Published in final edited form as:

J Vestib Res. 2011 ; 21(5): . doi:10.3233/VES-2011-0430.

The reliability and response stability of dynamic testing of the vestibulo-ocular reflex in patients with vestibular disease

Maha T. Mohammad, PT, PhD,

University of Pittsburgh, School of Health and Rehabilitation Sciences, Department of Physical Therapy, Pittsburgh, PA

Susan L. Whitney, PT, PhD, DPT, NCS, ATC, FAPTA,

University of Pittsburgh, School of Health and Rehabilitation Sciences, Departments of Physical Therapy and Otolaryngology, Pittsburgh, PA

Gregory F. Marchetti, PT, PhD,

Duquesne University, Rangos School of Health Sciences, Department of Physical Therapy, Pittsburgh, PA

Patrick J. Sparto, PT, PhD,

University of Pittsburgh, School of Health and Rehabilitation Sciences, Departments of Physical Therapy and Otolaryngology, Pittsburgh, PA

Bryan K. Ward, MD, and

Johns Hopkins University, Department of Otolaryngology- Head and Neck Surgery, Baltimore, MD

Joseph M. Furman, MD, PhD

University of Pittsburgh, School of Medicine, Departments of Otolaryngology and Physical Therapy, Pittsburgh, PA

Abstract

The purpose of the study was to investigate the test-retest reliability and response stability of the Dynamic Visual Acuity (DVA) and Gaze Stabilization Test (GST) in patients with vestibular disorders. Twenty-nine patients with vestibular disease (16 – 78 years) participated. Subjects performed the GST and DVA in pitch and yaw planes, twice in one session and once after 7 – 10 days. The GST output is the maximum head velocity at which the patient was able to identify orientation of the letter E. The DVA output is the change in visual acuity when moving the head compared to static acuity. Subjects indicated their level of dizziness and visual blurring using a visual analog scale. Within- and between-sessions intraclass correlation coefficients ranged between 0 – 0.5 for the DVA and GST measures, with better correlations for within-session assessments. Response stability (standard error of measurement / mean) of the GST ranged between 21 – 32% and the DVA ranged between 25 – 69% with vertical DVA being most influenced by measurement error. Subjects' symptoms did not correlate with performance on either test. The current test protocol needs refinement to enhance reliability and stability in persons with vestibular disorders.

Keywords

Vestibulo-ocular reflex; gaze stabilization test; dynamic visual acuity

1.Introduction

In order to maintain gaze on a stationary target while the head is moving, the vestibular system detects the velocity of the head movement and, by means of the vestibular-ocular reflex (VOR), induces eye movements that are equal in magnitude and opposite in direction to the head movements [8, 22, 30]. Maintaining fixed gaze during head movement insures that images are held steady on the retina, thus maintaining clear vision. Compared to the other systems that control the movement of the eyes, the VOR has the shortest latency (<5 msec) [22]. One method of measuring VOR function is the VOR gain, i.e. the ratio of the eye velocity to the head velocity. Reduced or increased VOR gain indicates that the velocity of the eyes' movements is not equal to that of the head. VOR gain should normally equal 1, however in low light conditions the VOR gain can be less than 1. Dysfunction of the VOR, may lead to deterioration of vision during head movements. Visual blurring indicates that compensation following a vestibular insult has not been completed by mechanisms such as VOR gain restoration or vestibular catchup saccades. [13, 30, 32].

The dynamic visual acuity test (DVA) [14–17, 31] and the gaze stabilization test (GST) [1, 7, 37] are two functional measures of the VOR. During DVA testing, visual acuity during head movement is compared with visual acuity with the head stationary to provide an indication about the amount of loss of visual acuity due to head movement. The DVA test was first introduced as the dynamic illegible E test [23, 24]. In the dynamic illegible E test, the subject attempts to determine the orientation of the letter E on a visual acuity chart while the head is moved passively by the examiner at predetermined frequencies (1 or 2 Hz). A computerized version of the test was later introduced in which the letters were presented on a computer screen when the head velocity reached a certain, predetermined value [17]. In the computerized version of the DVA, head movements are monitored using a velocity sensor mounted on the patient's head. The GST provides another way of examining the dynamic function of the VOR. In the GST, the maximum velocity a subject is able to achieve while still being able to correctly identify the orientation of an optotype (the letter E) presented at a specific size is measured [7, 36].

Patients with unilateral or bilateral vestibular lesions perform significantly worse on the DVA [2, 5, 14, 15, 31] and the GST [2, 7, 27, 37] when compared with healthy subjects. The performance of patients improves following a program of vestibular physical therapy which includes gaze stabilization exercises [1, 3, 10, 14, 15, 19, 28]. Accordingly, the DVA test and the GST results can provide valuable information about the function of the VOR, degree of compensation, and provide a method to track changes during the physical therapy intervention. The DVA test uses successively smaller optotypes, which may limit the DVA's use in patients with visual comorbidities,[27]whereas the GST measures the maximum head speed that the subject can achieve while being able to correctly identify an optotype[7]. It is not clear if both tests measure the same construct.

Earlier versions of the GST and DVA test were reliable [7, 17, 27, 29]. NeuroCom International recently introduced a new computerized version of the tests; the InVision™ Tunnel System.

The purpose of the current study was to examine the test-retest reliability and the response stability of the GST and the DVA test using the computerized InVision™ Tunnel system in patients with vestibular disorders. This study also examined the influence of patients' perceived levels of visual blurring and dizziness symptoms on their performance on the GST and DVA tests.

