1040-5488/96/7301-0049$03.00/0
OPTOMETRY AND VISION SCIENCE

Vol. 73, No. 1, pp. 49–53

instruments / techniques

The Freiburg Visual Acuity Test—Automatic Measurement of Visual Acuity

MICHAEL BACH
Universitäts-Augenklinik, Elektrophysiologisches Labor, Freiburg, Germany

ABSTRACT

The Freiburg Visual Acuity test is an automated procedure for self-administered measurement of visual acuity. Landolt-Cs are presented on a monitor in one of eight orientations. The subject presses one of eight buttons, which are spatially arranged on a response box according to the eight possible positions of the Landolt-Cs' gap. To estimate the acuity threshold, a best PEST (best Parameter Estimation by Sequential Testing) procedure is used in which a psychometric function having a constant slope on a logarithmic acuity scale is assumed. Measurement terminates after a fixed number of trials. With computer monitors, pixel-discreteness artifacts limit the presentation of small stimuli. By using anti-aliasing, i.e., smoothing of contours by multiple gray levels, the spatial resolution was improved by a factor of four. Thus, even the shape of small Landolt-Cs with oblique gaps is adequate and visual acuities from 5/80 (0.06) up to 5/1.4 (3.6) can be tested at a distance of 5 m.

Key Words: visual acuity, computer test, psychometric threshold estimation

A rapid and examiner-independent test of visual acuity is desirable for clinical studies. The idea of automatic presentation of optotypes on a computer screen dependent on subject response has been a subject of considerable interest.[1–14] Previous studies suffered from the inherently low resolution of standard visual display units. We have overcome this limitation by "anti-aliasing" (see below) and combined this with the "best PEST" (best Parameter Estimation by Sequential Testing), a psychometric procedure based on signal detection theory to estimate the acuity threshold that we had already used in a previous version of the Freiburg Visual Acuity Test (FAT).[13] Anti-aliasing also allows using oblique positions of the Landolt-C in addition to the four straight ones. Thus, eight different orientations are possible, which reduces measurement time as the guessing rate is lower compared to four different orientations. All calculations were based on the decimal representation of visual acuity (also known as the Snellen fraction) at an observer distance of 5 m.

Received March 21, 1995; revision received October 12, 1995.

METHODS

The schematic setup is depicted in Fig. 1. A Landolt-C optotype appears on a computer monitor and the subject responds by pressing one of eight buttons (eight alternative forced choice task). Immediately on button press, there is feedback by a gray, spatially-growing disk in place of the optotype; the growing motion attracts attention, the gray tint avoids an irritating afterimage. Optionally, this feedback can inform of the correct gap position by an appropriate missing sector in the gray disk (similar to a pie diagram). Then the next optotype is presented, the size of which is

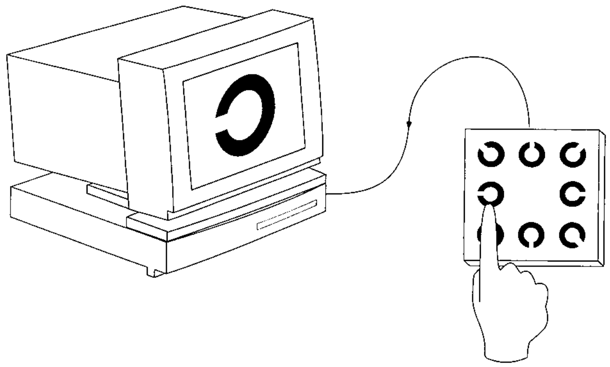

Figure 1. Schematic setup of the FAT. When a Landolt-C appears on the screen, the subject responds by pressing one of eight buttons, which are spatially arranged according to the direction of the C's gap. Immediately on button press there is visual feedback in the form of a growing "pie", where a missing piece indicates the correct position of the Landolt-C's gap. After that, the next optotype size is presented, the size of which is set according to an optimized strategy.

calculated using the best PEST procedure.[15] Eighteen trials are run to estimate the acuity threshold (see below for discussion of the number of trials). Finally, the result is calculated based on the viewing distance (which can be set to any value) and presented in large type on the screen, either in Snellen format or as decimal acuity (Snellen's fraction).

