Vanishing Optotype Letter Acuity: Repeatability and Effect of the Number of Alternatives

N. Shah[1], S.C. Dakin[1], T. Redmond[1], R.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, United Kingdom.

Abstract

Vanishing Optotype letters have a pseudo high-pass design so that the mean luminance of the target is the same as the background and the letters thus ‘vanish’ soon after the resolution threshold is reached. We wished to determine the variability of acuity measurements using these letters compared to conventional letters, and in particular how acuity is affected by the number of alternatives available to the subject.

Acuity was measured using high contrast letters of both conventional and Vanishing Optotype design for three experienced normal subjects. Thresholds were determined for central vision in a forced choice paradigm for two alternatives (2AFC; AU and OQ), 4AFC (AQUO), 6AFC (QUANGO) and 26AFC (whole alphabet) using a QUEST procedure. Three measurements were made for each condition.

Threshold letter size was always larger for the Vanishing Optotypes than conventional letters, although the size of this difference (0.11 – 0.34 logMAR) depended on the number of alternatives and what they were. The effect of the number of AFC, and the individual letters employed, was smaller for the Vanishing Optotypes, implying that they are more equally legible than conventional optotypes. Variability was also lower for the Vanishing Optotype sets (0.01 – 0.03 logMAR) than the conventional letter sets (0.03 – 0.06).

The smaller effect of the number of alternatives, combined with more equal discriminability and lower threshold variability, implies that Vanishing Optotypes may be appropriate targets from which to design letter charts to measure small clinical changes in acuity.

Introduction

Visual acuity measurements remain of the upmost importance in forming clinical decisions when monitoring disease progression and the efficacy of therapy. Any test measuring visual acuity should provide precise and repeatable measurements in order to reliably determine whether or not a significant change in performance has resulted from either abnormality or treatment. Variability can originate from a variety of sources including the observer, the clinician, the overall design of the test chart or the psychophysical testing procedure, and recommendations have been made to minimize at least some of these.[1, 2]

LogMAR acuity charts were designed to remove many of the recognized limitations of conventional Snellen charts[3, 4] and are becoming more widely used in both clinical and research settings. While the letter-by-letter scoring system in theory allows step sizes of 0.02 log units, test-retest variability remains a problem for these charts with reported 95% confidence intervals between 0.06 and 0.19 log units for normal, focused eyes,[5-13] increasing significantly with the presence of either optical defocus[14] or retinal disease.[15]

[16] found significant differences in logMAR scores as a result of different termination rules and numbers of alternatives during a forced choice test (AFC). He suggested that between-subject variability arises as a result of different patient criteria where a subject may not be forced to identify small letters, depending on testing rigour. For Bailey-Lovie or ETDRS charts, employing letter-by-letter scoring, Carkeet suggested termination of the test when four or more mistakes are made on a line.

Although the Sloan[17] letter set, employed by modern ETDRS charts, was originally devised to have closely similar discriminability, closer examination of the literature indicates that this may not be the case.[18] If a test-chart's within-line discriminability difference is greater than the between-line discriminability difference, the test will be very variable. But is discriminability the inherent property of an individual letter or does a letter's discriminability depend on what, and how many, other letters it is being discriminated from? Visual acuity results will be affected by the probability that the subject will be able to discriminate the optotype from any number of other alternatives available. Carkeet[16] found that the mean and standard deviation of logMAR scores was significantly affected by the number of forced choice alternatives. The increase in the mean is not surprising since, as the number of alternatives increases, the degree of letter uncertainty increases in that there are more likely to be other letters that look similar to the presented one, meaning that the letter must appear visibly different from all of the other possibilities before the subject ventures an identification. However, this greater letter uncertainty does not necessarily lead to greater threshold variability; in fact the opposite is likely true since the subject is less likely to guess correctly even when the letter is unresolvable.

Several studies have shown that the visual system relies on the lower object spatial frequency content for conventional letter recognition, in both foveal and peripheral vision.[19-25] Several of these studies also indicated large differences in the spatial frequency content at these low object frequencies[20, 24, 26] resulting in some letters remaining easily recognizable when small and blurred, while others do not. However, if these lower frequencies, where conventional letters differ substantially, are removed, the visual system must rely on the higher spatial frequency content and the letters may thus become more equally discriminable. If this is so, the effect of different numbers of alternatives may also become less.