2. Methods

2.1 Subjects

Twenty-eight subjects (mean age = 50 years, SD = 14, range: 16 – 78, 4 male) with vestibular dysfunction were recruited for the study. Participants were included in the study if they were diagnosed with a vestibular disorder by a neuro-otologist. Patients were referred to the study by their treating physical therapists when the subjects demonstrated that they could tolerate the head movements required for the DVA and GST, as demonstrated by their ability to perform the VORx1 exercise without a large increase in their symptoms. The study procedures were approved by the Institutional Review Board at the University of Pittsburgh. All participants signed the written consent before the start of the research activities. Exclusion criteria included: 1) a musculo-skeletal problem that limited neck mobility, 2) severe nausea, headache, or dizziness symptoms with head movement, 3) or static visual acuity (with corrective eye wear) worse than 20/40.

Sixteen subjects had peripheral vestibular disorders, 9 subjects had central vestibular disorders, and 3 subjects had mixed central and peripheral vestibular disorders. For the 16 subjects with peripheral loss, 7 had been previously treated for benign paroxysmal positional vertigo, and 12 had a clear unilateral hypofunction. In addition, 8 of the subjects with peripheral disorders had dysfunction in their horizontal semicircular canal only, 4 had otolith dysfunction only, and 2 subjects had both canal and otolith abnormalities. Of the 9 subjects with central vestibular dysfunction, 6 had post-concussive symptoms, 2 had rotational chair abnormalities, and 1 had oculomotor dysfunction in addition to rotational chair abnormalities.

2.2 Equipment

The InVision™ device (NeuroCom International, Inc., Clackamas, OR, USA) is composed of a CRT computer monitor that is placed inside a cabinet with a series of four mirrors that reflect the image on the screen; the subject was seated on a chair near the edge of the cabinet so that the total visual distance between the subject and the image displayed on the computer monitor was approximately 4 meters. Testing distance was standardized at 4 meters to better match other traditional visual assessments and testing was conducted with the subjects personal lens correction, if needed. The operator viewed a separate computer screen to control the InVision testing (Figure 1). A head mounted sensor (an InterSense Inertia Cube², 3-axis integrating gyro) was used to provide feedback about the speed and range of head movements. The head sensor is mounted on a headband that is secured around the head using an adjustable strap.

The InVision device was placed in a quiet room with white walls in the background. Room light was measured using a lux meter and fixed at 70 lux (per manufacturer recommendation).

2.3 Testing protocol

During all testing an optotype (the letter E) was presented on the computer screen. The open end of the E was pointing either up, down, right, or left and the subject was asked to determine the orientation of the E and verbally respond to the operator, who entered the response using the corresponding arrow key. The size of the E was reported in logMAR units. A logMAR is a logarithm of minimal angle of resolution [11] and is considered the standard psychophysical unit for assessing visual acuity [12, 29]. All trials started with a short practice to ensure that subjects understood the required tasks and were comfortable performing the eye/head movements.

2.3.1 Static visual acuity—Testing commenced with the assessment of static visual acuity (SVA). The SVA test determined the smallest size of the optotype E the subject could see and identify correctly while the head was stationary. For this test, an outline of a black square was presented for two seconds in the center of the screen to ensure that the subject fixed his/ her vision on the location where the E would appear. After two seconds, the square disappeared, then after a 200 msec delay, the E appeared. The E remained on the screen for 2 seconds. The best Parameter Estimate by Sequential Testing (best PEST) adaptive algorithm was used to calculate the subject's SVA score. The original PEST algorithm was developed by Taylor and Creelman as an adaptive procedure to measure psychometric properties [34]. Pentland proposed a modification of the original PEST algorithm, i.e. the best PEST adaptive algorithm [25]. Compared to other algorithms, the best PEST algorithm is a more efficient procedure, requiring the fewest trials to reach a certain precision [21]. The best PEST algorithm has been used in the InVision device to determine the SVA score, in addition to determining the minimum perception time (mPT), DVA, and GST scores as will be discussed. A new SVA score was established during each testing session.

2.3.2 Minimum perception time—After establishing the subject's SVA, the mPT was measured. In the mPT test, the shortest duration of optotype presentation while the subject could still correctly identify the orientation of the optotype was determined. The optotype size during the mPT was fixed at 0.25 logMAR above the subject's SVA. During the mPT test a square was presented at the center of the computer screen. After two seconds the square disappeared and 200 msec later an E appeared in its place. Based on the duration of presentation of the optotype and the subject's performance, the best PEST algorithm was used to calculate the mPT score.

2.3.3 Dynamic visual acuity and gaze stabilization tests—The outcomes of the SVA and mPT tests were used as the baseline for the GST and DVA tests. During the GST, the size of the E was held constant at 0.25 logMAR above the subject's SVA. A value of 0.25 logMAR above SVA has *been reported previously in the literature*. During the GST and the DVA test, the duration of the optotype presentation was calculated as follows:

- If the mPT was 40 msec or less, the optotype E was presented for 40 – 75 msec
- If the mPT was greater than 40 msec, the optotype E presentation time ranged between (mPT) and (mPT + 35) msec

According to the above criteria, an mPT score > 70 msec would lead to a presentation time >100 msec. Such long presentation time in the GST or the DVA test may allow for other eye movement systems, such as saccades, to contribute to the fixation of the eyes on the target during head movement. The InVision system flags the output in such cases as “unreliable”.

The DVA and the GST required the patients to move their heads in the pitch or yaw planes while watching the computer screen. Feedback was provided to the subject and the operator via two feedback bars that appeared on the screen before each task (Figure 2). One of the bars provided feedback regarding the range of head movement. The amount of head movement was reflected by the movement of a dot on the feedback bar. For all tasks, the subject was expected to move his/ her head approximately 20 degrees to the right, left, up, and down from the subject's neutral head position. The second feedback bar provided speed of head movement information which turned from grey to green whenever the required head velocity was achieved. Subjects were given a practice trial before each test to ensure they understood the task and the feedback provided via the two bars.

Each trial began with a presentation of the feedback bars on the screen to show the subject the required range and speed of head movement. Once the subject achieved the required speed, the bars disappeared and the subject was instructed to continue the head movement and focus his/ her vision on the square in the middle of the screen where the E appeared.

