



Article scientifique

Article

2006

Published version

Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

Computer-based measurement of letter and word acuity

Whatham, Andréw Robert; Mermoud, Christophe; Deruaz, Anouk; Goldschmidt, Mira; Zesiger, Pascal Eric; Safran, Avinoam B.

How to cite

WHATHAM, Andréw Robert et al. Computer-based measurement of letter and word acuity. In: Ophthalmic and physiological optics, 2006, vol. 26, n° 2, p. 156–168. doi: 10.1111/j.1475-1313.2006.00345.x

This publication URL: <https://archive-ouverte.unige.ch/unige:82283>

Publication DOI: [10.1111/j.1475-1313.2006.00345.x](https://doi.org/10.1111/j.1475-1313.2006.00345.x)

Computer-based measurement of letter and word acuity

Andrew R. Whatham¹, Christophe Mermoud¹, Anouk Déruaz¹, Mira Goldschmidt¹, Pascal Zesiger² and Avinoam B. Safran¹

¹Department of Clinical Neurosciences and Dermatology, Ophthalmology Clinic, Neuro-Ophthalmology Unit, Geneva University Hospitals, 1211 Geneva 14, and ²Faculty of Psychology and Educational Sciences, University of Geneva, Geneva, Switzerland

Abstract

Determining causes of poor reading ability is an important step in trying to ameliorate reading performance in low-vision patients. One important parameter is word acuity. The principal aim of the current study is to develop a method to reliably measure acuities for isolated lowercase letters and words of differing length that can be used to test low-vision patients. Using isolated stimuli means that testing is relatively free of potential crowding and/or distracting attentional effects from surrounding words, it is unambiguous which stimulus subjects are trying to read and response times can be recorded for each stimulus. Across a series of experiments, subjects with normal vision were asked to read isolated lowercase single letters and lowercase words of 4, 7 and 10 letters, in separate tests. Acuities for uppercase Sloan letters were also measured to provide a reference, as they are commonly used to measure visual acuity. Each test was based upon the design principles and scoring procedures used in the Bailey–Lovie and ETDRS charts. Acuities for uppercase Sloan letters were found to be equivalent whether measured using ETDRS charts or the computer-based method. Measurement of acuities for lowercase single letters and lowercase words of 4, 7 and 10 letters had a reliability that was no worse than acuities for uppercase Sloan letters. Lowercase word acuities were essentially independent of word length. Acuities for single lowercase letters and lowercase words were slightly better than uppercase Sloan letters acuity. Optimal processing of lowercase single letters and 4-, 7- and 10-letter words occurred at character sizes that were at least 0.2–0.40 log MAR above acuity threshold, i.e. between 1.5 and 3 times threshold acuity for that particular stimulus. In general, critical character sizes appear similar across word lengths as progressive increases or decreases in these values were not observed as a function of the number of letters in the stimulus. We conclude that a computer-based method of stimulus presentation can be used to obtain highly repeatable measures of acuity for lowercase single letters and lowercase words in normal vision.

Keywords: acuity reserve, critical character size, word acuity, word length

Introduction

One of the main goals in the management of low-vision patients is to maximise reading ability. Reading is particularly difficult for patients with a central scotoma

(Legge *et al.*, 1985b). Two reasons that make reading difficult for this low-vision population are the obstructive aspect of the scotoma and the need to rely exclusively on information coming through the eccentric visual field to accomplish this task. That visual field loss alone can inhibit reading simply by obstructing the visual field, even when spatial resolution is relatively normal, is apparent from studies of reading performance in individuals with peri-central visual field loss (Trauzettel-Klosinski and Tornow, 1996), peripheral field loss (Virgili *et al.*, 2004) and hemianopic visual field loss (Trauzettel-Klosinski and Brendler, 1998; Trauzettel-Klosinski and Reinhard, 1998). In these studies it has

Received: 8 November 2004

Revised form: 15 July 2005

Accepted: 10 August 2005

Correspondence and reprint requests to: Dr Andrew R. Whatham.

Tel.: +41 22 38 28379; Fax: +41 22 346 8405.

E-mail address: andrew.whatham@hcuge.ch

been shown that reading speed is reduced even though fixation remains at, or very close, to the centre of the fovea so that normal visual acuity can be attained. Also, the closer the visual field defect is to fixation, the greater is the negative impact of the region of visual field loss on reading speed. Visual field loss appears to reduce reading speed through reducing the field of view around the point of fixation. This may be through a mechanism of reducing the perceptual span [area around the fixation point which influences eye movements during reading (Rayner, 1994)] or the visual span [the number of letters around the fixation point that can be recognised at a glance (Legge *et al.*, 1997)].

It has been consistently demonstrated that there must be fundamental limitations of text processing, somewhere in the visual pathways, that prevent normal reading speeds being attained through the eccentric visual field. First thresholds for identifying words are raised. Word acuity thresholds increase with eccentricity in the normal visual field (Abdelnour and Kalloniatis, 2001), as presumably do thresholds for static text in a paragraphed format. Word acuity is also worse in central scotoma populations, many of whom have eccentric fixation, than it is in normal vision both in absolute terms as well as relative to letter visual acuity (Bullimore and Bailey, 1995; Lovie-Kitchin *et al.*, 2000; Wolffsohn and Eperjesi, 2004). This necessitates prescription of magnification to achieve accurate word identification when only the eccentric visual field will be used to read them. Second, reading speeds do not often, if ever, return to normal levels even for suprathreshold character sizes that are above the minimum acuity reserve (Whittaker and Lovie-Kitchin, 1993). This has been reported for both static text passages and text presented through a rapid serial visual presentation (RSVP) paradigm (Rubin and Turano, 1994). A variety of methodologies, including RSVP and scrolled presentation paradigms, have been used to show that maximum reading speed declines with eccentricity in subjects with normal vision (Culham *et al.*, 1992; Rubin and Turano, 1994; Chung *et al.*, 1998).

Understanding the limiting mechanisms on reading performance in the case of eccentric fixation is important with regard to providing an explanation for poor reading performance in the central scotoma population, and identifying what factor(s) need to be overcome in order to improve reading ability. It has been proposed that the visual span is reduced in the eccentric visual field (Legge *et al.*, 1997, 2001), which accounts for poor reading speed. This idea comes from the observation that under conditions of (1) low stimulus contrast or (2) restricting stimulus presentation to an eccentric location(s) in the visual field, there is a greater slowing in RSVP reading speed for longer words compared to shorter words (Legge *et al.*, 1997). As crowding or contour interaction

effects are more prominent in eccentric vision than they are in the central visual field (Jacobs, 1979; Leat *et al.*, 1999), their impact on word recognition and reading speed needs to be integrated into the idea of a reduced visual span, if the cause of poor reading performance in central scotoma populations is to be fully understood. While it is a theoretical possibility that increased crowding in eccentric vision causes a reduced visual span, whether and how these variables interact is not yet clear. A detailed study of thresholds for different word lengths in eccentric vision might shed some light on these issues, by adding new information to what is already known about word identification speed as a function of word length, and provide a more complete picture of word recognition in low vision.

In clinical research, word acuity and reading speed for lowercase words can be measured in English using the Bailey–Lovie word reading chart (Bailey and Lovie, 1980), in German using the Radner reading charts (Radner *et al.*, 1998, 2002) and in several languages including English using the MNRead tests (Legge *et al.*, 1989). These reading tests allow the measurement of several parameters of reading performance including threshold print size, critical print size [minimum print size to read at a maximum rate – e.g. (Mansfield *et al.*, 1996)] and maximum reading speed. Visual performance in age-related macular degeneration appears to be related to task complexity in that grating and letter acuities, for example, are relatively easier compared to reading crowded letters and words (Kitchin and Bailey, 1981; Bullimore and Bailey, 1995). As yet, word acuities have not been investigated as a function of the number of letters in a word, to our knowledge, in eccentric vision. Current clinical tests that are designed to measure word acuity contain words of differing length at each character size such as 4, 7 and 10 letters in the Bailey–Lovie word chart (Bailey and Lovie, 1980) or the MNRead test, for example. While the use of different word lengths is advantageous in obtaining a measure of word acuity that can be considered to be representative of the range of word lengths encountered in reading material, this method effectively combines acuities for different word lengths into a single measure so that any effect of word length on word acuities may be masked.

