LBL-32031
UC-350

To be presented at the Illuminating Engineering Society Annual
Conference, San Diego, CA, August 2–6, 1992,
and to be published in the Proceedings

Size as a Determinant of Reading Speed

I. Bailey, R. Clear, and S. Berman

March 1992

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

IES Annual Conference
San Diego, CA., August 1992

L-162
LBL-32031
UC-350

Size as a Determinant of Reading Speed

Ian Bailey, Robert Clear and Sam Berman

Lighting Systems Research Group
Energy & Environment Division
Lawrence Berkeley Laboratory
Berkeley, California 94720

March 1992

This work was supported by the Assistant Secretary for Conservation and Renewable Energy, Office of Building Technologies, Building Equipment Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Size as a Determinant of Reading Speed

Ian Bailey, Robert Clear, and Sam Berman

Lighting Systems Research Group
Energy & Environment Division
Lawrence Berkeley Laboratory
Berkeley, California 94720

Abstract

The speed of reading unrelated words as a function of luminance, size, and contrast, was measured with an eye movement monitor for fifteen young adults. Subjects read up to 5,000 words in a test session, with the exact number depending upon their acuity. The size of the smallest legible print at a given luminance and contrast for these subjects was found to fit well to the Blackwell-Taylor detection threshold data above about 1 minute of arc. At lower sizes inclusion of a resolution size term provided an excellent fit.

Reading speed was fit to a number of visual performance models. It was found that for most subjects that a ratio of the print size to an estimate of the threshold print size (a VL_{size}) gave the best fits to the data. The threshold size was computed with a fit to the Blackwell-Taylor detection threshold data, modified to include a resolution size term as above. For the sole remaining subject a slightly better fit was obtained with a $VL_{contrast}$ model, where again the thresholds were modified by a limiting size term. The implication of these results for visual performance modeling is discussed.

The reading speed for all subjects varied rapidly with size near the acuity limit, but became almost independent of visibility parameters as long as size is two times the acuity limit. These results show that size is a powerful determinant of reading speed, and suggest that minification of about 1/2 power could be used as a field test for adequate visibility.

Introduction

Background

In this paper we show that there is a significant advantage to size-based performance models for predicting reading speed. The reading task is a complex resolution and identification task, while common visual performance models used by engineers have been based on simple detection, which is fundamentally driven by contrast. Visibility Level (VL) visual performance models, such as CIE 19/2, or Clear and Berman's model, are based on measurements of the probability of detection of a disk target.[1,2] The independent variable VL is the ratio of a targets' actual contrast to the "threshold" contrast needed for 50% probability of detection at the measured luminance, size, and exposure time. Rea's visual performance model is based on the speed of detection of a square target.[3] His fitting function depends upon the difference of the tasks' actual contrast and its absolute threshold contrast for detection at the measured luminance and size.

VL based models have been applied to reading and other tasks which involve more than simple detection by using the measured contrast threshold of the task. The shape of the threshold curve versus luminance is assumed to be unchanged.[1] Rea applies his function to reading tasks by assuming that detection reaction times are linearly related to reading speed.[4]

Contrast plays a central role in detection models because there is nothing to detect unless there is contrast. For large targets with sharp outlines it is reasonable to assume that if the components of the target can be detected its shape will be recognizable. This leads to the generalization of applying detection models to predict performance of complex visual tasks.

For small targets, conditions for detection and resolution are not the same. Ricco's law says that for small targets, detectability is related to the total flux difference between the target and its background (ΔL x area = constant). Lack of proper focus, scattering in the eye, diffraction, aberrations, and receptor spacing, make little or no difference to detectability, because though these factors increase the size of the image the total flux difference remains constant. This is not true for resolution. Consider a target consisting of two bars. To resolve the image as two bars it is necessary to detect the presence of the space between the bars, and not just the presence of the bars. The width of this space is fixed, so as shown in figure 1, smearing of the image of the bars reduces the image contrast of the space without a compensating increase in size.

Overview

In the experiment described here, we measured and analyzed reading speed over a range of sizes, luminances, and contrasts.[5] At the larger print sizes we examined fits of both the Clear-Berman VL model and Rea's reaction time (RT) model. The two models differ in their manner of extrapolation to the reading task. The RT model assumes that accuracy and speed are related, so that only speed needs to be specified. In addition the RT model expresses size in solid angles (an area), while the VL model uses angular size. The VL model fit our data significantly better the RT model.

The VL model provides a good fit to the reading speed data for larger print sizes, but neither model extrapolates well to the smallest print sizes, especially when near a subject's acuity limits. We got significant improvements in the fits by using a function that models the reduced contrast available for resolution at the smaller print sizes. The function approaches 1.0 for large targets where thresholds for detection and resolution are essentially the same, and decreases smoothly with size to 0.0 at the resolution limit.

Although the "resolution" correction factor improved the fits, there were still patterns in the data that were not properly fit. These patterns led us to consider fits based on the ratio of size/threshold size, which is a VL_{size}. To develop a rationale for such a procedure consider the problem of distinguishing between two shapes. The retina is a mosaic of sensors. If the shapes are small, then the number of sensors stimulated is small, and it is very hard to distinguish between shapes. Increasing the contrast or luminance will not help, but increasing size will, as long as the contrast and luminance are sufficient. Similarly, limitations in reading speed may be primarily related to size, with the role of contrast and luminance being that of bringing as much of the shape above threshold as is possible.

A few of our subjects showed an optimum size for reading speed. Increasing size makes fixated letters individually more visible, but at the same time nearby letters are imaged on more peripheral parts of the retina, which decreases their visibility. Modeling performance as the sum of these two competing factors significantly improved the fit in 13 of 15 cases examined.

A size-based model is a sharp break from current practice. Individual performance curves vary, but all of them are flat at large sizes and then become very steep as their individual resolution limit is approached. Simple minification thus can be used to determine if subjects are sufficiently above their resolution limits to achieve near optimal performance. This suggests the use of a minifying lens as a field device for assessing the adequacy of illumination conditions.

Experiment and Apparatus

Performance was measured as the speed of reading unrelated words on Bailey-Lovie word reading charts (figure 2).[6] On these charts 11 rows of words are arranged so that each row of words on the chart is smaller than the row above. The ratio of sizes is approximately 0.1 log units (4/5). The subjects were tested on letter sizes ranging from 20 to 2 points (newspaper print is 8 point). A brow bar was used to give a viewing distance of 40 cm, which gives task detail sizes (1/5 the letter height) of 0.6 to 6 minutes of arc. On each of the 11 rows in this size range there are 2 four-letter words, 2 seven-letter words, and 2 ten-letter words. The words are arranged in an almost random fashion. Two different sets of 40 charts were used in the study. In one set the first letter of the words of each row has approximately the same frequency. In the second set the words were selected so that words in each row had approximately the same frequency of use. The charts were typeset with a Times-Roman font.

The charts were prepared as black-on-clear transparencies for retro-illumination. The light source behind the test charts was an opal plexiglass screen over a light box containing switchable, and dimmable, incandescent lights. An identical second light box was used in conjunction with a rotatable disk that contained one open window, and two partially reflecting mirrors to vary contrast by imaging a veiling luminance on the plane of the target. The layout is shown in figure 3.

As the subjects read from large to small their eye position was recorded using a spectacle mounted scleral reflection eye movement monitor. This type of eye-tracker uses a small IR light and a pair of small IR detectors directed towards the cornea/sclera junction. The difference between the two detectors provides a sensitive measure of eye position.

The subjects read silently. The time taken to read each row was determined by the time between the large leftward saccadic movement made to take up fixation at the start of the next row (see figure 4). With few exceptions, the subjects required only one fixation per word (6 fixations per row), and in general they also exhibited good accuracy in finding an appropriate eye position to begin reading the next line.

Methods

Of the initial seventeen subjects two were tested over a restricted set of conditions, and their results were not analyzed for this paper. Of the 15 remaining subjects one was 39, one 27, and the other 13 ranged in age from 20 to 23. Each subject had a vision screening examination. All subjects had visual acuities of 20/20 or better in both eyes, and none had a history of binocular vision disorders, or significant eye disease. Subjects who used glasses wore them during the testing. No subject had more than 3 diopters of spherical refractive error or 1 diopter of cylindrical refractive error. All subjects were tested monocularly with their preferred eye, while the other eye was occluded.

The testing session typically lasted 3 1/2 hours with the inclusion of a half hour break in the middle. In the first half session, 40 different charts were presented in a predetermined order. These same charts were used in the second half session, but with a different order. Subjects did not appear to learn the charts. Within each half session, the luminance levels were increased or decreased progressively, with different contrast conditions being presented in a random preselected order at each luminance level.

Table 1 shows how the 80 charts were distributed over the luminance and contrast conditions of the experiment. Each chart had 11 possible target sizes, but the smallest sizes were not readable under the lowest luminance-contrast conditions. The average subject read about 800 six-word rows of print.

Preliminary Data Analysis

Because the number of charts presented to each subject for each condition was small, an effort was made to remove obviously aberrant data before the performance analyses were made. Occasionally subjects would lose their place on the charts, so that the number of rows read on a given chart was anomalously low. The whole chart was eliminated from analysis if the number of rows read was 3 or more fewer than achieved at other attempts at the same visibility. Another occasional problem was that a subject might not recognize a word, or alternatively they might skim a word instead of actually reading it. To identify such anomalies the reading times for each relative size versus the estimated acuity limit were averaged over all 80 charts. Reading times at the acuity limit were excluded if they were substantially shorter than the times at larger sizes. Above the acuity limit reading time data was excluded when it was more than 3 standard deviations from the average for its particular relative size. The combination of these procedures resulted in a loss of the reading speed data from about 10 rows per subject, or slightly more than 1% of their total data.

VL Calculations and Resolution Limit

As a first test of how well detection threshold models predict reading performance data we compared our measured reading acuity limits to predicted detection threshold sizes. Reading acuity was estimated from linear interpolation (or extrapolation) to the print size that was read 50% of the time, for each combination of luminance and contrast. In the few cases where a subject read all words in the smallest row the acuity was estimated as being 1/2 line smaller. Acuities were determined both for each subject, and as an average over subjects.

We compared the reading acuities to a fit of a comprehensive set of contrast thresholds for detection that were measured by Blackwell and Taylor, and reported as averages over subjects by Blackwell.[7] For a fixed luminance, the data follow Ricco's law when the target size is small, while for larger sizes there is an abrupt change to a functional form suggested by Hills.[8] We used the following three equations to fit the Blackwell-Taylor data:

$$\Delta L_t = [A(L)/S]^2 \qquad S < S_r, \qquad (1)$$

$$\Delta L_t = \{[B(L) \times [S - S_r] + A(L)]/S\}^2 \qquad S \geq S_r, \qquad (2)$$

and

$$S_r = (4 + 2L)/(1 + L) \qquad (3).$$

Here ΔL_t is the threshold luminance difference between the task and its immediate background (contrast multiplied by adaptation luminance); A(L) and B(L) are functions of adaptation luminance, L; and S_r is the size below which Ricco's law applies. Luminances are expressed in cd/m^2, and size in minutes of arc. The functions A(L) and B(L) have the form $a_1(L^{1/n} + a_2)^{n/2}$.

Our fitted values for the parameters of A(L) and B(L) are listed in table 2.

The maximum deviation of the above fit from the Blackwell-Taylor data is 14%, and the standard deviation is only 7%. This new fit has about a 50% lower variance than the fit we reported previously.[9]

Solving equation 1 or 2 for size gives a threshold size for detection. In our reading experiment, the threshold sizes were all smaller than the maximum size of the Ricco law region (equation 3), so only equation 1 is used. There are three luminances needed for the solution of equation 1: the task, or print detail luminance, the background or surround luminance, and the adaptation luminance. In the Blackwell-Taylor experiment the target, a disk, was shown against a uniformly illuminated background. The background and adaptation luminances were thus essentially the same. In our experiment, the luminance increment between task and its immediate surround is seen against a background which includes other print. We computed L as an area-weighted average which for our charts was 25% task to 75% surround.