If the subject was unable to achieve the required range or velocity of head movement within 8 seconds from the start of the task or could not maintain the required speed for the minimum duration of the optotype presentation (see above), the trial was interrupted and a message appeared on the screen to indicate that the required task velocity was not achieved. The latter was considered a failed trial.

2.3.3.1 Dynamic visual acuity test: In the DVA test, the speed of required head movement was fixed while the size of the optotype E changed. The InVision system (version 8.3.0) developed by NeuroCom uses a velocity range between 85 – 120 deg/ sec for the horizontal DVA test and 60 – 85 deg/ sec for the vertical direction. Although the velocity we used (85 – 120 deg/ sec) is slower than the velocity chosen by the other laboratories (120 – 180 deg/ sec) [1, 10, 14, 15, 17, 29], early trials of the DVA test using the InVision system with patients with vestibular disease showed more symptom aggravation and reduced patient tolerance at head velocities greater than 120 deg/ sec (personal communication with NeuroCom, Feb 1, 2010). Other oculomotor systems for maintaining gaze stabilization, such as the smooth pursuit system, are unlikely to be functional at these velocities [4, 33].

The size of the optotype in each trial varied based on the subject's responses and the best PEST adaptive algorithm until the DVA score was achieved. The DVA scores were determined for the left and right head movements during the yaw head movement test. Likewise DVA scores for up and down directions were determined during the pitch head movement test. The DVA score is the size of the smallest optotype that the subject was able to correctly identify while moving his/ her head within the predetermined speed limits. This "dynamic" visual acuity was then compared to the "static" visual acuity to determine the amount of loss in visual acuity during head movement. The DVA loss is defined as the DVA minus the SVA ($DVA\ loss = DVA - SVA$). The DVA loss score is computed for the left, right, up, and down head movements. The outcome of the DVA test is reported as the logarithm of minimum angle of resolution (logMAR) units.

2.3.3.2 Gaze stabilization test: In the GST, the size of the optotype was fixed at 0.25 logMAR above the subject's SVA while the speed of head movements changed based on subject's performance and the best PEST algorithm. The test was stopped when the subject could no longer increase the head velocity while still being able to accurately identify the orientation of the optotype. Using the best PEST algorithm, if the subject was unable to perform the GST in one direction after the first few trials, even at very slow speed, testing for that direction was stopped and the subject was assigned a "No Score" for that direction. The outcome of this test is the maximum speed at which the subject can move his/ her head while still able to identify the orientation of the E. The outcome is reported for the left, right, up, and down directions in degrees/second.

2.4 Experimental procedure

In order to establish the test-retest reliability and response stability of the InVision test, patients with vestibular disorders were tested twice within one session with a 30 minute rest between testing and once again 7 – 10 days later. The order of testing, i.e. DVA vs. GST, and the direction of the head movement, i.e. horizontal vs. vertical, were randomly chosen for each subject. Each repetition of the whole battery of testing was completed in 20 – 25 minutes.

Before the start of head movement testing and after each head movement task, a visual analog scale (VAS) was used to indicate the subject's level of headache, dizziness, nausea, and visual blurring [9]. For each symptom, a 10-cm vertical line marked with "no symptom at all" at the bottom of the line and "as bad as it can be" on top of the line was used. Subjects were asked to place a mark on the line for each symptom based on how they felt [9]. In order to examine the change in subject's symptoms at any given test, their VAS score following that test was compared to the score following the immediately preceding test.

2.5 Statistical analysis

To determine the test-retest reliability of the measures, an intraclass correlation coefficient (ICC, two-way mixed-effects model) and 95% confidence intervals (95% CI) were calculated for repetitions 1 and 2 within the first session, repetitions 1 and 3 between the first and second sessions, which were 7–10 days apart, and repetitions 2 and 3, also between sessions 1 and 2. The ICC is the ratio of between-subject variability to total variability with values ranging between 0 (no reliability) and 1.0 (perfect reliability) [26]. Values of the ICC between 0.75 and 1.0 indicate excellent reliability, values between 0.40 and 0.74 indicate fair to good reliability, while values less than 0.39 indicate poor reliability [6]. Excellent reliability (i.e. high ICC values) indicates that subject's scores are consistent over repeated measures. Intraclass correlation coefficients were calculated for subjects' scores on the SVA test, GST, DVA loss (DVA – SVA), and DVA scores.

Response stability of the GST and the DVA test was examined using the standard error of the measurement (SEM) and the SEM% for the two within-session repetitions and the first and third repetitions (between-session). Response stability is a measure of the consistency of repeated responses over time [26]. The SEM is an estimate of the within-subject variability (standard deviation of an individual's score) after repeated measures. The SEM was calculated using the sample standard deviation (SD) and the ICC as follows: $SEM = SD \cdot (1 - ICC)$. The SEM% provides a standardization of the SEM to enable the comparison of measurement error across different tests. The SEM% equals $SEM / \text{mean} \cdot 100\%$. Using the SEM, the minimal detectable change (MDC) was calculated for those tests with ICC values that are significantly different from 0. The MDC is the amount of change in the measure that is greater than chance or measurement error [20]. The MDC is an important measure when examining the change in subjects' performance after repeated measurements on a test. The $95\% \text{ MDC} = SEM \cdot 1.96 \cdot 2$.

Correlational analysis between the GST and the SVA was performed using Pearson product moment correlation (PPMC). The goal was to examine whether high SVA scores (larger optotypes) correlated with faster GST scores. The correlation between the GST and the DVA test was calculated to examine the relationship between the two tests. In order to examine whether subjects' symptoms were associated with their performance on the GST or DVA test, correlational analysis between the GST or DVA loss scores and subjects' symptoms was performed using Spearman non-parametric correlations. Spearman correlations were chosen due to the measurement properties of the VAS which was used to report the subjects' symptoms. Due to the multiple comparisons performed in the correlational analyses, alpha levels (α) were set at 0.01. All analyses were performed using SPSS (version 17, SPSS Inc., Chicago, IL.)