It would be possible to use multiple word charts to obtain an acuity score for a single word length at each character size. However presenting words in isolation might be a better measurement protocol, because any possible crowding or distracting attentional effects from surrounding words that might interfere with word recognition could be avoided, in particular for patients with eccentric fixation (Chung, 2004). Also it is conceivably easier to score letter and word acuities using a system of isolated presentation, because it is never ambiguous which stimulus the subject is trying to read.

This can be an issue in the central scotoma population because the presence of a scotoma and crowding effects in eccentric vision can cause subjects to mislocalise stimuli or lose stimuli within the scotoma. Isolated presentation of stimuli is very easy to achieve using a computer-based test. A computer-based test would also allow automatic word randomisation, automatic calculation of acuity threshold, easy manipulation of stimulus parameters (such as reductions in stimulus contrast), and recording of response times.

Several computer-based systems for measuring visual acuity have already been developed, using various types of stimuli and scoring procedures (e.g. Arditi and Cagnello, 1993; Bach, 1996; French, 1997). These systems have not yet been extended to incorporate measures of word acuity, which might prove extremely valuable in assessment of visual function and low vision patients. The principal aim of the current study is to develop a methodology that can be used to reliably measure acuity for individual lowercase letters and words of differing length. If a reliable methodology can be found, this would potentially be useful for measuring acuity for any lowercase stimuli in any language. In the current study we investigated, in normal vision, the relationship between word acuity measures for lowercase letters and single (Sloan) uppercase letter acuity, and analysed the response times recorded during all acuity tests.

Methods

Subjects

All subjects participating in the study were aged between 20 and 44 years. No subjects had any evidence of ocular pathology and all were in good general health. Subjects wore their habitual optical correction, if they could attain ETDRS acuity of better than 0.1 log MAR in each eye with it. Any subject in whom this was not possible in either eye, underwent a full subjective refraction to determine the spherocylindrical correction that gave the best visual acuity. In these cases, subjects wore the new correction in a trial frame and performed the tests through full-aperture ophthalmic trial lenses. Each subject wore the same optical correction, if one was used, in all tests. Informed consent was obtained from each subject prior to their participation and all test procedures conformed to the Declaration of Helsinki and had been approved by the local ethical committee.

General study design

This study is divided into four parts.

Part 1. Uppercase letter acuity (Sloan set) was compared between ETDRS charts and the new

computer-based method (CB acuity) to obtain an idea of test calibration. Sloan letter acuity was measured in 30 subjects using ETDRS charts (charts 1, 2 or R at a background luminance of 250 cd m⁻²) and the computer-based method, using a background luminance of 100 cd m⁻². In an additional experiment eight subjects performed both tests but also read an ETDRS chart through a neutral density filter (Corion Corporation, Holliston, MA, USA) that reduced the background luminance to 100 cd m⁻², in order to equalise background luminance between the ETDRS and CB tests.

Part 2. The repeatability of uppercase Sloan letter acuity and separate acuities for lowercase words of 4, 7 and 10 letters was assessed using the computer-based method. Lowercase word groups were composed of the most frequent words at each word length.

Part 3. The relationship between acuities for lowercase stimuli at different word lengths was further investigated using lowercase single letters and word groups, balanced for frequency, across the 4-, 7- and 10-letter length words. The repeatability of lowercase letter acuity was first determined in ten subjects. A group of nine subjects, who did not participate in *Part 2*, underwent testing for uppercase Sloan letters, single lowercase letters and 4-, 7- and 10-letter lowercase words. Word frequencies were equalised across word groups. Acuity was measured at each word length in each subject using only one word group.

Part 4. Repeated acuity measures were obtained using a single word group. This was performed to determine whether a single word group can be used to obtain repeatable measures of word acuity, when word position in the stimulus sequence is randomised, or whether different word groups must be used. Five subjects, who did not participate in the other word acuity tests, were tested on five occasions in one 2-h session using only a single word group. This was one of the seven letter word groups from the First Words library. Subjects were told before testing that the same word group would be re-used each time.

The test design and methods for determining spatial resolution followed the general principles for acuity measurement outlined by Raasch *et al.* (1998).

Instrumentation and stimuli

Test stimuli were uppercase single letters, lowercase single letters and lowercase words of 4, 7 and 10 letters. All stimuli were presented on an 18 in. CRT monitor (P910, Hewlett Packard, Palo Alto, CA, USA) using a spatial resolution of 1280 × 1024 pixels and a temporal resolution of 75 Hz. All stimuli were composed of black

letters (luminance 0.1 cd m^{-2}) on a white background (luminance 100 cd m^{-2}), which resulted in a luminance contrast greater than 99% (Michelson contrast). Direct measurement of the phosphor decay time with an oscilloscope indicated that luminance reduced from the background luminance (100 cd m^{-2}) to 10% of the foreground luminance (0.1 cd m^{-2}) at 4 ms. The uppercase letter stimuli consisted of the 10 Sloan letters used in the ETDRS log MAR charts (C, D, H, K, N, O, R, S, V, Z). These stimuli were constructed in a 5×5 block format using Photoshop, identical in style and proportion to the letters in ETDRS log MAR charts. Anti-aliasing of stimuli was achieved using Photoshop 5.0 (Adobe). This allowed accurate representation of stimulus sizes and smooth text contours at smaller sizes. Each stimulus bitmap was created manually. The total size of the anti-aliased letters was defined using the letter 'O' as a reference. First, we fitted two inverse Gaussian functions, independently, across the upper and lower stroke widths of the letter 'O' along the central vertical axis of this letter. We then found the minima of these two functions using the pixel greyscale values converted to real luminance values, based upon the monitor's measured relationship between greyscale values and luminance output using a photometer (S370 Optometer, United Detector Technology, Baltimore, MD, USA). The distance between these two minima was taken to be 80% of the total letter size. Total letter size was calculated from the stimuli as $1.25 \times$ the centre-to-centre separation of the upper and lower stroke widths. Exactly the same methods that were used to create the different stimulus sizes for the letter 'O' were used to create all the other Sloan stimuli. Once the total letter size was known, the log MAR value of each letter was calculated as:

$$\log \text{MAR} = \log_{10} \frac{(\text{total letter size in min arc})}{5}, \quad (1)$$

where *letter size* refers to the outer dimensions of the letter in the vertical direction.

Fifteen letter sizes were created in 0.1 log MAR steps, producing a stimulus range from 1.0 log MAR to -0.4 log MAR at a 4 m test distance. The smallest character size stimuli (corresponding to -0.4 log MAR at 4 m) consisted of luminance information derived from 10 pixels in the vertical direction. Thus all stimuli had at least 10 pixels in the vertical direction.

The lowercase letter and word stimuli were composed in Courier Bold font. This font was used because it inherently contains equal inter-letter spacing and the horizontal angular extent of each word stimulus would be constant across all words at a single character size. Also, stroke width is approximately 1/5th of the common body height (also referred to as the *x-height*)

across all 26 lowercase letters. In this way the relationship between stroke width and total letter size (uppercase Sloan) and stroke width and *x-height* (lowercase) are similar. For this reason character size was defined in a similar manner for both types of stimuli in Equations (1) and (2).

Two word libraries were created from the online French lexical database *Lexique 2* (<http://www.lexique.org>), using the FastSearch frequencies that are derived from the number of web pages containing the particular word. In either of these libraries, if the same word family was present more than once, but with a different ending (e.g. singular and plural form of the same noun or different person or tense of the same verb) only the most frequently used word form was used. Proper nouns, names and recent words of English origin were also excluded from the test list.