Figure 5 shows the threshold print detail size averaged across subjects for each of the lighting conditions from our experiment plotted on the y-axis, versus the predicted threshold sizes from equation 1 plotted on the x-axis. Figure 6 is a similar plot for a subject whose whose resolution limits are substantially larger than our measurement limitations. The top horizontal line is the lowest detail size on the chart (0.6 minutes) and the bottom line is the extrapolation resolution limit (0.525 minutes). The diagonal line represents a fit to equation 1. The inverse of the slope, K, of the diagonal line shows the detection contrast sensitivity of the subject relative to the Blackwell-Taylor data. The curved lines shows a resolution limit fit that we will describe below. Separate symbols were used for the three contrasts used in the experiment to show that the fitted curve applies to all contrasts. The three contrast curves do not match well if the adaptation luminance is assumed to equal the background luminance, instead of an average over task and background areas as was discussed above.

At small sizes the discrepancy between the predicted disk detection thresholds and our reading acuity data is presumably due to the loss of effective contrast for the resolution task. We adopted a resolution contrast correction factor, f(S), which increases from zero at a resolution limit, S_L, to one for large targets, of the form:

$$f(S) = [(S^n - S_L^n)/S^n]^{2/n} \qquad (4).$$

We fit S_L as a constant, and thus are ignoring possible pupil size effects on image quality and retinal illuminance.[10] When target size is small enough to be within the Ricco's law region equation 4 leads to the following equation for luminance difference required for resolution:

$$\Delta L_R = [A(L)/S]^2 /f(S) = [A(L)]^2/(S^n - S_L^n)^{2/n} \qquad (5),$$

Equation 5 can be solved to give a size limit for resolution, S_R, as long as it is within the Ricco's law region:

$$S_R = \{[A(L)/(\Delta L)^{1/2}]^n + S_L^n\}^{1/n} = (S_D^n + S_L^n)^{1/n} \qquad (6)$$

Here S_D is the size limit for detection.

Fits of our reading acuity data with Equation 6 were fairly insensitive to the value of the exponent, n. A value of 4 was the best single (integer) exponent for the 15 subjects as a group, and was used for all subsequent analysis. The fitted resolution limits, S_L, ranged from 0.5 to 0.8 minutes. Because the decision as to when the print was too small to read may not have been consistent, the reading speed data described in the next section probably provides better estimates of S_L than the reading acuity data. While the average estimate of S_L from the time data was close to that of the acuity estimates, the range was shifted to 0.4 to 0.7 minutes of arc.

Equation 6 fit the reading acuity data for most subjects quite well. Interestingly, the goodness of fit was similar to the goodness of fit we later obtained for reading speeds.

Reading Speed

Figure 7 shows reading speeds as averaged over subjects, at each luminance, size, and contrast. This average is what is most relevant for lighting engineering, while individual data provide information on the quality of the fits and parameter variations between subjects.

For the conditions we examined, performance begins to be significantly affected by luminance and contrast when the letters were 8 point type or smaller. The resolution correction described in the previous section becomes insignificant for 5 to 6 point type and above. Reading speeds for these larger sizes were initially analyzed against two models based on detection. To fit the data for the smaller print sizes we first modified the VL detection model with the resolution correction from equation 4. These fits still did not capture all of the trends, so we turned finally to a size-based model.

Data Analysis

The averages and standard deviations of the reading times for each subject were computed for each luminance, contrast and size condition for which there was sufficient data. The calculated standard deviations were approximately proportional to time, so we used the geometric mean for the average. A linear fit of the standard deviations with respect to time was used to weight the data for fitting. This allowed us to include visibility conditions for which there was only one data point, and improves the reliability of the fits.

For some visibility conditions, the words were readable by some subjects, but not others. To account for this in the average over subjects, we converted the reading times to speeds, and assigned a speed of zero to visibility conditions which a subject did not read. The speeds were then averaged. Both the individual and average fits were done versus speed. For individuals, the time uncertainties were transformed to speeds before fitting. Visibility conditions for which no words were read were excluded from the fits of individual and group data.

In averaging over subjects we did not weight the mean or standard error over subjects. The χ^2 goodness of fit test, which is used in our fitting program, overestimates significance for the average data, because the variances between and within subjects are combined.

Rea's Reaction Time Model

Rea fit his measured reaction times, T_{RT}, with a model of the form:[3]

$$T_{RT} = (\Delta C^{0.97} + K^{0.97})/(R_{max} \times \Delta C^{0.97}) \qquad (7).$$

ΔC is the difference between the target's actual contrast and its threshold contrast, C_t; K is the half saturation function; and R_{max} is the maximum response function. Both C_t and K are functions of luminance and solid angle, while R_{max} is a function of luminance alone.

Rea assumes that speed of reading, RS, is linearly related to the reaction times fit above:[4]

$$RS = A \times (1 - B \times T_{RT}) \qquad (8).$$

In Rea's reading experiment B is 0.00094. Since our experiment differs from his, we treated A and B as free parameters.

The solid angles for the reading task is given by the area occupied by strokes of the letters.[4] We used a video camera to digitize a representative test chart and then measured print area. The printed stroke area averaged 5.2 μ-steradians for 8 point type. On the other hand, a typical small letter of 8 point print was only about 3.5 μ-steradians in size. Since it is not clear whether the average, or typical small size that determines performance, we tried fits using both assumptions. Area varies as the square of the point size. Since Rea has only tested his model to 2 μ-steradians, we limited our fits under the small letter assumption to 6 point type, and fits under the average letter assumption to 5 point type.

Table 3 gives information about the parameters and goodness of fit for the fitted functions. The fit to one subject was considerably worse than fits to the remaining subjects, so we have shown the means and standard deviations both with and without this subject. The fit parameters vary a great deal among subjects, and are also sensitive to the size range that is fit. When fit to the 6 point type data and above, the fits are slightly better using the typical small letter instead of average letter size. Nonetheless, in both analyses, only 3 out of 15 fits were statistically significant. Figure 8 shows the residuals, or deviations for the small letter fit to the average data, for 6 point type and above. The fit does not predict the diminished performance under the poorest visibility conditions (low size, luminance and contrast). There is also a distinct trend from negative to positive deviations as luminance increases for the medium and high contrast points. The fits get worse when average letter size is used, and the size range is extended to 5 point type.

The parameters used for the fits to large print do not predict the sharp loss of performance as size is reduced. Figure 9 shows the residuals from the fit in figure 8 as a function of size over the entire print range. It is clearly seen that the parameters that give the best fit to the larger print sizes do not give a good fit at smaller print sizes.

The Clear-Berman VL model

The VL-type model developed by Clear and Berman has the form:[2]

$$T_R = T_{NV} + T_c/(F \times VL/V - 1) \qquad (9).$$

VL is computed from equations 1 - 3:

$$VL = \Delta L_p \times S^2/A^2(L) \qquad S < S_r \qquad (10a)$$

$$= \Delta L_p \times S^2/[A(L) + (S - S_r)\, B(L)]^2 \qquad S \geq S_r \qquad (10b)$$

T_R is the reading time ($RS = 1/T_R$); the free parameter T_{NV} represents the non-visual component of the task; T_c is a time constant which describes how visibility changes with viewing time; $F = (T_c + 0.2)/0.2$, where 0.2 seconds is the presentation time in the Blackwell-Taylor experiment; the free parameter V is the visibility level that corresponds to the accuracy criterion of the subject; and ΔL_p is the physical luminance difference between the task and its immediate background. We used the value of 210 msec for T_c proposed by Blondel in place of the complicated function developed by Adrian as the latter did not noticeably improve our fits.[11] Given this value of T_c, the parameter F is equal to 2.05. Equation 9 has the same number of free parameters (2) as the fit to Rea's model.

Table 4 gives the parameters and goodness of fit for equation 9 for 5 point type and larger. Figure 10 shows the distribution of residuals from the fit of Equation 9 to the average data . The fit shown in figure 10 has about 1/3 the variance of the RT fit for figure 8, despite the inclusion of data for the next smaller size print. There is no luminance trend in figure 10, but the fit is poorer at the lowest visibilities. Equation 9 does reasonably well at fitting individual data, with 9 out of 15 fits being statistically significant. For 5 point type and above, it has a lower variance than the RT fits for 13 of the 15 subjects . The parameter values are reasonably constant for these fits, especially as compared to the RT fits. However, it should be noted that the residuals in the poorest fits do show patterns, indicating that Equation 9 is not applicable to all the subjects. As expected, the fits do not in general extrapolate well to smaller sizes.

The Resolution Correction

To handle the smaller print sizes we made two changes to our fit. The first was to multiply VL by the expression in Equation 4 to get a resolution visibility level, VL_{res}. The second was to relax the assumption that accuracy of reading is constant, and to allow for a decrease in accuracy as visibility declines. This is a much more natural assumption for which there is confirming evidence in other experiments.[2] We assumed a simple form for V:

$$V = a\, VL_{res}/(b + VL_{res}) \qquad (11).$$

Inserted into Equation 7 this gives:

$$T_R = T_{NV} + T_c/(C\ VL_{res} - D), \qquad C = 2.05/a, \qquad D = 1 - 2.05\ b/a, \qquad (12).$$

The two parameters C and D of equation 12 replace the single parameter V of equation 9 in these generalized fits. Including the parameter S_L of Equation 4 this yields a 4 parameter fit.

Fits to Equation 12 for all sizes are markedly better than such fits with equations 8 or 9. Equation 12 “explains” 96 percent of the variance for the averaged data, and from 50 to 90 percent for the individual data sets. However, with one exception, the parameters of equation 12 that give the best fit to the data as a whole worsen the fit to the larger print (5 point type and above). With this larger size print, the variance increases by an average of 30% for the individual fits, and by a factor of 2.5 times for the average data, over that of Equation 9. In addition, examination of the residuals show a pattern that becomes particularly evident at the smallest sizes. Figure 11 shows this pattern for the average data. For any given size, the fit using equation 12 predicts a greater change in performance than is in fact present.

This size trend is particularly evident for the subject with the worst fitting data. Figure 12 shows this subject's performance data versus visibility at every other print size to better illustrate the trend versus size and visibility. This subject's data fits poorly to Equation 9 for 5 point type and larger, which suggests that there is a problem with the basic model, instead of just problems extending it with the approximate methods to handle accuracy changes (Equation 11) or resolution problems (Equation 4).

A Size-based Model

Another size trend is visible in figure 13, which shows the data for a subject who appears to have an optimal size in the range from 4 to 8 point type. Not all subjects showed this optimum, and it was not obvious in the average data plotted in figure 7.

As mentioned in the overview, we feel that there are good reasons to believe both the above trends. To fit the tendency for some of the data to have an optima versus size we added a linear term in size, $T_S S$, to the model. Depending on how large T_S is this can give a performance maximum at one of the print sizes in our experiment, or larger. To fit the trend in figure 12 we wanted a model that is mainly dependent on size when the size is small, but becomes more dependent upon contrast as size increases. A model with these properties can be derived by replacing VL_{res} in Equation 12 with $VL^2(\text{size})$. VL(size) is defined as the ratio of an object's actual size to its threshold size ($VL(\text{size}) = S/S_R$). As long as the threshold size is less than Ricco's limit it can be computed from Equation 5.

Equation 13 is an expression for reading time that incorporates the above two changes:

$$T_R = T_{NV} + T_S\, S + T_c/(C\; VL^2(\text{size}) - D) \qquad (13).$$

In this expression T_S is a free parameter, S is size in minutes, and the remaining terms are as defined above.