3. Results

Table 1 shows the mean (SD) scores for subjects over the 3 repetitions of testing on the SVA, GST, DVA loss, and DVA. The GST and DVA scores are shown for left/right and up/down directions. Of the 28 subjects who participated in the study, 24 subjects returned for the second visit in which the third repetition of testing was performed. One of the subjects

who returned for the second visit did not repeat the DVA for the 3rd repetition due to limited session time (i.e. n for DVA on 3rd repetition = 23) but the subject did complete the GST. The overall mean of the SVA scores was -0.04 ± 0.12 logMAR. The mPT scores ranged between 20 and 50 msec in increments of 10 msec. The median of the mPT score was 20 msec.

On average, patients' scores ranged between 124 – 152 deg/ sec on the GST and 0.16 – 0.25 logMAR on the DVA loss test. On the DVA loss test, patients' scores in the yaw plane were worse than in the pitch plane, perhaps due to the higher head velocity in the yaw plane (85 – 120 deg/ sec) compared to the pitch plane (60 – 85 deg/ sec); the difference was not statistically significant. There were no significant differences between patient's performance on the left and right directions or between up and down directions in the GST and DVA test. *There were statistically significant improvements in the left DVA between repetition 1 and repetition 3, and improvements in the right DVA and DVA loss between repetitions 1 and repetitions 2 and 3. No significant differences were found in any of the other scores.*

In Table 2, test-retest reliability was examined using the ICC (model 3.1). Test-retest reliability for the SVA was excellent for all of the comparisons, ranging from 0.78 between repetitions 2 and 3, to 0.88 between repetitions 1 and 2. The reliability of the GST was generally poor (ICC values between 0 and 0.48). Estimates of the GST reliability in the horizontal direction (ICC: 0 – 0.48) were slightly better than the vertical direction (ICC: 0 – 0.35). In the DVA test, the reliability of the absolute DVA scores were good to excellent (ICC: 0.64 – 0.83), but the reliability of DVA loss was poor to fair in all four directions (ICC: 0 – 0.50).

The response stability of the GST and DVA loss was examined using the SEM and SEM% (Table 3). The SEM% for the GST ranged between 22 – 37%, and for the DVA loss between 38 – 75%. The vertical DVA loss score appeared to be the most influenced by measurement error (SEM%: 50 – 75%).

The GST and DVA scores of the 12 patients with a unilateral peripheral lesion (8 right and 4 left lesions) in terms of lesion vs. non-lesion sides are displayed in Table 4. There were no significant differences toward the lesion and non-lesion sides. There were no significant correlations between subjects' scores on the GST and their SVA scores ($R < 0.40$) (Table 5). The lack of significant correlation between GST and SVA indicated that, within the parameters used in the current test, patients' performance on the GST was not influenced by the size of the optotype presented during the test.

The correlations between the GST and the DVA test were examined (Table 6). The correlations were expected to be somewhat strong to reflect that the two tests measure two constructs that are similar (the function of the VOR) but not identical. In general, the GST and the DVA loss were inversely correlated. Higher GST scores (i.e. faster head velocities) were correlated with lower DVA loss scores (i.e. reduced loss of visual acuity with head movement). Most of the correlations (8 out of 12) were weak ($r < 0.38$). However, 4 of the 12 correlations were strong ($r: 0.52 – 0.72$). The correlations seemed to be slightly stronger in the yaw plane (left and right) compared to the pitch plane (down and up).

Patients were asked to report their symptoms of nausea, headache, dizziness, and visual blurring on a 10-cm VAS before and after each test. Most patients did not report any nausea or headache symptoms during testing. On average, patients' level of dizziness before testing was 3.7 cm out of 10.0 cm. The average increase in dizziness after testing was 1.5 cm (SD = 2.3). The level of visual blurring before testing was 2.6 / 10.0 cm and the average increase in the symptom after testing was 0.6 cm (SD = 1.5). Over the 3 repetitions of testing, patients

were asked about their symptoms 12 times (3 repetitions of 2 tests (GST and DVA) with 2 directions in each test (horizontal and vertical)). The change in dizziness was statistically significant 6 out of the 12 times. The change in visual blurring was statistically significant only once; after the 3rd repetition of the vertical GST. None of the correlations between patients' scores on the GST or the DVA loss test and their dizziness or visual blurring symptoms were significant.

4. Discussion

The magnitude of the GST velocity in this group of subjects with vestibular dysfunction (130 – 152 deg/s horizontal, and 124–142 deg/s vertical) is similar to the velocity reported for healthy subjects in several other studies [7, 18, 36]. In contrast with the current study, Goebel et al. (2007) reported a much lower GST velocity of subjects with unilateral vestibular loss (UVL, mean 84 deg/s)[7]. It is not clear what can account for the differences between our population and that of Goebel et al. (2007)[7], but perhaps it is because we had documented horizontal canal paresis in approximately 50% of our subjects, or that our subjects were less symptomatic at the time of testing. Therefore, it is not certain if the current version of the test is able to distinguish between people with and without disease. Regarding the DVA loss, our subjects had less impairment in the horizontal plane than the subjects with UVL as reported by Herdman et al. (1998)[17], and similar impairment in the vertical plane compared with subjects with UVL in the study of Schubert et al. [29].