First a word library was created from the 165 most frequently used French words for 4-, 7- and 10-letter words (*hereafter referred to as First Words*). A second library (*hereafter referred to as Balanced Words*) was created consisting of 165 common 4-, 7- and 10-letter words of equal frequency between groups (all words in each group equally spaced between 4000 and 40 000 occurrences per million web pages). Some words were common to both libraries. In both libraries, the 165 words, at each word length, were then randomly allocated to one of three lists of 55 words using random number generation software in each case (labelled groups A, B and C). Fifty-five words enabled a different word to be presented in the test format from 0.6 to -0.4 log MAR (11 levels of character size by five stimuli at each character size) in a single test if all words were presented. The anti-aliased lowercase single letters and words in Courier Bold font were also defined according to the letter 'o'. In this font, stroke width is slightly more than 1/5th of letter size. We measured this to be 0.2105. Therefore the centre-to-centre separation of the upper and lower stroke widths, along the vertical axis of this letter, constitute 78.95% of the total *x-height*. To obtain the total *x-height*, we multiplied the centre-to-centre separation by 1.267. Once the total *x-height* of the stimuli had been determined, the log MAR value for words was calculated, in the same way as for the uppercase Sloan letters:

$$\log \text{MAR} = \log_{10} \frac{(\text{letter } x\text{-height in min arc})}{5}, \quad (2)$$

where *letter x-height* refers to the vertical extent of a lowercase letter 'x', which corresponds to the common body of the lowercase letters excluding ascenders and descenders. The above equation has also been used by others to define character size e.g. (Mansfield *et al.*, 1996). The *x-height* of the smallest lowercase letter and

word stimuli ($-0.4 \log \text{MAR}$) was defined by luminance information from 10 pixels in the vertical direction.

The largest character size at which 10-letters could fit onto the monitor display was $0.6 \log \text{MAR}$. Therefore this value was the maximum test value that could be used for the 10-letter words at the 4 m test distance. The starting value for all lowercase stimuli was set to $0.6 \log \text{MAR}$ to make all word acuity tests equivalent in terms of the possible range of character sizes presented. Pilot studies showed that subjects with normal vision do not make errors between 0.6 and $1.0 \log \text{MAR}$ so that not presenting them would not affect acuity scores. However the starting value of $0.6 \log \text{MAR}$ was well above threshold for all stimuli presented, so that reducing the test distance to increase the starting value from 0.6 to $1.0 \log \text{MAR}$ was considered to unnecessarily increase testing time in the present study. During a test with a single word group, each word was allocated at random to one of the 55 positions in the stimulus sequence. This randomisation procedure was performed each time the same word group was reused in different subjects. In this way testing different subjects with the same word list resulted in a different word order as well as different words at each character size between tests.

A control experiment was performed to determine whether monitor characteristics limited spatial resolution measures. This experiment is described in Appendix 1. Specifically, uppercase and lowercase letter acuities were measured at 4 and 8 m and were found to be similar, indicating that the pixelisation characteristics of the stimuli did not limit acuity measurements.

Test procedure

All tests were performed monocularly, with the same eye being used for all tests that were performed by the same subject. The contralateral eye was occluded. Test order, and the eye used, were randomised and balanced across subjects in each case. Test distance was 4 m in all tests.

For all acuity programmes, one stimulus was presented at a time to each subject in the centre of the monitor. Five stimuli were shown, consecutively, at each character size, beginning with the starting (maximum) character size. Once subjects had attempted to identify each of the five stimuli, character size was reduced by $0.1 \log \text{MAR}$. Testing stopped once four or five errors were made at the same character size (Carkeet, 2001). Subjects were asked to respond to all stimuli and encouraged, but not forced, to guess if they were unsure – thus they were also allowed to respond ‘don’t know’. Subjects were told the stimulus length, in number of letters, before each test began to standardise this variable between tests. It was apparent in pilot studies that if we did not do this, subjects invariably asked us

midway through a test whether all stimuli had the same number of letters. Thus if their knowledge of test word length changed during a test – based either on their own intuition or our response to their question – this may have introduced additional variability in the measures.

Subjects were under no time pressure. However an upper limit of 30 s, known only to the experimenters, was set: stimuli would be scored as incorrect if they had not been correctly read by this time. This was rarely reached as subjects tended to respond ‘don’t know’ if they could not identify stimuli within about 10 s.

Determination of acuity measures

The calculation of acuity scores was identical for each test stimulus. The only difference between test procedures was that the uppercase letters began at $1.0 \log \text{MAR}$, so that these results could be compared to ETDRS acuity, while the presentation of all lowercase stimuli began at $0.6 \log \text{MAR}$. All acuities were calculated using letter-by-letter and word-by-word scoring with a value of $0.02 \log \text{MAR}$ assigned to each stimulus. This follows the principle in scoring visual acuity from $\log \text{MAR}$ charts outlined by Bailey (1980). A word was only scored as correct if it was completely correct. Even if a subject said a word that was only different from the stimulus by a single letter, regardless of the word length, it was scored as being incorrect. The acuity score in each test was calculated as:

$$\text{Acuity} = \text{SCS} + (E \times 0.02), \quad (3)$$

where SCS is the Smallest Character Size attempted (at which a minimum of four errors were made) in $\log \text{MAR}$ and E is the total number of errors made.

Response time measurement

As each stimulus was presented in isolation to the subject it was possible to record a response time for each stimulus. This duration was the time between the appearance of the stimulus on the monitor until the subject had finished reading the stimulus. This time was determined by the examiner’s key-press denoting in the data set that the response was correct or incorrect. Response times measured in this way were considered to incorporate a visual processing time, the subject’s verbalisation time and the examiner’s manual response latency. The logic behind analysing response times, as we measured them, was to keep the subject’s verbalisation time and the examiner’s manual response latency constant within all tests so that any variations in the measured response times across character sizes would reflect variations in visual processing. Even though the visual processing time might be quite small relative to the

other two components, especially for single letters, our logic was that non-random variations in visual processing can be distinguished from random variations in both the subject's verbalisation time and the examiner's response latency through averaging. Response times were analysed up to, and including, the smallest character size at which at least four correct responses were made.

Data analysis

Statistical testing was carried out using Sigmasat Version 3.1 statistical software. Comparisons between data groups were made using both parametric (repeated measures ANOVA and paired *t*-tests), and non-parametric tests (Friedman repeated measures ANOVA on ranks and Wilcoxon sign-ranked test) when the data were not normally distributed.

Results

Part 1

ETDRS acuity compared to computer-based acuity. Mean ETDRS acuity, at 250 cd m^{-2} , was -0.06 ± 0.017 (S.E.) log MAR compared to a mean CB acuity, at 100 cd m^{-2} of -0.05 ± 0.023 (S.E.) log MAR in the group of 30 subjects. There was no statistically significant difference between the two techniques (paired *t*-test, $t = 1.487$, d.f. = 29, $p = 0.15$). In the eight subjects that performed both the above tests as well as reading an ETDRS chart through a neutral density filter, there was no statistically significant difference in acuities between ETDRS and CB measures (paired *t*-test, $t = -1.768$, d.f. = 7, $p = 0.12$). Mean ETDRS acuity, at 100 cd m^{-2} , was 0.03 ± 0.034 (S.E.) compared to a mean CB acuity of 0.06 ± 0.033 (S.E.) for the eight subjects tested.

Part 2

Comparison of word groups. Acuities from all 10 subjects for each word group, within each word length set, are shown in Figure 1. Mean acuities (\pm S.E.) for the three word groups (referred to as A, B and C), of four-letter words were -0.15 ± 0.024 , -0.17 ± 0.020 and -0.15 ± 0.023 , respectively. Mean acuities for the three seven-letter word groups were -0.17 ± 0.020 (A), -0.16 ± 0.021 (B) and -0.15 ± 0.016 (C). Acuities were also similar for the three 10-letter word groups: (A) -0.17 ± 0.020 , (B) -0.17 ± 0.022 and (C) -0.16 ± 0.024 .

