Published in final edited form as:

Optom Vis Sci. 2011 October ; 88(10): 1153–1163. doi:10.1097/OPX.0b013e3182271638.

Grating Acuity and Contrast Tests for Clinical Trials of Severe Vision Loss

Ava K. Bittner^{*}, Pamela Jeter[†], and Gislin Dagnelie[‡]

Lions Vision Center, Wilmer Eye Institute, Johns Hopkins University, Baltimore, Maryland

Abstract

Purpose—To evaluate the reliability and validity of grating visual acuity (VA) and contrast sensitivity (CS) tests, which could be useful outcome measures to assess changes in severely reduced vision.

Methods—The Grating Acuity Test (GAT) and Grating Contrast Sensitivity (GCS) tests, which involve the detection of grating orientation in a four-Alternative Forced Choice paradigm on a liquid crystal display screen, were compared with the well-validated Early Treatment of Diabetic Retinopathy Study (ETDRS) and Pelli-Robson (PR) charts. Grating tests were repeated two or three times within-visit, across three or four sessions, in 20 legally blind subjects: 8 with retinitis pigmentosa (RP) (16 eyes) and 12 with other retinal diseases (OR) (16 eyes).

Results—VA determined by ETDRS and GAT was in good agreement and scaled very similarly, as shown by regression of the within-session difference between the two measures against their mean [RP group: slope (m) = 0.11; 95% confidence interval [CI]: −0.06, 0.29; p = 0.21; OR group: m = −0.07; 95% CI: 0.33, 0.20; p = 0.62]. On average, higher logCS levels were obtained using the GCS than the PR in both groups. The two CS measures scaled similarly in the RP group (m = 0.07; 95% CI: −0.09, 0.22; p = 0.39) but not in the OR group (m = 0.41; 95% CI: 0.12, 0.70; p = 0.005). The within- and between-visit 95% coefficient of repeatability (CR_{95}) were 0.11 to 0.17 log units for the ETDRS charts and GAT in both groups and 0.14 to 0.15 log units for the PR and GCS in the RP group, whereas the OR group demonstrated more variability in CS. Between-visit CR_{95} did not significantly change with mean VA or CS for the ETDRS, PR, or GCS tests, but RP patients' CR_{95} on the GAT increased significantly with decreasing VA. Floor effects occurred for some RP eyes with ETDRS and PR charts but not with the GAT and GCS.

Conclusions—Computer-driven grating tests appear to be reliable, capable of evaluating vision that may fall outside of the range of standard clinical tests and may be useful during clinical trials for advanced eye disease.

Keywords

visual acuity; contrast sensitivity; grating; reliability; variability; low vision

Clinical trials to arrest, slow, prevent, or reverse blinding retinal degenerative diseases are on the horizon and are already increasing in number. Given the possible high-risk nature of some therapies, legally blind subjects with advanced eye disease are likely to be the

Copyright © 2011 American Academy of Optometry

Ava K. Bittner, Lions Vision Center, Wilmer Eye Institute, Johns Hopkins University, 550 N. Broadway, 6th Floor, Baltimore, Maryland 21205-2020, abittne1@jhmi.edu.

^{*}OD, PhD, FFAO

[†]PhD

[‡]PhD, FFAO

participants in early phases of such trials; thus, these trials require reliable and well-validated measures of visual function that are particularly well suited to quantify low levels of vision.¹⁻³ Furthermore, evaluating remaining functional vision such as visual acuity (VA) and contrast sensitivity (CS) in persons with severe vision loss is central for determining rehabilitative strategies. The ability of a person with severe vision impairment to resolve details in daily living (e.g., reading) is assessed by measures of VA, whereas the ability of a person to detect objects relative to their background (e.g., walking down steps) is measured by CS.⁴ Accurately evaluating visual function in these subjects requires outcome measures of VA and CS that can accommodate subjects with vision at or beyond the limits of charts established for use in clinical practice and previous clinical trials.⁵ The Early Treatment of Diabetic Retinopathy Study (ETDRS) chart used to measure VA and the Pelli-Robson (PR) chart used to measure CS are both standardized and validated in clinical practice, yet are suitable for monitoring VA and CS changes only in patients with VA >20/1600 (at 0.5 m) and >20/700 (at 1 m), respectively. Beyond this range, clinicians typically default to estimates such as “counting fingers” and “hand motion”; outcome measures in clinical trials require quantification, however.

The range of functional deficits that accompany ocular disease may introduce other limitations for the use of letter charts. For patients with severely compromised visual fields or patients with large central or paracentral scotomas, the letters to be resolved might fall outside of the field of view or preferred retinal locus. For patients with poor acuity, resolving letters may be subject to the influence of nearby letters or high-contrast edges, and eye movements, collectively known as the “crowding effect,”⁶ and this effect is greater when letters are viewed in the periphery.⁷ The PR chart uses a single optotype size that may be at or beyond the patient’s level of VA in cases of severe vision loss, whereas it is possible for computer-based programs evaluating CS to adjust the test target size so that it is well below the patient’s threshold.

Letter discrimination is dependent upon recognition of several features of a complex optotype, whereas an alternative to quantify vision beyond the range of letter charts that only requires the recognition of a single feature is a grating-based test to assess very low levels of VA.⁸ Discrimination of very large letter optotypes at a close test distance (e.g., at 50 cm) requires patients to effectively scan across a large area, which may be a difficult task for those with severely constricted visual fields or large central scotomas, and therefore gratings may be a better alternative. For patients with cataracts or other opacities, the use of gratings has been demonstrated to provide a better estimate of resolution.⁹ VA measured with gratings is less affected by optical defocus, which might be found in patients with various ocular diseases.¹⁰⁻¹² Commercially available grating acuity tests, in particular the Teller acuity charts, have been extensively used among pediatric populations to quantify and monitor vision¹³ and as a vision screening test for cognitively impaired elderly in nursing homes.¹⁴ However, the use of such non-automated tests in clinical trials for low vision may be limited by interexaminer variability that could potentially influence acuity measures. The Berkeley Rudimentary Vision Test¹⁵ is a set of test cards that have been developed for subjects with more severe vision [using a single tumbling E up to 2.6 logMAR (20/8000) or gratings up to 2.9 logMAR (20/16,000) at 25 cm]. However, these tests depend on operator skill to vary the test stimuli according to patients’ responses (e.g., changing from optotype to gratings, changing or maintaining test distance, randomizing stimulus orientation order, and manual record keeping), and therefore do not lend themselves to standards of objectivity and rigorous statistics required in clinical research. Computer-based high-contrast square-wave grating tests presented on a liquid crystal display (LCD) monitor have been previously³ and are currently being used¹² to quantify and monitor VA in clinical trials of retinal prostheses in patients with advanced retinitis pigmentosa (RP),² and a grating CS test displayed on a

video monitor has also been used to test a retinal degeneration patient¹; however, these types of grating tests have not been independently validated outside of the clinical trials.

