Effect of Scoring and Termination Rules on Test–Retest Variability of a Novel High-Pass Letter Acuity Chart

Link to publication record in Ulster University Research Portal

Effect of Scoring and Termination Rules on Test–Retest Variability of a Novel High-Pass Letter Acuity Chart

Nilpa Shah,[1] Steven C. Dakin,[1] Heather L. Whitaker,[1] and Roger S. Anderson[1,2]

[1]NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust & UCL Institute of Ophthalmology, London, United Kingdom
[2]Vision Science Research Group, University of Ulster, Coleraine, Northern Ireland, United Kingdom

Correspondence: Nilpa Shah, UCL Institute of Ophthalmology, 11-43 Bath Street, London EC1V 9EL, UK; nilpa.shah@ucl.ac.uk.

PURPOSE. Test–retest variability (TRV) limits our ability to detect clinically significant changes in visual acuity (VA). We wanted to compare the effect of scoring and termination rules on TRV for logMAR charts, employing either conventional or pseudo high-pass (Vanishing Optotype) letters.

METHODS. VA measurements and TRV were compared in 50 uncorrected normal observers (17 male, mean age 42.8 ± 16.2 years) using both conventional logMAR-style charts and letter charts of the same layout but containing pseudo high-pass letters (Moorfields Acuity Chart [MAC]). Additional charts employing a different 10-letter alphabet to the Sloan set were also tested. Mean spherical refractive error was −0.93 diopters (D; range, −5.38 to +3.00 D). Acuity scores were calculated using three methods: letter-by-letter, with either line- or chart-based termination, and line-by-line scoring. Bland–Altman methods were used to calculate 95% ranges for TRV.

RESULTS. While acuity thresholds were higher for the MAC, they were less affected by termination criteria and displayed significantly lower 95% TRV values across all scoring techniques. Ordinary least squares regression analysis confirmed a proportional as well as systematic bias between conventional and MAC measurements ($r^2 = 0.217$, $P = 0.001$) such that the difference between the two was greater with better VA.

CONCLUSIONS. TRV was consistently lower for a logMAR chart employing high-pass rather than conventional letters in uncorrected refractive error and was less affected by termination and scoring methods. The MAC was also less affected by optical defocus. Further work is required to determine the usefulness of different charts to differentiate between optical and neural losses of vision.

Keywords: visual acuity, visual acuity charts, test–retest variability, vanishing optotypes, high-pass letters

Measurement of visual acuity is the most widely used test in ophthalmic care, and the single most important way of characterizing visual function when detecting or monitoring refractive error or ocular disease. In addition, it is the standard outcome measure of treatment efficacy in many clinical trials.[1–5]

A long-standing weakness of conventional measures of visual acuity is their test–retest variability (TRV). TRV quantifies the variability in test results when the test is repeatedly administered to the same individual, even in the absence of true clinical change. In order that a change in clinical status should not be masked by measurement error, tests of visual acuity must not only be valid but also *reliable*, the latter being quantified by TRV. In clinical measurements, this is commonly expressed as a Bland–Altman 95% limit of agreement.

The Snellen chart has high (i.e., poor) TRV with values ranging from ±5 to 16.5 letters in normal subjects[6] and up to 0.33 logMAR in a group of subjects with cataracts, pseudophakia, or early-stage glaucoma.[7] TRV is known to be further influenced by factors such as the chart design,[6,8,9] optical defocus,[10,11] the ocular status of the observer,[12,13] intra-examiner variability,[14] the number of alternative letter choices given to the observer,[15,16] and the scoring and test termination rules employed.[16,17] The large TRV associated with the Snellen chart has often been attributed to the line scoring method employed, in which the visual acuity score is the smallest row in which a specified proportion of letters are read correctly. LogMAR charts such as the ETDRS and Bailey–Lovie charts, were designed to overcome some of the limitations of Snellen charts and are now employed as the "gold standard." Vanden Bosch and Wall[18] showed that a single-letter scoring algorithm, in which a score is given to each letter read, reduces TRV. Using this method, Laidlaw et al.[19] demonstrated a reduction in TRV from 0.20 to 0.14 logMAR; Bailey and Lovie,[9] a change from 2 to 1 logMAR lines; and Arditi and Cagenello,[17] a change from 0.13 to 0.09 logMAR using the ETDRS chart. Applying this technique to Snellen charts is complicated by virtue of the unequal number of letters per line and the unequal jumps between lines, but a score for each letter can still be calculated.

Another factor potentially affecting acuity TRV is the termination criterion. Carkeet[16] combined exact calculation and Monte Carlo simulation to investigate the effect of termination rules on TRV. He found that Arditi and Cagenello's[17] recommendation that subjects attempt all symbols on a

chart should, in theory, give the lowest variance only if there is an infinite number of alternative letter choices. Carkeet[16] investigated the effect of using termination rules in which a set number of letters are read incorrectly on a line (a *line-by-line termination rule*) and the effect of using a set number of mistakes across the whole chart (a *letter-by-letter termination rule*). He suggested that, for the ETDRS chart and single-letter scoring, termination rules of four mistakes or more on a line should be used.