'Vanishing Optotype' targets, first described by Howland *et al.*,[27] have a pseudo 'high-pass' design in that they are typically constructed of a dark core surrounded by light edges (or vice versa), the mean luminance of which is the same as the background (Figure 1). While such stimuli are not truly high-pass, their construction means that the detection and resolution thresholds are closely similar in the fovea [28] and, unlike conventional letters, the characters 'vanish' almost as soon as the resolution threshold is reached.

The Vanishing Optotype target design has been employed in High-Pass Resolution Perimetry (HRP)[29] and is currently employed in tests such as the paediatric Cardiff Acuity Test which uses preferential looking techniques to determine visual acuity in children and in those unable to participate in conventional optotype identification tests.[30, 31] However, despite some academic interest, Vanishing Optotypes have, to date, received relatively little attention in clinical visual acuity testing. This study aims to determine the variability of acuity measurements using Vanishing Optotype letters relative to conventional letters to test the hypothesis that, if lower frequencies are removed, the letters become more equally discriminable. This being the case, the number of alternatives available to the subject should have less effect on acuity measurements with Vanishing Optotypes. The results of this would be valuable when thinking about new test chart designs.

Methods

Ethical approval for this study was obtained from the relevant UCL research ethics committee and all procedures adhered to the tenets of the Declaration of Helsinki. All tests were conducted on three experienced psychophysical observers (NS, RSA and TR), with no ocular abnormalities and corrected visual acuities of 6/5 or better. The refractive error was carefully corrected prior to the start of each testing session using trial lenses. Subjects NS and RSA were emmetropic while subject TR had a mean spherical refractive error of -3.00D.

Foveal visual acuity measurements were made monocularly in the right eye of all subjects using both conventional and Vanishing Optotype letters. The Vanishing Optotypes were constructed with an inner black 'core' flanked by a white border of half the width of the central section. This created a target with the same mean luminance as the background and thus had a pseudo high-pass design. For both stimulus types, the letter height and width were five times the 'stroke width', which in the case of the Vanishing Optotype consisted of the dark middle bar with its two white flanks. All optotype stimuli were generated using MATLAB v7.6 (Mathworks, Inc., Natick, MA, USA) and were presented at high contrast (94.6%) on a γ-corrected high-resolution (1280 x 1024 pixels) Dell Ultrascan P991 CRT monitor (Dell Corp. Ltd, Brackness, Berkshire, UK) driven by a Macintosh computer (Apple Computer Inc, Cupertino, CA, USA). Presentation time was 500ms and the CRT monitor had a background luminance of 53.9cd/m^2. All testing was conducted at 3.8m under low room illumination to avoid screen reflections; at this distance the screen subtended 4 x 5.3 degrees and one pixel subtended 0.25 minutes of arc. Scaling of stimuli was achieved using the OpenGL capabilities of the computer's built-in graphics card (ATI Radeon X1600; AMD, Sunnyvale, CA, USA). This (bilinear interpolation) procedure allowed us to display stimuli of arbitrary size with sub-pixel resolution while retaining accurate representation of their (balanced) luminance structure.

For each subject, threshold visual acuity was determined for both conventional and Vanishing Optotypes for differing numbers of AFC using QUEST, an adaptive psychometric procedure.[32] In this paradigm, the size of any displayed letter is determined by knowledge of the previous responses, with trials evenly spread on a decimal/log axis. The prior density function was limited by the maximum and minimum displayable letter size on the screen and an initial letter size of 115.8 x 115.8 minutes of arc was displayed. The slope (β) of the psychometric function used was set to 3.5 which is widely used in psychophysical literature. The final acuity threshold was determined by QUEST's built in maximum likelihood estimation procedure of threshold. Each test run involved 50 letter presentations. The alternative choices in each session were 2AFC (AU and QO), 4AFC (AQUO), 6AFC (QUANGO) or 26AFC (whole alphabet) (Figure 1).

The viewing distance was 3.8m and the subject's verbal letter identification was entered on the keyboard by the examiner. Responses were limited to the letter set available for each test. These were displayed in the corner of the screen to remind subjects of the choice of letters. The final threshold size under each AFC condition was recorded and converted to logMAR where, for the Vanishing Optotypes, the 'stroke width' includes both the central dark bar and its white flanks. Three repeat measurements were made for each condition for all subjects.

