

# VISUAL ACUITY ASSESSMENT: CHARTS VS. COMPUTERS

Aries Arditi, Ph.D., F.A.A.O.  
Lei Liu, Ph.D.

Arlene R. Gordon Research Institute  
The Lighthouse Inc.  
111 East 59th Street  
New York, NY 10022 USA

Visual acuity measurement with letter optotype charts is without doubt the most widespread vision test in the world. It is used nearly universally in ophthalmology and optometry for screening, refraction, and monitoring of disease progression; in licensing for vision-intensive tasks such as driving and flying; and in occupational testing. Optotype acuity testing is also used in both basic and clinical vision research, as a means of characterizing observers' visual resolving capacities. Given the importance of this measurement, it is surprising that very little attention has been devoted to assessing its accuracy. On the other hand, the many variations of test conditions, procedure, chart used, and population tested, leave most tests of accuracy open to criticism. A few years ago, using pristine, controlled, and admittedly unrealistic clinical conditions, we measured the test-retest reliability of the optotype acuity chart procedure<sup>1</sup> in order to assess the *best* possible accuracy likely to be achieved in any test situation.

**[Figure 1 is scatter.gif and probably will take up a full page]**

Figure 1 shows scatterplots of our test measurements for 0.05 log MAR size increments and the 0.1 log MAR size increments using letter-by-letter scoring recommended by Ferris et al.,<sup>2</sup> and for the more common 3 of 5 correct line scoring with 0.1 log MAR size increments. Surprisingly, even under tightly controlled laboratory conditions, the test is quite inaccurate. The 95% confidence intervals around the correlation coefficients, after scaling by the standard deviations of the scores, show that the 3 of 5 correct scoring method with 0.1 log MAR size increments yields acuity measurements that are only accurate to within  $\pm 0.13$  log units, or 1.3 lines on the chart. Using letter-by-letter scoring improves the accuracy somewhat but only to  $\pm 0.1$  log unit, or 1 line. Using 0.05 log MAR size steps between lines further increases accuracy to  $\pm 0.06$  log units, but these step sizes are generally not feasible to use in charts, because only a greatly restricted range of acuities can be tested with a single chart.

We have been using sequential adaptive psychophysical methods<sup>3</sup> on a computer to increase the accuracy of our acuity measurements. These methods concentrate testing near threshold—where responses are most informative—by modifying the stimuli based on data collected earlier in the experiment. In acuity testing, the sizes of the letters presented depend on the correctness of the response(s) in previously read lines of letters. Adaptive procedures were first described in 1948 by Dixon and Mood,<sup>4</sup> but were not commonly used in psychophysics until their popularization by Cornsweet.<sup>5</sup> Computers vastly simplify the process of implementing adaptive procedures, because they allow easy randomization (which

reduces learning effects) and can implement complex algorithms for experimentation. Adaptive procedures are now commonplace in psychophysics.

In our staircase method, lines of randomly chosen 5 Sloan letters are presented on a monitor beginning at a clearly suprathreshold size. For each line correctly read, the size is reduced until the subject makes an error. If there are one or more errors on a line the next line is increased in size. The size increment (or decrement) is 0.05 log unit throughout the experiment, except in the initial descending run of sizes, where, in order to produce rapid convergence to threshold, the step size is 0.1 log unit. Our technique is very similar to chart testing, except that lines are made up of randomly chosen (without replacement) letters, and the line tested is contingent on the response to the previously tested line.

**[Figure 2 is composed of fig2a.gif and fig2b.gif. These should be printed one above the other, and will together take up one page.]**

Figure 2 shows typical examples of our runs for a highly-practiced able-sighted (AA) subject and a highly variable subject with optic neuritis (JG). For both subjects, variability about the acuity estimates stabilizes in just a few minutes. These data illustrate only how the within-subject variability decreases throughout the test procedure. To determine the test accuracy of the method, test-retest data need to be collected, and this experiment is currently underway.

While our method has advantages over chart testing, there are a few disadvantages as well. First, because of the limited pixel resolution of the monitor, we can present well-formed characters having a size range of 0.6 log units, as compared with the 1.3 log unit size range of typical charts. In practice, we circumvent this problem by using different test distances. Also, charts are still cheaper than computers and may be the only affordable option in some developing countries. Finally, chart luminance can be much higher than CRT display technology currently allows.

