Published in final edited form as:
Vision Res. 2006 September ; 46(18): 2956–2967.

Development and evaluation of a linear staircase strategy for the measurement of perimetric sensitivity

Rizwan Malik[a,*], **William H. Swanson**[b], and **David F. Garway-Heath**[a]

a *Glaucoma Research Unit, Moorfields Eye Hospital, London, EC1V 2PD, UK*

b *SUNY State College of Optometry, New York, NY 10036, USA*

Abstract

Perimetric sensitivity of patients with glaucoma has traditionally been measured in logarithmic (dB) units, but linear sensitivity correlates better with conventional structural measures of glaucomatous damage. Monte Carlo simulations of perimetric algorithms were used to assess potential effects of logarithmic steps on bias and variability when perimetric sensitivity was represented in linear units, and to assess the potential benefits of algorithms using linear steps. Simulations predicted that linear staircases could reduce the sensitivity-dependence of bias, variability and efficiency. These predictions were supported by a perimetric study of 21 patients with glaucoma and 20 age-similar controls who made repeat visits over several weeks.

Keywords

Monte Carlo simulation; Differential light sensitivity; Linear units; Perimetry

1. Introduction

Standard automated perimetry plays a central role in the assessment of visual function of patients with glaucoma, as do structural assessments of the optic nerve and the retinal nerve fibre layer (Weinreb & Khaw, 2004). Converging empirical and theoretical analyses support the hypothesis that linear perimetric sensitivity correlates better with ganglion cell number than logarithmic sensitivity (Garway-Heath, Holder, Fitzke, & Hitchings, 2002; Hood et al., 2002; Reus & Lemij, 2004; Swanson, Felius, & Pan, 2004).

Evaluation of the hypothesis of linearity is confounded by the fact that conventional perimetric algorithms measure sensitivity in logarithmic steps, with the result that in linear units there are relatively few steps at high sensitivities and many steps near the lower limits of the testing apparatus (Harwerth, Carter-Dawson, Smith, & Crawford, 2005).

Logarithmic units (such as decibels) relate to the maximum stimulus luminance available for a given perimeter and represent certain attenuation from this maximum, so that 1 dB on a given instrument is not equivalent to 1 dB on another (under the same test conditions and background luminance) unless they have same dynamic range. For instance, the maximum stimulus on the Humphrey Field Analyzer is 10,000 apostilbs (asb), whilst that on the Goldmann perimeter is 1000 asb so that 1 dB represents a stimulus of 8000 asb on the Humphrey Field Analyzer but 800 asb on the Goldmann perimeter.

* Corresponding author. Fax: +44 20 7566 2826. E-mail address: rzmalik@doctors.org.uk (R. Malik).

The use of logarithmic algorithms for measuring sensitivity with conventional perimetry is a potential source of bias and variability when sensitivity is converted to linear units for comparison with structural and electrophysiological measurements. Conventional perimetric algorithms employ luminance increments at multiple locations in the visual field and report differential light sensitivity (DLS) in decibel (dB) units, where 10 dB = 1 log unit of attenuation. Such algorithms typically use 4–7 trials per location in visual space and return about 20 possible sensitivities in 0.2 log unit steps. In linear units of sensitivity, the upper half of the range of possible sensitivities is therefore represented by only a few values, with most of the sensitivities compressed into the bottom 10% of the range. This makes it likely that, when sensitivity is scaled in linear units, the accuracy and precision of conventional perimetric algorithms will be lower in regions of high sensitivity than in regions of low sensitivity, and that test-time will show the opposite tendency. Sensitivity-dependence for accuracy, precision, and efficiency are potentially serious complicating factors for comparisons of perimetric and structural measures of glaucomatous damage.

Computer simulations have been extensively used to predict properties of perimetric strategies (e.g., Chauhan & Johnson, 1994; Johnson, Chauhan, & Shapiro, 1992; Spry, Johnson, McKendrick, & Turpin, 2003; Turpin, McKendrick, Johnson, & Vingys, 2003). Analysis of simulations has certain advantages over the use of data from human subjects for this purpose: computers are not subject to the effects of fatigue or emotion, a large number of permutations can be modelled in a short amount of time and bias of sensitivity measurement can be computed since the ‘true’ sensitivity is known. Monte Carlo simulations, which involve generation of random numbers, are widely used to analyze situations where a large number of outcomes are possible, such as nuclear science, traffic flow systems, and cancer therapy, as well as perimetry (e.g., Anderson, 2003; Anderson & Johnson, 2003; Glass, Schaumberger, & Lachenmayr, 1995; Maloney, 1990).