The reliability of another newly developed VA test for patients with severe vision loss has been previously studied. The Freiburg Visual Acuity Test (FrACT) is a computerized VA test that presents Landolt C optotypes on a LCD screen using an adaptive staircase procedure and can quantify VA in the “hand motions” range between 20/2000 and 20/6060.¹⁶ However, the test-retest variability of this test in terms of a 95% confidence interval (CI) has been reported as being more variable than the ETDRS letter charts.¹⁷ The Frankfurt-Freiburg Contrast and Acuity Test is based on the FrACT and may have the potential to monitor CS in those with advanced vision loss, but this test’s capability has not yet been reported or validated. Variability of CS testing with the Frankfurt-Freiburg Contrast and Acuity Test in subjects without retinal disease has been reported to be about 1.5 to 2 times as high as with the PR letter charts.¹⁸ A computerized logMAR VA measurement system (COMPlog) was recently validated by Laidlaw et al.¹⁹ in amblyopic children and a group of ocular disease patients; however, this automated letter test only reaches a 1.68 logMAR (~20/1000) limit at 3 m and has not been validated to measure worse VA at closer test distances. Beyond this range, the program relies on scoring in terms of counting fingers or hand movements.

The Basic Assessment of Light and Motion, a computerized test battery developed to evaluate temporal resolution, light perception, and localization of light and motion, was designed to measure improvements in visual function with a visual prosthesis and does not target VA or CS specifically.²⁰ Furthermore, as VA and CS are complementary measures,⁸ validating tests for both VA and CS concurrently would be beneficial because ocular disease affects both measures to varying degrees. It is well known that both VA and CS are important predictors for the ability of low-vision patients to perform activities of daily living²¹ and mobility,²² thus validating both VA and CS tests together at more extreme ranges may be more clinically relevant. Previously described studies generally focus on validating one or the other. Clearly, a need still exists to establish reliable tests of VA and CS to monitor patients with severe vision loss.

The Grating Acuity Test (GAT) and Grating Contrast Sensitivity (GCS) tests were developed to quantify varying levels of severe vision loss for use during a multicenter clinical trial of the Optobionics’ Artificial Silicon Retina for patients with RP.³ The GCS was designed to assess CS in subjects who were unable to provide meaningful results with the PR standardized test of CS, as a result of severely advanced disease resulting in reduced VA and CS. Our aim was to determine whether the grating tests may be used to follow VA and CS stability, and thus the clinical course of progressive deterioration or improvement (e.g., with effective treatment) of overall vision in legally blind subjects. To accomplish this goal, we assessed validity by comparing results with the new tests to results obtained with established VA and CS tests and the ETDRS and PR letter charts. Reliability was assessed by repeating measurements over time, both within and across sessions separated by 6 to 50 days.

METHODS

The protocol for the study was approved by the Institutional Review Board of the Johns Hopkins University School of Medicine and followed the tenets of the Declaration of Helsinki. All participants provided informed consent.

Participants

Participants were recruited through a database of previous research subjects of the Lions Vision Center and from referrals by the Low Vision Clinic of the Wilmer Eye Institute at Johns Hopkins, located in an urban hospital setting in Baltimore, Maryland. The 20 participants (10 males and 10 females) were from local communities within 1.5 h driving from our center. Roughly one-third of the subjects were black ($n = 6$), nearly three-fourths were white ($n = 13$), and 1 subject was Hispanic. Subjects were aged between 39 and 90 years, with a mean age of 69 years. Subjects were offered a modest remuneration for their participation, in addition to lunch, parking validation, and a limited reimbursement for travel expenses. Test results were obtained from April to October 2006 for the first 18 participants. Because of insufficient data from RP patients with advanced vision loss, data from two additional participants were collected between July and August 2010. The subjects were divided into two groups defined on the basis of the type of eye disease: 8 with RP and 12 with other retinal diseases (OR). The OR group consisted of persons with optic neuropathy ($n = 2$ subjects; 3 eyes), cone-rod dystrophy ($n = 1$ subject; 2 eyes), retinal vein occlusion ($n = 1$ subject; 1 eye), glaucoma ($n = 1$ subject; 2 eyes), diabetic retinopathy ($n = 1$ subject; 1 eye), and age-related macular degeneration ($n = 5$ subjects; 7 eyes). Sixteen RP eyes met the study criteria for legal blindness on the basis of visual field diameter $<20^\circ$ as determined by Goldmann and/or Humphrey visual field tests with the III4e target. Sixteen OR eyes met the study criteria for legal blindness on the basis of best-corrected VA $<20/200$ as determined with the ETDRS chart (Lighthouse International, New York, NY).

Subjects who were not undergoing treatment or surgery for their eye disease and whose vision was likely to remain stable throughout a 1- to 3-mo study period were enrolled; their visual status was monitored during every visit. Any significant changes in the subjects' visual condition were detected through either the subjective medical and ocular history taken at the start of each return visit or an unanticipated shift in the results of several of the tests. Data for one eye in the OR group were excluded from analysis involving VA because no measures could be acquired; however, data for this eye were included for CS.