Results

The mean of the three repeat thresholds measurements obtained for each subject was plotted in logMAR values for each AFC condition for conventional letters and for Vanishing Optotypes (Figure 2). Error bars represent the standard deviation of the three repeat measures. For all AFC conditions, threshold letter size was significantly larger for the Vanishing Optotypes than for the conventional letters at the 0.05 significance level, except for QO ($p=0.08$). However, the actual difference in performance between the two stimulus types (0.11 – 0.34 logMAR) was not only dependent on the number of alternatives but also on what they were. Interestingly, both the smallest and largest between-optotype difference occurred under 2AFC conditions for the letters OQ and AU respectively. The mean threshold acuity for conventional optotypes ranged from -0.33 (AU) to 0.06 (QO), a 0.39 log difference. Significant differences in discrimination thresholds ($p<0.05$, paired t-test) were found between AU and all other AFC combinations. Significant differences were also found between QO and AQUO, QO and QUANGO, and AQUO and QUANGO (all $p<0.05$).

The Vanishing Optotype discrimination thresholds were less affected by the number of AFC, and the individual letters employed, compared to the conventional letters. The discrimination thresholds ranged only from 0.01 (AU) to 0.17 (QO), a 0.16 log difference. Significant differences ($p<0.05$, paired t-test) were again found between AU and all other AFC combinations, but not between any other AFC combinations. The effect of the differing numbers of AFC is thus less overall for the Vanishing Optotypes.

Figure 3 shows a plot of the mean standard deviation as a percentage of the logMAR thresholds for each of the letter types. It can be seen that the variability was lower for the Vanishing Optotypes (0.01 – 0.03 log units), compared to the conventional letters (0.03 – 0.06 log units).

Discussion

As previously stated, visual acuity measurements contribute significantly to clinical decision making with regard to disease progression and treatment efficacy. A measured deterioration in visual acuity often forms one of the criteria for further intervention, but only if it is deemed clinically significant. For this reason, any test of visual acuity should be both precise and repeatable.

The aims of this study were to determine the repeatability of acuity measurements using Vanishing Optotype letters and to investigate how acuity using these is affected by the number of alternatives available to the subject. The results suggest that, overall, visual acuity measured using Vanishing Optotypes is ‘poorer’ than conventional letter acuity, regardless of the number of AFC. This conclusion is the logical result of directly comparing the threshold letter heights of the two letter types. However, as previously mentioned, the letter types are composed of different spatial frequencies in the Fourier domain and several studies have shown that the visual system relies on the lower spatial frequency content for conventional letter recognition in both foveal and peripheral vision.[19-25] If lower spatial frequency information is removed, as in the Vanishing Optotypes, the visual system must rely on the high frequencies for identification,[33] hence the ‘poorer’ performance observed for these characters. However, the aim of this study was not so much to compare absolute differences in threshold letter size between the two target types, but to determine the effects of different numbers of AFC and threshold variability. From a clinical perspective, this is more important.

Vanishing Optotypes are less affected overall by the number of alternatives available and what they are, likely because, as hypothesized, they are more equally discriminable than conventional optotypes. As mentioned, several studies have indicated that the visual system utilizes the low spatial frequencies for conventional high contrast letter acuity. Some of these studies also indicated large differences in the spatial frequency content at these low object frequencies.[20, 24, 26] If two letters are very different in their low spatial frequency content, they should remain discriminable down to very small sizes. Two letters that are more similar in their low spatial frequency content force the visual system to rely on higher spatial frequencies for discrimination, thus their acuity threshold will be larger. This would explain why AU is much more discriminable than OQ in conventional form (Figure 2). Under 4AFC conditions (AQUO) performance fell in between the two 2AFC conditions. As the AFC number rises to 6 and 26 the letters become more ‘similar’ (on average), increasing letter uncertainty and leading to larger discrimination thresholds, i.e. each letter must begin to look more ‘like itself’ rather than ‘not the others’ in order for the subject to confidently identify it.