Figure 14 shows the deviations from this new fit versus the average reading speed data over the same size region plotted in figure 11. The scatter is now mainly confined to the lowest visibility points, and the trend towards positive residuals at the higher visibilities shown in figure 11 is essentially gone. The size fits are superior to the contrast fits for 14 of the 15 subjects, and average about 20 percent smaller variance. The fit to the average data has about 40 percent lower variance. For 5 point type and above the fits show about 10 percent more variance than the simple contrast fits (equation 9) for most subjects, but there were no longer any really bad fits.

Table 5 provides summary information on the fits. Only two of the individual fits are actually statistically significant over the entire data set, and only five are significant for 5 point type and above, but many are very close. Most of the deviations from the model are at the lowest visibilities. The low visibility points have the least reliable reading speeds as they are the most likely to suffer from changes in reading strategy. Equation 13 fits the basic trends quite well.

Figure 15 makes this claim visual. Subtracting the calculated linear size term,$T_S S$, from the measured reading speeds, T_R, gives transformed points that are a function of VL(size) alone. Figure 15 plots these transformed points (for the average data), together with their predicted values from Equation 13 against VL(size).

Minification as a field test for Visibility

For field testing, individual variability becomes important. Figure 16 shows that different individuals share the same basic performance trend. The vertical axis gives the relative task performance (RTP) which was computed by dividing each subject's actual reading speeds under a given visibility condition by their predicted reading speeds from equation 13 with 20 point print (S = 6 minutes of arc) and VL = ∞ . The horizontal axis is a relative size calculated by dividing each subject's VL(size) by the VL(size) which gives them RTP = 2/3 with 20 point type. We are assuming that at RTP = 2/3 subjects will be able to report that they are slowing down. The figure shows a fairly abrupt knee at a relative size of two, which suggests that minification may provide a good field test for adequate visibility.

For example, if a person uses a minifying optical system with a minification of 1/2 on a reading task, figure 16 tells us that there will be no difficulty in performing the task as long as it was originally 4 times its reference size for that person (the size at which RTP = 2/3). This value is comparable to the factor of 5 quoted by Legge from his work.[12] Minification of less visible tasks will cause most people to begin to slow down, and in the range of relative sizes from 2 down people using the minifier should report that the task has become either very slow or impossible. Thus a minifier of minification 1/2 provides a simple test of whether a particular person has adequate visibility for the specific task that they are performing.

For field test use the minifier should be combined with a reference task. Complaints about visibility from the field can be examined by a tester with normal vision using both the worker's normal task and the reference task to determine if the problem is the lighting, the task, or the worker's vision.

Discussion

The present work does not explain why the VL models gave substantially better fits than the Rea model, nor does it guarantee that the VL models will fit better on other data sets. However, there are two aspects of Rea's generalization of his reaction time fits to resolution tasks that appear unjustified. The first is the use of the total stroke area of the letters. Simply detecting the presence of the letter as a whole does not necessarily mean that the critical details that allow identification can be resolved. We tried using smaller areas as better matches to critical detail areas, and found that using a value of 1/3 the average letter stroke area reduced χ^2 by an average of 12 percent. The VL model still fit better however, for the average data, and for 13 of the 15 subjects.

The reader should note that both Rea's model and our models use nominal measures of size to evaluate visibility. Thickening the stroke of letters reduces the space between details and hence eventually reduces the ability to identify letters, as they ultimately all become blobs. Using either 1/5 the letter height or some multiple of letter stroke area as the measure of size clearly fails in this situation.

The second difficulty with Rea's model is the assumption that reading <u>speeds</u>, instead of reading times, are linearly related to reaction <u>times</u>. Rea describes it as a simple scaling, but the formula he presents reduces to the form given in our equation 8.[4]

The size model represents a major departure from the standard method of modeling performance. We propose that the performance of any moderately complex task represents the successful completion of a number of subtasks. In any given situation, one or more different subtasks may be the limiting tasks that define the performance level. For reading tasks, at least over some of the range of conditions which we studied, we are suggesting that it is relative size, not relative contrast, that most directly limits speed, as long as there is sufficient contrast to permit resolution. Our relative size model provided the best fit to the reading speed data over the entire data range, but at the larger sizes fit using a relative contrast model was equally good for most subjects.

It is possible that the fits we have made represent a compromise to a situation where size is the dominant factor over part of the range, and contrast dominates over the remainder of the range. The inclusion of the linear term in size (T_sS) should be verified with other experiments. Our experiment presented the largest print first, and it is possible, although we feel it is unlikely, that this effects the data.

Extrapolation outside the range of conditions studied should be avoided if possible. It is unlikely that the term in T_s remains linear for very large targets, and the parameters C and D represent an empirical fit to a change in criteria which may be meaningless for different conditions.

Conclusion

Above a task detail size of about 1.5 minutes of arc (5 point print), detection contrast models can be made to fit the reading speed data presented here. For our experimental conditions the residual variance from fits to the simple VL model averaged about 70 - 80% of that of fits to the RT model for the individual data sets, and about 20% for the averaged data. Extrapolation of either model to smaller sizes does not work well. Resolution and accuracy corrections improve the VL model, but do not completely eliminate all the trends in the residuals. Changing to a relative size model eliminates most of the trend errors, and is thus preferred, even though the fits are still noisy. It appears that size is a more relevant parameter than contrast for predicting reading speeds. A simple minifying system can be used as a field test for adequate visibility.

Acknowledgement

This work was supported by the Assistant Secretary for Conservation and Renewable Energy, Office of Building Technologies, Building Equipment Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098. We also thank the National Vision Research Institute of the University of Melbourne, Australia, for laboratory facilities used for some of this work.

References

[1] Publication Commission Internationale de L'Eclairage, 19/2. 1981. "An Analytic Model for Describing the Influence of Lighting Parameters upon Visual Performance". International Commission on Illumination, Paris, France.

[2] R. Clear, and S. Berman. 1990. "Speed, Accuracy, and VL". *J of the IES* Vol. 19(2): 124-131.

[3] M.S. Rea and M. J. Ouellette. 1988. "Visual Performance Using Reaction Times". *Lighting Research & Technology,* Vol. 20(4): 139 - 153.

[4] M. S. Rea and M. J. Ouellette. 1991 "Relative visual performance: A basis for application". *Lighting Research & Technology,* Vol. 23(3): 135-144.

[5] I. L. Bailey. 1987. "Mobility and Visual Performance under Dim Illumination in Night Vision - Current Research and Future Directions". Symposium Proceedings - Committee on Vision: National Research Council American Academy of Sciences. National Academy Press. Wash. D.C. pages 220 -230.

[6] I. L. Bailey, and J. E. Lovie. 1976. "New design principles of visual acuity letter charts". *Am. J. Ophthal.,* Vol. 53: 740 - 745.

[7] H. R. Blackwell. 1981. "Description of a Comprehensive Family of Relative Contrast Sensitivity (RCS) Functions of Luminance to Reflect Differences in Size of Task Detail, Task Eccentricity, and Observer Age". *J of the Illuminating Engineering Society*, Vol. 11(1).

[8] B. L. Hills. 1975. "Visibility under Night Driving Conditions". *Lighting Research & Technology*, Vol. 7(3): 179 - 184.

[9] R. Clear and S. Berman. 1985. "Target Size, Visibility, and Roadway Performance". *J. of the Illuminating Engineering Society*, Vol. 15(1): 167 - 180.

[10] S. Berman. 1992. "Energy Efficiency Consequences of Scotopic Sensitivity". *J. of the IES,* Vol. 21(1): 3 - 14.

[11] W. Adrian. 1989. "Visibility of targets: Model for calculation". *Lighting Research & Technology,* Vol. 21(4): 181 - 188.

[12] G. E. Legge. 1991. "Glenn A. Fry Award Lecture 1990: Three Perspective on Low Vision Reading". *Optometry and Vision Science,* Vol. 68(10): 763 - 769.

Table 1:
Number of charts presented at each luminance and contrast condition

Luminance[1]	Contrast[2]		
	0.29	0.78	0.985
11.0	2	4	4
34.3	4	-	8
75.4[3]	-	4	4
158	-	4	8
343	4	-	4
857	-	4	8
2160	4	-	4
3430	-	4	-
4280	2	-	-
5480	-	-	4

Notes

1) Luminance, L, of letter background in cd/m^2.

2) Contrast defined as: [L(background) - L(letter)]/L(background).

3) For two of the 15 subjects this luminance level was 130 cd/m^2.

Table 2:
Parameters for the functions A(L) and B(L) in the threshold detection fit

parameter	A(L)	B(L)
a_1	0.40743	.0840155
a_2	1.6074	.43369
n	2.8723	1.7188

Table 3:

Parameters & degree of fit - RT model

Range of fit	Size	χ^2	R^2	A	B
			Average Data		
≥ 5 pt.	5.2	1.44	.667	5.96	0.00153
≥ 6 pt.	5.2	0.88	.567	5.19	0.00129
≥ 6 pt.	3.5	0.78	.618	4.92	0.00117
			Individual Fits		
≥ 5 pt.	5.2	2.1 ± 1.5	.49 ± .20	6.17 ± 2.31	.0015 ± .0004
≥ 6 pt.	5.2	1.8 ± 0.7	.39 ± .23	5.91 ± 2.49	.0014 ± .0005
≥ 6 pt.	3.5	1.6 ± 0.6	.40 ± .24	5.41 ± 2.09	.0012 ± .0005
			Individual Fits without worst subject		
≥ 5 pt.	5.2	1.7 ± 0.4	.47 ± .20	5.86 ± 2.06	.0014 ± .0004
≥ 6 pt.	5.2	1.6 ± 0.4	.36 ± .21	5.60 ± 2.25	.0013 ± .0005
≥ 6 pt.	3.5	1.5 ± 0.4	.37 ± .22	5.15 ± 1.90	.0012 ± .0004

Notes: The size column gives the assumed size of 8 point type in μ-steradians. χ^2 is the sum of squares deviations of data from the fit relative to the standard deviations of the data. R^2 is the fraction of variance of the data that is "explained" by the fit.

Table 4:

Parameters & degree of fit - Simple VL model

Range of fit	χ^2	R^2	V	T_{nv}(msec)
		Average Data		
≥ 5 pt.	0.28	0.936	1.78	314.7
		Individual Fits		
≥ 5 pt.	1.7 ± 1.8	.59 ± .22	1.87 ± .86	331.7 ± 85.7
		Individual Fits without worst subject		
≥ 5 pt.	1.2 ± 0.4	.59 ± .23	1.70 ± .56	335.6 ± 87.6

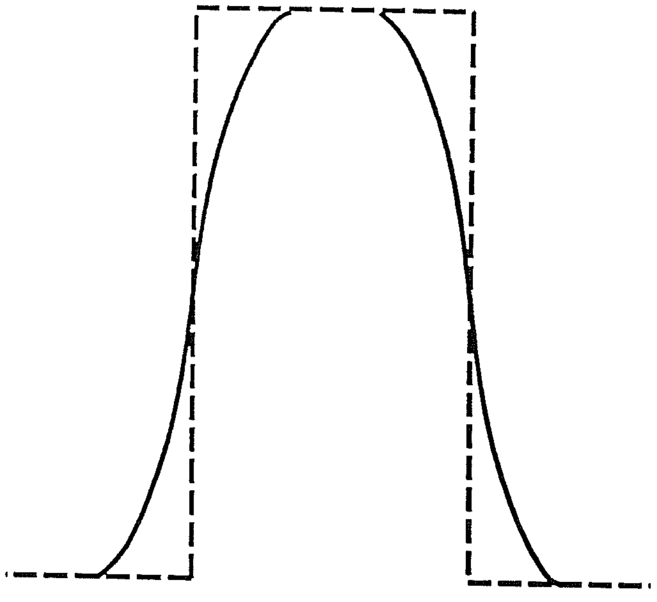
Table 5:

