A Computer-Based Optotype Acuity Test System Suitable for Evaluation of Acuity Charts*

Aries Arditi and Ron Cagenello
Vision Research Laboratory
The Lighthouse Inc.
800 Second Avenue
New York, NY 10017

Introduction

Although computerized visual acuity testing is psychophysically more flexible than "manual" testing using printed or projected optotype charts, the quality of currently available and affordable pixel-oriented computer images makes computers poorly suited for clinical testing. This is mainly because low spatial resolution in one or both dimensions limits the range of sizes available for testing at a single optical distance.

Charts and slides, on the other hand, do not suffer from this limitation, and with them presentation of a wide range of letter sizes at uniform high image quality and at high intensity is easy. For example, the chart developed by Ferris et al.[2] and now available as the Lighthouse Distance Acuity Chart, displays letters ranging a factor of 1.4 log units in size, in a chart of some 62 cm in height, with unnoticeable degradation of even the smallest letters.

Despite their superior imagery, chart systems have serious disadvantages. Most importantly, they are difficult to evaluate, and important specifications that should affect their use, such as their statistical reliability, are virtually unknown.

The main reason why chart evaluation is so difficult is that stimulus randomization is awkward if not impossible. Even though several forms of charts are available (e. g. for each eye and binocular testing), chart tests are highly susceptible to practice effects with repeated testing, ruling out the within-subjects experimental designs that are critical to proper evaluation and to experiments that can improve their design.

We have developed a computer-based visual acuity testing system that provides somewhat smaller optotype size range (0.8 log units) but double the number of size increments of typical chart-based acuity systems, and allows randomization of optotypes and hence within-subjects designs. Furthermore, the system is flexible enough to permit study of many other important variables that affect visual acuity. We illustrate the usefulness of this type of system with preliminary data on test-retest reliability of a slightly modified version of the Lighthouse chart.

*Abstract submitted to the Optical Society of America for inclusion in the 1992 Topical Meeting on Noninvasive Assessment of the Visual System

Figure 1: A line of 5 Sloan characters, similar to that used in the Lighthouse chart and in the experiment.

Method

Optotypes were Sloan characters displayed on a 15" Mitsubishi color monitor driven by an X-windows (X11) server residing on a Silicon Graphics IRIS 4D/25 graphic workstation via a StereoGraphics Corporation stereo display system, with a final pixel resolution of 1280 horizontal X 512 vertical, and a luminance of 88.3 cd/m^2. The stereo display system, which is responsible for the unusually low vertical resolution, allows us to present separate stimuli to each eye and is used in experiments other than those described here. Client software that actually ran the experiments, via a local network, resided on a Sun SparcStation 1.

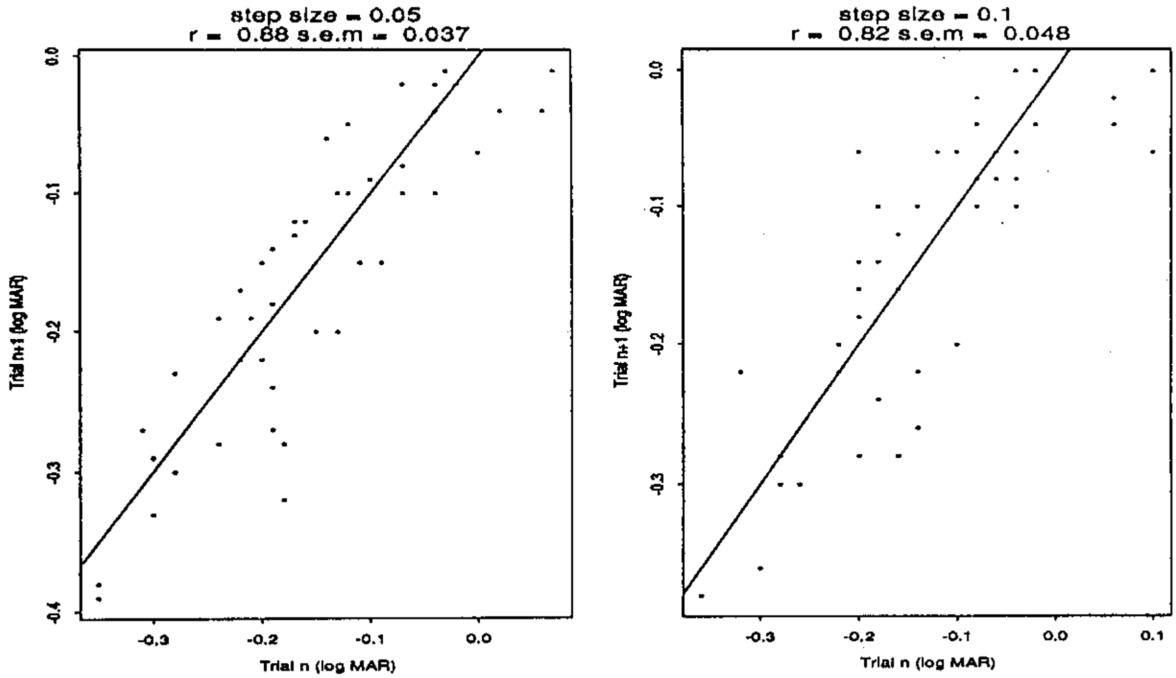
We intended to mimic as closely as possible the Lighthouse chart, which incorporates accepted principles established by Bailey and Lovie[1], and to adhere as closely as possible to accepted standards for visual acuity testing[3]. Letters were displayed in lines of 5 characters, separated from each other and from the edge of the screen by one character width, as shown in Figure 1. Only a single line was displayed and tested at a time.