However, if these lower frequencies, which give rise to large inter-letter discriminability differences for conventional letters, are removed, the between-letter differences should become smaller and much more uniform. This is borne out in Figure 2 where, except for AU, there is no significant difference in performance with different AFC conditions. Using the higher frequencies there seems to be closer similarity and greater letter uncertainty, even under low AFC conditions. It may even be that, on filtering out the low frequencies, the visual system switches to a strategy based less on spatial frequency content and more on localized features.

In addition, measurement variability was found to be lower using Vanishing Optotypes (Figure 3). This has been attributed to the fact that conventional letters have two distinctly different thresholds for detection and resolution. [34] point out that variability can arise as a result of the transitional zone between these two points, as it is known that subjects can learn to recognize blurred images that are close to the detection threshold. Any ability to recognize blurred images relies on the presence of different low spatial frequencies in the targets that permit discrimination (e.g. 'A' from 'U') even though they no longer resemble the actual letters. With conventional letters, under greater AFC conditions, different low spatial frequency content will lead to large inter-letter legibility differences. If this difference *within* steps is significantly greater than *between* steps, increased variability in any staircase threshold measure will result.

In conclusion, the smaller effect of the number of alternatives, combined with more equal discriminability and better repeatability, at least in normal subjects, suggests that Vanishing Optotypes may be promising targets from which to design clinical letter charts. More work remains to be done to understand the differences in how the visual system resolves the Vanishing Optotypes compared to conventional letters. In addition, we have yet to examine the effects of optical defocus and ocular abnormality on Vanishing Optotype acuity to determine whether these stimuli are appropriate to measure clinically significant changes in vision.

Acknowledgements

This work is supported by an award from the NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust & UCL Institute of Ophthalmology.

References:

1. Ferris FL, 3rd, Bailey I. Standardizing the measurement of visual acuity for clinical research studies: Guidelines from the Eye Care Technology Forum. Ophthalmology. 1996;103(1):181-2. Epub 1996/01/01.
2. Raasch TW, Bailey IL, Bullimore MA. Repeatability of visual acuity measurement. Optom Vis Sci. 1998;75(5):342-8. Epub 1998/06/13.
3. Bailey IL, Lovie JE. New design principles for visual acuity letter charts. Am J Optom Physiol Opt. 1976;53(11):740-5. Epub 1976/11/01.
4. McGraw P, Winn B, Whitaker D. Reliability of the Snellen chart. BMJ. 1995;310(6993):1481-2. Epub 1995/06/10.

5. Elliott DB, Sheridan M. The use of accurate visual acuity measurements in clinical anti-cataract formulation trials. Ophthalmic Physiol Opt. 1988;8(4):397-401. Epub 1988/01/01.
6. Lovie-Kitchin JE. Validity and reliability of visual acuity measurements. Ophthalmic Physiol Opt. 1988;8(4):363-70. Epub 1988/01/01.
7. Reeves BC, Wood JM, Hill AR. Vistech VCTS 6500 charts--within- and between-session reliability. Optom Vis Sci. 1991;68(9):728-37. Epub 1991/09/01.
8. Brown B, Lovie-Kitchin J. Repeated visual acuity measurement: establishing the patient's own criterion for change. Optom Vis Sci. 1993;70(1):45-53. Epub 1993/01/01.
9. 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):411-7. Epub 1997/01/01.
10. Arditi A, Cagenello R. On the statistical reliability of letter-chart visual acuity measurements. Invest Ophthalmol Vis Sci. 1993;34(1):120-9. Epub 1993/01/01.
11. Bailey IL, Bullimore MA, Raasch TW, Taylor HR. Clinical grading and the effects of scaling. Invest Ophthalmol Vis Sci. 1991;32(2):422-32. Epub 1991/02/01.
12. Hazel CA, Elliott DB. The dependency of logMAR visual acuity measurements on chart design and scoring rule. Optom Vis Sci. 2002;79(12):788-92. Epub 2003/01/07.
13. 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(8):3278-81. Epub 2003/07/29.
14. 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(4):432-6. Epub 2001/03/27.
15. 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(10):4347-52. Epub 2008/06/21.