Parameters & degree of fit - Size VL model

	Average Data		Individual fits	
sizes	all	≥ 5 pt.	all	≥ 5 pt.
χ^2	0.49	0.42	1.75 ± 0.54	1.44 ± 0.49
R^2	0.976	0.902	0.786 ± 0.145	0.583 ± 0.232
S_L	0.553	NA	0.554 ± 0.088	NA
C	0.949	NA	0.887 ± 0.665	NA
D	0.622	NA	0.259 ± 0.373	NA
T_{NV}(msec)	298.6	NA	286.8 ± 84.3	NA
T_S(msec)	3.54	NA	9.04 ± 5.97	NA

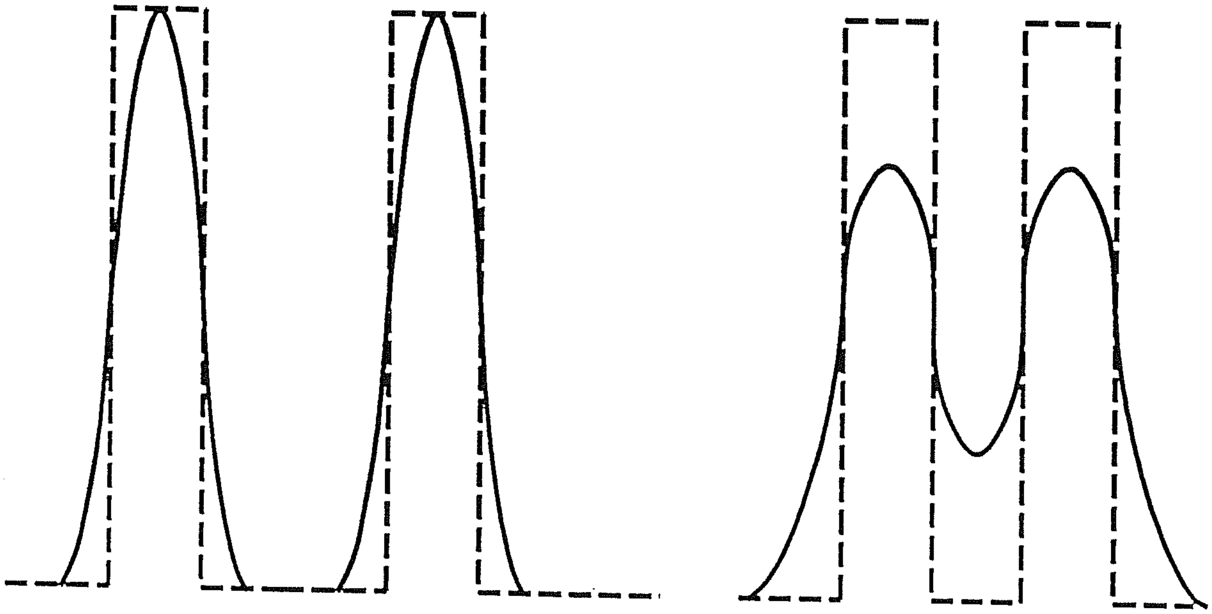
NA = not applicable. The χ^2 and R^2 values shown for 5 point type print and above were calculated with the fits to the entire data set.

Figure 1: Physical luminance (dashed line) vs retinal luminance (solid line). The retinal luminance has rounded corners due to blur and scatter



a) Detection task - wide image

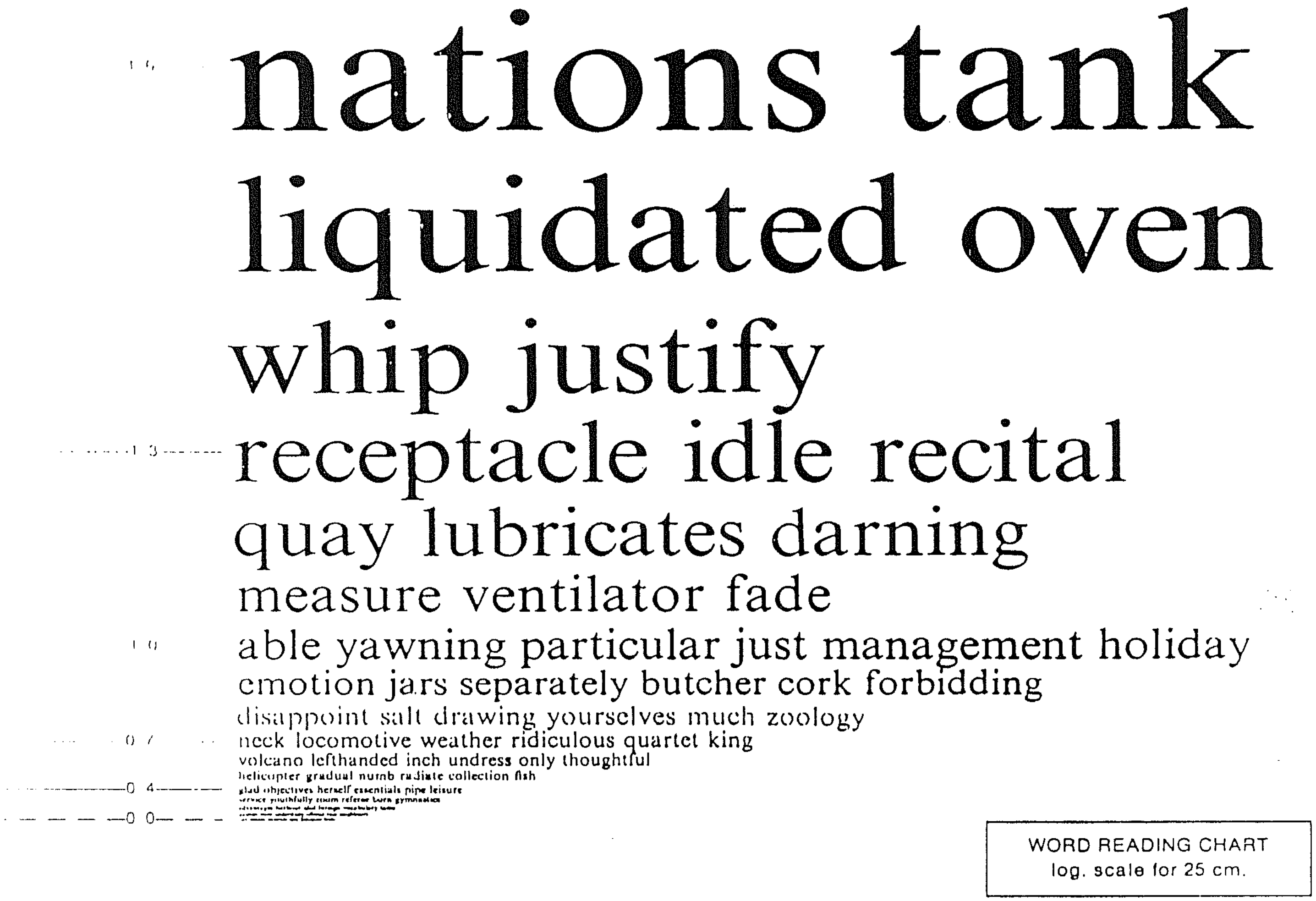
b) Detection task - narrow target



c) Resolution task - large targets widely separated compared to blur.

d) Resolution task - small targets close together compared to blur.

FIGURE 2: BAILEY - LOVIE READING CHART



LOGMAR

1.6 nations tank

liquidated oven

whip justify

1.3 receptacle idle recital

quay lubricates darning

measure ventilator fade

1.0 able yawning particular just management holiday

emotion jars separately butcher cork forbidding

disappoint salt drawing yourselves much zoology

0.7 neck locomotive weather ridiculous quartet king

volcano lefthanded inch undress only thoughtful

helicopter gradual numb radiate collection fish

0.4 glad objectives herself essentials pipe leisure

service youthfully exam referee born gymnastics

[illegible]

[illegible]

0.0 [illegible]

WORD READING CHART
log. scale for 25 cm.