Stimulus and Display

Stimuli for the GAT and GCS tests were displayed on a gamma-corrected LCD high-definition television monitor (LNS3251D; Samsung, Tokyo, Japan). Fig. 1 shows the monitor displaying the GAT stimulus in the diagonal right orientation (A) and a dot pattern that is presented between gratings (B). Square wave grating stimuli within a circular area of 37.5 cm diameter were displayed at a test distance of 1, 2, or 4 m for the GAT, subtending 21.2° , 10.7° , or 5.4° , respectively, and either 0.5 or 1 m for the GCS, subtending 41.1° or 21.2° , respectively. The test distance was based on the subject's VA or CS at the initial visit; for consistency, the same test distance was maintained at each follow-up visit. Stimuli were presented in a dimly lit room.

The average screen luminance of the ETDRS chart used in this study was initially calibrated at 80 cd/m^2 , but by the end of the study, it was found to be 50 cd/m^2 , presumably due to aging of the fluorescent light source; the average luminance of the PR chart is determined by room illumination and was always between 100 and 150 cd/m^2 . The average screen luminance for the GAT and GCS was about 115 cd/m^2 for the gratings and about 50 cd/m^2 for the dot pattern. Luminance was reduced from the observer's viewpoint, especially toward the lower edge of the screen where it fell to 50 cd/m^2 . The monitor had a pixel resolution of 1366×768 and a screen size of $39 \times 70 \text{ cm}^2$. Dithering with adjacent gray levels within each 4×4 pixel block within grating elements was used to expand the available number of gray levels for the GCS test from 256 to 2048. Screen luminance was calibrated with a spot meter (Minolta CS-100; Konica Minolta, Ramsey, NJ) for grayscale

levels 0 to 255 to provide the gamma function used for contrast linearization in the GCS presentation. The software was developed by Optobionics Corporation for use in clinical trials evaluating the Artificial Silicon Retina device for RP.²³

Design

The subjects were tested during three or four visits, each of which lasted between 1 to 3 h, depending on whether one eye or both were tested. During these visits, the subjects completed ETDRS letter VA, PR letter CS, and grating VA and CS tests, in that order. The grating tests involved the detection of grating orientation in a four-alternative forced-choice (4-AFC) paradigm on a large LCD screen. Visits were spaced by at least 6 days and not more than 50 days, with a mean $\pm$ standard deviation of 20 ± 9 days between visits. Subjects were given regular breaks between tests, as deemed necessary by the subject or tester, to minimize the effects of fatigue. All subjects were able to maintain attentiveness throughout the test series. Two examiners, initially working together and later individually, administered the test procedures across subjects and visits and ensured that all procedures were executed in exactly the same manner at all visits. On each visit for every subject, the same examination room and equipment were used to ensure that all test conditions were consistent.

Procedure

The procedures for ETDRS VA and PR CS measurement used in this study were identical to those described previously.⁵ All subjects wore appropriate refractive correction during testing. Subjects were asked to guess at the letters on the next line whenever they correctly identified at least one letter on a previous line, until they were no longer able to correctly identify any letters. The grating tests are capable of assessing acuities between 20/32 and 20/4000, vs. the ETDRS charts, which can measure acuity down to 20/1600 at 50 cm. The PR charts can measure CS up to 2.3 logCS, whereas the GCS test can measure CS over a 3 log unit range. At the standard 1 m distance, the PR optotype size corresponds to 20/700, whereas much coarser stimuli can be presented with the GCS. The paradigm involved a 4-AFC orientation identification task in which the grating was oriented horizontally, vertically, or obliquely at 135° (top left) or at 45° (top right). The grating orientations were presented in a pseudorandom order. A tone was presented to indicate the beginning of the grating presentation, whereas two tones indicated the end of each grating presentation. Each grating was presented for 5 s, and the program waited for the subject's response (about 2–3 s per response). Participants responded by pressing one of four textured buttons each representing an orientation on a button box. The button box was intuitively designed for ease of use by elderly visually impaired subjects. Each button was fitted with a protruding oriented bar that facilitated a tactile response. Participants were given a few initial trials to practice using the response box and confirm their understanding of the buttons. Participants were required to respond to each stimulus even if it involved guessing, and they were told that the test could not continue until they provided a response for each grating shown. No feedback was given. Participants were also encouraged to scan and find their area of best possible vision.

Gratings in the GAT were initially presented at a size larger than the expected resolution acuity, as determined by ETDRS VA testing. Gratings in the GCS test were presented at a size four times (0.6 log units) larger than the GAT resolution limit. The GAT and GCS test required about 15 min to evaluate both eyes. During each visit, two VA and CS threshold procedures were performed. In some cases, usually because of appreciable discrepancy between the first two threshold values, a third threshold was obtained. In these cases, the two measures that were most similar were used in the final analysis.

Staircases

Square wave gratings were presented in 0.1 logMAR steps for the GAT or 0.075 logCS steps for the GCS threshold on the LCD screen, using a staircase algorithm: 2 correct = 1 step down or 1 wrong = 1 step up (tracking an accuracy level of 70.7%), until a total of six grating orientations were correctly identified at any level.

Data Analysis

To account for correlations within subject, these repeated measures data were fitted using multilevel, mixed effects linear regression models with clustering by subject and eye. Models assuming random intercepts were explored, and the log likelihoods were approximated by adaptive Gaussian quadrature using STATA/IC version 10.0 (Stata Corp., College Station, TX).

Coefficient of Repeatability—The use of multiple test repetitions per subject in this study allowed for the establishment of an individual mean and 95% coefficient of repeatability (CR_{95}), defined as 1.96 times the standard deviation of the differences in threshold within- or between-sessions using all test-retest combinations: $n \times (n - 1)/2$ values for n test repetitions. These values specify a range around a mean baseline measure obtained from several visits outside which the measure must fall on repeat testing for the change to be regarded as statistically significant, i.e., not attributable to chance. Exceeding this CR_{95} in either direction indicates a high degree of confidence that an individual's visual function has changed between sessions. In this study, CR_{95} measures were established for each subject, as well as a mean for each of the groups defined on the basis of ocular pathology. Participants with VA and CS measures that may have been affected by floor effects on either measure were excluded from the CR_{95} calculations. Floor effects were defined as eyes that saw fewer than five letters on all visits for the ETDRS and fewer than three letters on all visits, or no letters on any visit, for the PR.