Although logMAR visual acuity charts have corrected many of the deficiencies ascribed to the Snellen chart, TRV nevertheless remains high using these charts. Even with letter-by-letter scoring, values of up to 2 logMAR lines have been reported,[8,17,18,20-26] and this value increases in the presence of optical defocus[10] or retinal disease.[13]

The 10 Sloan letters that are used in the ETDRS chart are chosen in combinations so as to give each line the same average difficulty.[27] However, if a test chart's within-line legibility differences are greater than its between-line legibility differences, the test may display higher variability.[28] While by-letter scoring may help to mitigate this, our previous study,[29] which examined the individual recognition thresholds of letters, suggested that it may be possible to select 10 different letters that have much more individually similar recognition thresholds that could further reduce TRV.

Another potential way to reduce TRV is to employ a different letter design. "Vanishing Optotypes," first described by Howland et al.,[30] have a pseudo high-pass construction consisting of a black "core" with white edges (or vice versa) with the letter being presented on a gray background. Letters are constructed so that their mean-luminance is equal to the background, resulting in thresholds for detection and recognition being almost identical (under foveal viewing), and to the illusion that the letters have "vanished" soon after their discrimination threshold has been reached. One possible advantage of using such letters to measure visual acuity is that previous studies have shown that conventional letters differ significantly in their low spatial frequency content.[31-33] If the low frequencies are removed, then letters may become more similar, more closely equating their legibility and so reducing the TRV. Initial work by our group[15,29] supports this notion and has demonstrated that the effect of the number of alternatives available and the actual letter identity on the visual acuity score and variability is greatly reduced for high-pass letters compared to conventional letters.

Here we investigated the effects of different test termination criteria and scoring methods on estimates of visual acuity and associated TRV, using charts composed of either conventional or high-pass letters in normal subjects with differing levels of uncorrected refractive error. Additionally, we compared the performance (in terms of threshold and TRV) of charts using the set of 10 Sloan letters to identically constructed charts employing an alternative alphabet set chosen from a previous study[29] to contain no "circular" letters (which were found to behave as a separate subset) and be more equally legible.

This study should allow us to determine the vulnerability of the different chart designs to refractive losses of vision, thus permitting future comparison with performance in neural visual loss, in order to better adopt appropriate tests to discriminate neural and optical losses of vision.

METHODS

Ethical approval for this study was obtained from the UCL Research Ethics Committee and all procedures adhered to the tenets of the Declaration of Helsinki.

Eight different charts were constructed. The first two, termed Conventional Sloan letter set 1 and 2 (CS1 and CS2), were almost identical in format to ETDRS charts 1 and 2; that is, they used the 10 Sloan letters (C D H K N O R S V Z) of conventional 5 × 5 matrix black-on-white letter design. Letters on a line were spaced a letter width apart and each line was spaced a letter height from the line above. The third and fourth charts, termed the Moorfields Acuity Chart Sloan letter set 1 and 2 (MAC S1 and MAC S2) were of identical construction except that they employed a high-pass letter design. The visual acuity range on the charts (based on stroke width) was 1.20 to −0.20 logMAR. The letter sizes of all charts commenced at 1.2 logMAR instead of 1.0. The range was extended at the poorer acuity end in order to avoid having to use different test distances between charts, since our previous studies had demonstrated overall "poorer" logMAR visual acuities with high-pass compared to conventional letters.

The fifth and sixth charts had the same design as CS1 and CS2 but used a different 10-letter alphabet (B E H K N P R S X Z) whose legibility was more similar in conventional form.[29] These charts were referred to as Conventional New letter set 1 and 2 (CN1 and CN2). The seventh and eighth charts, the Moorfields Acuity Chart New letter set 1 and 2 (MAC N1 and MAC N2) used this same alphabet in high-pass form. The appearance of charts CS1, MAC S1, CN1, and MAC N1 is shown in Figure 1.

Fifty normal observers were recruited from a primary care optometric practice and from the staff of Moorfields Eye Hospital. Ages ranged from 20 to 76 years (mean age 42.8 years) and 17 participants were male. The right eye was tested in half of the subjects (25) and the left eye in the other half. Each observer underwent preliminary tests to screen for their eligibility to participate in the study. These tests included baseline refraction (retinoscopy and subjective) at 4 m and ocular examination (including binocular indirect ophthalmoscopy) to exclude any ocular pathology. The mean spherical refractive error was −0.93 diopters (D; range, −5.38 to +3.00 D). If, on initial screening, a potential subject's acuity fell outside the measurement range of the charts at 4 m they were excluded, with the result that the mean CS1 unaided letter acuity for the 50 eligible subjects was 0.44 logMAR (range, −0.04 to 1.16 logMAR).

Unaided acuity was then measured on each subject using each of the eight different charts in one single visit. Subjects were required to read from the top of each chart and encouraged to guess if they were unsure of a letter's identity. Viewing time was not restricted. The test was terminated when a whole line was guessed incorrectly and responses for each letter recorded by the examiner on a pro forma data sheet. The eye to be tested was randomized, as was the order in which the test charts were administered, to control for both learning and fatigue effects.