Visual Acuity as a Psychometric Threshold

It is well-known that statistical fluctuations play an important role in the estimation of visual thresholds. [16] The psychometric function relates percent correct responses to stimulus intensity (size of the optotype, Fig. 2). For very small optotypes, the hit rate corresponds to the guessing rate of ⅛ = 12.5%. "Visual acuity" is defined as the reciprocal of the Landolt-C's gap size, measured in minutes of arc, at the steepest part of the psychometric function. For the logistic function this amounts to

$$\frac{100\% - 12.5\%}{2} + 12.5\% = 56.25\%.$$

At this point small changes in optotype size translate to the largest change in the hit rate.

How does reduced visual acuity affect the psycho metric function? Because the slope of the psychometric function is fully determined by the variance of acuity measurement, one way to answer this question is to look at the variance obtained when acuity measurement sare performed across the visual field where acuity decreases with eccentricity. Westheimer did this and found the standard deviation of the minimum angle of resolution (MAR) to be proportional to the MAR.[17] Because log(acuity) = −log(MAR), the slope of the psychometric function is independent of acuity changes occurring with changes in eccentricity. Petersen, using the method of constant stimuli, verified the constancy of the psychometric function's slope across arange of physiological and pathological values of visual acuity.[18]

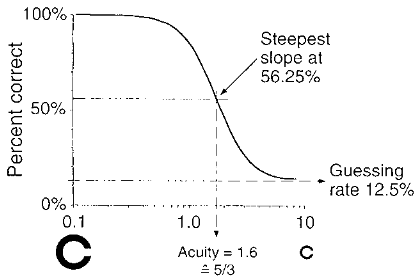

Figure 2. The psychometric function used by the test algorithm. The probability of a correct response is assumed to depend on the optotype size in the form of a sigmoid function. The threshold is set to be at the point of steepest slope. The example depicted corresponds to a visual acuity of 5/3 (1.6).

Threshold Estimation Using Best PEST

As an accurate and fast procedure to estimate psychophysical thresholds we chose the best PEST (Best Parameter Estimation by Sequential Testing).[15] Several parametric representations (cumulative Gaussian, Weibull's, and the logistic function) fit the psychometric function about equally well. In the best PEST procedure the logistic function has been chosen.[19] It describes the hit rate P depending on visual acuity as follows:

$$P_{v_o}(v) = 0.125 + \frac{1 - 0.125}{1 + \left(\frac{v_o}{v}\right)^s}$$

where s = slope, ν_o = threshold acuity.

As discussed above, we assumed the slope to be constant. Its value was set to 2 to fit the results of Petersen, who measured the psychometric function over a wide range of visual acuities.[18] Thus, the threshold is the only free parameter. On a probability scale, the usual least squares fit should not be used, so a maximum likelihood procedure is used.[20] The likelihood $L_{\nu 0}$ of the threshold being at ν_0, given a set of answers i with $n_{correct}(\nu_i)$ correct and $n_{incorrect}(\nu_i)$ incorrect responses at acuity value ν_i, is given by

$$L_{v_o} = \prod_i (p_{v_o}(v_i))^{n_{correct}(v_i)} \cdot (1 - p_{v_o}(v_i))^{n_{incorrect}(v_i)}.$$

Thus, the task is to find the value of ν_i that maximizes this expression. After each trial, the best PEST procedure calculates this maximum on the basis of all previous answers and the next stimulus is presented at that value. Because this value is, at the same time, expected to be at the steepest slope of the psychometric function, most information is gained by the outcome at this point. For the present application the test size is rounded to the nearest physically realizable value as determined by observer distance and pixel size (see below). Fig. 3 shows how the best PEST brackets the threshold first in big, then in smaller steps.