The reliability of the NeuroCom InVision™ GST and DVA test in patients with vestibular disorders is poor to fair. Overall, the reliability of the GST was better within-session than between-sessions. In addition, the reliability of the absolute DVA scores was better than that of the GST or DVA loss scores. To our knowledge, no other study has examined the reliability of the DVA or the GST between-session in persons with vestibular disorders; all reliability studies available in the literature repeated testing in one session on the same day [7, 17, 29]. Because of the typical variation in symptom improvement or worsening that is characteristic of vestibular dysfunction, the DVA and GST scores may have been better or worse in the follow-up assessment 1 week later. Most of the subjects were undergoing vestibular rehabilitation, which may have resulted in improvement in scores in some subjects. There was a statistically significant improvement in the horizontal DVA and DVA loss between the first and third repetitions, but no improvements were found in the vertical DVA, DVA loss or GST scores. Nonetheless, a change in subject status can be a contributing factor in the lower reliability scores reported between sessions.

The DVA loss is defined as DVA – SVA. Although the DVA scores were reliable (ICC between 0.64 and 0.83) and the SVA test was also reliable (ICC between 0.78 and 0.88), the DVA loss scores were not as reliable (ICC between 0 and 0.50). The drop in reliability raises the question that maybe the patient's SVA scores should be established at the beginning of the intervention and the same SVA scores should be used for all follow up DVA tests. The standardization of the test procedure could increase the reliability of the test and ensure that any changes in the DVA loss scores are attributed to the changes in DVA and not SVA scores.

In a previous GST study, the reliability of the NeuroCom InVision system was examined in healthy young (25 ± 3 years) and older (76 ± 5 years) subjects [36]. The device was found to be reliable for both the GST (ICC values between 0.54 – 0.75) and the DVA test (ICC values between 0.10 and 0.60)[36]. In the study of healthy subjects, right and left test scores were averaged into a single yaw plane score. Similarly, up and down scores were averaged into a single pitch plane score. The statistical analyses in the current study were conducted differently since the nature of the patients' vestibular diseases renders the unidirectional

performance asymmetrical, thus averaging the scores was not plausible in the patient population. However, when patients' scores in the yaw plane were averaged, the reliability values did not improve; the reliability within-session was 0.44 and between-session 0.09. It appears that the low reliability in the current study was not due to the different statistical methods used.

Herdman et al [17] examined the reliability of an earlier version of the computerized DVA test in the horizontal direction while Schubert et al [29] examined the vertical direction. Their results indicate that the computerized DVA test was reliable when used with patients with vestibular deficits (ICC = 0.83 for horizontal and 0.94 for vertical DVA). Differences exist between the current study and the ones conducted by Herdman et al. and Schubert et al. that render the comparison between the studies difficult. The testing distance in our study was fixed at 4 meters and the other two studies used a testing distance of 2 m. Although the VOR gain changes as a function of target distance, especially for near targets less than 1 m away, at distances of 2 and 4 meters the change in VOR gain is negligible [35], and therefore unlikely to explain differences in reliability. We chose 4 m to better match other traditional visual assessments. In addition, the head velocity used for the DVA test in the Herdman et al. and Schubert et al. studies was 120 – 180 deg/ sec for both horizontal and vertical DVA. In our study, we used a slower velocity range (85 – 120 deg/ sec for horizontal and 60 – 85 deg/ sec for vertical). Finally, while the InVision system uses the best PEST algorithm to calculate the final DVA score, the algorithm used by Herdman et al and Schubert et al always used 5 presentations of the optotype (the letter E) at each acuity level; such an algorithm can be time consuming when considering that the DVA test is done for right, left, up, and down directions. One of the advantages of the best PEST adaptive algorithm is that it uses fewer trials to calculate the final score but the increased efficiency may come at a cost of reduced reliability.

Schubert et al [29, 30] reported that when examining visual acuity of their subjects, testing was stopped when a subject scored 0 logMAR on the SVA or the DVA. A score of 0 logMAR corresponds to 20/20 vision on the standard Snellen chart. In our experience, subjects, especially those with corrected vision, can score below 0 on visual acuity testing, especially the SVA. For example, in the current study, 19 out of the 28 patients had an SVA score below 0 logMAR. Stopping testing at 0 logMAR means that the DVA loss score will not reflect the real difference between subjects' DVA and SVA scores but the difference between their DVA score and 0. Assessing reliability of DVA loss when using a lower limit of 0 for SVA would be similar to assessing reliability for DVA (not DVA loss). Our study found that reliability for DVA was higher than that for DVA loss and more similar to the findings of Schubert et al.

When examining the DVA test or the GST in the vertical direction, it is important to document if the patient is wearing bifocal glasses. During the testing of the downward direction, the optotype is presented shortly after the subject starts moving downward after reaching the maximum upward excursion; thus the optotype may be presented while the patient is looking through the lower half of the lens which is typically used for reading. Since the optotype is presented at a distance of 4 meters, the subjects' inability to identify the orientation of the optotype may be due to the inappropriate correction rather than a VOR deficiency. The DVA test would be expected to reflect this problem more than the GST since the size of the optotype in the GST is held constant above the subject's SVA. Eight patients were tested while wearing bifocals; the mean score on the DVA loss test for the bifocal wearers in the down direction was 0.16 logMAR vs. 0.13 logMAR in the up direction. On the GST, the mean score of the bifocal wearers was 148 deg/ sec for both up and down directions.

Goebel et al [7] examined the reliability of the GST in the horizontal direction only on an earlier version of the InVision system. The ICC for the GST in the horizontal direction was 0.91. The optotype used by Goebel et al was 0.30 logMAR above the subjects' SVA and the testing distance was 3 meters. In the current study the optotype size was 0.25 logMAR above SVA at a 4-meter distance. Goebel et al [7] used larger optotypes at a shorter testing distance. In our clinical experience, such a seemingly minor change in optotype increment and testing distance could significantly affect the patients' performance on the test, rendering the comparison between the results of the two studies difficult.

Ward et al. [36] calculated the MDC for the GST and DVA loss test using a sample of healthy subjects. The MDC for the GST was 51 – 55 deg/ sec and for the DVA loss test 0.14 – 0.24 logMAR. The MDC values for the GST in the current study are larger than those of Ward et al. The difference is mainly due to the larger ICC value for the horizontal GST in the sample of healthy subjects (ICC = 0.69). A possible explanation for the larger MDC value in patients is that there was greater variability in the patients' performance.

