)	Older (n = 21)
SVA	0.434	0.940	0.499	0.948
Left DVA	0.417	0.696	0.414	0.720
Right DVA	−0.025	0.856 (n = 21)	0.108	0.792
Up DVA	0.598	0.847 (n = 21)	0.690	0.835
Down DVA	0.412	0.801	0.463	0.895 (n = 20)
Left DVA loss	0.152	0.157	0.365	0.260
Right DVA loss	0.008	0.579 (n = 21)	0.065	0.466
Up DVA loss	0.329	0.546 (n = 21)	0.418	0.485
Down DVA loss	0.240	0.552	0.175	0.621 (n = 20)

DVA indicates dynamic visual acuity; ICC, intraclass correlation coefficients; SVA, static visual acuity.

sample, equipment, algorithm, and settings used. The purpose of the current study was to determine both inter-rater reliability and test–retest reliability of a commercially available computerized DVA protocol in healthy younger and older adults.

When examining inter-rater reliability, results ranged from poor to excellent depending upon the variable examined. Overall, the raw scores demonstrated fair-to-excellent inter-rater reliability; whereas inter-rater reliability of the DVA loss scores was considerably worse. When stratified by age group, inter-rater reliability was better for the older adults than the younger adults. To the best of our knowledge, inter-rater reliability has not previously been reported.

Test–retest reliability results followed a similar pattern to the inter-rater reliability test results. Overall test–retest reliability ranged from poor to excellent depending upon the variable examined. In the current study, raw DVA scores were more reliable than DVA loss scores and reliability was better in older adults compared with younger adults. Similar findings were reported by Ward et al. (11) who found poor-to-fair test–retest reliability for healthy adults evaluated 7 to 10 days apart when using older Neurocom instrumentation with layered mirrors to achieve a testing distance of 4 m. Ward et al. did not report raw DVA scores, so a direct comparison of results is not possible; however, the authors do report older adults' scores were more stable than younger adults' scores, which was also found in the present study. Another study with similar methodology and instrumentation examined patients with vestibular dysfunction and reported test–retest reliability measures were better for raw measurement scores as opposed to the difference scores (9). Although the populations are different (healthy controls versus patients with vestibular dysfunction), both the current study and the Mohammad et al. study (9) found poor reliability for the DVA loss scores. Overall, our reliability results are similar to studies that used similar instrumentation and scoring algorithms (9,11). The only difference in instrumentation is that Ward et al. (11) and Mohammad et al. (9) used layered mirrors to reflect the image onto the screen for a distance of 4 m.

Conversely, our reliability results differ compared with studies using different instrumentation and protocols. For example, Rine et al. (10) reported reliability for the recently developed NIH toolbox DVA test for horizontal head rotation for subjects across the lifespan.

TABLE 6. *Mean (SD) of DVA scores by age group (collapsed across tester (n = 2) and time (n = 2) resulting in 88 and 89 data points for younger and older adults, respectively*

Variable	Younger Adults	SEM	MDC_{95}	Older Adults	SEM	MDC_{95}	*p*
SVA	−0.238 (0.095)	0.069	0.192	0.167 (0.238)	0.056	0.156	<0.001
Left DVA	−0.060 (0.119)	0.091	0.251	0.356 (0.233)	0.126	0.349	<0.001
Right DVA	−0.058 (0.120)	0.118	0.326	0.368 (0.236)	0.099	0.274	<0.001
Up DVA	−0.049 (0.150)	0.089	0.247	0.386 (0.242)	0.096	0.267	<0.001
Down DVA	−0.029 (0.146)	0.109	0.303	0.391 (0.233)	0.091	0.251	<0.001
Left DVA loss	0.179 (0.104)	0.090	0.248	0.189 (0.155)	0.138	0.383	0.615
Right DVA loss	0.180 (0.104)	0.102	0.283	0.202 (0.164)	0.113	0.314	0.274
Up DVA loss	0.189 (0.136)	0.108	0.298	0.224 (0.162)	0.113	0.313	0.117
Down DVA loss	0.209 (0.127)	0.113	0.312	0.225 (0.163)	0.105	0.291	0.461

DVA indicates dynamic visual acuity; MDC_{95}, 95% confidence interval for minimal detectable change (smallest amount of change that can be detected by a measure); SEM, standard error of measurement (an estimate of measurement error); SVA, static visual acuity.

$SEM = 1SD \times \sqrt{(1 - ICC)}$. Note: ICC from test–retest reliability (averaged for Testers 1 and 2) used in this calculation.