The difference in mean acuity between any of the word groups, at a particular word length, was always 0.02 log MAR or less. There were also no statistically

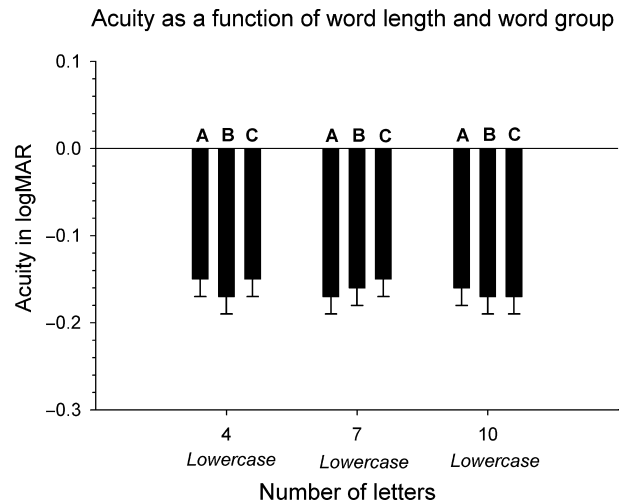


Figure 1. Mean word acuity (\pm S.E.) for each of the three word groups (A, B and C) from the First Words library for each word length. Acuity is in log MAR and word length is expressed according to the number of letters. Note that acuity is similar for each of the word groups used at each word length and also similar across word lengths.

significant differences in acuity between the word groups for four letter words (repeated measures ANOVA on ranks, $\chi^2 = 2.552$, d.f. = 2, $p = 0.28$), seven letter words (repeated measures ANOVA, $F = 0.947$, d.f. = 2, $p = 0.41$), or 10-letter words (repeated measures ANOVA, $F = 0.61$, d.f. = 2, $p = 0.55$).

Repeatability of acuity measures for uppercase Sloan letters and lowercase words. Test-retest reliability was calculated for the uppercase single letters and each word length by analysing the difference in acuity measures between the first and second test in the case of the single letters, and the word groups A and B for the 4-, 7- and 10-letter words. Test-retest reliability for the single letters was ± 0.13 log MAR compared to ± 0.11 log MAR for four-letter words, ± 0.06 log MAR for seven-letter words and ± 0.06 log MAR for 10-letter words. The variance of test-retest measures was equivalent for uppercase single letters and four-letter words (F -test, $p = 0.31$). The test-retest variance was significantly lower for seven-letter words (F -test, $p < 0.05$) and for 10-letter words (F test, $p < 0.01$) compared to the uppercase single letter measures.

Word acuity as a function of the number of letters – First Words library. Mean acuities \pm S.E. for all lowercase word lengths and the uppercase Sloan letters are shown in Figure 2a. Sloan letter visual acuity averaged -0.09 ± 0.029 , while mean acuity for the four-letter words was -0.15 ± 0.019 , for the seven-letter words was -0.16 ± 0.017 and for the 10-letter words was -0.17 ± 0.021 . There was no significant difference between acuities for 4-, 7- and 10-letter words (repeated

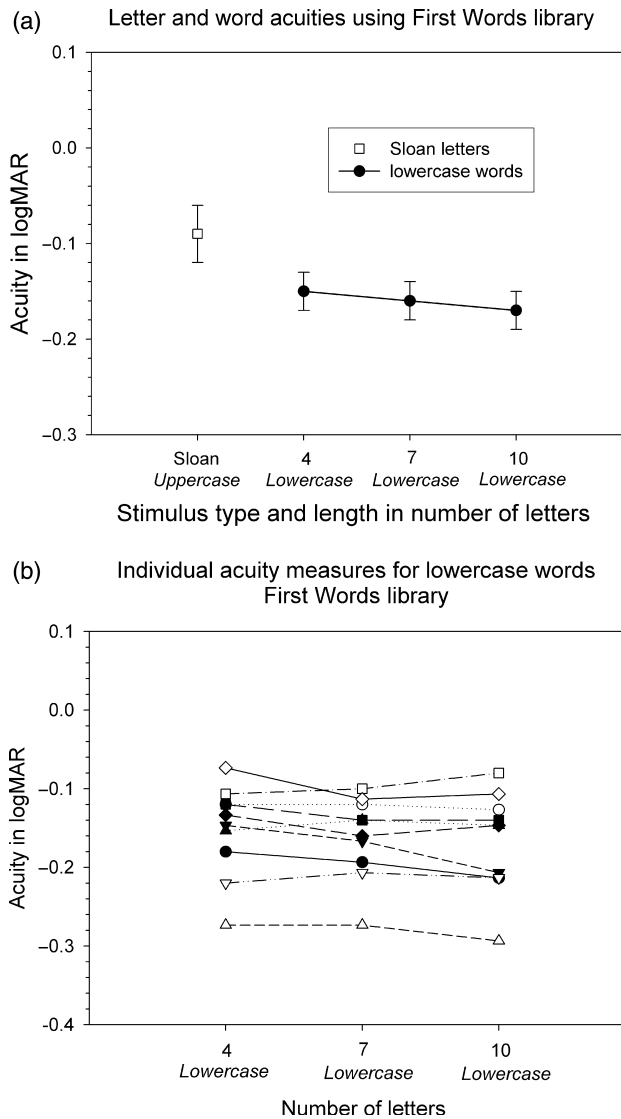


Figure 2. (a) Mean acuity (\pm S.E.) for single uppercase Sloan letters (hollow-square) and lowercase words of 4, 7 and 10 letters (filled circles) for all 10 subjects. The mean value for single letter acuity is calculated from the average of three acuity tests in each subject, while mean word acuity for each word length is calculated using the average of the three word groups. Acuity is in log MAR and word length is in number of letters. (b) Individual acuity measures for lowercase words of 4, 7 and 10 letters for each of the 10 subjects. Note the low variability in acuity measures across word lengths in individual subjects.

measures ANOVA on ranks, $\chi^2 = 2.324$, d.f. = 2, $p = 0.31$). However, there was a statistically significant difference between acuity for the uppercase Sloan letters and 4-, 7- and 10-letter words (repeated measures ANOVA on ranks, $\chi^2 = 13.206$, d.f. = 3, $p < 0.01$). Sloan letter acuity was 0.06 log MAR worse than acuity for four-letter words but not statistically significant (Tukey multiple comparison test; $p > 0.05$), yet 0.07 log MAR

significantly worse than acuity for seven-letter words (Tukey multiple comparison test; $p < 0.05$) and 0.08 log MAR significantly worse than acuity for 10-letter words (Tukey multiple comparison test; $p < 0.05$).

Individual mean acuity measures for each text stimulus are shown in *Figure 2b*. It is evident in this figure that acuity (average of the three acuity tests for each subject) did not differ by more than 0.1 log MAR between any of the three word lengths measured in any subject.

Response time during acuity testing. The response times recorded during all acuity tests are shown in *Figure 3* for (a) single letters and (b) 4, 7 and 10-letter words. There was a significant effect of character size on response time for all text stimuli, so that response time increased as stimuli approached threshold (Friedman repeated measures ANOVA on ranks, $p < 0.001$ for uppercase Sloan letters and 4-, 7- and 10-letter lowercase words). The threshold values in *Figure 3* are the smallest character size at which at least four correct responses were made – equating to an 80% correct size. These were within 0.1 log MAR of the thresholds calculated using letter-by-letter scoring. The variability of response times also increased as stimuli approached threshold. In each graph there is a similar relationship between response time and character size above threshold, in which response time is slowest at threshold and follows an exponential decline, asymptoting towards a minimum value. There was a significant difference between response times for the 4-, 7- and 10-letter words (repeated measures ANOVA, $F = 47.75$, d.f. = 2, $p < 0.001$), with response time increasing as word length increased. However there was no interaction between word length and character size on response time (repeated measures ANOVA, test by word length, $F = 0.76$, $p = 0.71$), indicating that the relationship between character size and response time, including character sizes close to and at threshold, is not different for words of different length. For each stimulus length a critical character size was calculated which was considered to be the minimum size needed for optimum processing (i.e. read in the shortest time). The steps taken to make this calculation are described in Appendix 2. Calculated critical character sizes were 0.33, 0.21, 0.21 and 0.28 log MAR for the uppercase Sloan letters and lowercase words of 4, 7 and 10 letters, respectively.