FIGURE 3: EXPERIMENTAL LAYOUT

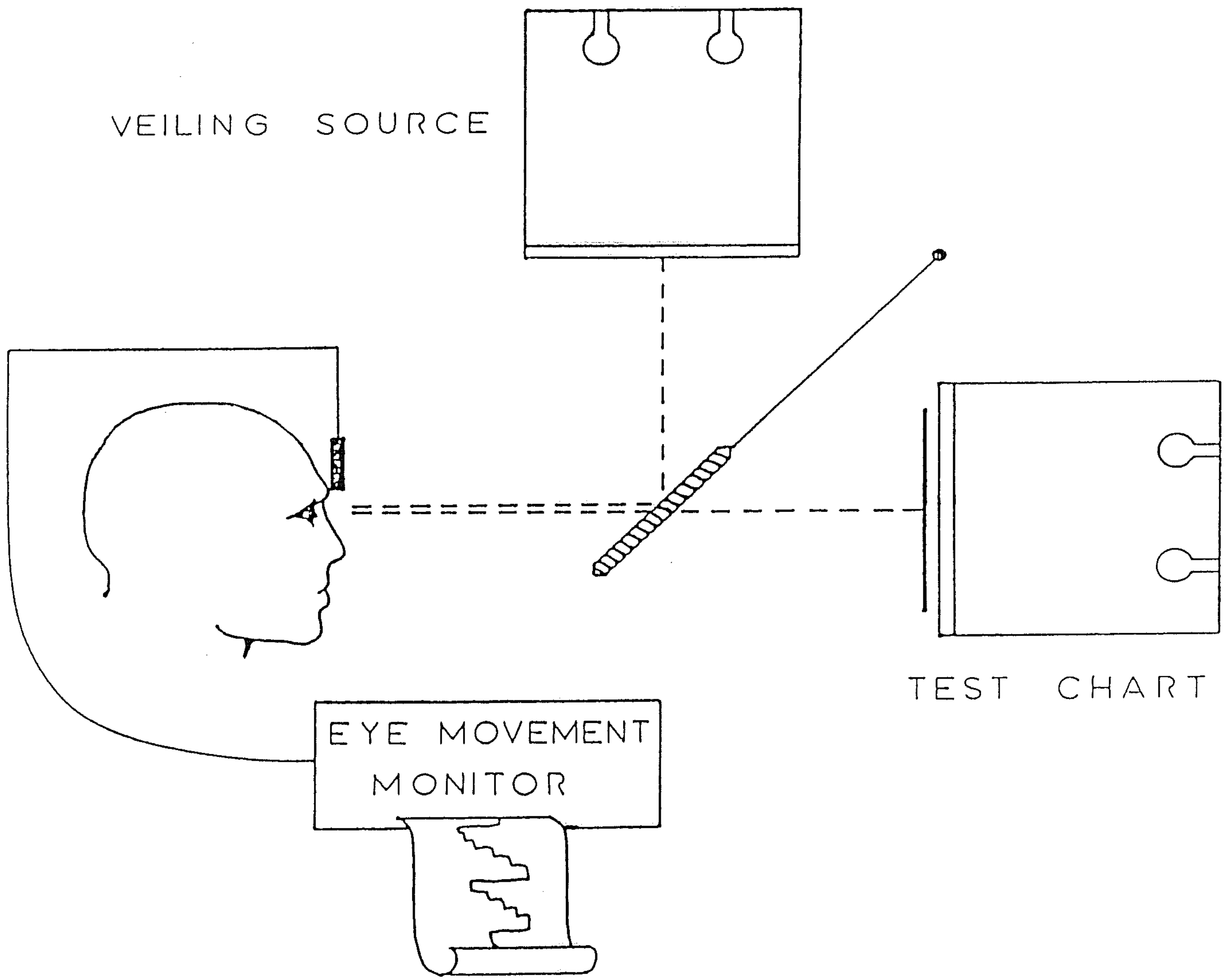


Figure 4: Schematic of strip chart recording of signal from eye-movement monitor

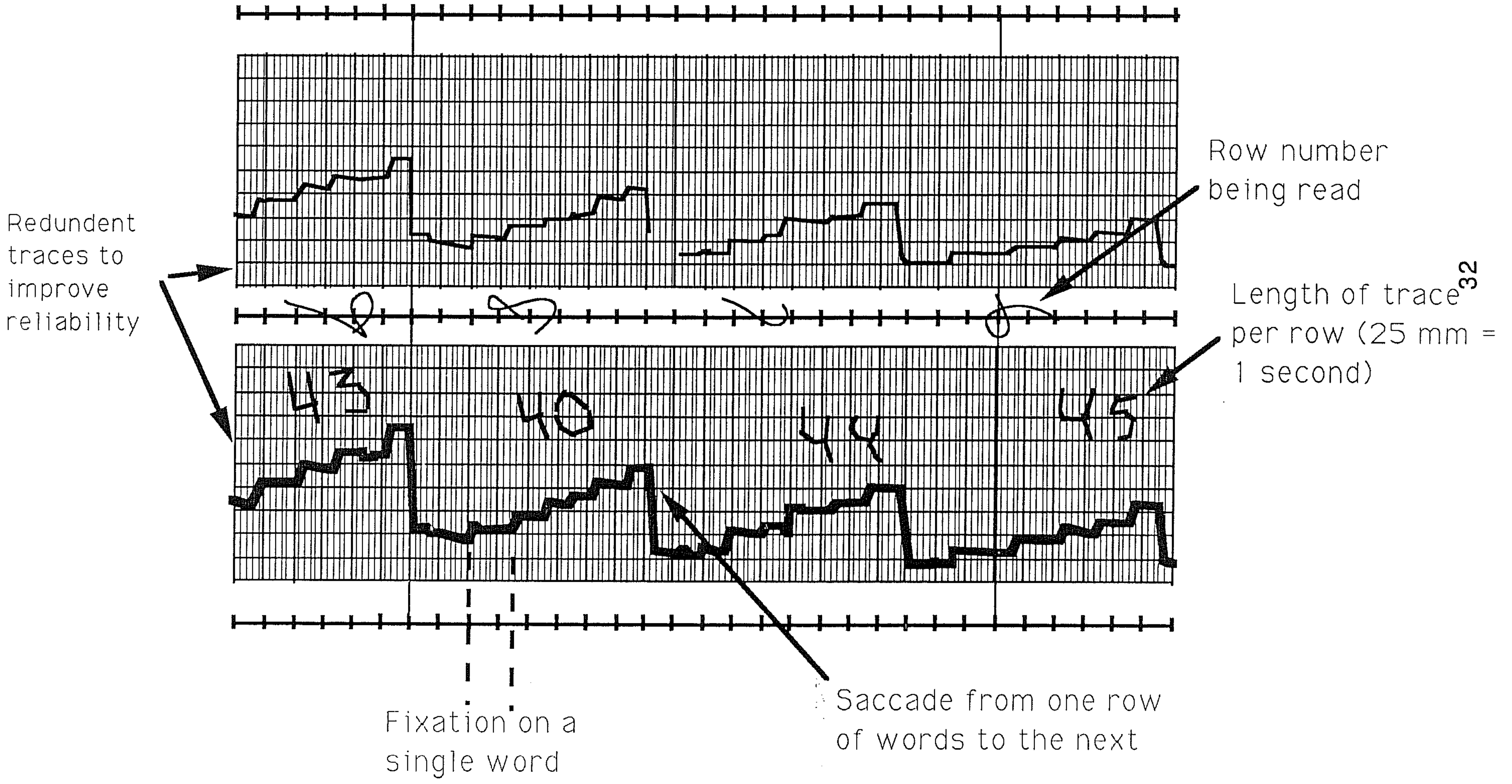


Figure 5: Measured versus predicted threshold sizes for the average data (Slope = 1, Resolution limit = 0.577)

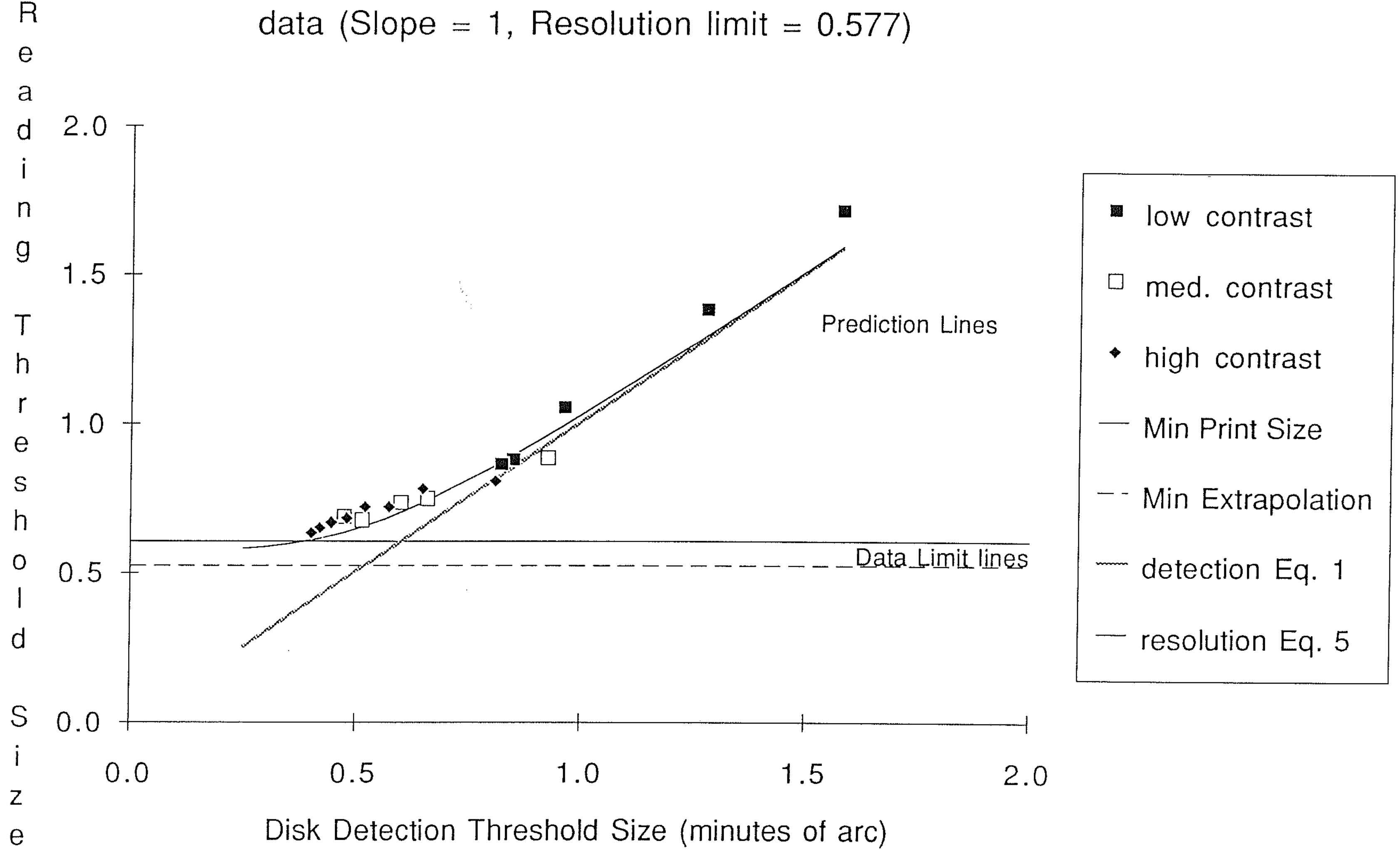


Figure 6: Measured versus predicted threshold sizes
(Subject LC: Slope = 1.25, Resolution Limit = 0.716)

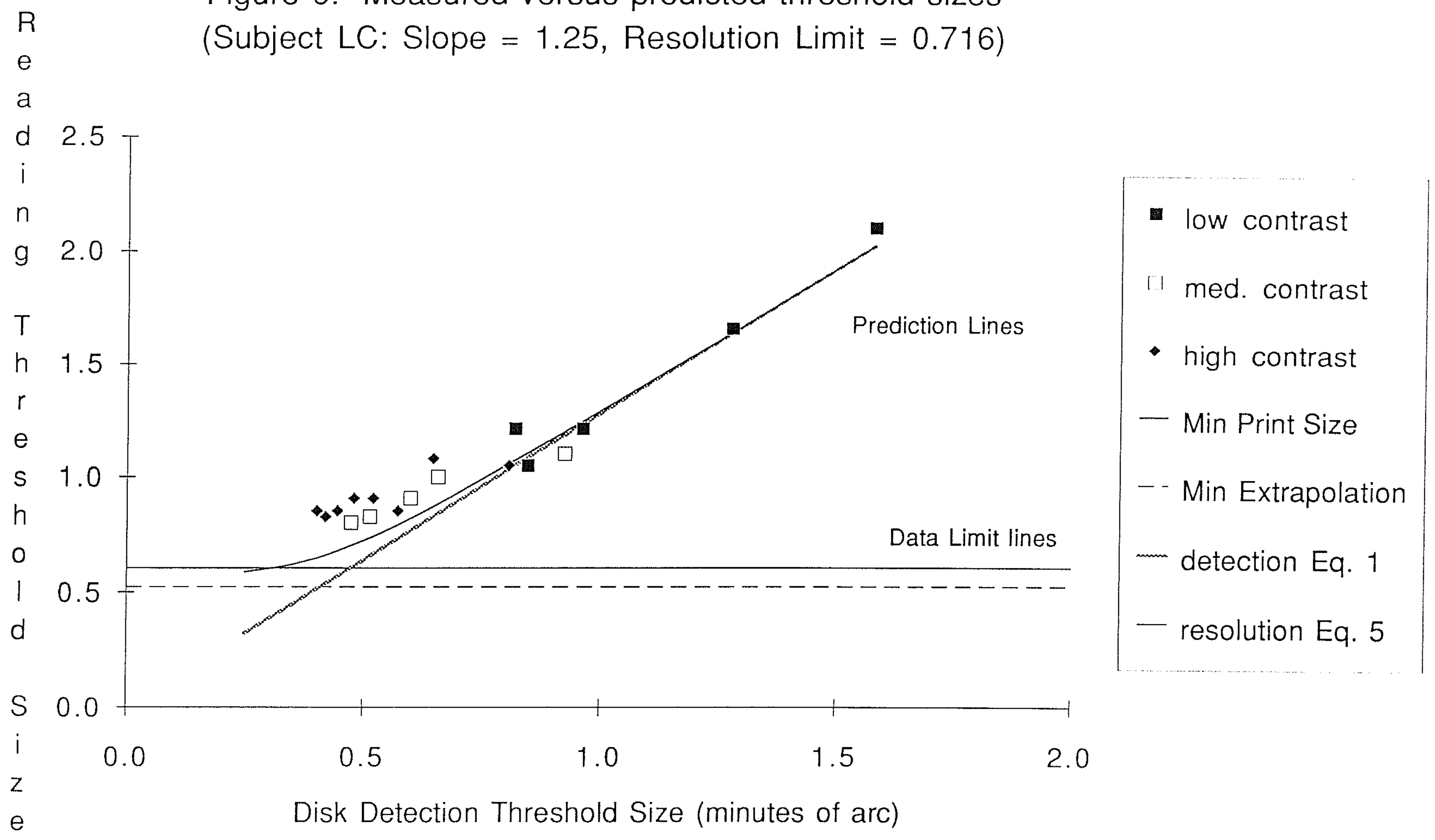


Figure 7: Reading Speed (words/sec) vs. Print Size at three Contrast levels
(15 subject Average)