## References

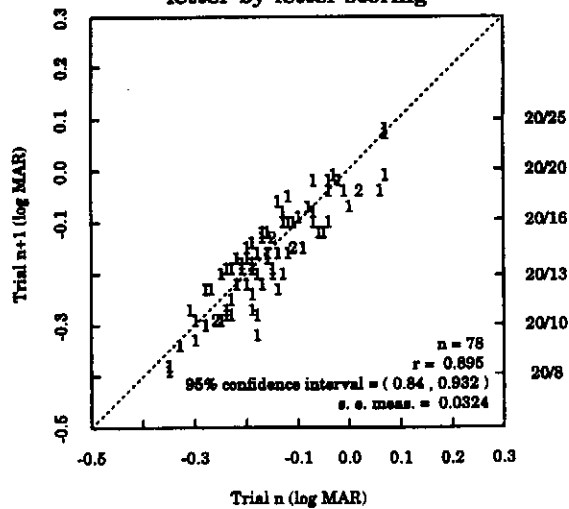
1. Arditi, A. and Cagenello, R. (1993). On the statistical reliability of letter-chart visual acuity measurements. *Investigative Ophthalmology and Visual Science*, 34, 120–129.
2. Ferris, F.L., Kassoff, A., Bresnick, G.H., and Bailey, I.L. (1982). New visual acuity charts for clinical research. *American Journal of Ophthalmology*, 94, 91–96.
3. Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research*, 2503–2522.
4. Dixon, W.J. and Mood, A. (1948). A method for obtaining and analyzing sensitivity data. *Journal of the American Statistical Association*, 43, 109–126.
5. Cornsweet, T.N. (1962). The staircase-method in psychophysics. *American Journal of Psychology*, 75, 485–491.

## Figure Captions

Figure 1: Scatterplots of acuity scores on repeated trials in an experiment assessing test-retest reliability of visual acuity testing using letter chart procedures. The numbers at each plotted point represent frequency of observations at each location. The dashed lines have unity slope and pass through the origin.  $r$  is the Pearson product moment correlation between test and retest scores.

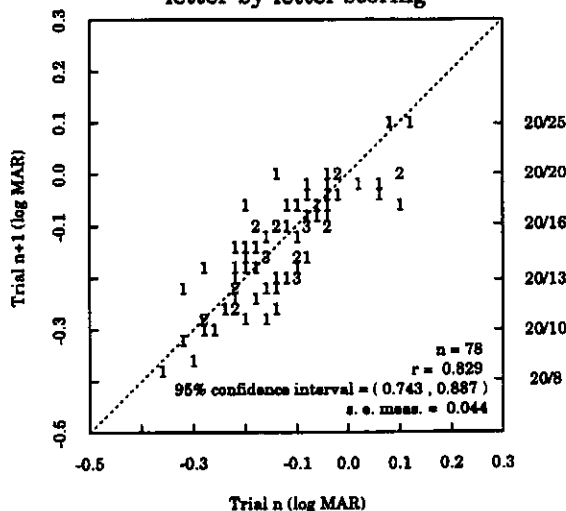
Figure 2: (a) Able-sighted subject. (b) Subject with low vision (wearing a +2D add). These graphs plot the history of the experiment beginning at the trial at which an error is first made. The points in the upper graph of (a) and (b) indicated the sizes of the letters presented, in log MAR (left axis) and Snellen fractions (right axis); the line through the points is a running (geometric) mean of levels tested thus far in the experiment, and thus represents the best instantaneous estimate of visual acuity. The lower graphs in (a) and (b) indicate the size of the 95% confidence interval about the means plotted in the upper graphs in log units (left axis) or standard 0.1 log MAR chart lines. The horizontal axis is shown in units of number of letters read (top scale), approximate test time (middle scale; assuming 20 letters read per minute), and trials (lower scale).

0.05 log MAR size increment  
letter-by-letter scoring



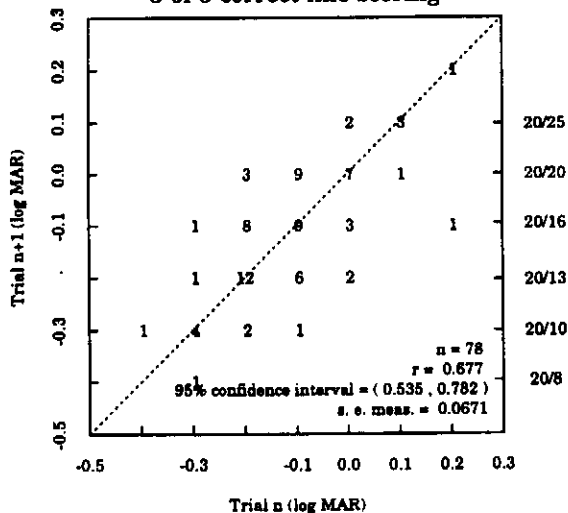
(a)