Part 3

Word acuity as a function of word length – Balanced Words library. Acuities for the uppercase Sloan letters, single lowercase letters and the 4-, 7- and 10-letter words

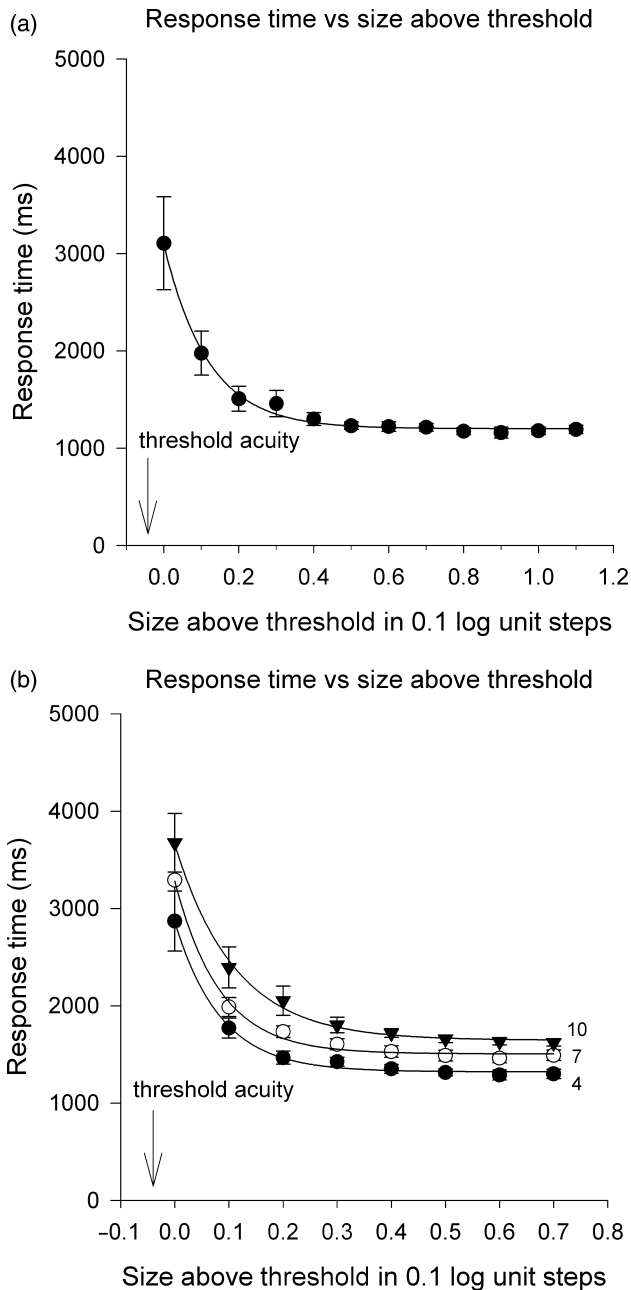


Figure 3. Mean (\pm S.E.) response time measures as a function of character size above threshold (threshold here is the smallest character size at which at least four correct responses were made equating to an 80% correct size) for (a) single uppercase letters and (b) four (filled circles), seven (open circles) and 10 (triangles) lowercase letter words. Response time is in milliseconds and character size in 0.1 log unit steps (above threshold). Also plotted is the best fitting exponential decay function of the form $y = y_0 + ae^{-bx}$. Each data point represents the mean of at least 100 response time measures. Data from different tests and different subjects were averaged according to the smallest character size at which four stimuli were correctly identified. The first data point is not more than 0.1 log MAR greater than the mean threshold acuity value for each stimulus type. This offset is denoted in the figure. Note that for all text stimuli the variance of response time increases as threshold size is approached and that response time is a minimum for character sizes at least 0.4 log MAR larger than threshold.

words were significantly better (Holm–Sidak method for multiple comparisons). There was no significant difference in acuity across all the lowercase stimuli (repeated measures ANOVA, $F = 0.925$, d.f. = 3, $p = 0.44$).

Repeatability of lowercase letter acuity. There was no significant difference in lowercase letter acuity between the first and second tests (paired t -test, $t = 0.307$, d.f. = 9, $p = 0.77$). Mean acuity was -0.06 ± 0.025 (S.E.) log MAR for the first test and -0.07 ± 0.030 (S.E.) log MAR for the second test. There was also no difference between uppercase Sloan acuity in the two tests (paired t -test, $t = 0.39$, d.f. = 9, $p = 0.49$). Mean Sloan acuity was -0.01 ± 0.022 (S.E.) log MAR for the first test and 0.00 ± 0.021 (S.E.) log MAR for the second test. Test–retest reliability was not different between the two letter acuity tests (F -test, $p = 0.91$), where confidence limits (± 1.96 S.D.) were ± 0.12 - log MAR for the lowercase letters and ± 0.14 log MAR for the uppercase Sloan letters.

Response times for lowercase letters. The response times, as a function of character size above threshold, for the lowercase letters are shown in Figure 4b. Response time also clearly varied as a function of stimulus size above threshold in this group (Friedman repeated measures ANOVA on ranks, $\chi^2 = 37.286$, d.f. = 5, $p < 0.001$). Calculated critical character size for the isolated lowercase letters was 0.40 log MAR above threshold.

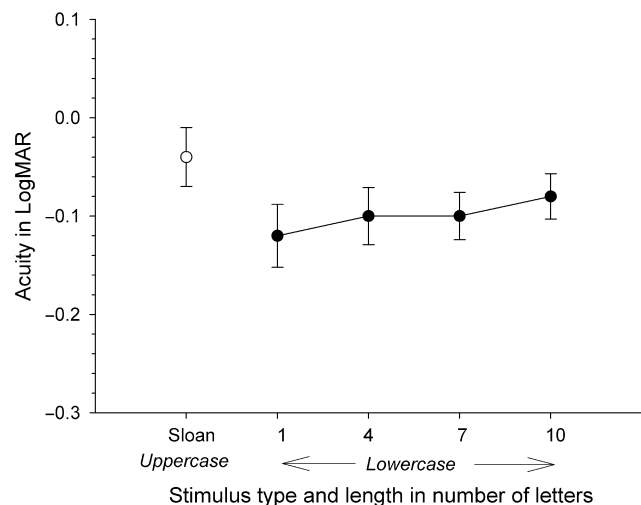
Part 4

Repeated use of the same word group. Changes in word acuity measures for one of the seven-letter word groups over five consecutive tests in five subjects over 2 h are

for the Balanced Words library are shown in Figure 4a. Acuities (mean \pm S.E.) were -0.04 ± 0.030 log MAR for the uppercase Sloan letters, -0.12 ± 0.032 log MAR for the single lowercase letters, -0.10 ± 0.029 log MAR, for the four-letter words, -0.10 ± 0.024 log MAR for the seven-letter words and -0.08 ± 0.023 log MAR for the 10-letter words.

Mean acuities were not equal across all groups (repeated measures ANOVA, $F = 4.192$, d.f. = 4, $p < 0.01$). Acuities were slightly better for all the lowercase stimuli compared to the uppercase Sloan letters, but only the single letters and seven-letter

(a) Letter and word acuities using Balanced Words list



(b) Response times for lowercase letters

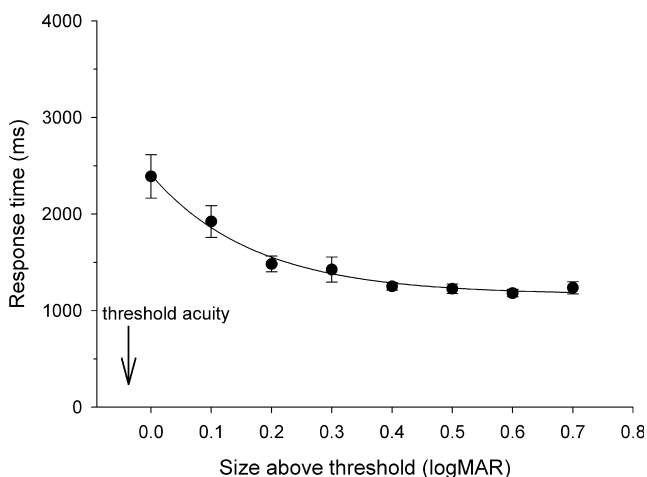


Figure 4. (a) Mean acuity (\pm S.E.) for nine subjects with normal vision for uppercase Sloan letters, as well as for lowercase single letters and words of 4, 7 and 10 letters using the balanced frequency word groups, all measured using the computer-based method. All acuities were calculated from a starting character size of 0.6 log - MAR. (b) Mean response times (\pm S.E.) for isolated lowercase letters for 10 subjects. The first data point is not more than 0.1 log MAR greater than the mean threshold acuity value. This offset is denoted in the figure.

shown in *Figure 5*. Mean acuity for these words significantly improved with increasing test experience (repeated measures ANOVA, $F = 14.195$, $d.f. = 4$, $p < 0.001$). Acuity progressively improved from the first to the fourth test – after which no further improvement was seen. Mean acuity (\pm S.E.) at the first measurement was -0.09 ± 0.033 log MAR, improving to -0.16 ± 0.030 log MAR at the fourth measurement, and regressing slightly to -0.13 ± 0.023 log MAR in the fifth measurement.