PR Acuity Limit—Participants had a wide range of VA levels and therefore the PR letter size at 1 m (20/700) was much larger or smaller than many of the subjects' VA, whereas the GCS was always performed at 0.6 logMAR above the GAT level for each subject. Therefore, we dichotomized measures into two subgroups according to the difference between their mean ETDRS VA and the PR letter size (1.54 logMAR, i.e., 20/700). Subjects' eyes demonstrating a ≤ 0.3 logMAR difference (i.e., with Snellen VA $< 20/350$) were considered to be in the "poorer acuity" range, whereas those with > 0.3 logMAR difference (VA $> 20/350$) were considered as have "better acuity." Data analysis was conducted for each subgroup and designated accordingly in corresponding figures.

Contrast Correction—The PR chart is characterized by optotype contrast values, defined as $K = (H - L)/H$, where H is the high (background) luminance and L is the low (letter) luminance, whereas the GCS is characterized by Michelson contrast values, $M = (H - L)/(H + L)$. To evaluate the CS values on an equal scale, a conversion of GCS (M -based) contrast sensitivities to PR (K -based) contrast sensitivities was performed. One can easily derive the following relationship expression from the definitions of K and M , above:

$$K = 2 \times M / (1 + M)$$

Note the similarity with the relationship between Weber and Michelson contrast values,²⁴ which is not surprising, because Weber contrast is the opposite of optotype contrast. Substituting CS for contrast and taking the 10-base log, we obtain

$$\begin{aligned} \text{Klog} &= \log(1/K) = -\log(K) = -\log(2) - \log((1/M)/(1+1/M)) = -\log(1/(M+1)) - \log(2) \\ &= \log(M+1) - \log(2) = \log(1+10 \wedge M\log) - \log(2) \end{aligned}$$

where the Klog is the log of the optotype CS and Mlog is the log of the Michelson CS. Thus, all logCS values presented for GCS in the Results section are converted values.

RESULTS

Measures were limited in some cases for participants with more severe vision loss using the standard letter charts due to the restricted measuring range of the ETDRS ($\geq 20/1600$ at 50 cm) chart compared with the GAT (20/32 to 20/6400). The range of visual acuities measured were comparable across the two tests for the RP group (ETDRS 0.322 to 2.003 logMAR and GAT 0.398 to 2.204 log-MAR), whereas the ranges obtained for the OR group were narrower for the GAT (ETDRS 0.98 to 1.903 logMAR and GAT 0.398 to 1.35 logMAR). The CS ranges were wider when measured with the GCS test than the PR for both the RP (converted GCS 0.02 to 1.07 logCS, PR -0.10 to 0.60 logCS) and OR groups (converted GCS 0.31 to 1.8 log CS, PR -0.05 to 1.20 logCS).

Fig. 2 demonstrates the relationship between the within-session mean GAT and ETDRS acuities obtained at each visit for each eye for all subjects. The better eye is designated by open symbols and the worse eye with filled symbols. In the RP group, the relationship between the GAT and ETDRS establishes equivalence of the GAT test (Fig. 2A), whereas in the OR group, the relationship is poorer for patients with more severe pathology (Fig. 2B). To compare whether the GAT was in good agreement with the ETDRS, we plotted the difference between measures obtained with the two methods in a given session against the mean of both measures in Fig. 3. The slopes (m) of the regression lines for these Bland-Altman²⁵ plots for RP (Fig. 3A) and OR (Fig. 3B) were not statistically significantly different from zero, when using multilevel, mixed effects models to account for correlations within subjects and eyes (RP: $m = 0.11$; 95% CI: $-0.06, 0.29$; $p = 0.21$; OR: $m = -0.07$; 95% CI: $-0.33, 0.20$; $p = 0.62$).

Fig. 4A, B demonstrates the relationship between the GCS and PR measures for each eye at each visit. We dichotomized the measures into two subgroups according to the difference between their ETDRS VA and the PR letter size (1.54 logMAR or 20/700 Snellen). Subjects with poorer acuity ($VA \leq 20/350$; i.e., not more than 0.3 logMAR better than the PR letter size) are presented with open symbols, whereas subjects with better acuity ($VA > 20/350$, i.e., more than 0.3 logMAR better than the PR letter size) are presented with filled symbols in Fig. 4.

The RP subgroup in the upper acuity range fared better with a stronger relationship between methods ($m = 0.59$, $r = 0.45$, $p = 0.02$, $n = 26$, excluding measures at floor from two subjects with the poorest acuity levels). However, there were more measures ($n = 31$) at floor in the RP group with poorer acuity than the OR group. The OR subgroup with better acuity ($> 20/350$) yielded a better relationship between measures ($m = 0.53$, $r = 0.39$, $p = 0.06$, $n = 24$) than the poorer acuity subgroup ($m = 0.27$, $r = 0.31$, $p = 0.08$, $n = 34$); however, both were marginally significant. On average, higher logCS levels were obtained using the GCS than the PR in both groups. Some subjects with VA close to or worse than 20/700 demonstrated very poor CS when measured with the PR chart but were able to obtain much better CS levels with the GCS. This result was likely obtained because the resolution of the GCS test was set at 0.6 logMAR above their GAT threshold, while the resolution of the PR chart is limited and fixed.

The linear regression slope (m) for the Bland-Altman analysis for differences between the two CS measures was not statistically significantly different from zero for the RP group ($m = 0.07$; 95% CI: $-0.09, 0.22$; $p = 0.39$), when using multilevel, mixed effects models to account for correlations within subjects and eyes. This slope was statistically significantly different from zero for the OR group for subjects with acuity $\leq 20/350$, however ($m = 0.41$; 95% CI: $0.12, 0.70$; $p = 0.005$).