0.1 log MAR size increment  
letter-by-letter scoring



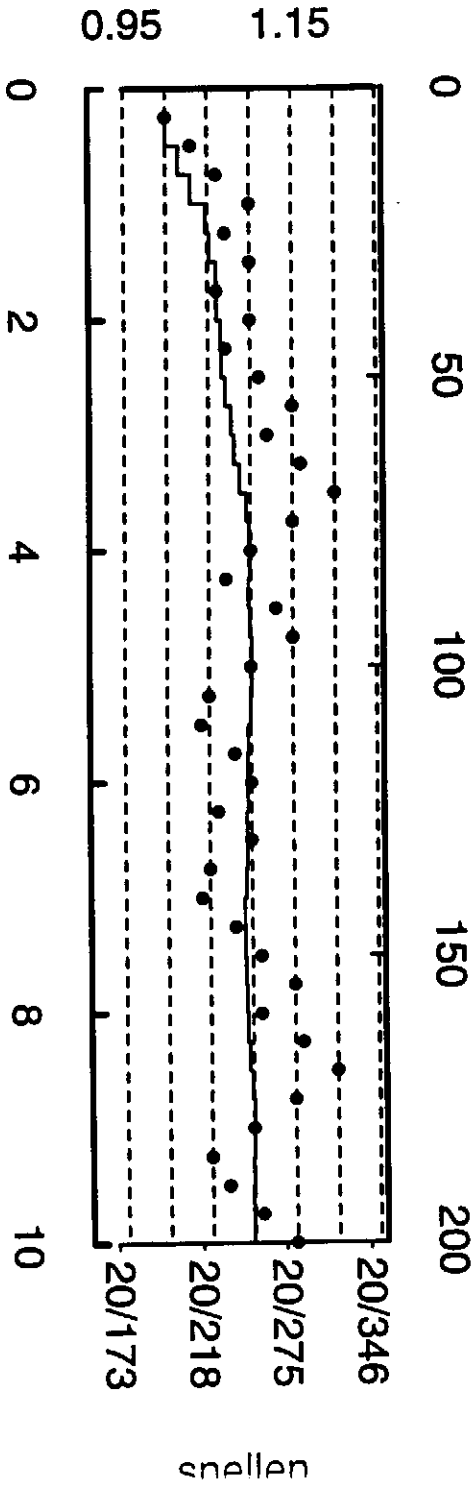
(b)

0.1 log MAR size increment  
3 of 5 correct line scoring

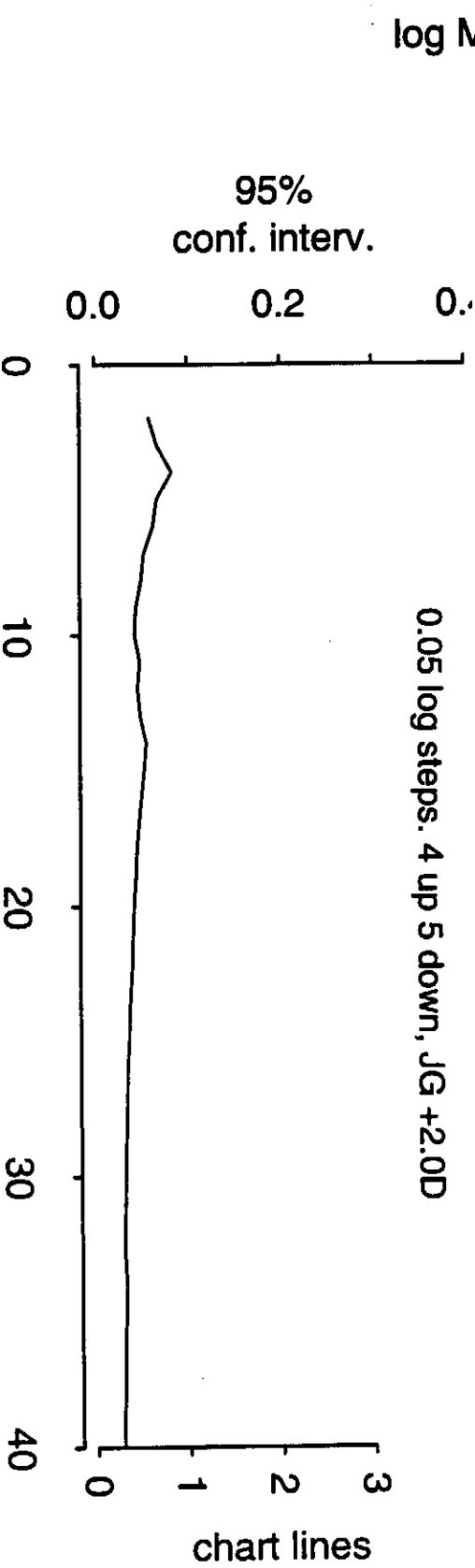


(c)