Repeated testing using a single word group

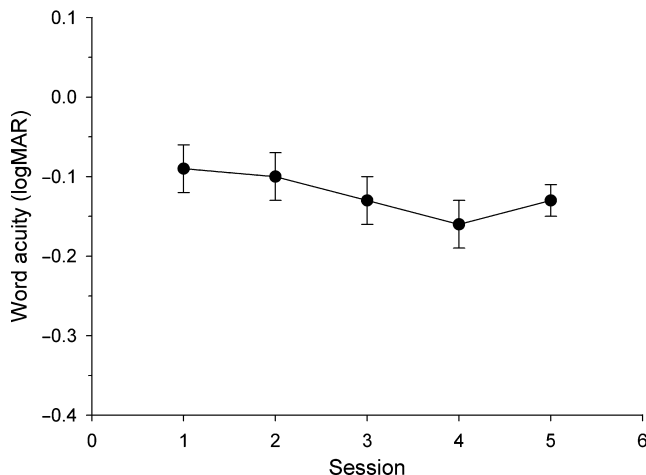


Figure 5. Word acuity measures as a function of repeated testing using the same seven-letter word group. The word group used was one of the seven letter groups from the First Words library. Note a small but clear improvement in acuity with repeated testing, which arrested at the fourth test occasion.

Discussion

Letter and word acuity can be reliably measured using a computer-based test

Despite the great variability in the global shape of the lowercase words we observed a good repeatability of acuity measures for these stimuli. Test–retest reliability for uppercase letter acuity was ± 0.13 log MAR in *Part 2* and ± 0.14 log MAR in *Part 3*, which are similar to the repeatability reported using ETDRS, Bailey–Lovie or Snellen charts (Lovie-Kitchin, 1988; Raasch *et al.*, 1998; Siderov and Tiu, 1999; Lovie-Kitchin and Brown, 2000; Ruamviboonsuk *et al.*, 2003; Rosser *et al.*, 2004), and similar to the repeatability observed for four-letter words in the present study. We observed a better test–retest reliability for the 7- and 10-letter words compared to the uppercase Sloan letter reliability.

It was especially evident, though, during word acuity testing that subjects used the visual form (shape) of the word in trying to identify stimuli close to threshold. If a word close to threshold was read incorrectly, subjects frequently guessed a word that, although wrong, had the correct number and position of letters with elements outside the *x*-height (ascenders and descenders) such as guessing the word ‘sait’ when the word ‘voit’ was presented. The importance of word shape as a clue in accurate identification has been demonstrated by others (e.g. Nazir *et al.*, 1992). However, although lowercase stimuli have varying global shapes that can aid in word identification, this does not make acuity testing unreliable. This is probably due to the fact that the global

shape of a lowercase word is rarely, if ever, unique to a word.

Overall, this indicates that word acuity can be reliably measured using a computer-based stimulus presentation method *without* needing to consider the visual form of test words. This might be applicable to measuring acuities for lowercase words in any language. It may have been possible to get better letter and word acuities, and improved test–retest reliabilities, if we had refracted all subjects and managed to find, and correct for, any uncorrected residual refractive errors. Recent data has emphasised that small uncorrected refractive errors can lead to poorer repeatability of ETDRS acuity measures (Rosser *et al.*, 2004). However as visual acuity was very good in our subjects these were unlikely to be very large. Also, each subject used the same optical correction in all tests and the experiments were designed to look at comparisons between acuity tests, under the same optical conditions.

Statistically significant improvements in acuity were measured when using the same word group on repeated occasions (Figure 5). These effects can be attributed to memorising some of the words in the test, thereby reducing the number of possible words that could be on the stimulus screen in subsequent tests. This indicates that although repeated use of the same words with a single subject would not be wildly variable, different word groups should be used where statistical comparisons will be made between multiple word acuity measures in the same individual.

It was also observed that highly repeatable measures of lowercase single letter acuity, using all 26 letters of the alphabet, could be obtained using a randomised presentation format (Part 3). This is interesting because even though some individual letters have been well demonstrated to have better legibility than others in acuity tests, randomly selecting a letter from the alphabet at each stimulus presentation consistently gives similar results when using a Bailey/Lovie or ETDRS design. Thus the design used in the present study can be used to reliably measure acuity for lowercase single letters and words of differing length, each of which has equivalent repeatability that is also comparable to repeatability for acuity measures using uppercase Sloan letters.

Acuity for lowercase words is independent of the number of letters in a word in normal vision

Acuities were equivalent for lowercase stimuli whether they were single letters or 4-, 7- or 10-letter words and whether word groups based on the most frequent words at each word length (*First Words library*) or on words of equal frequency (*Balanced Words library*) were used. There was a slight tendency for better acuity for the longer words using the *First Words* library (Figure 2a)

but worse acuity for longer words using the *Balanced Word* library (Figure 4a). However in neither case were there statistically significant effects. In the normal reading process, some characteristics of word identification depend upon word length where response latencies, for example, increase with an increasing number of syllables in a word (e.g. Ferrand and New, 2003). However such an effect of word length on lexical processing, in the French language, is evident for low-frequency words but not high-frequency words (Ferrand and New, 2003).

The fact that equivalent acuities were obtained for words of 4, 7 and 10 letters using either the *First Words* library or the *Balanced Words* library indicates that word acuity is relatively insensitive to frequency changes, at least for the range of relatively common words used in the present study. It may have been possible to observe an effect of word frequency on acuity measures if words of much lower frequency were used.

Our findings indicate that acuity can be considered to be relatively invariant across word lengths in normal vision. Therefore it appears that extra letters in a word do not make it more difficult to resolve a word. To correctly decipher a word requires, at the lowest level, serial identification of each letter so that a word can be read. So for stimuli close to threshold, when the probability of correctly identifying a word is less than the maximum value for suprathreshold stimuli, one is more likely (in probability terms) to correctly guess a single letter than to correctly guess 4, 7 or 10 single letters correctly. However the increased burden of having to correctly guess a longer string of letters (in longer words) appears to be offset by the fact that the combination of individual letter guesses must form a known word and, for lowercase letters as used in the present study, must conform to the shape of the word in terms of the number and position of letters with elements above or below the *x*-height.

As only Courier font was used in the present study it is not yet known whether this finding generalises across font types – particular for proportionally spaced fonts such as Times New Roman. Font type has been previously shown in several studies to affect reading performance (Arditi *et al.*, 1990; Mansfield *et al.*, 1996).

Comparison between lowercase word acuity and upper and lowercase single letter acuity

We did not expect to find word acuity to be as good as that for single lowercase letters, or as good or better than acuity for single uppercase letters, given that words could be conceived as crowded stimuli and crowding effects in normal vision have been reported when stimuli are close to letter acuity threshold (Simmers *et al.*, 1999; Hess *et al.*, 2000; Levi *et al.*, 2002). Also, Arditi reported

that acuity for uppercase single letters was 0.1 log MAR better than acuity for uppercase five letter words (Arditi, 1994). This indicates that for Courier Bold font, surrounding letters do not have an obvious deleterious effect on letter identification within words and subsequent word recognition in normal vision through contour interaction, or crowding effects.