Between- and within-session reliability for the VA and CS tests are shown in the box plots in Fig. 5. The between-visit median $CR_{.95}$ was 0.11 to 0.16 log units for the ETDRS charts and GAT in both groups (Fig. 5A). The between-visit median $CR_{.95}$ for the converted GCS test in the RP group was 0.15 logMAR and 0.14 for the PR chart. For the most part, measures were consistent across all eyes/subjects except for two extreme values in the CS scores. The OR group showed more variability for the CS tests: 0.24 median logCS for PR and 0.34 and 0.41 median logCS for the worse and better acuity subgroup, respectively (Fig. 5A, right). The within-session median $CR_{.95}$ was 0.11 and 0.17 logMAR for the OR and RP GAT, respectively (Fig. 5B). The median converted GCS was 0.13 logCS for the RP Group and 0.13 and 0.15 for the worse and better acuity subgroups, respectively (shown in Fig. 5B, right). GAT measures on average were more consistent than the GCS measures for within- and between-session measurements for all subjects. The OR subgroup that was limited by poorer acuity ($\leq 20/350$) on their PR scores showed more within-session variability in their GCS scores.

The increased variability observed in the GCS between-group $CR_{.95}$ may be due to the inclusion of participants with severe vision loss who experienced floor effects when tested with the PR. To assess this potential effect, paired samples t-tests between $CR_{.95}$ values were conducted for tests with and without participants with values at floor. No significant differences were found for the RP group when including participants with values at floor (9 eyes) on the PR [$t(15) = -1.37$, $p = 0.19$] compared with those without [$t(6) = 2.35$, $p = 0.06$]. Although the observed difference between GCS and PR was greater in the OR group, no significant differences were found when including the one subject with values at floor [$t(15) = 1.75$, $p = 0.10$] vs. without this subject's data [$t(14) = 1.86$, $p = 0.09$]. There was a slight difference in $CR_{.95}$ values observed between the ETDRS and GAT measures for the RP group. A paired samples t-test indicated a significant difference [$t(15) = 2.26$, $p = 0.04$] when including those eyes with floor measures on the ETDRS (three eyes) relative to the between test comparisons without floor values [$t(14) = 0.22$, $p = 0.83$]. No participants were excluded from the VA measures in the OR group. At least for the RP group, it appears that the increased variability in GAT measures may have been influenced in part by including subjects with floor values in the ETDRS.

The between-visit $CR_{.95}$ did not depend on the magnitude of the VA or CS results obtained with the ETDRS, PR, or GCS tests. The multilevel linear regression slopes for the between-visit $CR_{.95}$ as a function of mean VA or CS were not statistically significantly different from zero for RP ETDRS ($m = -0.003$, $p = 0.93$, $n = 13$), as shown in Fig. 6A, left; RP PR ($m = -0.17$, $p = 0.40$, $n = 7$) and RP GCS tests ($m = -0.13$, $p = 0.06$, $n = 16$) in Fig. 6B, left; OR ETDRS ($m = 0.07$, $p = 0.34$, $n = 15$), as shown in Fig. 6A, right; and OR PR ($m = 0.04$, $p = 0.74$, $n = 15$) and OR GCS tests ($m = 0.10$, $p = 0.51$, $n = 16$) in Fig. 6B, right. Fig. 6B, right subdivides the data by acuity limit according to Fig. 4. Filled symbols represent data in the lower acuity range for both tests and open symbols represent the upper acuity range. Graphically, it appears that variations in VA do not contribute to the observed $CR_{.95}$ values. The multilevel linear regression slope was statistically significantly different from zero for the GAT and is shown in the scatterplot of the between-visit $CR_{.95}$ as a function of mean VA for RP GAT ($m = 0.10$, $p = 0.001$, $n = 13$) (Fig. 6A, left). Therefore, it appears that there may be a slight increase in variability for the GAT for RP subjects with severely reduced

mean VA. A statistically significant slope was also found for the GAT in the OR group ($m = -0.25$, $p = 0.01$, $n = 15$) (Fig. 6A, right). However, we also performed a sensitivity analysis in which we excluded the OR subject with the most variability who also had the best VA in the group and found that the slope was no longer statistically significant ($m = -0.19$, 95% CI: 0.004, -0.38 , $p = 0.055$). Because it is difficult to assess the variability according to the level of acuity loss in the OR group due to the limited range of the GAT results we obtained for that group and because the linear regression was no longer statistically significant after removing the single most variable subject, the regression line is not shown in Fig. 6A, right.

DISCUSSION

In this study, we have shown that a software-driven grating VA test measuring up to 2.2 logMAR (20/3200), presented on an LCD screen to legally blind retinal degeneration subjects has very similar scaling and between-sessions reliability when compared with ETDRS VA. On average, the GAT measures higher VA (mean logMAR difference of about -0.5 for GAT – ETDRS) in OR patients, presumably because the remaining (peripheral) visual field locations can more effectively detect the simpler pattern of parallel edges in a grating stimulus than the more complex optotype shapes in the ETDRS. On average, the between-session variability of the GCS test was greater than that for the PR charts; however, we were able to obtain measurable CS results using the GCS test in RP patients with severe vision loss who demonstrated floor effects with the PR charts. These two grating tests are capable of evaluating VA and CS that may fall outside of the range of standard clinical tests and would be easy to implement during clinical trials for advanced eye disease.

The correlation between the GAT and ETDRS VA found in our study was similar to another study in which grating acuity using Teller acuity cards among elderly nursing home residents was reported to be well correlated with recognition optotype acuity.¹⁴ In addition, the Teller acuity cards only test resolution as two alternative forced-choice orientations, whereas the GAT assesses four grating orientations. A lower number of alternatives may lead to better acuity but also to greater variability as a result of correct guessing in letter charts.^{26,27} The benefit of using the GAT was that we used a staircase that tracked a fixed accuracy and an automated termination rule that was not known to test administrator or subject, thus reducing variation or bias in subject and experimenter decision criteria. Subjects performed a 26-AFC task with the ETDRS which may have resulted in greater variability or a systematic offset; however, the scores were on average equivalent, at least in the RP group. The GAT also showed better repeatability than the FrACT for patients with severe vision loss.²⁸

Another advantage of the GAT, such as the COMLog system¹⁹ and the Electronic Visual Acuity,^{29,30} is that they are unbiased, automated systems that do not require a test administrator to determine seeing vs. non-seeing subject responses and record preferential looking. Unlike the COMLog system that spans a range of -0.12 to 1.3 logMAR at a fixed 3 m distance and the Electronic Visual Acuity system that spans -0.22 to 1.6 logMAR,²⁹ the GAT was able to evaluate VA in a group of subjects with more diminished VA ranging between 0.40 and 2.2 logMAR. Although the Laidlaw group reported good repeatability (± 0.12 logMAR) in a group of 70 adults with mixed ocular disease, the group included eight normal adults in the analysis making it difficult to get a clear picture of the COMLog performance in a population with compromised acuity. In addition, the COMLog system presents black letters on a gray screen that may be difficult to distinguish for patients with severely reduced CS, such as in RP.