95% CI minimal detectable change (MDC_{95}) = $1.96 \times SEM \times \sqrt{2}$ (calculations from www.rehabmeasures.org accessed on September 17, 2015).

Copyright © 2016 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.

Focusing on the adult DVA loss scores, test–retest ICCs for the NIH toolbox DVA test are better than the current study. Differences between the NIH and the current study's DVA protocol include test distance (12.5 ft versus 6 ft), minimum head velocity (180 degrees/s versus 120 degrees/s), optotype (multiple optotypes from Early Treatment of Diabetic Retinopathy Study versus single optotype, ''E,'' with four orientations), and algorithm (larger letters progressing incrementally to smaller letters versus a convergence algorithm). Further, the Rine study included individuals with normal and abnormal vestibular function, which likely increased the between-subject variability of data compared with the current study, which included only healthy individuals (10). Greater between-subject variability in the Rine dataset may have contributed to higher ICCs than the current study (13). In addition, the current study's test–retest reliability is poorer than published values by Herdman et al. (6) for the horizontal plane and Schubert et al. (7) for the vertical plane. In both of those studies, the algorithm used to determine visual acuity was different from the current protocol but equivalent to the algorithm used in the NIH DVA test.

These studies indicate that one important difference between various DVA protocols that may contribute significantly to reliability is the algorithm used to determine threshold values. The early work by Herdman et al. (6) using a research system that is not commercially available demonstrated excellent reliability (ICCs = 0.83–0.94). The DVA portion of testing was initiated at a LogMAR of 0.500 above SVA and the optotype (''E'') was presented five times at each acuity level. Testing was terminated when all five optotypes were missed at a particular acuity level or reached a LogMAR of 0.000 (Snellen equivalent of 20/20). The convergence algorithm used in the present study is adaptive and allows an individual to reach threshold quickly. Initial testing of DVA begins at 0.300 logMAR above SVA, but then quickly adjusts according to the participant's response. Testing is terminated when three optotypes are correct at a particular level of acuity.

Another important difference in these algorithms is the use of a ''forced choice paradigm'' by Herdman and colleagues (6,7) in which the participant had to give an answer versus the current paradigm in which guessing is discouraged and the participant is encouraged to inform the tester when unsure of the response so that the response can be coded as incorrect for the algorithm.

Other differences in DVA protocol that may contribute to reliability include head velocity, lower limit of optotype size, and testing distance (see table, supplemental digital content 3, http://links.lww.com/MAO/A377, for a summary of DVA study protocols). Minimum targeted head velocities required to trigger the optotype in the current protocol were 120 degrees/s for horizontal DVA and 100 degrees/s for vertical DVA. Minimum head velocities to trigger stimulus presentation in previous studies have ranged from 85 to 180 degrees/s for the horizontal plane and 60 to 120 degrees/s for the vertical plane (6–11). The rationale for the target velocities in the current study was based on work that demonstrated that individuals can produce smooth pursuit eye movements up to velocities of 100 degrees/s; thus, during slower head movement, smooth pursuit eye movements may contribute to gaze stability (15). Ensuring that head velocities exceed a minimum standard is critically important to isolate contribution from the vestibular system. Size of the smallest optotype displayed varies across studies and may also contribute to reliability. In the current study, the smallest optotype displayed was 20/10 (consistent with NIH DVA protocol; 10), whereas some previous studies have limited the smallest optotype to 20/20 (6,7,16). Another difference between studies is the distance at which the test was administered. It is well established that VOR gain is modulated by viewing distance (17–19); however, it is not expected that testing distance is a significant factor in DVA reliability since all testing distances were at least 2 m which likely results in minimal change in VOR gain.

Age Differences

It is interesting that, in general, YA were less reliable than OA for both inter-rater reliability and test–retest reliability. This difference is most pronounced in the test–retest reliability ICCs (Table 5). Across all variables (other than left DVA loss), OA were more reliable than YA. In the case of left DVA loss, both groups demonstrated poor reliability. For inter-rater reliability, the differences are less systematic; although in general inter rater-reliability was better for older adults versus younger adults (Table 3). The largest increase in reliability between sessions is for SVA in YA with a significant improvement from Sessions 1 to 2; although both increases and decreases in reliability were noted between sessions for the other variables. It is not clear the reason why YA were less reliable than OA. Several potential reasons for the difference may exist. First, the YA data exhibited minimal variability: YA had excellent SVA with significantly smaller standard deviation compared with the OA and YA exhibited less loss of visual acuity with head movements. Minimal variability in data can produce lower ICCs (13). Second, although participants were instructed not to guess, it is possible that YA were more inclined to guess. Last, it may be possible that a learning effect occurred; however, this seems unlikely as no systematic changes were noted for test–retest reliability.