Optimal processing of letters and words of different lengths can be safely assumed to occur when character size is at or greater than three times threshold size for a range of text stimuli

The relationship between character size and response time, for all acuity stimuli that were used, resembled that between reading speed and character size (but inverted due to time rather than speed being plotted on the ordinate) as described, for example, by Legge *et al.* (1985a,b) and Whittaker and Lovie-Kitchin (1993). Because every effort was made to keep the subjects' and examiner's response characteristics constant, we consider the variations in response time to reflect variations in visual processing in all cases. The minimum character size at which optimal processing for uppercase and lowercase single letters, and for lowercase 4-, 7- and 10-letter words were found to be between 0.2 and 0.4 log MAR above acuity threshold for each of these stimuli. These values were always less than three times acuity threshold in visual angle (0.45 log MAR elevation). This value represents the elevation from the lowest character size at which at least four out of five stimuli were correctly recognised. This stimulus size was found to be slightly greater than the mean of actual acuity threshold, but always within 0.1 log MAR of this value.

It is easy to appreciate from the method outlined in Appendix 2 that calculated critical character size is dependent upon the criterion used to define this parameter – being greater for a criterion of 95 or 99% of maximum processing speed, for example, than if a value of 90% is used. Nevertheless, the values reported in the present study are, in general, similar to the minimum elevation for optimal processing of text passages above threshold print size, which is often referred to as a minimum acuity reserve (Whittaker and Lovie-Kitchin, 1993). The result that optimal processing of letters and words of different lengths occurs at some minimum value above, rather than at, acuity threshold is hardly surprising, but what is interesting is that these values for acuity reserve appear identical to the acuity reserve values for reading continuous text passages. Thus an acuity reserve of approximately 1.5–3 times acuity threshold is required for each type of text stimulus – whether that be single letters, words of (probably) any length or text passages.

Therefore neither the concept of acuity reserve, nor the often quoted minimum acuity reserve value of two times acuity threshold (0.3 log MAR elevation in size) for reading continuous text (Whittaker and Lovie-Kitchin, 1993; Cheong *et al.*, 2002) are specific to text passages.

As an approximately similar acuity reserve value was also found for the lowercase words of different lengths, as well as for the uppercase and lowercase single letters, it appears that the acuity reserve for continuous text passages does not depend upon, nor is influenced by, a rapid declining of the visual span close to threshold in a readily predictable way. The term 'visual span' is often used to describe the number of letters that can be recognised in a single fixation (e.g. Legge *et al.*, 2001). As visual acuity rapidly declines with eccentricity from the centre of the fovea, it is certainly conceivable that the visual span would shrink close to letter acuity threshold. If this was the case, then visual processing time should be proportionally longer for longer vs shorter stimuli close to threshold. We did not find an obvious progressive change in acuity reserve values for each word length with greater values for longer words. Acuity reserve values appear to be similar across a range of different types of acuity stimuli. Instead, the requirement for an acuity reserve might reflect the need for increased temporal summation to successfully read stimuli that are close to threshold.

The concept of allowing for an acuity reserve in the level of magnification required by low-vision patients is almost always spoken about, in research literature, for near vision and in the context of reading continuous text passages (Whittaker and Lovie-Kitchin, 1993; Cheong *et al.*, 2002). The results reported in the present study indicate that an acuity reserve value, of a similar degree, is needed in normal vision when reading isolated letters or words. As low-vision patients, like subjects with normal vision, require an acuity reserve to read text passages as efficiently as possible (Legge *et al.*, 1985b), they presumably would need to also have a sufficient acuity reserve for reading tasks that would simply involve letter or word identification, such as reading bus numbers and street signs through a distance telescope, for example. While the difference in processing time between reading a word successfully at threshold and that to read a word more than three times larger than threshold may be short – only 1–2 s for example – there may be a considerable increase in the amount of attentional resources that must be allocated to the task for reading stimuli at or close to threshold. If distance sighting tasks need to be frequently made in a short period of time, then increased attentional demand in each task might hasten the onset of fatigue.

Appendix 1: Validation of the usability of the smallest character size stimuli

We performed a validation test to determine whether the smallest letters (nominal $-0.4 \log \text{MAR}$ at 4 m, 10 pixels) in our stimulus set could be used for computer-based acuity measurement, without concern that the monitor resolution either limits or influences the measures. The logic of our approach is that if monitor resolution does not affect the measures, acuities obtained at 4 and 8 m should be equivalent. We used single letters for this test. Specifically, we measured acuities for uppercase Sloan letters and lowercase Courier Bold letters at 4 and 8 m (Figure 6).

Mean Sloan letter acuity ($\pm \text{S.E.}$) at 4 and 8 m was -0.11 ± 0.029 and $-0.11 \pm 0.034 \log \text{MAR}$, respectively. Mean Courier Bold letter acuity ($\pm \text{S.E.}$) at 4 and 8 m was -0.14 ± 0.014 and $-0.11 \pm 0.037 \log \text{MAR}$, respectively. The mean difference ($\pm \text{S.D.}$) in Sloan letter acuity measured at the two distances was $0.00 \pm 0.005 \log \text{MAR}$ (4–8 m measures), which was not statistically significant (Wilcoxon sign ranked test, $p = 0.69$). The mean difference ($\pm \text{S.D.}$) in lowercase Courier Bold letter acuity was $-0.03 \pm 0.008 \log \text{MAR}$ (4–8 m measures), which was also not statistically significant (Wilcoxon sign ranked test, $p = 0.22$).

Appendix 2: Determination of the critical character size

Critical character size was calculated in the following way. First an exponential decline-to-a-minimum curve,

of the form $y = y_0 + ae^{-bx}$, was fitted to each function relating character size (x -axis) and response time (y -axis) using Sigmaplot 2001. Minimum response time, or maximum response speed, was taken to be the value of y_0 given by the above function. Critical character size was operationally defined to be the amount in $\log \text{MAR}$ above threshold (x in the above equation) at which response time was 110% of the y_0 value. The above equation was solved for x , using the modified value just mentioned for y . This calculated value is also the character size at which response speed is 90% of the maximum value.

Acknowledgement

This study received generous financial support from the Pro Visu Foundation and the Association Bien des Aveugles (Geneva).

References

- Abdelnour, O. and Kalloniatis, M. (2001) Word acuity threshold as a function of contrast and retinal eccentricity. *Optom. Vis. Sci.* **78**, 914–919.
- Arditi, A. (1994) In: *Low Vision: Research and New Developments in Rehabilitation* (eds A. C. Kooijman, P. L. Looijestijn, J. A. Welling and G. J. van der Wildt), IOS Press, Amsterdam, The Netherlands, pp. 38–45.
- Arditi, A. and Cagenello, R. (1993) On the statistical reliability of letter-chart visual acuity measurements. *Invest Ophthalmol. Vis. Sci.* **34**, 120–109.
- Arditi, A., Knoblauch, K. and Grunwald, I. (1990) Reading with fixed and variable character pitch. *J Opt. Soc. Am. A.* **7**, 2011–2015.
- Bach, M. (1996) The freiburg visual acuity test – automatic measurement of visual acuity. *Optom. Vis. Sci.* **73**, 49–53.
- Bailey, I. L. (1980) Designation of visual acuity in logarithmic units. *Optom. Mon.* **71**, 80–85.
- Bailey, I. L. and Lovie, J. E. (1980) The design and use of a new near-vision chart. *Am. J. Optom. Physiol. Opt.* **57**, 378–387.
- Bullimore, M. A. and Bailey, I. L. (1995) Reading and eye movements in age-related maculopathy. *Optom. Vis. Sci.* **72**, 125–138.
- Carkeet, A. (2001) Modeling $\log \text{MAR}$ visual acuity scores: effects of termination rules and alternative forced-choice options. *Optom. Vis. Sci.* **78**, 529–538.
- Cheong, A. C., Lovie-Kitchin, J. E. and Bowers, A. R. (2002) Determining magnification for reading with low vision. *Clin. Exp. Optom.* **85**, 229–237.
- Chung, S. T. (2004) Reading speed benefits from increased vertical word spacing in normal peripheral vision. *Optom. Vis. Sci.* **81**, 525–535.
- Chung, S. T., Mansfield, J. S. and Legge, G. E. (1998) Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vis. Res.* **38**, 2949–2962.
- Culham, L. E., Fitzke, F. W., Timberlake, G. T. and Marshall, J. (1992) Use of scrolled text in a scanning laser