The correlation between the GST and the SVA was examined to investigate whether higher SVA scores were correlated with better/ faster GST scores. Higher SVA scores mean that larger optotypes would be presented during the GST. We wanted to confirm that the larger optotypes did not make the GST easier for subjects with poor vision simply due to the size of the optotype used. There were no significant correlations between the GST and the SVA suggesting that the larger optotypes used did not affect the findings.

Several studies have examined the correlation between patients' symptoms during testing, specifically visual blurring, and their GST or DVA scores [2, 14, 15, 29]. Our results were consistent with previous findings; there were no significant correlations between patients' symptoms and their functional VOR test results. The only study that found a significant correlation between DVA (specifically in the up direction) and visual blurring used a passive DVA test and had more stringent inclusion criteria (>50% asymmetry on the caloric test compared to >25% asymmetry in other studies) [2]. The lack of significant correlation between the GST and DVA test and the patients' symptoms may be because symptom provocation and the GST/DVA tests may measure different constructs. In addition, clinicians generally referred persons with vestibular disorders whom they thought could tolerate the GST and DVA testing protocol. If persons with greater impairment had been included, it may have resulted in a stronger relationship between the functional tests and magnitude of symptoms.

The algorithm used for the GST appears to need further refinement. While most available studies in the literature used an algorithm where the optotype was presented five times at each head velocity and testing advanced in increments of 10 deg/ sec [7, 27, 37], the InVision system used in the current study utilized the best PEST adaptive algorithm to control the progress of the test and calculate final GST scores. The earlier algorithm is time consuming and may provoke patients' symptoms, especially when considering that testing is usually done for the right, left, up, and down directions. At the same time, while the best PEST algorithm uses fewer trials to calculate patients' scores, its use may not be optimal for the GST but appears to be acceptable for the DVA.

5. Conclusion

Future research will be aimed at improving reliability and stability of DVA testing and of the GST. To accomplish this, the best PEST algorithm may need to be replaced with an algorithm that can be tailored for each patient. The GST and DVA tests provide valuable information about the functional performance of the VOR. If the SVA used for the DVA test

is standardized, the InVision system's DVA test appears to have good reliability. As for the GST test, the current protocol needs to be altered to improve the reliability of the test scores and the ease of comparison across different laboratories. Testing distance, size and presentation time of the optotype, and the testing protocol should all be examined further in an attempt to enhance the reliability of the GST in persons with vestibular dysfunction.

Acknowledgments

Supported in part by NIH grants P30 AG024827 and DC005205

References

1. Badaracco C, Labini FS, Meli A, De Angelis E, Tufarelli D. Vestibular rehabilitation outcomes in chronic vertiginous patients through computerized dynamic visual acuity and Gaze stabilization test. *Otol Neurotol*. 2007; 28:809–813. [PubMed: 17577127]
2. Badaracco C, Labini FS, Meli A, Tufarelli D. Oscillopsia in labyrinthine defective patients: comparison of objective and subjective measures. *Am J Otolaryngol*. 31:399–403. [PubMed: 20015789]
3. Badke MB, Shea TA, Miedaner JA, Grove CR. Outcomes after rehabilitation for adults with balance dysfunction. *Arch Phys Med Rehabil*. 2004; 85:227–233. [PubMed: 14966706]
4. Buizza A, Schmid R. Velocity characteristics of smooth pursuit eye movements to different patterns of target motion. *Exp Brain Res*. 1986; 63:395–401. [PubMed: 3758256]
5. Dannenbaum E, Paquet N, Chilingaryan G, Fung J. Clinical evaluation of dynamic visual acuity in subjects with unilateral vestibular hypofunction. *Otol Neurotol*. 2009; 30:368–372. [PubMed: 19318888]
6. Fleiss, J. Design and analysis of clinical experiments. New York: John Wiley and Sons; 1986.
7. Goebel JA, Tungsiripat N, Sinks B, Carmody J. Gaze stabilization test: a new clinical test of unilateral vestibular dysfunction. *Otol Neurotol*. 2007; 28:68–73. [PubMed: 17106431]
8. Grossman GE, Leigh RJ, Bruce EN, Huebner WP, Lanska DJ. Performance of the human vestibuloocular reflex during locomotion. *Journal of Neurophysiology*. 1989; 62:264–272. [PubMed: 2754477]
9. Hall CD, Herdman SJ. Reliability of clinical measures used to assess patients with peripheral vestibular disorders. *J Neurol Phys Ther*. 2006; 30:74–81. [PubMed: 16796772]
10. Hall CD, Heusel-Gillig L, Tusa RJ, Herdman SJ. Efficacy of gaze stability exercises in older adults with dizziness. *J Neurol Phys Ther*. 34:64–69. [PubMed: 20588090]
11. Hazel CA, Elliott DB. The dependency of logMAR visual acuity measurements on chart design and scoring rule. *Optom Vis Sci*. 2002; 79:788–792. [PubMed: 12512687]
12. Herdman SJ. Role of vestibular adaptation in vestibular rehabilitation. *Otolaryngol Head Neck Surg*. 1998; 119:49–54. [PubMed: 9674514]
13. Herdman, SJ. Vertigo and Imbalance: Clinical Neurophysiology of the Vestibular System. Amsterdam, Elsevier: 2010.
14. Herdman SJ, Hall CD, Schubert MC, Das VE, Tusa RJ. Recovery of dynamic visual acuity in bilateral vestibular hypofunction. *Arch Otolaryngol Head Neck Surg*. 2007; 133:383–389. [PubMed: 17438254]
15. Herdman SJ, Schubert MC, Das VE, Tusa RJ. Recovery of dynamic visual acuity in unilateral vestibular hypofunction. *Arch Otolaryngol Head Neck Surg*. 2003; 129:819–824. [PubMed: 12925338]
16. Herdman SJ, Schubert MC, Tusa RJ. Role of central preprogramming in dynamic visual acuity with vestibular loss. *Arch Otolaryngol Head Neck Surg*. 2001; 127:1205–1210. [PubMed: 11587600]
17. Herdman SJ, Tusa RJ, Blatt P, Suzuki A, Venuto PJ, Roberts D. Computerized dynamic visual acuity test in the assessment of vestibular deficits. *Am J Otol*. 1998; 19:790–796. [PubMed: 9831156]