Threshold Bias in Conventional Tests Like the DIN Procedure

Although the choice of the steepest slope of the psychometrical function as definition for the threshold is a justified choice from a signal analysis point of view, the results of such an algorithm need to be related to conventional methods for measuring visual acuity. One procedure is outlined in the German DIN 58220 (industry norm) for expertises.[22] Starting at low values on the visual acuity scale, up to 10 Landolt-Cs of a given size are presented. If 6 are correctly identified, the next higher visual acuity step is presented (advancing with a factor of $\sqrt[10]{10} = 1.26$) until the criterion 6 of 10 is no longer satisfied. The previous Landolt size is then taken to represent the

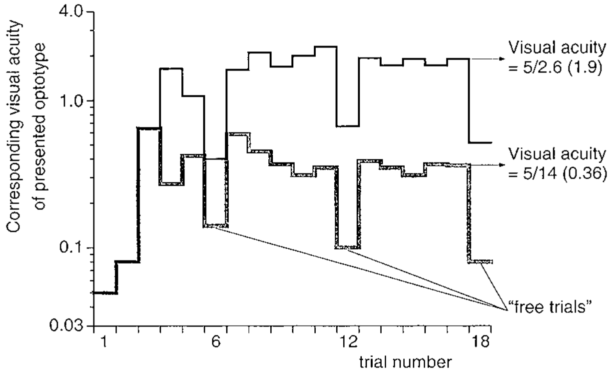

Figure 3. Estimation of visual acuity in adaptive steps based on the best PEST algorithm. Starting with large optotypes, the algorithm homes in on the threshold in decreasing step sizes, depending on the response of the subject. Free trials, where the optotype size is four times the current threshold estimate, are presented three times to motivate the subject. Two example sequences are depicted, which end in an acuity estimate of 5/2.6 (1.9) and 5/14 (0.36), respectively.

visual acuity of the subject. Although the criterion "6 of 10", or 60%, seems to correspond closely to the steepest point of the psychometric function, the DIN procedure introduces a bias toward lower acuities; whenever the subject fails to reach the 60% level, the process is terminated. The probability to pass a given level depends on (1) the level of the psychometric function at this level and (2) on the conditional probability to have passed the preceding levels. This iterative process can be described by binomial distributions. The net effect is to shift the effective hit rate for the 6 of 10 criterion to 67.56%. Fortunately, we can correct for this effect. As the slope of the psychometric function does not depend on the individual acuity, a correction factor translating between the best PEST results and those from the DIN procedure can be calculated on the basis of the mean population slope.[18] The correction factor for acuity was found to be 0.892 when translating from 56.25% to the 67.56% level; it depends very little on the individual slope: varying the slope by a factor of 2 changes the correction factor only by 2%. This insensitivity to the slope can be intuitively understood as follows. With a steeper slope, the acuity difference between the point of steepest slope and any higher hit rate will decrease. This is counteracted by an increase of the bias introduced by the DIN procedure, as the probability to fail at the next higher acuity level declines more sharply. In the FAT the correction factor can optionally be applied.

The Subject's Point of View

Performing the FAT is very simple and intuitive through the spatial arrangement and pictorial demarcation of the response box. The forced choice situation, however, is a less desirable experience for the subject. We try to help with comments like "You are doing so well, the computer makes it more difficult." It is useful to suggest rapid response, "Pondering doesn't help." Because the best PEST rapidly aims for the threshold, the subjects complain that they cannot see any gaps. Still, a "don't know" button was deliberately not used because it would lead to threshold bias by the subject's criterion.[16] Rather, the subject is asked to make her or his best guess. To reduce the level of discomfort near threshold, every six trials there is a trial with an optotype size four times higher than the current threshold estimate (Fig. 3). Although these "free trials" increase the test time, they are welcomed by the subjects.

Overcoming the Resolution Limits of Conventional Computer Monitors

What pixel size is required to display optotypes with sufficient resolution? The pixel size of current monitors is slightly below 0.4 mm. If a single pixel made up the gap in the Landolt-C, a 5 m distance would correspond to a visual acuity of 5/1.4 (3.6), which seems sufficiently high. However, a Landolt-C with 1 pixel gap size has an insufficiently defined shape, especially for oblique gap orientations; at least 2 pixels are necessary. This reduces the maximum acuity that can be presented to 5/2.8 (1.8), which is below the value reached by young normal subjects under forced choice conditions; Rassow et al. found a median acuity of 5/2.5 (1.99).[22] Furthermore, physically realizable optotypes are limited to integer gap widths (in pixels), which leads to coarse sampling in the high acuity range. Lastly, oblique gap positions still are insufficiently resolved with a 2 pixel gap.