The results for the OR group were consistent with a previous report which found that Landolt-C optotype acuity was generally lower than grating acuity with board panels among

patients with retinal disease.³¹ This previous study also found that patients with greater levels of acuity loss had a greater disparity between optotype and grating acuities, a finding that was confirmed in the OR subject group in this study. A comparison between near letter acuity and Teller grating acuity among patients with age-related macular degeneration has revealed that all patients tested with VA <20/100 had better acuity with gratings than with letters,³² as in our study.

There are several possible explanations to account for the overall better acuity and contrast measured with the grating tests when compared with the letter recognition tests in RP and OR. This effect may be a result of the larger area available for scanning during the grating tests, which would make the test easier for those with a central scotoma or patchy visual field, because the grating tests may assess, more than ETDRS, simultaneous central and pericentral vision. Most likely, however, it is related to the fact that optotype recognition and/or discrimination requires the detection of multiple features, whereas orientation discrimination only requires the identification of a single feature: any single edge is decisive. Administering letter charts at a closer test distance (e.g., 50 cm) would place the area to be resolved outside of the patient's narrow field of view or preferred retinal locus and thus will require extensive and complex scanning. Therefore, gratings are likely more easily recognized than optotypes.

Although letters may be familiar and easy to detect, the patchy vision due to scotomas in this population may make their multiple features more difficult to recognize. The higher (150 vs. 50 cd/m², i.e., ~0.5 log unit) luminance level of the LCD screen compared with the ETDRS chart (retro-illuminated in a dark room) used in our study may have had a minor effect: the best estimates allowed by Fosse et al.³³ is that for most RP and OR patients, such a difference would entail a resolution increase of GAT relative to ETDRS by 0.05 logMAR. Coincidentally, this corresponds to the offset of the regression line for RP patients in Figs. 2A and 3A, but it hardly reduces the pronounced difference between the two tests observed for OR subjects in Figs. 2B and 3B.

We found that the GCS was less reliable than the PR CS, but only for the OR eyes we tested; RP eyes that completed both tests without floor effect showed equal reliabilities on either test (Fig. 6B, left). The greater variability seen for OR eyes cannot be attributed to either a difference in test luminance (which were similar for the LCD screen and the PR chart) or to inadequate contrast specification of the LCD screen (which was increased to 12 bits through dithering). It is possible that this result was due to theoretical limitations of using LCD screens for vision testing,³⁴ such as luminance inhomogeneity across the screen, but our findings do not stand alone. A previous study³² using LCD-based letter CS tests found poor agreement with the PR chart possibly due to a suboptimal display of low-contrast levels on a LCD screen leading to reduced reliability. On the other hand, Bahrami et al.³⁵ found equal reliability for a screen-based letter CS test that used cathode ray tube (CRT) screens and the PR chart in RP subjects. Another study showed a differential effect on performance due to increased light levels for subjects with advanced retinal disease³³; however, this situation does not apply to our patient population with reduced CS and does not agree with our finding that the reliability for both CS tests did not depend on the magnitude of CS among our subject population. In our RP population, the variability of the GCS in Fig. 5A (right) was greater than of the PR because there were several subjects with very severe vision loss who had floor effects with the PR and were not included when calculating the between-session variability of that test.

During the GCS test, OR subjects showed much more variability between-sessions than the RP subjects. This variability could potentially be due to changes in eccentric fixation with a central scotoma during the test or to a loss of concentration or fatigue in this elderly

population, although the variability of the GCS test (between-sessions $CR_{.95}$) did not increase statistically significantly with age across subjects. The duration of the test or familiarity of gratings vs. letters was not a likely factor contributing to increased variability as the GAT reliability in the OR group was similar to or even slightly less than the ETDRS charts. Functional visual deficits differed across subjects in the OR group due to the varied eye diseases and could have also contributed to the between-subject variability observed. For some of the OR subjects, there was a large difference in vision between eyes; therefore, they were most likely not accustomed to using the vision in the worse eye and may not have developed a consistent preferred retinal locus. Some limitations of using the PR chart were that measures were taken at 1 m, the contrast step size (0.15 log units) is bigger than the GCS step size (0.075 log units), and letters are presented in groups of three, where subjects almost always stop responding between groups leading to reduced variability. Furthermore, the duration of the stimulus presentation was not matched between the two methods. Participants viewed the GCS stimuli for 5 s while standard procedures were used for the PR test, in which they were given up to 10 s to respond to each letter. Another possible explanation could be a disparate loss of CS for high spatial frequencies consistent with loss of VA relative to low or middle spatial frequencies.^{1,36}

It should be emphasized that the $CR_{.95}$ should be established for each individual participating in an experimental trial by recording multiple baseline measures. Our results show that published average values, even in similar populations, are likely to be inadequate, because substantial interindividual differences in variability may occur. In our opinion, any set value for a clinically significant treatment response in terms of VA change is likely to be too high for some patients and too low for others; the same is true for the magnitude of CS change, although to our knowledge, no clinical trial has adopted such a measure for legally blind subjects. Therefore, we recommend that future clinical trials should establish individual $CR_{.95}$ levels as demonstrated here, to allow detection of significant change.

CONCLUSIONS

This article presents grating VA and CS tests that may be potentially useful for monitoring outcomes during future clinical trials, in addition to other tests used to evaluate changes in visual field areas or other aspects of vision. These grating tests may be especially useful for monitoring changes in visual resolution in those with severe impairment beyond the limits of the standard letter eye charts. We have demonstrated the initial reliability of the grating tests, which scaled similarly to the letter charts. The software and equipment for the GAT and GCS tests would be easily adoptable by other centers for future testing.