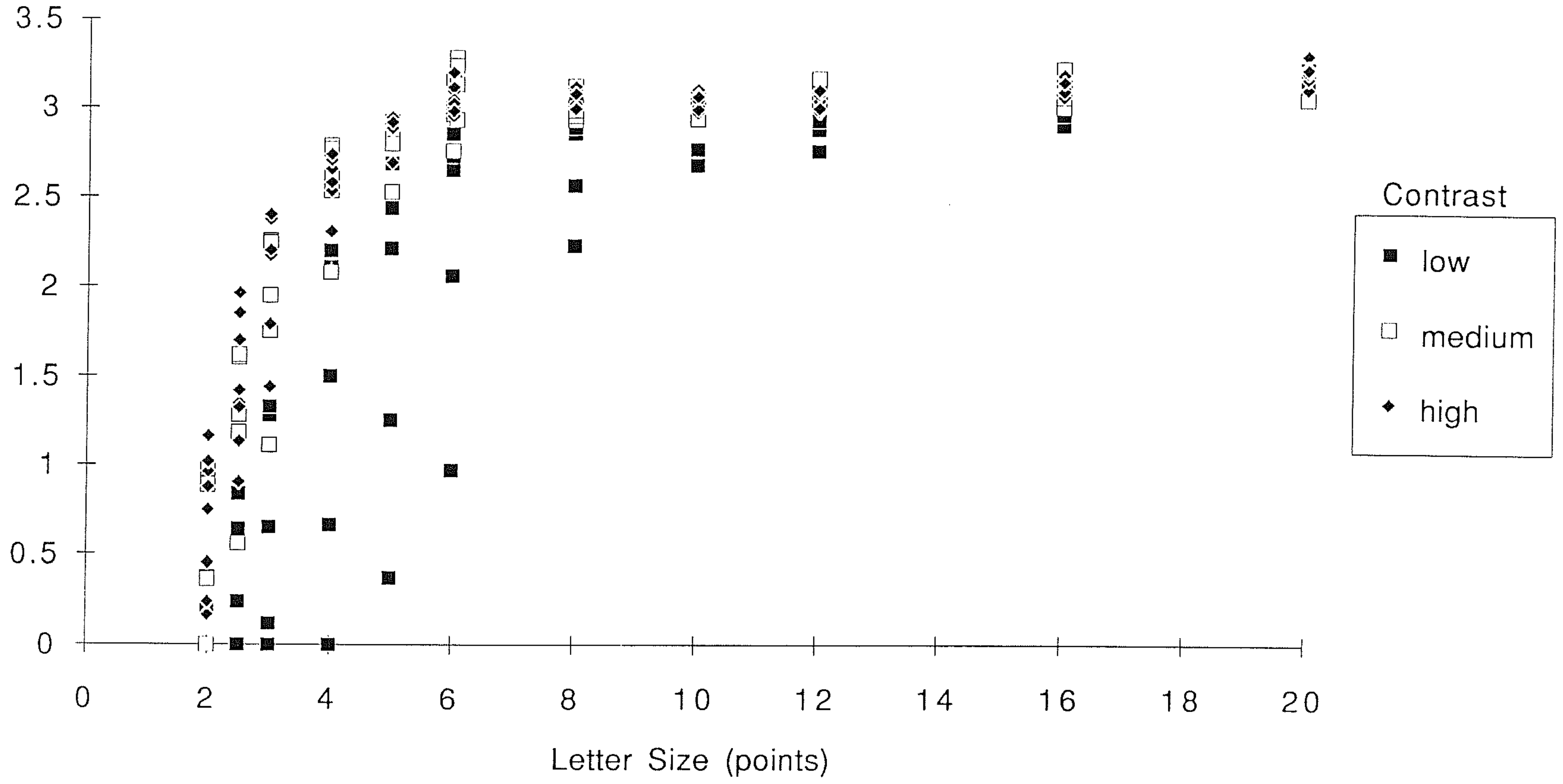


Figure 8: Normalized residuals vs. luminance for the reaction time model using the small letter assumption (6 point type & above)

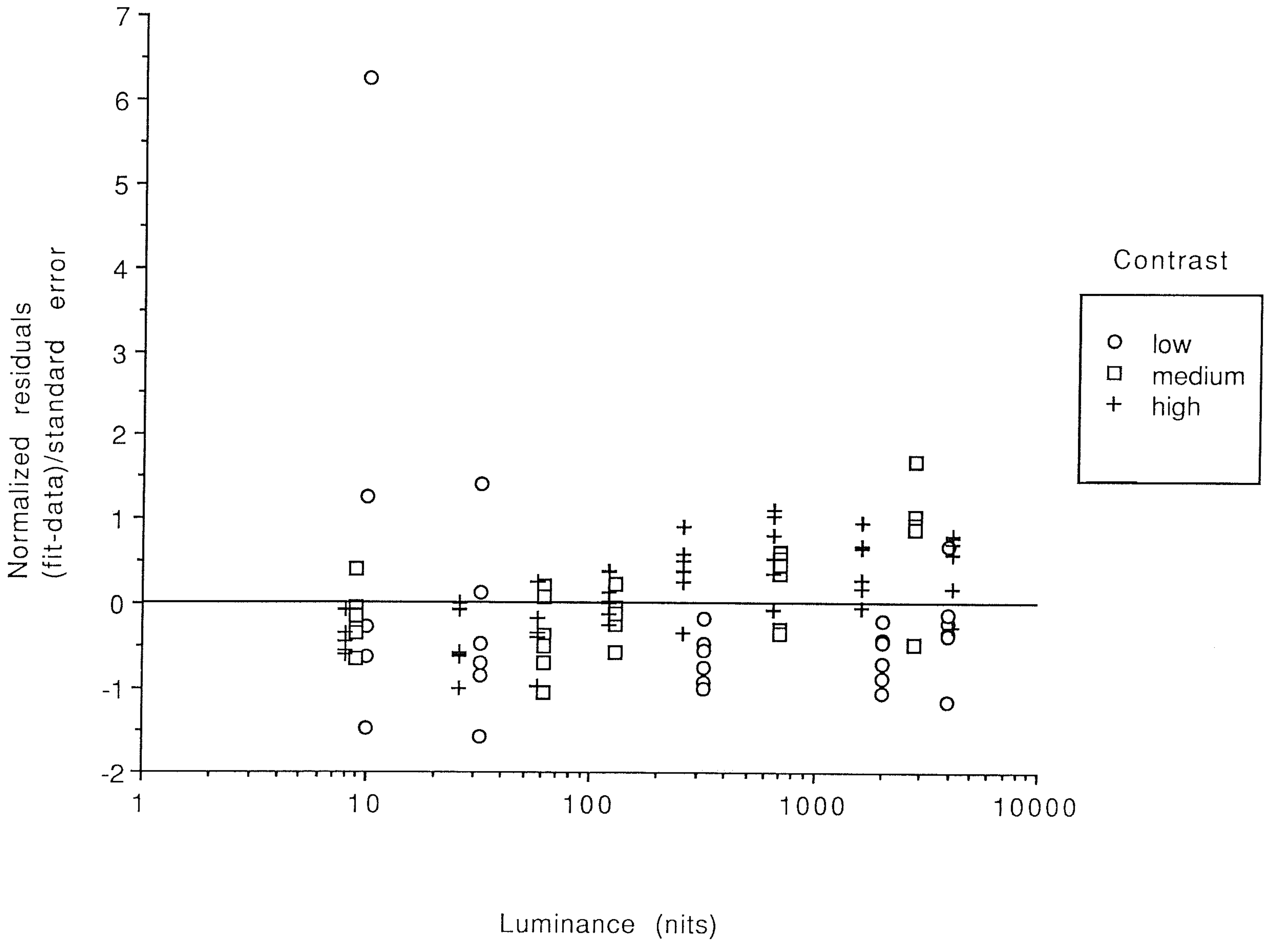


Figure 9: Residuals versus print size for the reaction time model.
The fit does not extrapolate well below 6 point type

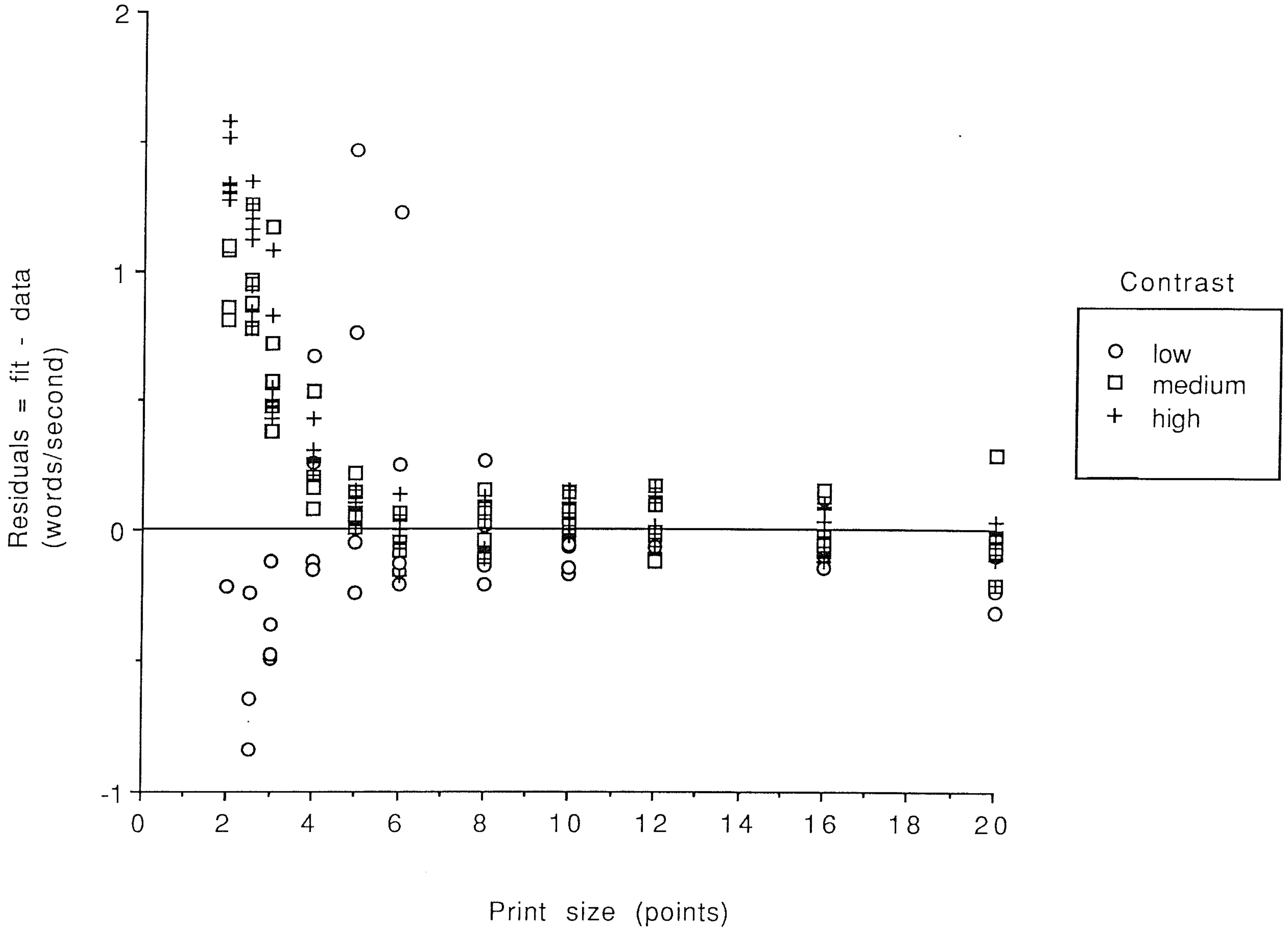


Figure 10: Normalized residuals vs. luminance for the VL(contrast) fit (5 point type & above)