Visual acuity scores were then determined in three different ways[16,18]:

a) *Letter-by-letter scoring with line-based termination* in which each letter was assigned a value of 0.02 logMAR and the test was terminated when a specified number of letters per line was read incorrectly (i.e., five, four or more, three or more, two or more, or one or more wrong per line). If, for example, in the case of one or more letters being read on a line, the full line being attempted was included in the score, and acuity was recorded as $1.3 - (0.02 \times$ number of correct letters).

b) *Letter-by-letter scoring with whole-chart termination* in which each letter was assigned a score of 0.02 logMAR, but test termination was taken as a specified number of letters wrong across the whole chart (i.e., five or more,
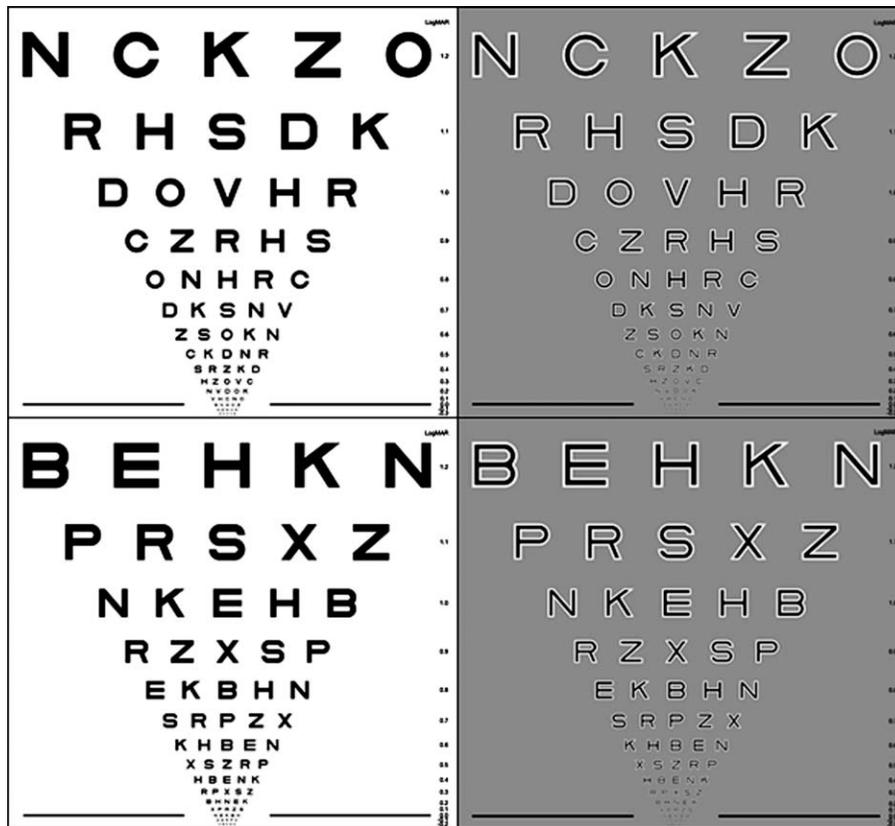
FIGURE 1. Appearance of charts CS1 and MAC S1 (*top*), which use the Sloan alphabet, and CN1 and MAC N1 (*bottom*), which use the New alphabet.

four or more, three or more, two or more, one or more letters wrong across the chart). Acuity was again recorded as $1.3 - (0.02 \times$ number of correct letters).

c) *Line by line scoring,* where the visual acuity score was taken as the last line in which more letters were read correctly than incorrectly.

The first two of these techniques employ the same scoring method but differ in their termination criteria. The third technique has a different scoring method.

Statistical Analysis

For each individual, the difference in the two acuity measurements made with each chart was calculated for each scoring and test termination rule, and the methods of Bland and Altman[34] and ordinary least squares regression analysis were employed. The Shapiro–Wilk W-test was used to confirm that the differences were normally distributed. The GraphPad Prism statistical analysis package (GraphPad Software, Inc., La Jolla, CA) was employed for this purpose.

RESULTS

Bland and Altman scatter plots were constructed to graphically present the spread of results for the CS1 and CS2, MAC S1 and MAC S2, CN1 and CN2, and finally MAC N1 and MAC N2 charts (Figs. 2a–d). These plots show the TRV of each measurement technique (here, $1.96 \times$ SD) alongside the mean difference. The data presented here are for the letter-by-letter scoring line-based termination for five letters wrong. There was no evidence of any systematic association between the level of

agreement and the underlying acuity. The differences were assumed normally distributed using the Shapiro–Wilk W-test ($P = 0.251$ for CS1 and CS2, $P = 0.309$ for MAC S1 and MAC S2, $P = 0.412$ for CN1 and CN2, $P = 0.286$ for MAC N1 and MAC N2). On this basis we are justified in using Bland and Altman summary statistics of mean bias and 95% limits of agreement, which are presented in the Table.

The F-test was used to test the null hypothesis that the standard deviation for CS1 is equal to that of MAC S1 and this was rejected (one-tailed $P = 0.05$, $F_{49,49} = 1.61$). No difference was found between the TRVs for the Sloan letter set and the new alphabet set in either conventional (±0.14 logMAR for both) or high-pass (±0.10 logMAR for both) format.