16. Carkeet A. Modeling logMAR visual acuity scores: effects of termination rules and alternative forced-choice options. Optom Vis Sci. 2001;78(7):529-38. Epub 2001/08/16.
17. Sloan LL. Measurement of visual acuity; a critical review. AMA Arch Ophthalmol. 1951;45(6):704-25. Epub 1951/06/01.
18. Reich LN, Bedell HE. Relative legibility and confusions of letter acuity targets in the peripheral and central retina. Optom Vis Sci. 2000;77(5):270-5. Epub 2000/06/01.
19. Alexander KR, Xie W, Derlacki DJ. Spatial-frequency characteristics of letter identification. J Opt Soc Am A Opt Image Sci Vis. 1994;11(9):2375-82. Epub 1994/09/01.
20. 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(10):2321-33. Epub 1999/10/12.
21. Anderson RS, Thibos LN. The filtered Fourier difference spectrum predicts psychophysical letter discrimination in the peripheral retina. Spat Vis. 2004;17(1-2):5-15. Epub 2004/04/14.
22. Bondarko VM, Danilova MV. What spatial frequency do we use to detect the orientation of a Landolt C? Vision Res. 1997;37(15):2153-6. Epub 1997/08/01.
23. Chung ST, Legge GE, Tjan BS. Spatial-frequency characteristics of letter identification in central and peripheral vision. Vision Res. 2002;42(18):2137-152. Epub 2002/09/05.
24. Gervais MJ, Harvey LO, Jr., Roberts JO. Identification confusions among letters of the alphabet. J Exp Psychol Hum Percept Perform. 1984;10(5):655-66. Epub 1984/10/01.
25. Parish DH, Sperling G. Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination. Vision Res. 1991;31(7-8):1399-415. Epub 1991/01/01.
26. 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(10):2334-42. Epub 1999/10/12.
27. Howland B, Ginsburg A, Campbell F. High-pass spatial frequency letters as clinical optotypes. Vision Res. 1978;18(8):1063-6. Epub 1978/01/01.

28. Anderson RS, Ennis FA. Foveal and peripheral thresholds for detection and resolution of vanishing optotype tumbling E's. Vision Res. 1999;39(25):4141-4. Epub 2001/02/07.
29. Frisen L. High-pass resolution targets in peripheral vision. Ophthalmology. 1987;94(9):1104-8. Epub 1987/09/01.
30. Adoh TO, Woodhouse JM. The Cardiff acuity test used for measuring visual acuity development in toddlers. Vision Res. 1994;34(4):555-60. Epub 1994/02/01.
31. Johansen A, White S, Waraisch P. Screening for visual impairment in older people: validation of the Cardiff Acuity Test. Arch Gerontol Geriatr. 2003;36(3):289-93. Epub 2003/07/10.
32. Watson AB, Pelli DG. QUEST: a Bayesian adaptive psychometric method. Percept Psychophys. 1983;33(2):113-20. Epub 1983/02/01.
33. Majaj NJ, Pelli DG, Kurshan P, Palomares M. The role of spatial frequency channels in letter identification. Vision Res. 2002;42(9):1165-84.
34. Koskin SA, Boiko EV, Sobolev AF, Shelepin YE. Mechanisms of recognition of the outlines of "vanishing" optotypes. Neurosci Behav Physiol. 2007;37(1):59-65. Epub 2006/12/21.

Figure legends.

Figure 1. a) the 2, 4 and 6 Alternative Forced Choice Vanishing Optotype letter set and b) the 26 Alternative Forced Choice Vanishing Optotype letter set.

Figure 2. LogMAR values for all three subjects under each AFC condition for a) conventional letters and b) Vanishing Optotypes. Error bars represent standard deviation of three repeat threshold measurements.

Figure 3. Mean standard deviation of the logMAR thresholds for conventional letters (filled symbols) and Vanishing Optotypes (open symbols).

Figure 1.