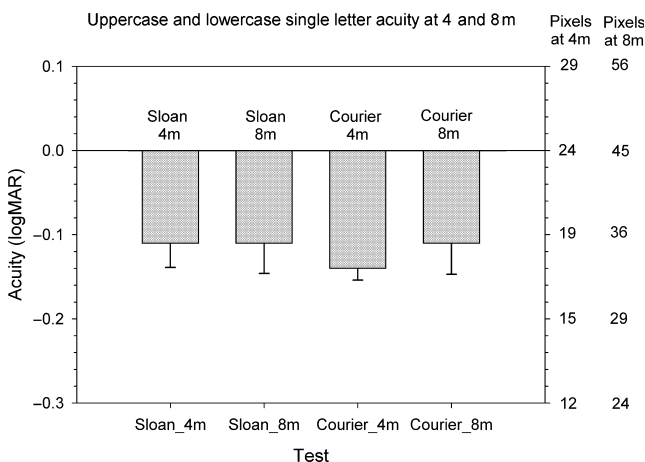


Figure 6. Acuities for uppercase Sloan and lowercase Courier Bold letters, measured at test distances of 4 and 8 m. The bars represent mean acuity in $\log \text{MAR}$ ($\pm \text{S.E.}$) for six subjects across the four tests. The number of pixels that correspond to each of the character sizes listed in the graph at both test distances is shown on the right-hand side of the figure. This denotes the number of pixels contributing luminance information to the total stimulus size for the uppercase Sloan letters and the x -height for the lowercase Courier Bold letters.

- ophthalmoscope to assess reading performance at different retinal locations. *Ophthalm. Physiol. Opt.* **12**, 281–286.
- Ferrand, L. and New, B. (2003) Syllabic length effects in visual word recognition and naming. *Acta. Psychol. (Amst.)* **113**, 167–183.
- French, C. N. (1997) The UMIST eye system. *Spat. Vis.* **11**, 91–93.
- Hess, R. F., Dakin, S. C., Kapoor, N. and Tewfik, M. (2000) Contour interaction in fovea and periphery. *J. Opt. Soc. Am. A. Opt. Image. Sci. Vis.* **17**, 1516–1524.
- Jacobs, R. J. (1979) Visual resolution and contour interaction in the fovea and periphery. *Vis. Res.* **19**, 1187–1195.
- Kitchin, J. E. and Bailey, I. L. (1981) Task complexity and visual acuity in senile macular degeneration. *Aust. J. Optom.* **64**, 235–242.
- Leat, S. J., Li, W. and Epp, K. (1999) Crowding in central and eccentric vision: the effects of contour interaction and attention. *Invest Ophthalmol. Vis. Sci.* **40**, 504–512.
- Legge, G. E., Ahn, S. J., Klitz, T. S. and Luebker, A. (1997) Psychophysics of reading – XVI. The visual span in normal and low vision. *Vis. Res.* **37**, 1999–2010.
- Legge, G. E., Mansfield, J. S. and Chung, S. T. (2001) Psychophysics of reading. XX. Linking letter recognition to reading speed in central and peripheral vision. *Vis. Res.* **41**, 725–743.
- Legge, G. E., Pelli, D. G., Rubin, G. S. and Schleske, M. M. (1985a) Psychophysics of reading – I. Normal vision. *Vis. Res.* **25**, 239–252.
- Legge, G. E., Ross, J. A., Luebker, A. and LaMay, J. M. (1989) Psychophysics of reading. VIII. The minnesota low-vision reading test. *Optom. Vis. Sci.* **66**, 843–853.
- Legge, G. E., Rubin, G. S., Pelli, D. G. and Schleske, M. M. (1985b) Psychophysics of reading – II. Low vision. *Vis. Res.* **25**, 253–265.
- Levi, D. M., Klein, S. A. and Hariharan, S. (2002) Suppressive and facilitatory spatial interactions in foveal vision: foveal crowding is simple contrast masking. *J. Vis.* **2**, 140–166.
- Lovie-Kitchin, J. E. (1988) Validity and reliability of visual acuity measurements. *Ophthalm. Physiol. Opt.* **8**, 363–370.
- Lovie-Kitchin, J. E., Bowers, A. R. and Woods, R. L. (2000) Oral and silent reading performance with macular degeneration. *Ophthalm. Physiol. Opt.* **20**, 360–370.
- Lovie-Kitchin, J. E. and Brown, B. (2000) Repeatability and intercorrelations of standard vision tests as a function of age. *Optom. Vis. Sci.* **77**, 412–420.
- Mansfield, J. S., Legge, G. E. and Bane, M. C. (1996) Psychophysics of reading. XV: font effects in normal and low vision. *Invest. Ophthalmol. Vis. Sci.* **37**, 1492–1501.
- Nazir, T. A., Heller, D. and Sussmann, C. (1992) Letter visibility and word recognition: the optimal viewing position in printed words. *Percept. Psychophys.* **52**, 315–328.
- Raasch, T. W., Bailey, I. L. and Bullimore, M. A. (1998) Repeatability of visual acuity measurement. *Optom. Vis. Sci.* **75**, 342–348.
- Radner, W., Obermayer, W., Richter-Mueksch, S., Willinger, U., Velikay-Parel, M. and Eisenwort, B. (2002) The validity and reliability of short German sentences for measuring reading speed. *Graefes Arch. Clin. Exp. Ophthalmol.* **240**, 461–467.
- Radner, W., Willinger, U., Obermayer, W., Mudrich, C., Velikay-Parel, M. and Eisenwort, B. (1998) A new reading chart for simultaneous determination of reading vision and reading speed. *Klin. Monatsbl. Augenheilkd.* **213**, 174–181.
- Rayner, K. (1994) In: *Eye Movements in Reading* (eds Y. Ygge, and G. Lennerstrand), Elsevier, Oxford, pp. 205–218.
- Rosser, D. A., Murdoch, I. E. and Cousens, S. N. (2004) The effect of optical defocus on the test-retest variability of visual acuity measurements. *Invest Ophthalmol. Vis. Sci.* **45**, 1076–1079.
- Ruamviboonsuk, P., Tiensuwan, M., Kunawut, C. and Masayaanon, P. (2003) Repeatability of an automated Landolt C test, compared with the early treatment of diabetic retinopathy study (ETDRS) chart testing. *Am. J. Ophthalmol.* **136**, 662–669.
- Rubin, G. S. and Turano, K. (1994) Low vision reading with sequential word presentation. *Vis. Res.* **34**, 1723–1733.
- Siderov, J. and Tiu, A. L. (1999) Variability of measurements of visual acuity in a large eye clinic. *Acta. Ophthalmol. Scand.* **77**, 673–676.
- Simmers, A. J., Gray, L. S., McGraw, P. V. and Winn, B. (1999) Contour interaction for high and low contrast optotypes in normal and amblyopic observers. *Ophthalm. Physiol. Opt.* **19**, 253–260.
- Trauzettel-Klosinski, S. and Brendler, K. (1998) Eye movements in reading with hemianopic field defects: the significance of clinical parameters. *Graefes Arch. Clin. Exp. Ophthalmol.* **236**, 91–102.
- Trauzettel-Klosinski, S. and Reinhard, J. (1998) The vertical field border in hemianopia and its significance for fixation and reading. *Invest Ophthalmol. Vis. Sci.* **39**, 2177–2186.
- Trauzettel-Klosinski, S. and Tornow, R.-P. (1996) Fixation behavior and reading ability in macular scotoma. *Neuro-ophthalmology* **16**, 241–253.
- Virgili, G., Pierrottet, C., Parmeggiani, F., Pennino, M., Giacomelli, G., Steindler, P., Menchini, U. and Orzalesi, N. (2004) Reading performance in patients with retinitis pigmentosa: a study using the MNREAD charts. *Invest Ophthalmol. Vis. Sci.* **45**, 3418–3424.
- Whittaker, S. G. and Lovie-Kitchin, J. (1993) Visual requirements for reading. *Optom. Vis. Sci.* **70**, 54–65.
- Wolffsohn, J. S. and Eperjesi, F. (2004) Predicting prescribed magnification. *Ophthalm. Physiol. Opt.* **24**, 334–338.