Figure 2e displays the comparison results for charts CS1 and MAC S1 for a letter-by-letter line-based termination for five letters wrong criterion. A proportional as well as a systematic bias can be inferred from this in that a greater level of disagreement between the two chart types existed at the better acuity end. The methods described by Bland and Altman do not further inform about such a potential proportional bias, so we performed ordinary least squares regression analysis on the data, confirming the presence of the bias ($r^2 = 0.217$, $P = 0.001$). The difference in visual acuity between CS1 and MAC S1 was approximately -0.20 logMAR at the 0.00 logMAR visual acuity level (indicating that conventional visual acuity measurements were, on average, 2 logMAR lines "better" than high-pass acuity measurements), compared to a difference of only -0.05 logMAR (half a line) at the 1.00 logMAR acuity level.

Figures 3a and 3b demonstrate how TRV for letter-by-letter scoring changed with different line and chart termination rules, respectively. Looking at Figure 3a, it can be seen that the TRV for CS1–CS2 was ±0.14 logMAR for five and four or more
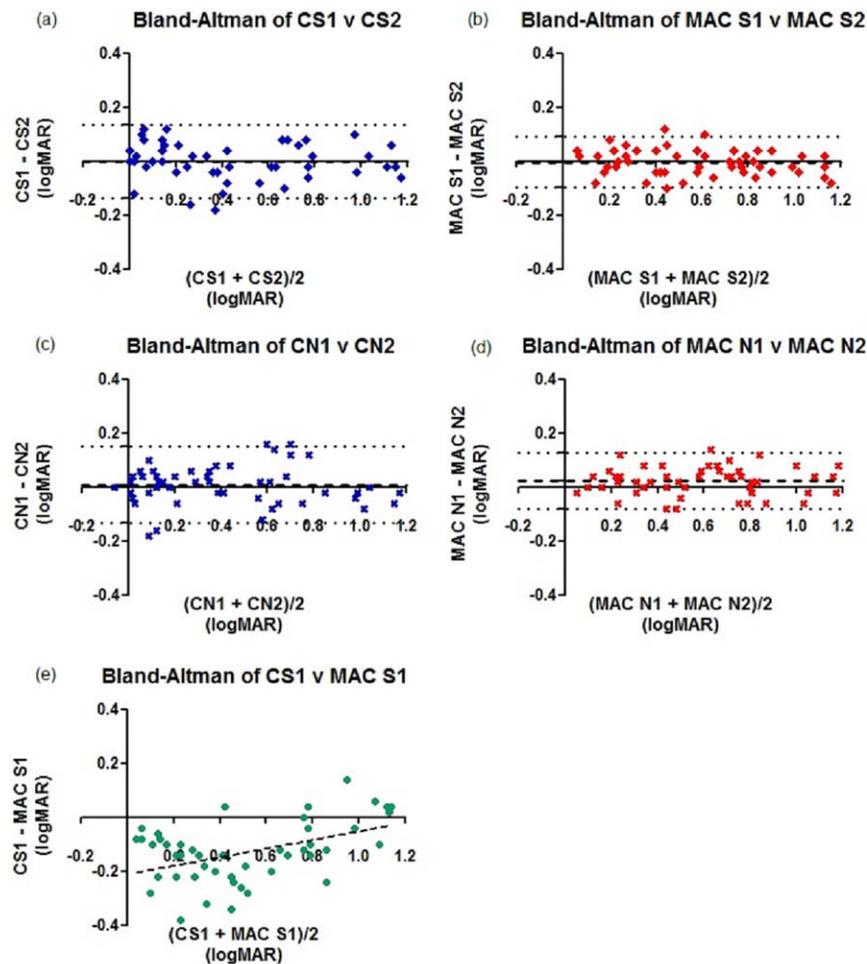
FIGURE 2. Bland–Altman plot for test and retest measurements using letter-by-letter line-based termination for five letters wrong for (**a**) CS1 versus CS2, (**b**) MAC S1 versus MAC S2, (**c**) CN1 versus CN2, (**d**) MAC N1 versus MAC N2, and (**e**) CS1 versus MAC S1. In each instance (**a–d**), the mean difference and 95% limits of agreement are plotted and for (**e**) the slope fitted using ordinary least squares regression analysis $r^2 = 0.217$, $P = 0.001$.

letters wrong per line termination criteria but increased (to ±0.17 logMAR) for three or more letters wrong. For MAC S1–MAC S2, the TRV was lower and remained at ±0.10 logMAR for five, four or more, and three or more letters wrong per line, increasing (to ±0.13 logMAR) for two or more letters wrong per line. The New alphabet set yielded no systematic improvement in TRV in either CN1 or MAC N1 form under line-based termination criteria.