Herdman et al. (6) demonstrated a significant effect of age on DVA loss scores, which was not confirmed with the current data. The difference in findings is likely because of the artificial threshold of 20/20 (LogMAR of 0.000) for SVA in the Herdman study. On average, our YA had SVA scores of 20/12 (LogMAR of −0.238). If we use this SVA score, Herdman's results for 20-year olds (DVA = 0.02 − −0.238 = 0.258 LogMAR) would result in similar scores to the current results (DVA = 0.180 LogMAR). On the other hand, our older participants had average SVA scores of 20/30 (LogMAR of 0.167); thus, DVA loss scores for OA would not be

Copyright © 2016 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.

impacted by Herdman's threshold for SVA. Li et al. (20) published normative data for the NIH Toolbox DVA and found age-related differences. The mean DVA LogMAR values for 18 to 29-year olds were 0.084 and for 60-year olds 0.215 versus the current values of 0.180 and 0.196 (left and right DVA loss averaged), respectively. It is not clear why the values for the younger group in the current study are worse than those obtained in the NIH normative study. The threshold for SVA was equivalent (20/10) in both the NIH and the current protocol. The main difference between the two study protocols is the convergence and forced choice paradigms.

Moreover, there was no significant difference between 60- and 70-year olds for any of the variables in the current study. In part this may be because of the lack of power; the study was not powered to examine differences in subgroups. However, the values for DVA loss for horizontal (left and right) head turns for 60-year olds (0.182 and 0.194, respectively) and 70-year olds (0.176 and 0.191, respectively) in the current study are similar to mean normative values of 0.215 and 0.228 for 60-year olds and 70- to 85-year olds (respectively) published by Li et al. (20). It may be that there is a gradual change in DVA with age that was not captured in the current study.

LIMITATIONS

A potential limitation of the current study was testing older individuals without the use of multifocal lenses. This protocol may have introduced error. Researchers have shown that the VOR is adapted on the basis of the degree of magnification in single-focal lenses (21–23) and further, there is a period of time needed to adapt VOR gain to the magnification. It is less clear how the VOR adapts in individuals using multifocal lenses. Testing individuals while wearing multifocal lenses likely would introduce error because multifocal lenses have varying degrees of magnification within the lens. These different degrees of magnifications would necessitate the VOR gain to vary relative to the viewing location in the lens (24). There is no consensus in the literature as to which testing method (i.e., with or without multifocal lenses) is best.

In the current study, volitional (active) head movements were used for testing DVA because this is the standard clinical protocol. Furthermore, active head movements reflect some activities of daily living (e.g., looking both ways to cross a street). This testing paradigm has been used across several studies examining DVA (e.g., 6–11, 20); whereas other studies have used a nonvolitional (passive) head movement for assessment of DVA (16,25,26). The ability to preprogram eye movements for gaze stability is different for active versus passive head movements; thus, comparison of the current study's results is limited to studies using active head movements.

CONCLUSIONS

In conclusion, the test–retest reliability of the current version of a commercially available computerized DVA system has poor-to-fair reliability for DVA loss scores, which are the primary outcome measures for vestibular rehabilitation. Herdman et al. (8) demonstrated a change in DVA loss scores of approximately 0.200 LogMAR in patients with unilateral vestibular hypofunction and 0.130 LogMAR in patients with bilateral vestibular hypofunction. Using the current system those changes are within measurement error and would potentially be missed. Additionally, the current study did not replicate age-related changes in DVA loss scores that have been demonstrated using different paradigms. The convergence paradigm and not using a forced choice paradigm may contribute to poor reliability.

Acknowledgments: Special thanks to Hana Gardner, Cody Keck, Justin Kerley, Andrew King, Brent Kington, Kirsten Lamie, and Ashley Tice for assistance with participant testing. The authors appreciate the participants who volunteered time to contribute to this study.