18. Honaker JA, Shepard NT. Age effect on the Gaze Stabilization test. *J Vestib Res.* 20:357–362. [PubMed: 20826934]
19. Horning E, Gorman S. Vestibular rehabilitation decreases fall risk and improves gaze stability for an older individual with unilateral vestibular hypofunction. *J Geriatr Phys Ther.* 2007; 30:121–127. [PubMed: 18171496]
20. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials.* 1989; 10:407–415. [PubMed: 2691207]
21. Leek MR. Adaptive procedures in psychophysical research. *Percept Psychophys.* 2001; 63:1279–1292. [PubMed: 11800457]
22. Leigh R, Zee D. *The Neurology of Eye Movements.* 2006
23. Longridge NS, Mallinson AI. The dynamic illegible E-test. A technique for assessing the vestibulo-ocular reflex. *Acta Otolaryngol.* 1987; 103:273–279.
24. Longridge NS, Mallinson AI. The dynamic illegible E (DIE) test: a simple technique for assessing the ability of the vestibulo-ocular reflex to overcome vestibular pathology. *J Otolaryngol.* 1987; 16:97–103. [PubMed: 3599163]
25. Pentland A. Maximum likelihood estimation: the best PEST. *Percept Psychophys.* 1980; 28:377–379. [PubMed: 7465322]
26. Portney, LG.; Watkins, MP. *Foundations of Clinical Research. Upper Saddle River, NJ: Applications to Practice, Prentice Hall; 2000.*
27. Pritcher MR, Whitney SL, Marchetti GF, Furman JM. The influence of age and vestibular disorders on gaze stabilization: a pilot study. *Otol Neurotol.* 2008; 29:982–988. [PubMed: 18667938]
28. Scherer M, Migliaccio AA, Schubert MC. Effect of vestibular rehabilitation on passive dynamic visual acuity. *J Vestib Res.* 2008; 18:147–157. [PubMed: 19126985]
29. Schubert MC, Herdman SJ, Tusa RJ. Vertical dynamic visual acuity in normal subjects and patients with vestibular hypofunction. *Otol Neurotol.* 2002; 23:372–377. [PubMed: 11981398]
30. Schubert MC, Migliaccio AA, Clendaniel RA, Allak A, Carey JP. Mechanism of dynamic visual acuity recovery with vestibular rehabilitation. *Archives of Physical Medicine & Rehabilitation.* 2008; 89:500–507. [PubMed: 18295629]
31. Schubert MC, Migliaccio AA, Della Santina CC. Dynamic visual acuity during passive head thrusts in canal planes. *J Assoc Res Otolaryngol.* 2006; 7:329–338. [PubMed: 16810569]
32. Schubert MC, Zee DS. Saccade and vestibular ocular motor adaptation. *Restorative neurology and neuroscience.* 2010; 28:9–18. [PubMed: 20086279]
33. Sharpe JA, Sylvester TO. Effect of aging on horizontal smooth pursuit. *Invest Ophthalmol Vis Sci.* 1978; 17:465–468. [PubMed: 640792]
34. Taylor MM, Creelman CD. PEST: Efficient estimates on probability functions. *J Acoust Soc Am.* 1967; 41:782–787.
35. Viirre E, Tweed D, Milner K, Vilis T. A reexamination of the gain of the vestibuloocular reflex. *J Neurophysiol.* 1986; 56:439–450. [PubMed: 3489820]
36. 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; 20:363–372. [PubMed: 20826935]
37. Whitney SL, Marchetti GF, Pritcher M, Furman JM. Gaze stabilization and gait performance in vestibular dysfunction. *Gait Posture.* 2009; 29:194–198. [PubMed: 18815040]

Figure 1.

The InVision test setup.

The InVision device (NeuroCom International, Inc., Clackamas, OR, USA). A CRT computer monitor is placed inside a cabinet (Center). Four mirrors reflect the image from the screen so that the distance between the subject and the computer screen is 4 meters. The operator views a different screen to control the test (Left monitor). A head mounted sensor (an InterSense Inertia Cube², 3-axis integrating gyro) provides information about the range and the speed of head movement.

Figure 2.

Screen view of the feedback bars on (A) patient screen and (B) operator screen utilized during the gaze stabilization and the dynamic visual acuity tests to provide feedback about head movement. The upper feedback bar reflected the range of the head motion (degrees). The lower feedback bar reflected the velocity of the head movement (degrees/ sec).

Table 1

Mean (standard deviation) of static visual acuity (SVA, logMAR), gaze stabilization test (GST, deg/sec), dynamic visual acuity (DVA, logMAR), and DVA loss (logMAR) for all 3 repetitions of testing in patients with vestibular disease. Repeated measures ANOVA tested for differences among the 3 repetitions.