One solution to this problem would be to increase observer distance, possibly using mirrors. This could limit the low acuity range though. Or one might use a monitor with smaller pixel size. These, however, are very expensive and are limited by electron optics, as pixels cannot become smaller than the image of the cathode. Too small an emitting surface of the cathode in turn limits electron current and consequently screen brightness. "Liquid crystal displays"[10] are currently not better, and are limited in luminance, contrast, and pixel size.

We have been able to overcome these limitations by using anti-aliasing. Anti-aliasing trades spatial resolution for luminance resolution. The principle is depicted in Fig. 4; any graphical shape must be approximated by pixels, which can be imagined like squares on calculus paper. To display a black shape that partially touches a pixel, the pixel is black if covered by more than half by the shape and is white otherwise. This produces the typical "jaggies" in computer graphics. With anti-aliasing, pixel size is the same but luminance is used to carry additional spatial information. Instead of just being black or white,

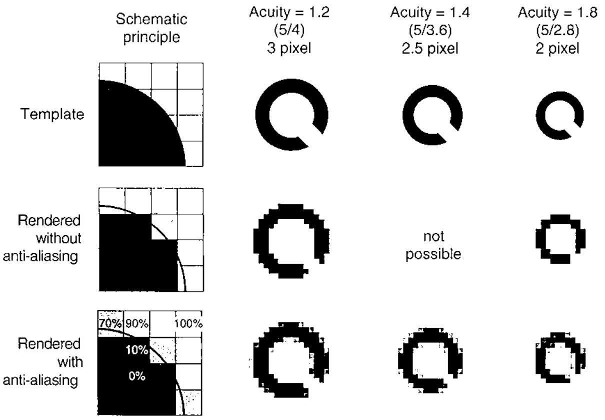

Figure 4. Anti-aliasing as used in the FAT. Top: ideal shape; center: shape rendered on a pixel raster without anti-aliasing; bottom: with anti-aliasing. With anti-aliasing, the shape is more accurately rendered and, after low-pass filtering through the eye's optics, the retinal image corresponds to one with ¼ the pixel size. The effect is difficult to present in print, and may be better appreciated by blurring the figure, e.g., viewing it from a large distance.

the pixel gray value is set according to the amount of area covered by the desired shape. In effect, shapes rendered with anti-aliasing appear smooth but slightly blurred. However, after the image has passed through the optics of the eye, the retinal image of an anti-aliased image cannot be discerned from an image rendered with a certain smaller pixel size.

The actual implementation of anti-aliasing is especially easy on Macintosh computers, as anti-aliasing is a built-in feature of the Macintosh operating system. Internally, the computer program renders the optotypes with four-fold magnification on a black-and-white pixelmap. This pixelmap is shrunk by a factor of four, simultaneously applying anti-aliasing, to a pixelmap with 16 gray levels, which is presented to the subject. Thus, the size of the Landolt-Cs can be chosen in steps of 0.25 pixels and a standard computer monitor is sufficient. An increase to 256 gray levels did not noticeably affect the results.

Number of Trials

The aim is to determine visual acuity as rapidly as possible without losing accuracy. To assess the required number of trials, one strategy is to analyze the confidence limit after each trial and stop the sequence when a given criterion is reached. However, to do so, the psychometric function must be sampled at points distant to the threshold to assess its slope. These trials provide little information on the threshold itself. Thus, in order to stop the sequence as soon as possible more trials may be necessary.