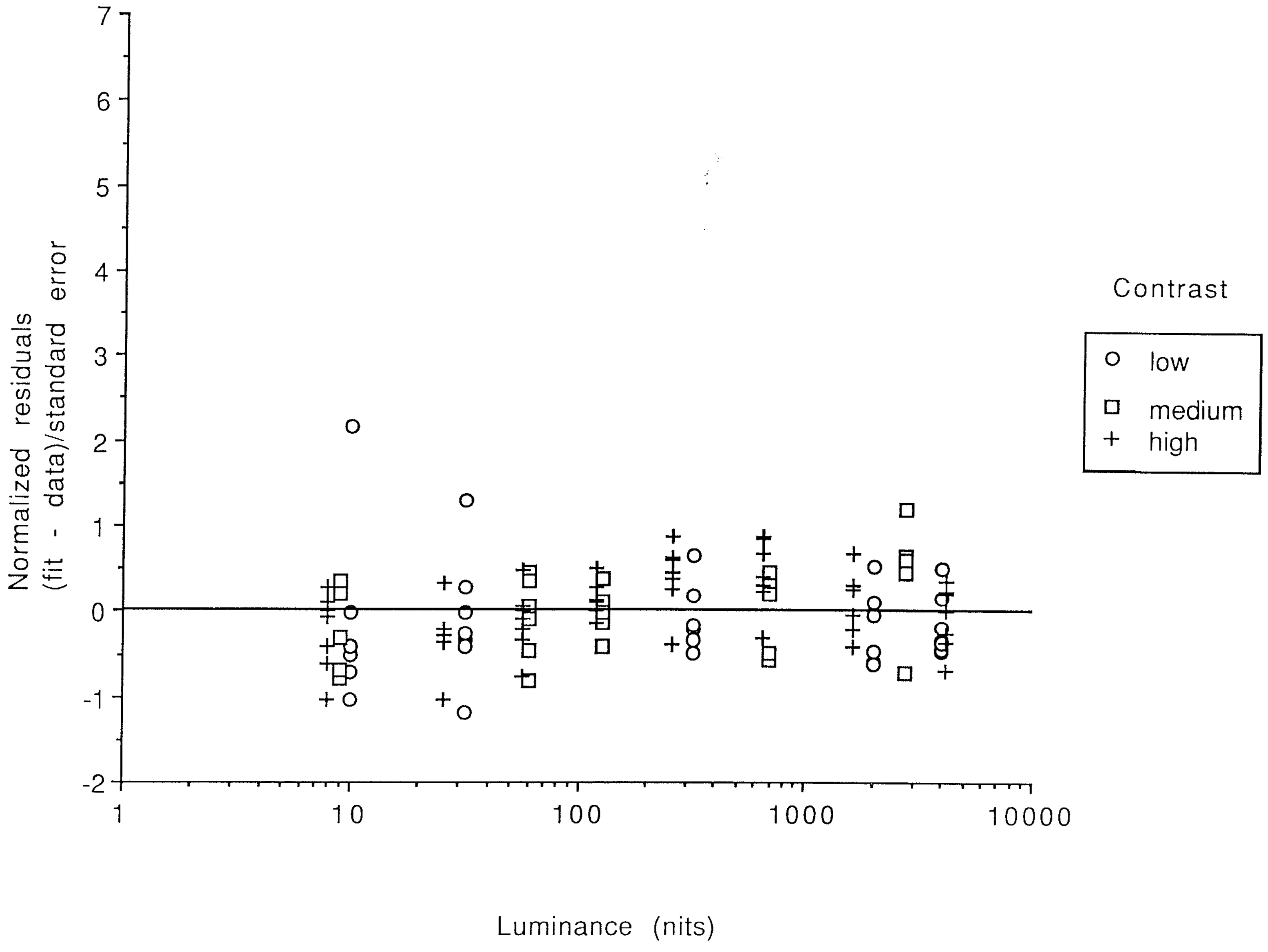


Figure 11: Normalized residuals vs. VL(contrast) for the generalized fit (4 point type & below)

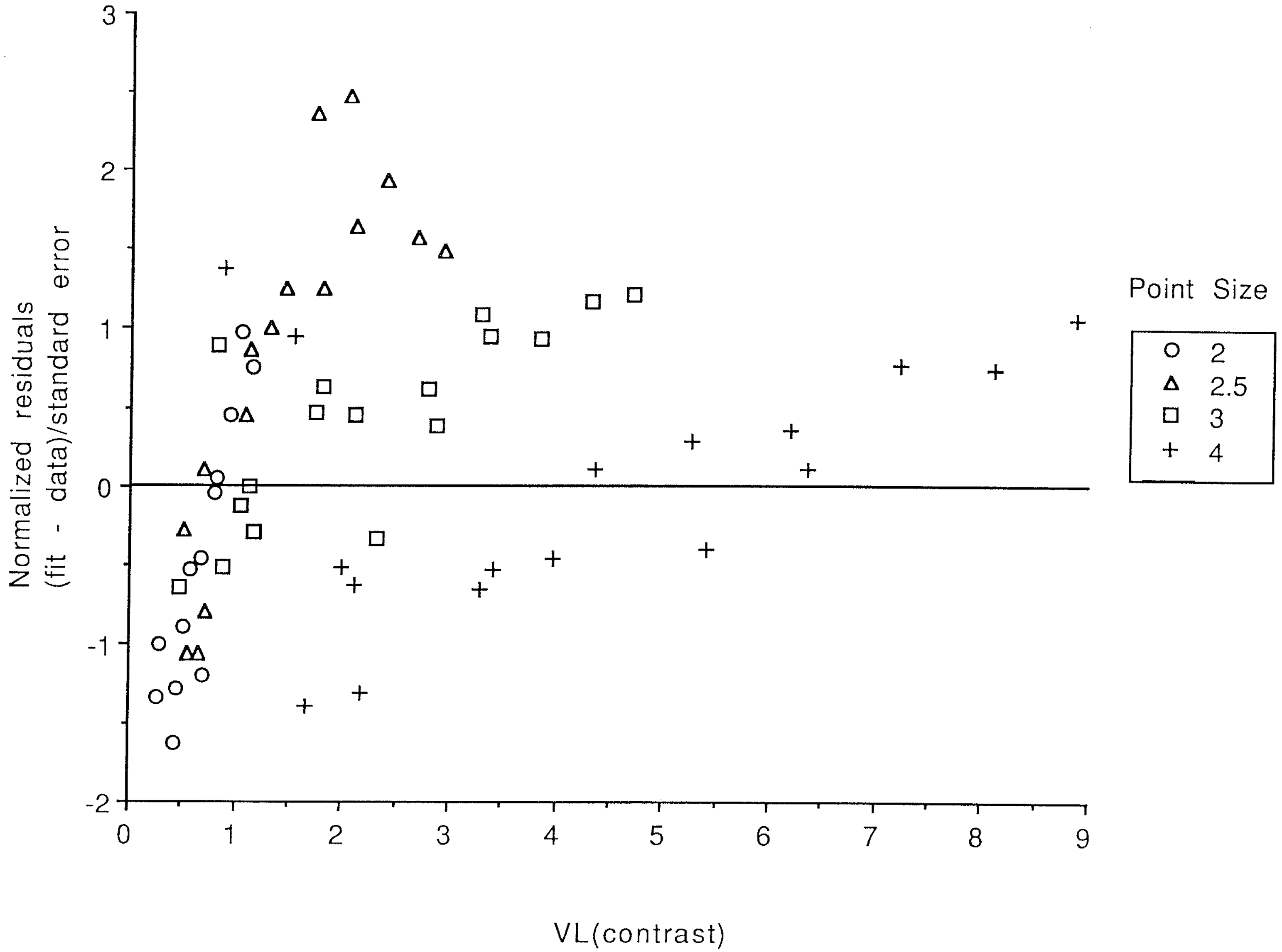


Figure 12: Reading Speed (words/second) versus VL for four print sizes - subject PS

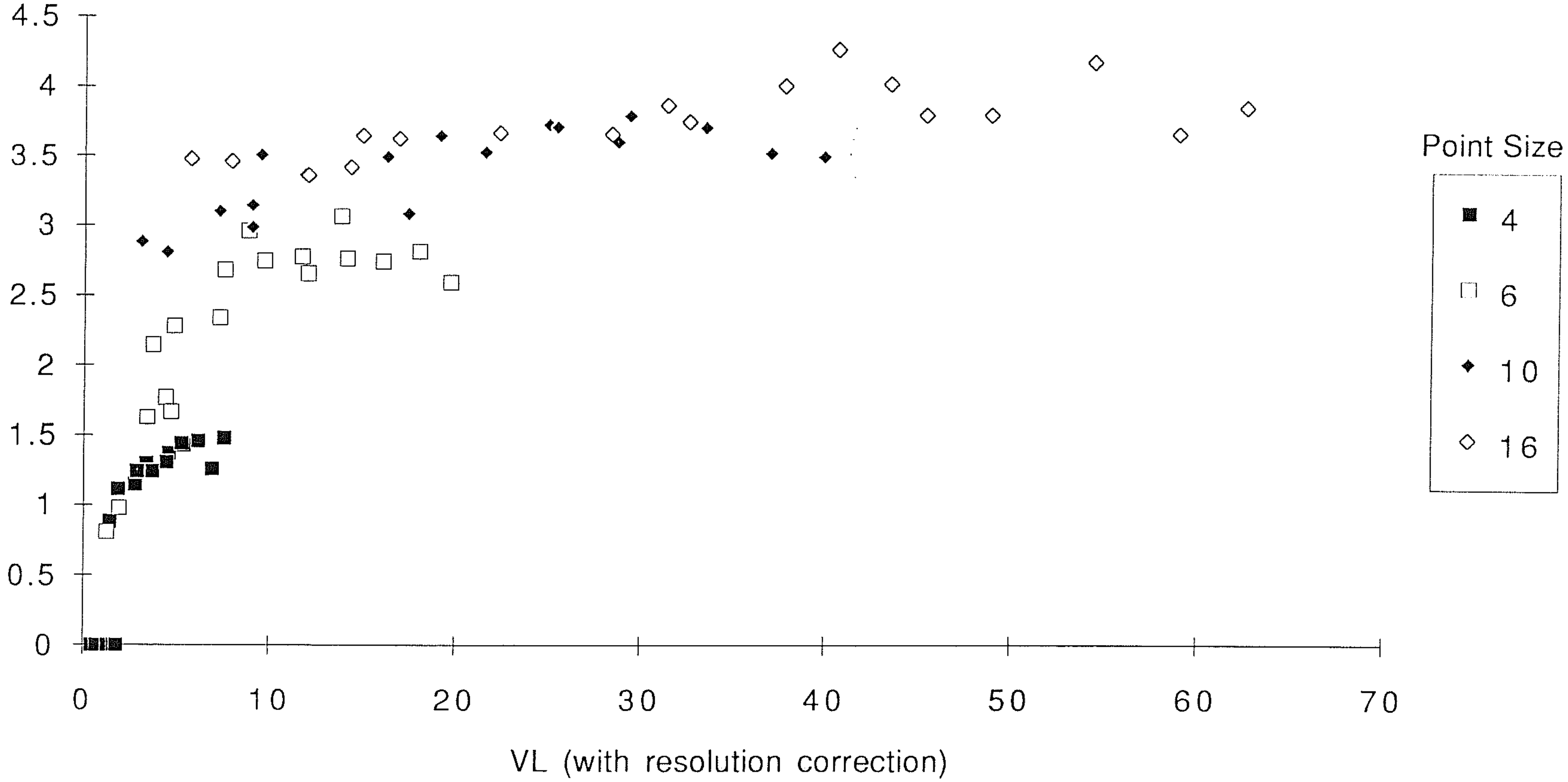


Figure 13: Reading Speed minus Reading Speed for 20 point type versus Print Size and Contrast (subject CB)