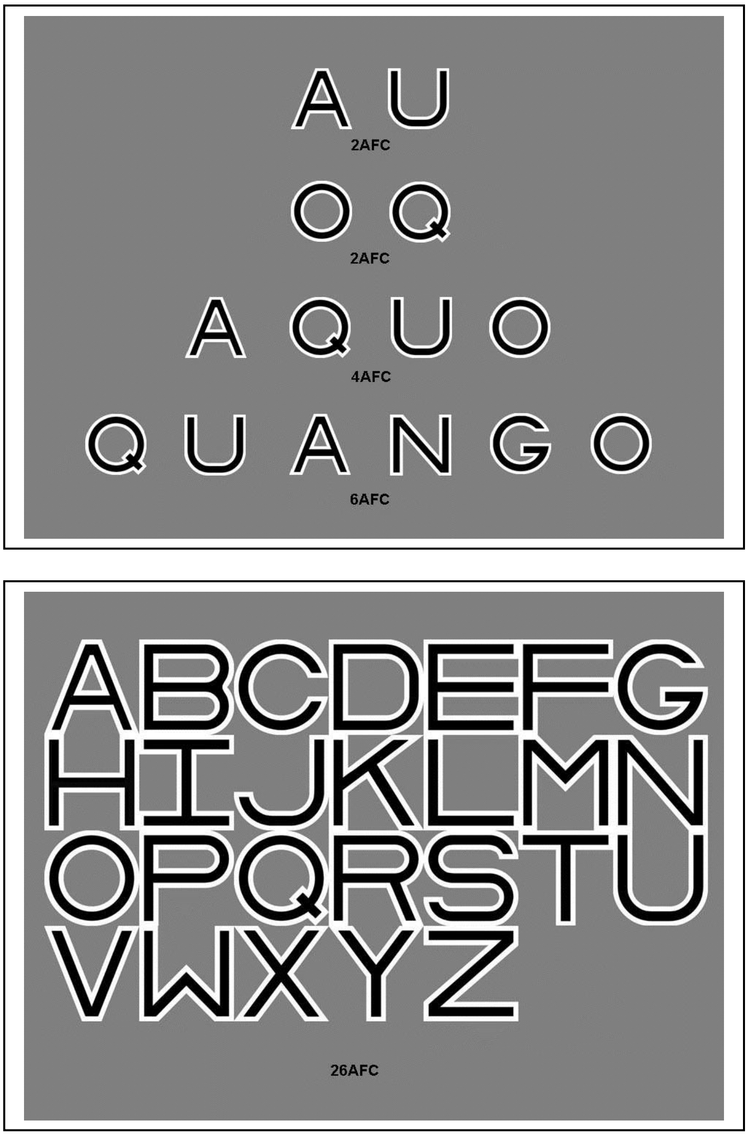
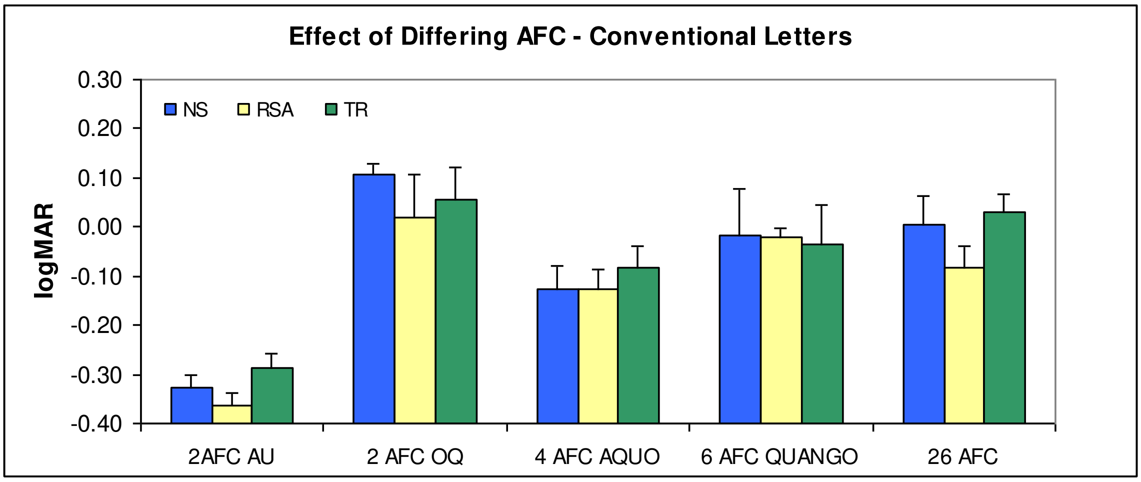

A U
2AFC
O Q
2AFC
A Q U O
4AFC
Q U A N G O
6AFC
ABCDEFG
HIJKLMN
OPQRSTU
VWXYZ
26AFC

Figure 2.