Acknowledgments

We would like to acknowledge the contributions of Jacek Kotowski, MD, Alan Y. K. Chow, MD, and Ron Schuchard, PhD, to the development of the grating tests and to an early version of this manuscript. We also thank research assistant Karthik Hiremath for his contributions during patient testing and data analysis.

This work was supported in part by a center grant module from the Foundation Fighting Blindness (to GD).

References

1. Genead MA, Fishman GA, McAnany JJ. Efficacy of topical dorzolamide for treatment of cystic macular lesions in a patient with enhanced S-cone syndrome. *Doc Ophthalmol*. 2010; 121:231–40. [PubMed: 20842519]
2. Caspi A, Dorn JD, McClure KH, Humayun MS, Greenberg RJ, McMahon MJ. Feasibility study of a retinal prosthesis: spatial vision with a 16-electrode implant. *Arch Ophthalmol*. 2009; 127:398–401. [PubMed: 19365014]

3. Chow AY, Bittner AK, Pardue MT. The artificial silicon retina in retinitis pigmentosa patients (an American Ophthalmological Association thesis). *Trans Am Ophthalmol Soc.* 2010; 108:120–54. [PubMed: 21212852]
4. Geruschat, DR.; Smith, AR. Low vision and mobility. In: Blasch, B.; Wiener, W.; Welsh, R., editors. *Foundations of Orientation and Mobility*. 2. New York, NY: American Foundation for the Blind Press; 1997. p. 60-103.
5. Kiser AK, Mladenovich D, Eshraghi F, Bourdeau D, Dagnelie G. Reliability and consistency of visual acuity and contrast sensitivity measures in advanced eye disease. *Optom Vis Sci.* 2005; 82:946–54. [PubMed: 16317369]
6. Flom M. Contour interaction and the crowding effect. *Problems Optometry.* 1991; 3:237–57.
7. Jacobs RJ. Visual resolution and contour interaction in the fovea and periphery. *Vision Res.* 1979; 19:1187–95. [PubMed: 550578]
8. Norton, TT.; Corliss, DA.; Bailey, JE. *The Psychophysical Measurement of Visual Function*. Boston, MA: Butterworth-Heinemann Medical; 2002.
9. Goldmann H. Examination of the fundus of the cataractous eye. *Am J Ophthalmol.* 1972; 73:309–20. [PubMed: 5013245]
10. Thorn F, Schwartz F. Effects of dioptric blur on Snellen and grating acuity. *Optom Vis Sci.* 1990; 67:3–7. [PubMed: 2308749]
11. Herse PR, Bedell HE. Contrast sensitivity for letter and grating targets under various stimulus conditions. *Optom Vis Sci.* 1989; 66:774–81. [PubMed: 2616138]
12. Humayun MS, da Cruz L, Dagnelie G, Mohand-Said S, Stanga P, Agrawal RN, Greenberg RJ. the Argus II Study Group. Interim performance results from the Second Sight ArgusTM II Retinal Prosthesis Study. *Invest Ophthalmol Vis Sci.* 2010; 51:E-Abstract 2022.
13. Quinn GE, Berlin JA, James M. The Teller acuity card procedure. Three testers in a clinical setting. *Ophthalmology.* 1993; 100:488–94. [PubMed: 8479705]
14. Friedman DS, Munoz B, Massof RW, Bandeen-Roche K, West SK. Grating visual acuity using the preferential-looking method in elderly nursing home residents. *Invest Ophthalmol Vis Sci.* 2002; 43:2572–8. [PubMed: 12147587]
15. Bailey IL, Jackson J, Minto H, Greer RB. Visual acuity measurement beyond the limits of the letter chart. *Invest Ophthalmol Vis Sci.* 2009; 50:E-Abstract 4724.
16. Schulze-Bonsel K, Feltgen N, Burau H, Hansen L, Bach M. Visual acuities “hand motion” and “counting fingers” can be quantified with the Freiburg visual acuity test. *Invest Ophthalmol Vis Sci.* 2006; 47:1236–40. [PubMed: 16505064]
17. Bach M. The Freiburg Visual Acuity Test-variability unchanged by post-hoc re-analysis. *Graefes Arch Clin Exp Ophthalmol.* 2007; 245:965–71. [PubMed: 17219125]
18. Bühren J, Terzi E, Bach M, Wesemann W, Kohnen T. Measuring contrast sensitivity under different lighting conditions: comparison of three tests. *Optom Vis Sci.* 2006; 83:290–8. [PubMed: 16699441]
19. Laidlaw DA, Tailor V, Shah N, Atamian S, Harcourt C. Validation of a computerised logMAR visual acuity measurement system (COMPlog): comparison with ETDRS and the electronic ETDRS testing algorithm in adults and amblyopic children. *Br J Ophthalmol.* 2008; 92:241–4. [PubMed: 17993577]
20. Bach M, Wilke M, Wilhelm B, Zrenner E, Wilke R. Basic quantitative assessment of visual performance in patients with very low vision. *Invest Ophthalmol Vis Sci.* 2010; 51:1255–60. [PubMed: 19850846]
21. Richman J, Lorenzana LL, Lankaranian D, Dugar J, Mayer JR, Wizov SS, Spaeth GL. Relationships in glaucoma patients between standard vision tests, quality of life, and ability to perform daily activities. *Ophthalmic Epidemiol.* 2010; 17:144–51. [PubMed: 20455843]
22. Lord SR, Clark RD, Webster IW. Visual acuity and contrast sensitivity in relation to falls in an elderly population. *Age Ageing.* 1991; 20:175–81. [PubMed: 1853790]
23. Chow AY, Chow VY, Packo KH, Pollack JS, Peyman GA, Schuchard R. The artificial silicon retina microchip for the treatment of vision loss from retinitis pigmentosa. *Arch Ophthalmol.* 2004; 122:460–9. [PubMed: 15078662]