Hence, we turned to an empirical determination of the appropriate number of trials. In 14 eyes of 14 subjects, both with normal and reduced visual acuity, we measured the visual acuity for a varying number of trials (n = 8, 11, 13, 16, 19, 23, 32). For any of these 7 sequence lengths, 6 measures were obtained in an interleaved block design; for further analysis we used the mean at each sequence length. In analogy of the standard deviation, we defined a measure D_n as estimate of the deviation from the "true" visual acuity at n trials, where the true acuity was estimated as the mean at 32 trials:

$$D_n = \sum_{i=1}^{14} (v_{n,i} - v_{32,i})^2$$

where n: number of trials, n = 8, 12, . . . ; i: repeat count over the 14 subjects; $\nu_{n,i}$: average acuity for subject n at sequence length i.

With this definition, D_n catches both inter-trial variance as well as absolute deviation from the estimated true visual acuity. The value of D_n is plotted in Fig. 5. It drops sharply with increasing n and seems to reach a temporary plateau above n = 16. The further drop at n = 32 is artificial, as there D_n degenerates to the normal expression for the standard deviation and half of the inter-trial variance drops from the equation. Based on these empirical findings, we set the standard number of trials at 18. This includes one "free trial" (enlarged optotype for motivation) as the last trial, leaving the subjects satisfied with the run.

Comparison to a Conventional (DIN) Test

We compared the visual acuity obtained with a conventional acuity test (following the DIN 58220[22] procedure as described above) in 25 subjects both with their best correction and also defocused to cover an acuity range of 5/16 (0.32) to 5/1.6 (3.2). The quotient between the two test results averaged to 0.991 after applying the factor correcting for the difference between the steepest point on the psychometric function and the con-

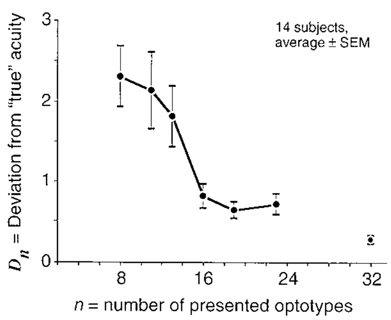

Figure 5. Deviation from the true visual acuity as a function of the number of optotypes presented (sequence length). As a measure of deviation from true visual acuity, D_n (see text) is depicted. At n = 32, D_n degenerates to the standard deviation. Based on the finding that there is a plateau above 16 trials, the standard number of trials was set to n = 18, including one "free" trial at the end of the sequence.

ditional probability-bias introduced by the DIN procedure as detailed above, indicating a high degree of agreement between the two methods.

DISCUSSION

The FAT yields the same results as conventional forced choice chart testing (e.g., DIN 58220[22]) while being more rapid.[13] In our experience, it is well liked by the patients, which agrees with results by Reading and Weale.[14] These authors found high readiness for self-paced testing in all but very old patients.

The FAT improves on other automated procedures in the following respects.

1. With best PEST the FAT implements an accurate and rapid psychometric procedure to estimate the threshold. The automated procedure reduces examiner bias.

2. The FAT runs on readily available Macintosh computers without technical modification.

3. Screen resolution is improved by anti-aliasing, which allows oblique Landolt-C orientations even at high visual acuities. This reduces the measurement time as the guessing rate is at ⅛ as compared to previous work, which was limited to the 4 straight orientations because of pixel resolution limits.

4. Operation of the test is delightfully simple.

5. A number of options satisfy special requests: —The test can be set to use only four as opposed to eight Landolt-C positions to reduce the possibility of confusion in pediatric applications. The parameters of the best PEST are then automatically adjusted and the number of trials is increased.—The results can be optionally represented as Snellen ratio or decimal value.—The results can be rounded to $\sqrt[10]{10}$ DIN steps.—An adjustable time-out can be used.—Flanking bars can be added to induce a crowding situation.—A Landolt-C contrast sensitivity test is included, using dithering to achieve high contrast resolution.

The results of the FAT are in high agreement (to within 1%) with those obtained following the DIN 58220 procedure (for expertises, in Germany acuity must be measured according to this standard). The current version of the FAT is available free of charge from the author.