The Lighthouse chart uses lines of letter combinations determined to be of equal confusability and not containing words or acronyms[2]. Since with repeated measurement subjects could, in this case, memorize letter sequences, in our procedure the computer generated sequences of five letters randomly for each line. We used an algorithm that ensured that no letter was repeated more than once in a line and that letters were displayed with roughly equal frequency in the set of lines comprising a single trial.

Letters were generated using the font design language METAFONT, which translates geometric definitions of type into pixel representation. The advantage of using METAFONT is that a single program can be used to generate optotypes of any size. Our optotypes were intended to vary in size by 0.05 log unit between 'lines,' as compared with the 0.1 log unit step size used by the charts emulated. Digitization error caused sizes to depart slightly from their intended values, of course, but for all letters the size along the (vertical) dimension of lowest resolution differed from its intended value by no more than 3%. Our smallest letters were 40 X 20 pixels.

In order to reduce the size of the letters at the eye, the display was viewed at a distance of 95 cm through the *objective* lenses (i. e. in reverse) of a pair of Minolta 10 X 25 Wide Angle Pocket Binoculars.

Three highly practiced subjects, including the authors, were tested repeatedly over a period of several weeks. Two of these were emmetropic, and one slightly myopic. The myopic subject wore his correction. All subjects were aware of the identity and number of characters in the Sloan set and used in testing. Subjects identified letters verbally to the



(a) (b)

Figure 2: Scatterplots of acuity scores on repeated trials illustrating reliability of letter acuity testing under conditions of the present experiment. (a) shows data scored using all 15 lines, while (b) shows the same data rescored using only the 8 lines separated by 0.1 log MAR unit. The solid lines have unity slope and pass through the origin.

experimenter, who typed responses at the keyboard. Tests were scored and data analyzed by computer.

Acuity was scored by the method recommended by Ferris et al.[2] that does not penalize subjects who fail to guess, wherein the subject reads the entire 'chart' of 15 lines, and is given 0.01 log minimum angle of resolution (MAR) unit for each letter read correctly. The log MAR acuity score for a single trial, then, was $0.25 - 0.01T_c$ where T_c is the total number of letters read correctly of the 75 in the test. The 0.25 constant is a bias term to adjust for our maximum acuity score of -0.5 log MAR, which corresponds to 20/6.3 in Snellen notation.

Although data were collected from only three practiced observers with normal and supranormal acuities, the range of acuity scores was artificially increased by testing monocularly and binocularly at luminance contrasts ranging from 0.055 to 0.989, where contrast between foreground L_{fg} and background L_{bg} was defined as $\frac{L_{bg}-L_{fg}}{L_{fg}+L_{bg}}$.

Results and Discussion

Figure 2(a) shows scatterplots of acuity scores of even-numbered repeats against odd-numbered repeats of tests within the same test condition (e. g. tested eye, contrast), using size steps of 0.05 log unit between lines. Each data point thus represents two entire readings of the chart. Test-retest reliability as measured by the correlation between tests was 0.88. Using Fischer's z' transformation, we find that if the population reliability coefficient is equal to our observed value, samples of our sample size ($n = 45$) should yield correlation coefficients within the range 0.786 to 0.931 95% of the time. The standard error of measurement was 0.037 log MAR, which is equivalent to nearly 4 letters, or 4/5 of a line. Assuming our observed r is accurate, and assuming the underlying variable is normally distributed, this means that under our conditions of testing we can conclude that our test is accurate only to within about 7 letters, or 1.2 lines 95% of the time.

Since our size steps between lines were one half those of the Lighthouse chart, we decided to rescore the data using only lines corresponding to 0.1 log MAR steps. Acuity in log MAR units in this case is $0.30 - 0.02T_c$. Data are shown in Figure 2(b) and are highly similar to those with the 0.05 log unit step sizes. Although coarser steps might be expected to produce a higher r, r in fact decreased, but not significantly so. If this sample is representative of the underlying population, in this case 95% confidence limits around acuity scores are only +/-0.094, equal to nearly a line of the Lighthouse chart.

Our experiment, of course, departs from typical conditions of acuity testing in several significant ways. Our subjects were highly motivated and practiced, were familiar with the set of stimulus letters, and had extremely good acuity. Our test luminance was lower than typical, though above the minimum specified by a commonly accepted standard[3]. Our testing used a scoring algorithm not typically used by most vision practitioners. All of these departures from typical testing conditions, with the possible exception of lower luminance, however, can reasonably be expected to decrease variability and increase test-retest reliability. Hence our data may be usefully viewed as an upper limit to test-retest reliability of the Lighthouse chart, achievable when strict testing conditions are met.

Although these data are preliminary, they suggest that accuracy of measurements in this commonly used test of visual function may be limited, and raise the question of its appropriateness in cases where a high degree of accuracy is desired.

References

[1] I. L. Bailey and J. E. Lovie. New design principles for visual acuity letter charts. *American Journal of Optometry and Physiological Optics*, 53:740–5, 1976.

[2] F. L. Ferris, A. Kassoff, G. H. Bresnick, and I. L. Bailey. New visual acuity charts for clinical research. *American Journal of Ophthalmology*, 94:91–6, 1982.

[3] Consilium Ophthalmologicum Universale. Visual acuity measurement standard. Visual Functions Committee, International Council of Ophthalmology, 1984.