24. Strasburger H. Unfocused spatial attention underlies the crowding effect in indirect form vision. *J Vis.* 2005; 5:1024–37. [PubMed: 16441200]
25. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986; 1:307–10. [PubMed: 2868172]
26. Carkeet A. Modeling logMAR visual acuity scores: effects of termination rules and alternative forced-choice options. *Optom Vis Sci.* 2001; 78:529–38. [PubMed: 11503943]
27. Shah N, Dakin SC, Redmond T, Anderson RS. Vanishing optotype acuity: repeatability and effect of the number of alternatives. *Ophthalm Physiol Opt.* 2011; 31:17–22.
28. Lange C, Feltgen N, Junker B, Schulze-Bonsel K, Bach M. Resolving the clinical acuity categories “hand motion” and “counting fingers” using the Freiburg Visual Acuity Test (FrACT). *Graefes Arch Clin Exp Ophthalmol.* 2009; 247:137–42. [PubMed: 18766368]
29. Beck RW, Moke PS, Turpin AH, Ferris FL 3rd, SanGiovanni JP, Johnson CA, Birch EE, Chandler DL, Cox TA, Blair RC, Kraker RT. A computerized method of visual acuity testing: adaptation of the early treatment of diabetic retinopathy study testing protocol. *Am J Ophthalmol.* 2003; 135:194–205. [PubMed: 12566024]
30. Cotter SA, Chu RH, Chandler DL, Beck RW, Holmes JM, Rice ML, Hertle RW, Birch EE, Moke PS. Reliability of the electronic early treatment diabetic retinopathy study testing protocol in children 7 to <13 years old. *Am J Ophthalmol.* 2003; 136:655–61. [PubMed: 14516805]
31. Stiers P, Vanderkelen R, Vandenbussche E. Optotype and grating visual acuity in patients with ocular and cerebral visual impairment. *Invest Ophthalmol Vis Sci.* 2004; 45:4333–9. [PubMed: 15557440]
32. Thayaparan K, Crossland MD, Rubin GS. Clinical assessment of two new contrast sensitivity charts. *Br J Ophthalmol.* 2007; 91:749–52. [PubMed: 17166891]
33. Fosse P, Valberg A, Arnljot HM. Retinal illuminance and the dissociation of letter and grating acuity in age-related macular degeneration. *Optom Vis Sci.* 2001; 78:162–8. [PubMed: 11327678]
34. Strasburger H, Wustenberg T, Jancke L. Calibrated LCD/TFT stimulus presentation for visual psychophysics in fMRI. *J Neurosci Methods.* 2002; 121:103–10. [PubMed: 12393166]
35. Bahrami H, Melia M, Dagnelie G. Lutein supplementation in retinitis pigmentosa: PC-based vision assessment in a randomized double-masked placebo-controlled clinical trial [NCT00029289]. *BMC Ophthalmol.* 2006; 6:23. [PubMed: 16759390]
36. Alexander KR, Barnes CS, Fishman GA. Characteristics of contrast processing deficits in X-linked retinoschisis. *Vision Res.* 2005; 45:2095–107. [PubMed: 15845241]

FIGURE 1.
Photographs of the LCD monitor displaying a GAT stimulus in the diagonal right orientation (A) and the dot pattern that is presented between gratings (B).

FIGURE 2.

Scatter plot demonstrating the relationship between VA with the ETDRS charts and the GAT for the RP (A) and OR (B) groups. Each subject's ETDRS acuity (y-axis) was plotted against his/her within-session mean GAT acuity (x-axis) at each visit for each eye and each group. Snellen equivalents are shown on top. The lines represent the least-squares bivariate regressions of the two measures by subject group. The better eye is designated by open symbols, the worse eye with filled symbols, with a unique symbol for each subject.

FIGURE 3.

Bland-Altman plots of the difference between the two VA tests (grating test minus letter chart) vs. the mean of both tests for the RP (A) and OR (B) groups.

FIGURE 4.

(A, B) Scatter plot demonstrating the relationship between CS obtained with the PR charts and the GCS test. Each subject's PR acuity was plotted against his/her within-session mean GCS at each visit for each eye and each group. The GCS values presented are a Weber conversion to match the contrast values of the PR. Measures are dichotomized into two subgroups according to the difference between the individual's mean ETDRS VA and the PR letter size (1.54 logMAR or 20/700). Subjects demonstrating ≤ 0.3 (20/350) difference, i.e., poorer acuity limits, are presented with open symbols, whereas subjects demonstrating better acuity are presented with filled symbols. The lines represent the least-squares bivariate regressions of the two measures. No line was fit for the RP group in the poorer acuity range (designated by open squares) because of the small number of data points in this category ($n = 2$). Triangles represent measures obtained at floor for the both groups and were not included in the regression fit.

FIGURE 5.

Box plots of the 95% coefficients of repeatability (CR_{95}) between test sessions (A) and within test sessions (B) for each test in each group. The bottom and top of the box are the 25th and 75th percentile (the lower and upper quartiles, respectively), and the band near the middle of the box is the 50th percentile (the median). The ends of the whiskers represent the lowest datum within 1.5 times the interquartile range of the lower quartile, and the highest datum still within 1.5 times the interquartile range of the upper quartile. Any data not included between the whiskers are plotted as an outlier indicated by a plus sign. The OR group GCS data were subdivided according to their acuity limits as either ≤ 0.3 or > 0.3 (equivalent to 20/350 per Fig. 4). Subjects with measures at floor were excluded from the analysis.

FIGURE 6.

Scatter plots indicating the relationship between the 95% coefficient of repeatability (CR_{95}) and mean VA measured with the ETDRS charts and GAT for each eye tested for each group (A), including a regression line for the GAT in the RP subjects indicating a statistically significant relationship. The scatter plots indicating the relationship between CR_{95} and mean CS measured with the PR charts and GCS test for each eye tested and each group are presented on the bottom (B). Triangles represent measures obtained for one method (e.g., GAT) but excluded from the plot in the alternate method (e.g., ETDRS) due to floor effects. The OR group GCS values (6B, right) are subdivided according to their acuity limit as either ≤ 0.3 (filled symbols) or > 0.3 (open symbols).