number of letters read

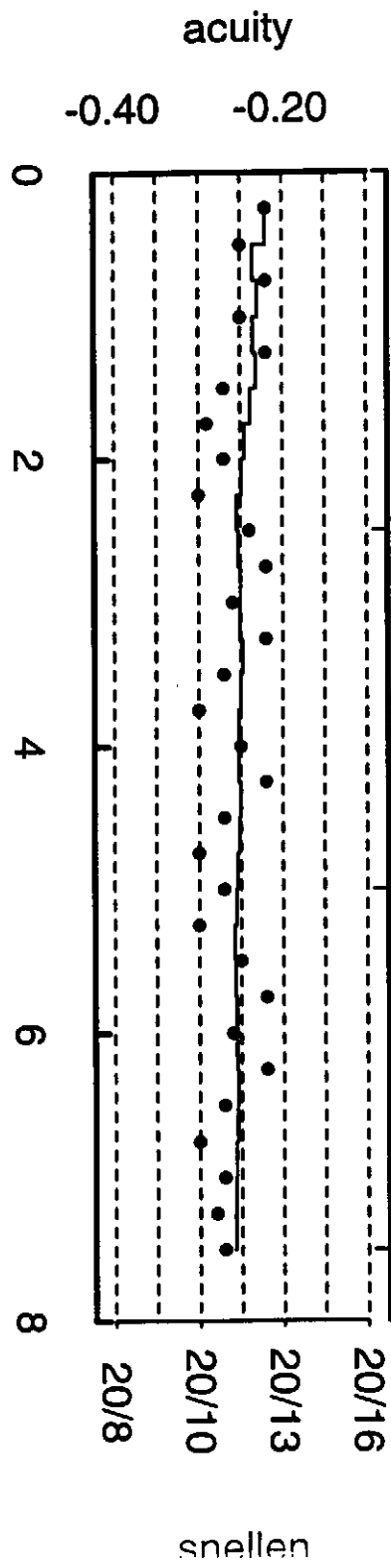


approximate time (min)



number of letters read

0 50 100 150

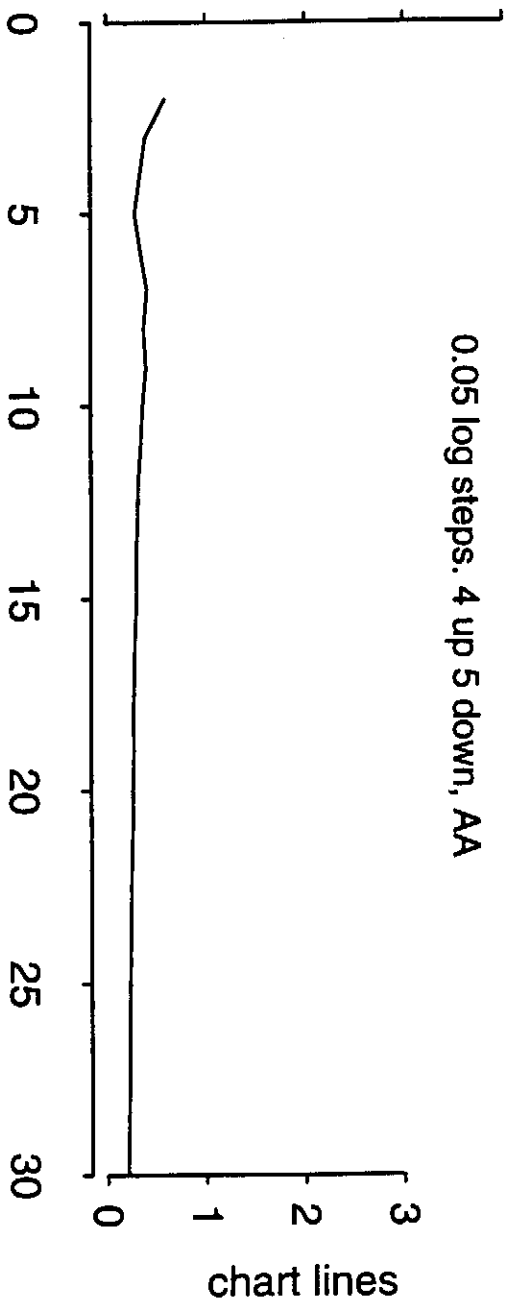


approximate time (min)

log MAR

95%  
conf. interv.

0.0 0.2 0.4



trials (post 1st reversal)