For chart-based termination (Fig. 3b), TRV for CS1–CS2 was even higher than with line-based termination with a value of ±0.17 logMAR for five or more errors and four or more errors across the chart, increasing slowly thereafter. For MAC S1–MAC S2 and chart-based termination, the TRV was ±0.10 logMAR (similar to that for line-based termination) with five or more letters wrong across the whole chart, increasing slowly

thereafter by around half a letter with four or more errors, and by half a letter again for three or more errors across the chart. Once again, no systematic improvement was found in TRV with the MAC N1 compared to MAC S1 chart with chart based termination rules. A small improvement of one letter was found with the CN1 compared to CS1 chart for five, four, and three or more errors across the chart.

For line-by-line scoring (not plotted) TRV was 0.23 logMAR for both CS1–CS2 and CN1–CN2, and much lower (0.14 logMAR) for both MAC S1–MAC S2 and MAC N1–MAC N2.

Figures 4a and 4b demonstrate how the mean logMAR visual acuity (based on stroke width) changed with letter-by-letter scoring for line and chart-based termination rules, respectively. Figure 4a shows how the mean visual acuity score for CS1 with line-based termination was 0.44 logMAR for

TABLE. Bland–Altman Summary Statistics of Mean Bias and TRV Expressed as 95% Limits of Agreement

	Mean Difference (SE)	95% CI Mean Difference	Range of Observed Differences	TRV, 95% Limits of Agreement
CS1–CS2	−0.001 (0.010)	−0.020, 0.019	−0.18, 0.12	±0.135
MAC S1–MAC S2	−0.004 (0.007)	−0.017, 0.010	−0.10, 0.12	±0.095
CN1–CN2	0.008 (0.010)	−0.012, 0.029	−0.18, 0.16	±0.142
MAC N1–MAC N2	0.022 (0.008)	0.007, 0.037	−0.08, 0.14	±0.105

These data are presented graphically in Figure 2a through 2d and visual acuity scores were calculated using letter-by-letter line-based termination for five letters wrong.
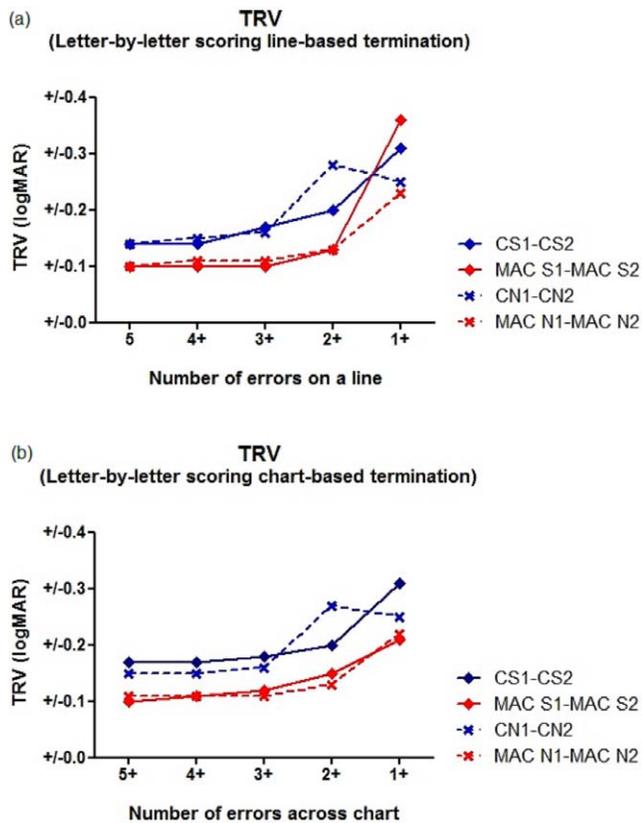
Investigative Ophthalmology & Visual Science

FIGURE 3. TRV for letter-by-letter scoring and (**a**) line- and (**b**) chart-based termination with different numbers of letters wrong.

five letters wrong per line, increasing to 0.45 logMAR with four or more letters wrong, and 0.46 logMAR with three or more letters wrong. The visual acuity threshold then increased to a value of 0.49 logMAR for a termination rule of two or more letters wrong per line. As previously mentioned, the mean logMAR score was larger with MAC S1 (0.57 logMAR) than CS1 for five letters wrong per line but little change in acuity was observed until a termination criterion of one or more letters wrong per line was reached, when threshold rose to 0.63 logMAR.

Chart-based termination (Fig. 4b) showed a similar pattern, with a larger mean threshold score for MAC S1 (0.58 logMAR) than for CS1 (0.46 logMAR) for five or more letters wrong across the whole chart. CS1 visual acuity thresholds increased by around half a letter with each criterion change until three or more letters wrong across the whole chart, rising to 0.51 logMAR for two or more letters wrong and 0.55 for one or more letters wrong. Once again, little change was noticed in thresholds with the MAC chart until a termination criterion of one or more letters wrong across the chart when the threshold was 0.63 logMAR.

For line-by-line scoring (not plotted), mean visual acuity was 0.44 logMAR for CS1, 0.41 logMAR for CN1, 0.56 logMAR for MAC S1, and 0.59 for MAC N1.