REFERENCES

1. Colledge NR, Wilson JA, Macintyre CC, MacLennan WJ. The prevalence and characteristics of dizziness in an elderly community. *Age Ageing* 1994;23:117–20.
2. Tinetti ME, Williams CS, Gill TM. Dizziness among older adults: A possible geriatric syndrome. *Ann Intern Med* 2000;132:337–44.
3. Kroenke K, Lucas CA, Rosenberg ML, et al. Causes of persistent dizziness. A prospective study of 100 patients in ambulatory care. *Ann Intern Med* 1992;117:898–904.
4. Cohen HS, Wells J, Kimball KT, et al. Driving disability and dizziness. *J Safety Res* 2003;34:361–9.
5. Yardley L, Verschuur C, Masson E, et al. Somatic and psychological factors contributing to handicap in people with vertigo. *Br J Audiol* 1992;26:283–90.
6. Herdman SJ, Tusa RJ, Blatt P, et al. Computerized dynamic visual acuity test in the assessment of vestibular deficits. *Am J Otol* 1998;19:790–6.
7. Schubert MC, Herdman SJ, Tusa RJ. Vertical dynamic visual acuity in normal subjects and patients with vestibular hypofunction. *Otol Neurotol* 2002;23:372–7.
8. Kaufman DR, Puckett MJ, Smith MJ, et al. Test–retest reliability and responsiveness of gaze stability and dynamic visual acuity in high school and college football players. *Phys Ther Sport* 2014;15:181–8.
9. Mohammad MT, Whitney SL, Marchetti GF, Parto PJ. The reliability and response stability of dynamic testing of the vestibule-ocular reflex in patients with vestibular disease. *J Vestib Res* 2011;21:277–88.
10. Rine RM, Roberts D, Corbin BA, et al. New portable tool to screen vestibular and visual function—National Institutes of Health Toolbox initiative. *J Rehabil Res Dev* 2012;49:209–20.
11. Ward BK, Mohammad MT, Whitney SL, Marchetti GF, Furman JM. The reliability, stability, and concurrent validity of a test of gaze stabilization. *J Vestib Res* 2010;20:363–72.
12. Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. *J Clin Epidemiol* 2005;58:595–602.
13. Portney LG, Watkins MP. *Foundations of Clinical Research: Application to Practice*, 3rd ed. Philadelphia, PA: F. A. Davis Company; 2015.
14. Fleiss JL. *The Design and Analysis of Clinical Experiments*. New York: John Wiley & Sons; 1986.
15. Meyer CH, Lasker AG, Robinson DA. The upper limit of human smooth pursuit velocity. *Vision Res* 1985;25:561–3.
16. Viciana D, Ferrer J, Palma MJ, Zapata C, Lopez-Escamez JA. Dynamic visual acuity during head-thrust test in canal planes in healthy subjects and patients with vestibular neuritis. *Acta Otolaryngol* 2010;130:1260–6.

Copyright © 2016 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.

17. Viirre E, Tweed D, Milner K, Vilis T. A reexamination of the gain of the vestibuloocular reflex. *J Neurophysiol* 1986;56:439–50.
18. Crane BT, Demer JL. Human horizontal vestibulo-ocular reflex initiation: Effects of acceleration, target distance, and unilateral deafferentation. *J Neurophysiol* 1998;80:1151–66.
19. Paige GD, Telford L, Seidman SH, Barnes GR. Human vestibuloocular reflex and its interactions with vision and fixation distance during linear and angular head movement. *J Neurophysiol* 1998;80: 2391–404.
20. Li C, Beaumont J, Rine R, Slotkin J, Schubert M. Normative scores for the NIH Toolbox dynamic visual acuity test from 3 to 85 years. *Front Neurol* 2014;5:223.
21. Cannon S, Leigh R, Zee D, Abel L. The effect of the rotational magnification of corrective spectacles on the quantitative evaluation of the VOR. *Acta Otolaryngol* 1985;100:81–8.
22. Demer J, Porter F, Goldberg J, Jenkins H, Schmidt K. Dynamic visual acuity with telescopic spectacles: Improvement with adaptation. *Invest Ophthalmol Vis Sci* 1988;29:1184–9.
23. Gauthier G, Robinson D. Adaptation of the human vestibuloocular reflex to magnifying lenses. *Brain Res* 1975;92:331–5.
24. Michaelides E, Schutt C. The correlation between the vestibulo-ocular reflex and multi-focal ocular correction: Implications for vestibular compensation. *Am J Otolaryngol* 2014;35:572–6.
25. Schubert M, Migliaccio A, Della Santina C. Dynamic visual acuity during passive head thrusts in canal planes. *J Assoc Res Otolaryngol* 2006;7:329–38.
26. Vital D, Hegemann S, Straumann D, et al. A new dynamic visual acuity test to assess peripheral vestibular function. *Arch Otolaryngol Head Neck Surg* 2010;136:686–91.

Copyright © 2016 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.