ACKNOWLEDGMENTS

Financial support by the German Ophthalmological Society (DOG) and the "Gesellschaft zur Förderung der Feinmechanik und Optik" is gratefully acknowledged. I thank Andreas Jedynak, who programmed intermediate versions; Thomas Meigen, who ran simulations to help in understanding the properties of the best PEST, Michael Tan, who performed the measurements for determining the optimal number of trials, and Mark W. Greenlee for reviewing the manuscript. Special thanks are due to Hans Strasburger for extensive critical review of the paper.

REFERENCES

1. Crossman ERFW, Goodeve PJ, Marg E. A computer-based automatic method for determining visual acuity. Am J Optom Physiol Opt 1970;47:344–55.
2. Timberlake GT, Mainster MA, Schepens CL. Automated clinical visual acuity testing. Am J Ophthalmol 1980;90:369–73.
3. Huber C. Computergrafik für den Augenarzt. Klin Monatsbl Augenheilkd 1984;184:439–41.
4. Friendly DS, Weiss IP. An automated visual acuity testing computer program using the Apple II system. Am J Ophthalmol 1985;99:188–92.
5. Huber C. Sehtest-Programm für Augenärzte auf Apple II, IIe und IIc Computer. Klin Monatsbl Augenheilkd 1985;186: 227–30.
6. Wong D, Plumb A. Computer automated visual acuity testing for visual screening. Trans Ophthalmol Soc UK 1986; 105:498–503.
7. Raymond JE, Shapiro KL. Usefulness of computerized visual acuity testing in a pediatric ophthalmology clinic. Can J Ophthalmol 1988;23:213–5.
8. Effert R, Steinmetz H, Jansen W, Rau G, Reim M. Automatisierte Messung des Fernvisus in Anlehnung an die DIN-Strategie. Klin Monatsbl Augenheilkd 1989;195:39–43.
9. Diepes H. Erzeugung von Sehproben mit dem Computerbildschirm—Möglichkeiten und Grenzen. Optometrie 1990; 1:3–11.
10. Ettelt B, Brandl H, Zrenner E, Lund OE. Visustestgerät: Möglichkeiten und Grenzen der LCD-Technik. Fortschr Ophthalmol 1991;88:59–63.
11. Strobel J, Jacobi F. Automatische Prüfung von Sehfunktionen durch Computereinsatz. Fortschr Ophthalmol 1989; 86:370–3.
12. Huber C. Prüfung der Sehfunktion mit Bildschirmgraphik auf dem Macintosh II Rechner. Klin Monatsbl Augenheilkd 1992; 200:589–91.
13. Mittelviefhaus K, Bach M, Jedynak A, Kommerell G. Der Freiburger Visustest. Ophthalmologe 1993;90:132–5.
14. Reading VM, Weale RA. Self-administered automatic sight-testing. A feasibility study. Doc Ophthalmol 1993;83:43–54.
15. Lieberman HR, Pentland AP. Microcomputer-based estimation of psychophysical thresholds: the Best PEST. Behav Res Methods Instrum Comput 1982;14:21–5.
16. Green DM, Swets JA. Signal Detection Theory and Psychophysics. New York: Wiley, 1966.
17. Westheimer G. Scaling of visual acuity measurements. Arch Ophthalmol 1979;97:327–30.
18. Petersen J. Zur Fehlerbreite der subjektiven Visusmessung. Fortschr Ophthalmol 1990;87:604–8.
19. Harvey LO. Efficient estimation of sensory thresholds. Behav Res Methods Instrum Comput 1986;18:623–32.
20. Collett D. Modelling Binary Data. London: Chapman & Hall, 1991.
21. Rassow B, Cavazos H, Wesemann W. Normgerechte Sehschärfenbestimmung mit Buchstaben. Augenärztliche Fortbildung 1990;13:105–14.
22. Sehschärfenbestimmung, Prüfung für Gutachten (DIN 58220). DIN Deutsches Institut für Normung e.V. (ed). DIN Taschenbuch 177. Berlin: Beuth-Verlag, 1991.

AUTHOR'S ADDRESS:
Michael Bach
Universitäts-Augenklinik
Elektrophysiologisches Labor
D-79106 Freiburg
Germany