Effect of Differing AFC - Conventional Letters
logMAR
0.30
0.20
0.10
0.00
-0.10
-0.20
-0.30
-0.40
NS
RSA
TR
2AFC AU
2 AFC OQ
4 AFC AQUO
6 AFC QUANGO
26 AFC

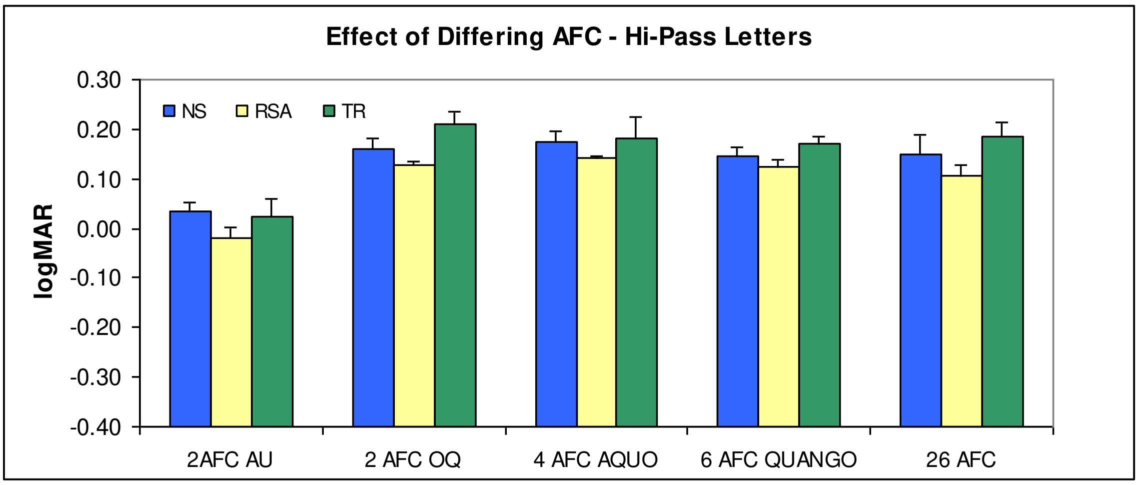
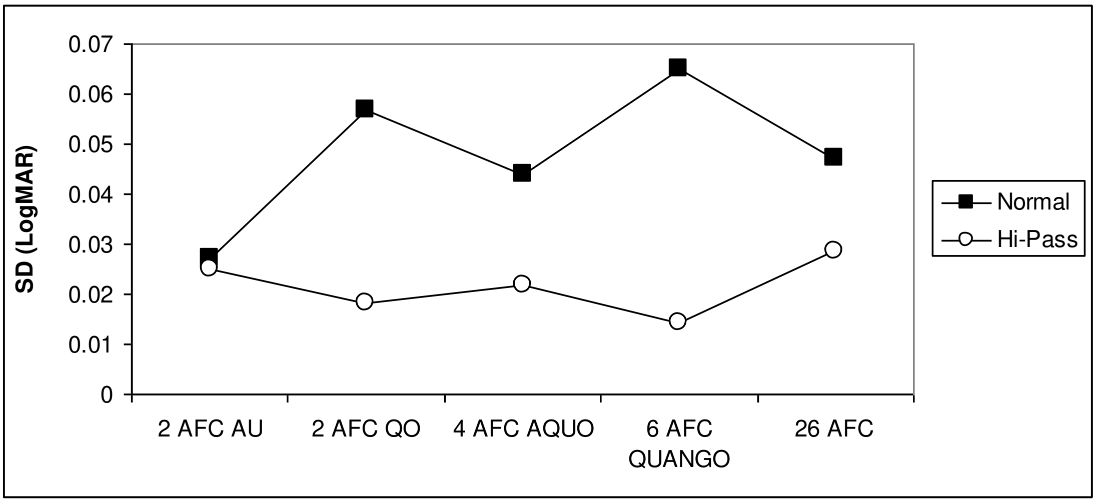

Effect of Differing AFC - Hi-Pass Letters
logMAR
0.30
0.20
0.10
0.00
-0.10
-0.20
-0.30
-0.40
NS
RSA
TR
2AFC AU
2 AFC OQ
4 AFC AQUO
6 AFC QUANGO
26 AFC

Figure 3.


SD (LogMAR)
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
2 AFC AU
2 AFC QO
4 AFC AQUO
6 AFC QUANGO
26 AFC
Normal
Hi-Pass