Discussion

TRV fundamentally limits the smallest detectable change in visual acuity that can reliably be attributed to a change in clinical status rather than just measurement error. Previous studies have concluded that logMAR charts are superior to Snellen charts since they achieve lower TRV by using improved

test design features and scoring techniques.[9,27] Although Carkeet[16] has investigated the effect of different termination rules on visual acuity using logMAR charts and single-letter scoring, that work was based on simulation results, whereas we set out to explore these issues by measuring human performance. Furthermore, our previous work with high-pass letters[15,29] suggested that it may be appropriate to create a chart using these letters (the Moorfields Acuity Chart) to further reduce TRV.

We found that, regardless of scoring technique (letter-by-letter or line-by-line) or termination criteria (line-based or chart-based), the TRV values for the Moorfields Acuity Charts, employing high-pass letters, are lower than for logMAR charts, composed of conventional letters, in uncorrected normal subjects. Furthermore, while the conventional letter charts were found to have higher TRV values using chart-based (0.17 logMAR at best) as opposed to line-based termination (0.14 logMAR at best) with letter-by-letter scoring, and higher still using line-by-line scoring techniques (0.23 logMAR), the high-pass charts were found to have more similar values across all three methods. In addition, TRV for high-pass charts appears to be less affected by different letter-by-letter termination number criteria compared to the conventional letter charts. While some of the differences may seem modest (0.10 vs. 0.14 logMAR at five errors on a line), they are still of similar magnitude to the improvements in TRV reported by previous studies comparing logMAR charts with Snellen charts[35] or different termination criteria.[16,17,19]

These observations apply for both the Sloan alphabet and the New alternative alphabet, which excludes circular letters.
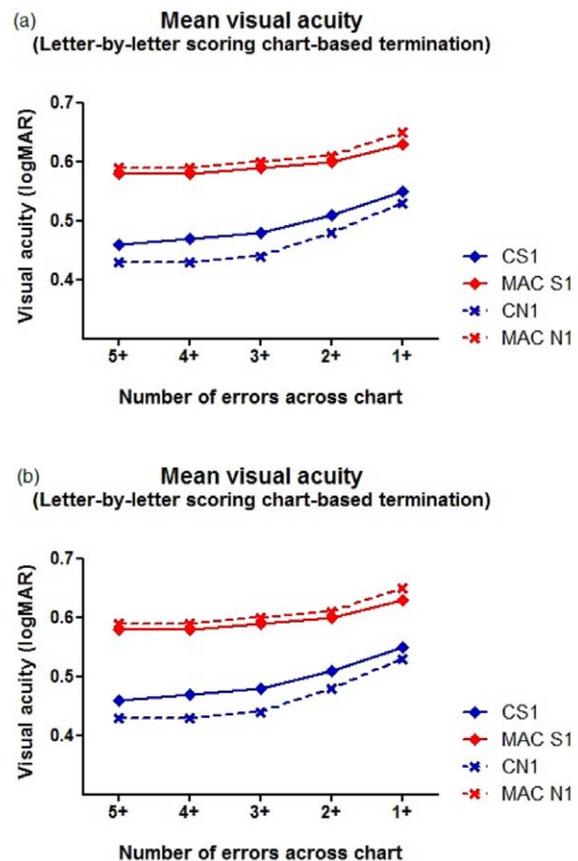

FIGURE 4. Mean visual acuity for letter-by-letter scoring with (**a**) line- and (**b**) chart-based termination with different number of letters wrong.

The 10 Sloan letters used on the ETDRS chart are arranged in combinations such that each line has the same average difficulty.[27] Over the years, it has been commonly thought that using a set of letters with more similar legibility would improve the repeatability of acuity measurements.[9] Arditi and Cagenello[17] investigated the effect on reliability of randomly arranging the 10 Sloan letters and found no difference between the two. Raasch et al.[36] selected a new letter set in which the probability of identification curves was supposedly identical for all the letters and compared the variability scores attained with these to those for the Sloan letters. They found minimal difference between the two, suggesting that the difference in identifiability of the Sloan letters is not a significant factor. In agreement with this, we found minimal differences in TRV between the Sloan set of 10 letters and our New alternative alphabet, indicating that the improvement mostly results from using the high-pass letter design rather than a different alphabet. TRV has previously been shown to increase with optical defocus in conventional letter charts, and it may be that the lower TRV found with the high-pass letters is a result of their greater resistance to blur.[29]

The data in Figure 3a support Carkeet's[16] recommendation of using a line-by-line termination criterion of four or five letters wrong per line with the ETDRS chart because these gave a TRV value of ±0.14 logMAR, increasing to ±0.17 logMAR with a termination of three or more letters wrong. The results for the conventional letter design charts are typical of those published in other studies of TRV for logMAR acuity measurements in which TRV values of up to 0.2 logMAR can be observed in subjects with unchanged acuity.[8,17,18,20–26]