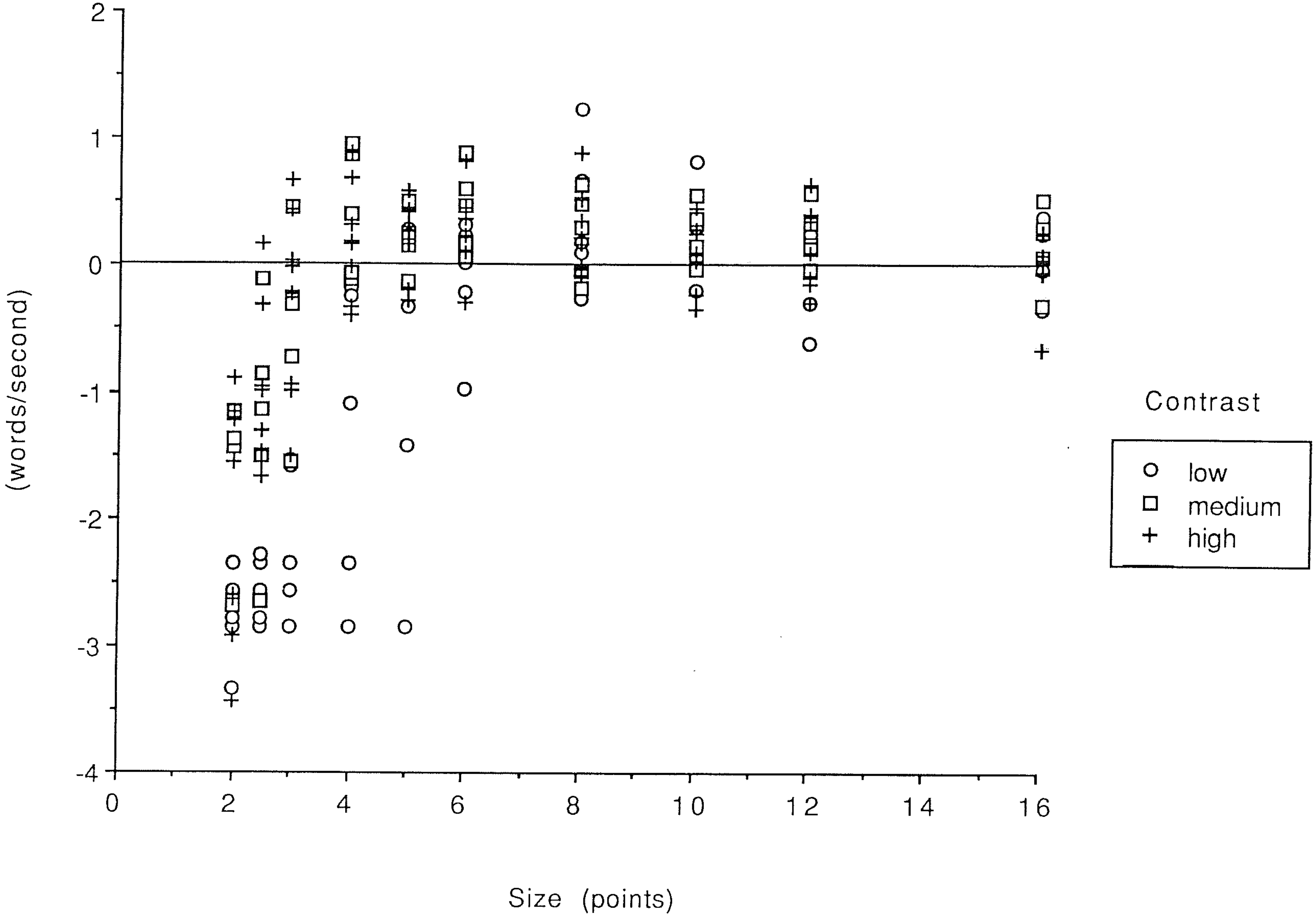


Figure 14: Normalized residuals vs. VL^2(size) for the VL(size) generalized fit
(4 point type & below)

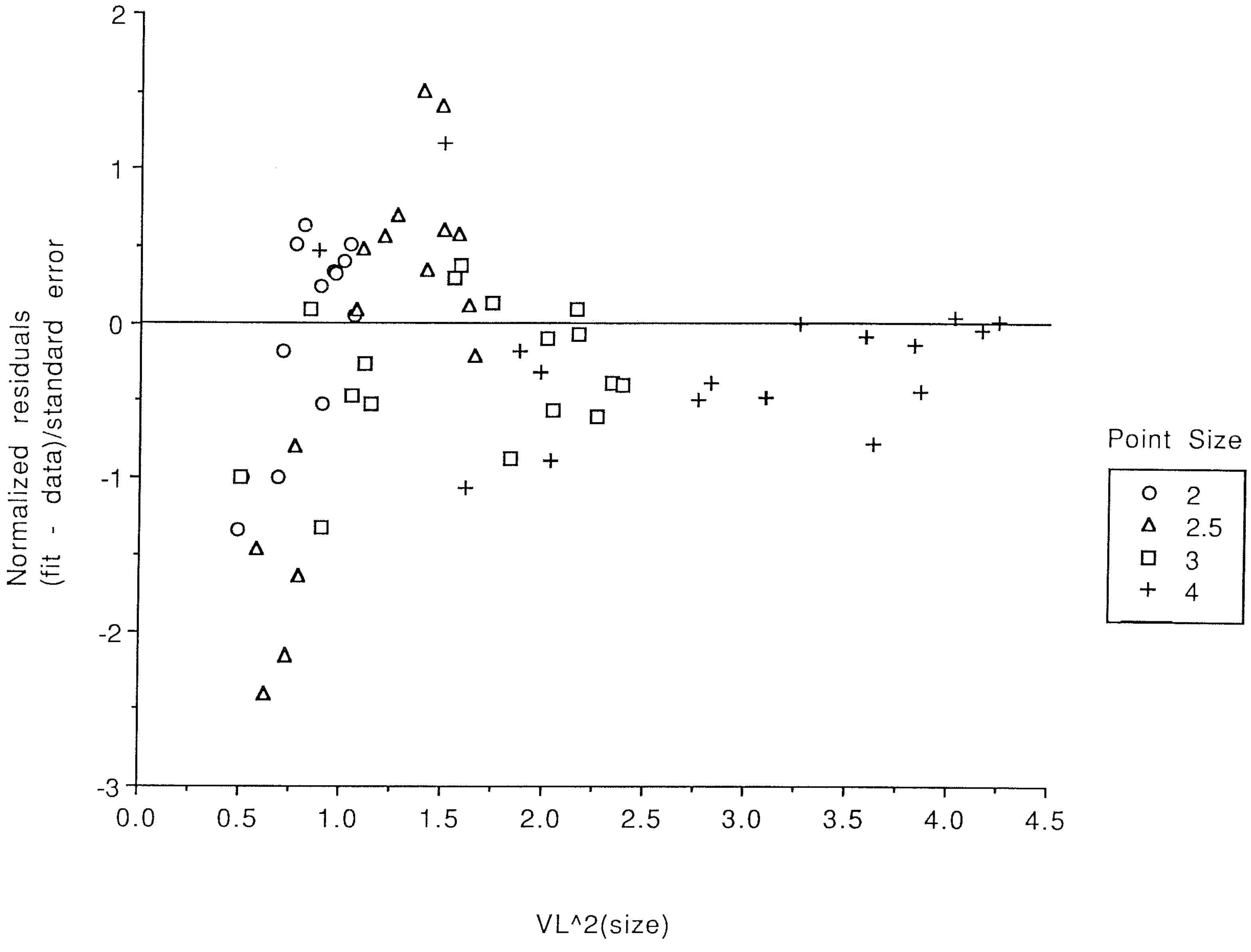


Figure 15: Performance in words/second versus VL(size). Calculated linear term added to measured values instead of subtracted from calculated values to allow plot vs. VL(size)

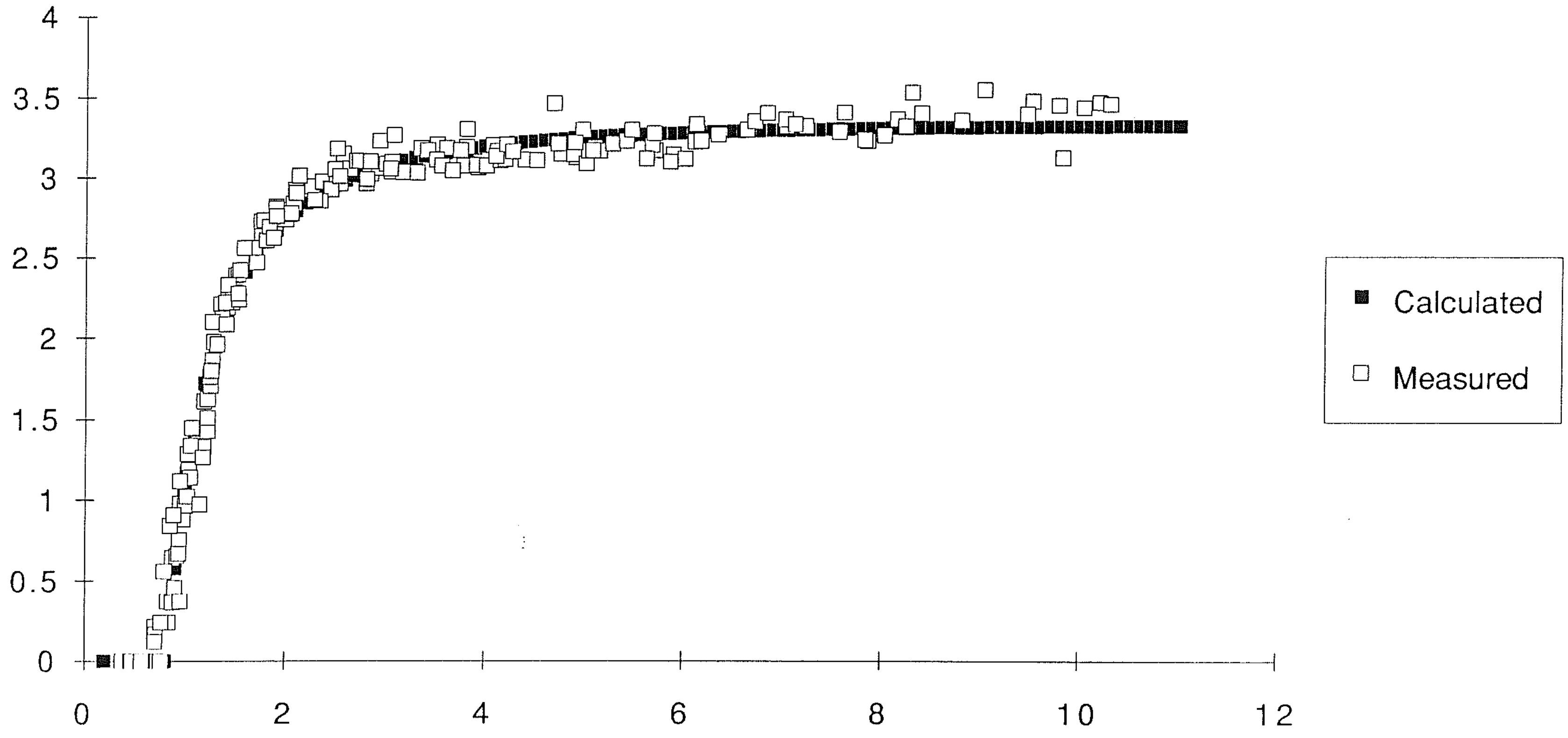


Figure 16: Relative Task Performance (RTP) versus size to give RTP = 2/3.
Data sorted and binned into groups of 15.
Error bars give standard deviation of distribution