Test	Direction	Session 1		Session 2	ANOVA
		1 st repetition	2 nd repetition	3 rd repetition	p-value
SVA		-0.03 (0.11)	-0.04 (0.13)	-0.06 (0.14)	0.10
GST	Left	130 (48)	139 (50)	152 (39)	0.33
	Right	136 (38)	143 (50)	143 (42)	0.94
	Up	131 (45)	138 (42)	124 (37)	0.46
	Down	138 (35)	142 (37)	128 (36)	0.61
DVA	Left	0.21 (0.14)	0.17 (0.13)	0.12 (0.13)	0.001
	Right	0.22 (0.14)	0.14 (0.16)	0.12 (0.17)	0.001
	Up	0.12 (0.18)	0.15 (0.17)	0.10 (0.17)	0.22
	Down	0.12 (0.15)	0.12 (0.16)	0.10 (0.15)	0.23
DVA loss	Left	0.24 (0.09)	0.20 (0.12)	0.19 (0.10)	0.05
	Right	0.25 (0.09)	0.18 (0.12)	0.18 (0.10)	0.003
	Up	0.16 (0.12)	0.18 (0.11)	0.17 (0.09)	0.76
	Down	0.16 (0.10)	0.16 (0.11)	0.17 (0.09)	0.85

DVA loss = DVA – SVA.

Table 2

Intraclass correlation coefficient (95% confidence intervals) for the static visual acuity (SVA, logMAR), gaze stabilization test (GST, deg/sec), dynamic visual acuity (DVA, logMAR), and DVA loss (logMAR) within-session (1st and 2nd repetitions) and between-sessions (1st and 3rd repetitions, 2nd and 3rd repetitions).

Test	Direction	Within-session (1 st and 2 nd reps)	Between-session (1 st and 3 rd reps)	Between-session (2 nd and 3 rd reps)
SVA		0.88 (0.74 – 0.94)	0.84 (0.67 – 0.93)	0.78 (0.55–0.89)
GST	Left	0.48 (0.13 – 0.72)	0 (NS)	0.11 (NS)
	Right	0.38 (0.01 – 0.66)	0.25 (NS)	0.33 (NS)
	Up	0.25 (NS)	0.35 (NS)	0.31 (NS)
	Down	0.29 (NS)	0 (NS)	0.28 (NS)
DVA	Left	0.79 (0.59 – 0.90)	0.76 (0.52 – 0.89)	0.78 (0.55–0.90)
	Right	0.65 (0.37 – 0.82)	0.79 (0.57 – 0.91)	0.69 (0.40–0.86)
	Up	0.77 (0.56 – 0.88)	0.64 (0.32 – 0.83)	0.69 (0.39–0.86)
	Down	0.77 (0.56 – 0.89)	0.79 (0.56 – 0.90)	0.83 (0.63–0.92)
DVA loss	Left	0.50 (0.16 – 0.73)	0.10 (NS)	0.60 (0.26–0.81)
	Right	0.22 (NS)	0.50 (0.12 – 0.75)	0.09 (NS)
	Up	0.26 (NS)	0 (NS)	0.03 (NS)
	Down	0.47 (0.13 – 0.72)	0.23 (NS)	0.29 (NS)

NS: ICC value not statistically different from 0 ($\alpha = 0.05$)

DVA loss = DVA – SVA

Table 3

Standard error of the measurement (SEM), SEM% and minimal detectable change (MDC) values for the gaze stabilization test (GST, deg/sec) and the dynamic visual acuity loss (DVA loss, logMAR). MDC was calculated for measures with significant ICC values only

	Within-session (1 st and 2 nd reps)			Between-session (1 st and 3 rd reps)			Between-session (2 nd and 3 rd reps)		
	SEM	SEM%	MDC	SEM	SEM%	MDC	SEM	SEM%	MDC
GST	Left	36	26	100	48	37	47	34	
	Right	39	28	109	36	25	41	29	
	Up	39	34		36	28	35	25	
	Down	31	26		36	28	31	22	
DVA loss	Left	0.08	42	0.24	0.09	50	0.08	38	0.21
	Right	0.11	59		0.07	39	0.11	64	
	Up	0.10	65		0.12	75	0.11	60	
	Down	0.08	50	0.22	0.09	55	0.09	58	

SEM = SD (1 - ICC)

SD = standard deviation

SEM% = SEM/ mean * 100%

MDC = SEM * 1.96 * 2

Gaze stabilization test (GST, deg/sec) and dynamic visual acuity loss (DVA loss, logMAR) scores (means and standard deviations) in the horizontal plane for patients with *unilateral* peripheral lesions (n = 12) toward the lesion vs. non-lesion side.

Table 4

	Session 1				Session 2	
	1 st repetition		2 nd repetition		3 rd repetition	
	Non-lesion side	Lesion side	Non-lesion side	Lesion side	Non-lesion side	Lesion side
GST	136 (42)	129 (29)	124 (47)	116 (50)	152 (51)	151 (39)
DVA loss	0.25 (0.08)	0.25 (0.07)	0.22 (0.14)	0.18 (0.14)	0.18 (0.09)	0.13 (0.12)

Table 5

Pearson product moment correlation (r) between subjects' static visual acuity (SVA, logMAR) and their gaze stabilization test (GST, deg/sec) scores. $\alpha = 0.01$

		Session 1		Session 2
		1 st repetition	2 nd repetition	3 rd repetition
SVA and GST	Left	r	0.17	-0.15
		p-value	0.39	0.50
		n	28	24
	Right	r	-0.06	0.34
		p-value	0.78	0.10
		n	28	24
	Up	r	-0.003	-0.38
		p-value	0.99	0.08
		n	27	22
	Down	r	-0.06	0.03
		p-value	0.77	0.89
		n	28	23

Table 6

Pearson product moment correlation between each gaze stabilization test (GST, deg/ sec) score and the dynamic visual acuity loss (DVA loss, logMAR) score in the same direction for the three repetitions of testing. $\alpha = 0.01$

		Session 1		Session 2
		1 st repetition	2 nd repetition	3 rd repetition
GST and DVA loss	Left	r	-.30	-.22
		p-value	.12	.006
		n	28	23
	Right	r	.10	-.61
		p-value	.60	.001
		n	28	23
	Up	r	-.38	-.72
		p-value	.05	.001
		n	27	21
	Down	r	-.28	-.38
		p-value	.15	.05
		n	28	22