In assessing the performance of the different chart designs in this study, we have directly compared the logMAR values based on "stroke-width," by which we mean the total (outside edge to outside edge) angular value of the limbs of the letters in question. The high-pass letters are, of course, constructed of black and white lines that are considerably narrower than the overall stroke-width, meaning that the visual system must rely on finer features (in the spatial domain) or higher letter frequencies (in the frequency domain) and the letter must be made larger to be correctly resolved. It is hardly surprising, therefore, that a systematic bias was observed between CS1 and MAC S1 measurements in that acuity for the high-pass letters appeared worse overall than for conventional letters. Previous studies have shown that the acuity value, if based on stroke or gap width, can vary considerably depending on the design of the letter chart,[6,37] the number of alternatives,[15,16] the targets employed to measure it (e.g., gratings,[31] Landolt rings,[38,39] vernier[40]), and even the choice of letter subset[15,41] depending on the features or spatial frequencies made available to the visual system. In fact, it is increasingly clear that conventional letters should not be regarded as a set of uniformly discriminable stimuli, but really as a group of stimuli, all of very different construction, for which the visual system uses different spatial frequencies to recognize different letters.[31,33,41,42]

The pattern of the points in the Bland–Altman plot in Figure 2e also suggests that the magnitude of the difference between measurements decreases with worsening acuity. Such a substantial *proportional* bias is not unusual in acuity chart comparisons and has also been previously demonstrated when comparing Snellen and ETDRS charts.[43] In a Snellen versus ETDRS comparison the reason for the bias must lie in the design of the chart (since the letter design is the same). In the current study, the reason must lie in the design of the letters since the chart layouts were identical. Our previous study[29] found that high-pass letters were more robust to induced optical defocus than conventional letters in that acuity for conventional letters was "better" with zero blur, was roughly equal at 1 D blur, but became "worse" at 2 D blur and beyond. The proportional bias observed in the current study is in agreement with this previous finding in that, as acuity owing to uncorrected refractive error declines, the acuity with the high-pass chart was affected less. This means that those with good visual acuity as a result of small refractive error, will achieve a score of approximately −0.20 logMAR lines better with conventional compared to high-pass letters. Those with poorer acuity resulting from uncorrected refractive error will attain a score of only −0.05 logMAR better.

In conclusion, it would appear that the new Moorfields Acuity Chart, which employs high-pass letters, displays a lower TRV with uncorrected refractive error than an ETDRS style chart employing conventional letters. It also seems to be somewhat less affected by different termination criteria and scoring methods. Further work is required to test the chart's performance in conditions resulting in neural rather than optical acuity loss to determine its ability to separate these two factors in patients suffering from diseases such as age-related macular degeneration. It may be that different charts, employing letter designs that are differently vulnerable to optical or neural losses of vision, could assist in differentiating between the two.

References

1. Bai HQ, Yao L, Meng XX, Wang YX, Wang DB. Visual outcome following intraocular foreign bodies: a retrospective review of 5-year clinical experience. *Eur J Ophthalmol*. 2011;21:98–103.
2. Daien V, Navarre S, Fesler P, Vergely L, Villain M, Schneider C. Visual acuity outcome and predictive factors after bevacizumab for central retinal vein occlusion. *Eur J Ophthalmol*. 2012;22:1013–1018.
3. Goh YW, Ang GS, Azuara-Blanco A. Lifetime visual prognosis of patients with glaucoma. *Clin Experiment Ophthalmol*. 2011;39:766–770.
4. Gonnermann J, Klamann MK, Maier AK, et al. Visual outcome and complications after posterior iris-claw aphakic intraocular lens implantation. *J Cataract Refract Surg*. 2012;38:2139–2143.
5. Shona O, Gupta B, Vemala R, Sivaprasad S. Visual acuity outcomes in ranibizumab-treated neovascular age-related macular degeneration; stratified by baseline vision. *Clin Experiment Ophthalmol*. 2011;39:5–8.
6. Kaiser PK. Prospective evaluation of visual acuity assessment: a comparison of Snellen versus ETDRS charts in clinical practice (an AOS thesis). *Trans Am Ophthalmol Soc*. 2009;107:311–324.
7. Rosser DA, Laidlaw DA, Murdoch IE. The development of a "reduced logMAR" visual acuity chart for use in routine clinical practice. *Br J Ophthalmol*. 2001;85:432–436.
8. Hazel CA, Elliott DB. The dependency of logMAR visual acuity measurements on chart design and scoring rule. *Optom Vis Sci*. 2002;79:788–792.
9. Bailey IL, Lovie JE. New design principles for visual acuity letter charts. *Am J Optom Physiol Opt*. 1976;53:740–745.
10. Rosser DA, Murdoch IE, Cousens SN. The effect of optical defocus on the test–retest variability of visual acuity measurements. *Invest Ophthalmol Vis Sci*. 2004;45:1076–1079.

11. Carkeet A, Lee L, Kerr JR, Keung MM. The slope of the psychometric function for Bailey-Lovie letter charts: defocus effects and implications for modeling letter-by-letter scores. *Optom Vis Sci*. 2001;78:113–121.

12. 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;92:241–244.

13. Patel PJ, Chen FK, Rubin GS, Tufail A. Intersession repeatability of visual acuity scores in age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2008;49:4347–4352.

14. Gibson RA, Sanderson HF. Observer variation in ophthalmology. *Br J Ophthalmol*. 1980;64:457–460.

15. Shah N, Dakin SC, Redmond T, Anderson RS. Vanishing Optotype acuity: repeatability and effect of the number of alternatives. *Ophthalmic Physiol Opt*. 2011;31:17–22.

16. Carkeet A. Modeling logMAR visual acuity scores: effects of termination rules and alternative forced-choice options. *Optom Vis Sci*. 2001;78:529–538.

17. Arditi A, Cagenello R. On the statistical reliability of letter-chart visual acuity measurements. *Invest Ophthalmol Vis Sci*. 1993;34:120–129.

18. Vanden Bosch ME, Wall M. Visual acuity scored by the letter-by-letter or probit methods has lower retest variability than the line assignment method. *Eye (Lond)*. 1997;11(Pt 3):411–417.

19. Laidlaw DA, Abbott A, Rosser DA. Development of a clinically feasible logMAR alternative to the Snellen chart: performance of the "compact reduced logMAR" visual acuity chart in amblyopic children. *Br J Ophthalmol*. 2003;87:1232–1234.

20. Elliott DB, Sheridan M. The use of accurate visual acuity measurements in clinical anti-cataract formulation trials. *Ophthalmic Physiol Opt*. 1988;8:397–401.

21. Lovie-Kitchin JE. Validity and reliability of visual acuity measurements. *Ophthalmic Physiol Opt*. 1988;8:363–370.

22. Reeves BC, Wood JM, Hill AR. Vistech VCTS 6500 charts—within- and between-session reliability. *Optom Vis Sci*. 1991;68:728–737.

23. Brown B, Lovie-Kitchin J. Repeated visual acuity measurement: establishing the patient's own criterion for change. *Optom Vis Sci*. 1993;70:45–53.

24. Bailey IL, Bullimore MA, Raasch TW, Taylor HR. Clinical grading and the effects of scaling. *Invest Ophthalmol Vis Sci*. 1991;32:422–432.

25. Rosser DA, Cousens SN, Murdoch IE, Fitzke FW, Laidlaw DA. How sensitive to clinical change are ETDRS logMAR visual acuity measurements? *Invest Ophthalmol Vis Sci*. 2003;44:3278–3281.

26. Manny RE, Hussein M, Gwiazda J, Marsh-Tootle W. Repeatability of ETDRS visual acuity in children. *Invest Ophthalmol Vis Sci*. 2003;44:3294–3300.

27. Ferris FL III, Kassoff A, Bresnick GH, Bailey I. New visual acuity charts for clinical research. *Am J Ophthalmol*. 1982;94:91–96.

28. McMonnies CW, Ho A. Letter legibility and chart equivalence. *Ophthalmic Physiol Opt*. 2000;20:142–152.

29. Shah N, Dakin SC, Anderson RS. Effect of optical defocus on detection and recognition of vanishing optotype letters in the fovea and periphery. *Invest Ophthalmol Vis Sci*. 2012;53:7063–7070.

30. Howland B, Ginsburg A, Campbell F. High-pass spatial frequency letters as clinical optotypes. *Vision Res*. 1978;18:1063–1066.

31. Anderson RS, Thibos LN. Relationship between acuity for gratings and for tumbling-E letters in peripheral vision. *J Opt Soc Am A Opt Image Sci Vis*. 1999;16:2321–2333.

32. Gervais MJ, Harvey LO Jr, Roberts JO. Identification confusions among letters of the alphabet. *J Exp Psychol Hum Percept Perform*. 1984;10:655–666.

33. Anderson RS, Thibos LN. Sampling limits and critical bandwidth for letter discrimination in peripheral vision. *J Opt Soc Am A Opt Image Sci Vis*. 1999;16:2334–2342.

34. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–310.

35. Lim LA, Frost NA, Powell RJ, Hewson P. Comparison of the ETDRS logMAR, 'compact reduced logMar' and Snellen charts in routine clinical practice. *Eye (Lond)*. 2010;24:673–677.

36. Raasch TW, Bailey IL, Bullimore MA. Repeatability of visual acuity measurement. *Optom Vis Sci*. 1998;75:342–348.

37. McMonnies CW. Chart construction and letter legibility/readability. *Ophthalmic Physiol Opt*. 1999;19:498–506.

38. Plainis S, Kontadakis G, Feloni E, et al. Comparison of visual acuity charts in young adults and patients with diabetic retinopathy. *Optom Vis Sci*. 2013;90:174–178.

39. Becker R, Teichler G, Graf M. Comparison of visual acuity measured using Landolt C and ETDRS charts in healthy subjects and patients with various eye diseases. *Klin Monbl Augenheilkd*. 2011;228:864–867.

40. Bedell HE, Patel S, Chung ST. Comparison of letter and Vernier acuities with dioptric and diffusive blur. *Optom Vis Sci*. 1999;76:115–120.

41. Anderson RS, Thibos LN. The filtered Fourier difference spectrum predicts psychophysical letter discrimination in the peripheral retina. *Spat Vis*. 2004;17:5–15.

42. Bondarko VM, Danilova MV. What spatial frequency do we use to detect the orientation of a Landolt C? *Vision Res*. 1997;37:2153–2156.

43. Falkenstein IA, Cochran DE, Azen SP, et al. Comparison of visual acuity in macular degeneration patients measured with snellen and early treatment diabetic retinopathy study charts. *Ophthalmology*. 2008;115:319–323.