The purpose of the present study was to evaluate the sensitivity-dependence of accuracy, precision, and efficiency of traditional logarithmic staircases when the results are expressed as linear perimetric sensitivity, and to develop strategies for which bias and precision are less dependent on sensitivity. Monte Carlo simulations of perimetric strategies were used to assess accuracy, precision, and efficiency for both conventional logarithmic staircases and alternative linear staircases. The simulations were used to identify a linear strategy similar in overall efficiency to conventional logarithmic staircases. The predictions from the simulations were then evaluated by using both linear and logarithmic algorithms to measure perimetric sensitivity and variability in a group of patients with glaucoma and an age-similar control group.

2. Methods

2.1. Monte Carlo simulations of perimetry

2.1.1. Linear units—DLS values were expressed as linear perimetric sensitivity, using units of L^{-1} (1/Lambert), where 1.0 Lambert is the maximum stimulus luminance available on the Humphrey Field Analyzer. We used

$$L^{-1} = 10^{(dB/10)},$$

where dB represents sensitivity in conventional perimetric units (number of 0.1 log unit steps from the maximum stimulus) and L^{-1} represents sensitivity in linear units. The high end of the normal range is about 34 dB (Heijl, Lindgren, & Olsson, 1987), which in linear sensitivity is 2512 L^{-1}, so we simulated responses for a sensitivity range of 0 to 2500 L^{-1}. A sensitivity of 1.0 L^{-1} corresponded to a threshold equal to the maximum stimulus luminance and sensitivities lower than this were scored as ‘0.’

2.1.2. Description of Monte Carlo computer simulation—For each perimetric strategy evaluated, Monte Carlo methods were used to simulate perimetric observers performing large numbers of staircases, for both normal-seeing and damaged regions of the visual field, using summary statistics to estimate accuracy, precision and efficiency. For a given stimulus presentation, the simulation yielded a response of either 'seen,' or 'not seen.' When the response was 'seen,' the staircase proceeded to a stimulus of lower luminance, and when the response was 'not seen' the staircase proceeded to a stimulus of higher luminance. A *reversal* occurred when the change in stimulus luminance yielded a change in response from 'seen' to 'not seen' or vice versa. Until the first reversal, the *step size* in changing from one luminance to another was constant in either logarithmic units (conventional staircase) or in units of linear sensitivity (new staircases). After a reversal, the staircase either terminated (if the criterion number of reversals was met) or else the step size was reduced. When a staircase terminated, the simulated output sensitivity was set to the reciprocal of the last stimulus luminance with a response of 'seen.' If a response of 'seen' was not given even for the maximum stimulus, the simulated output sensitivity was set to zero.

For a given stimulus presentation, the response was generated by comparing a random number from the uniform distribution [0, 1] with the *probability of responding function*, $R(x)$, where x is the luminance of the stimulus in Lamberts (since one Lambert is the maximum stimulus, values of x never exceeded 1.0). When the random number was greater than $R(x)$, the response was 'seen,' otherwise the response was 'not seen.' The function $R(x)$ was defined by four parameters: threshold, α (the luminance, in Lamberts, that is seen 50% of the time); slope of Weibull function, β (determining intrinsic variability); false negative rate, FN (the fraction of trials on which the stimulus was seen but not responded to); false positive rate, FP (the fraction of trials on which the stimulus was not seen yet a response was generated). The probability of seeing the stimulus, $P(x)$, was defined as

$$P(x) = 1 - 2^{-(x/\alpha)\wedge\beta}, \tag{1}$$

and the probability of responding, $R(x)$ was defined as

$$R(x) = (1 - \mathrm{FN})P(x) + \mathrm{FP}(1 - P(x)) = \mathrm{FP} + (1 - \mathrm{FP} - \mathrm{FN})P(x). \tag{2}$$

It is well-established that the variability of logarithmic perimetric sensitivity tends to increase as sensitivity decreases. The simulations used two different models for this increase: intrinsic noise and heterogeneous damage. Intrinsic noise reflects the noise within the signal used for psychophysical sensitivity, and was modelled by varying the slope parameter, β, using the equation of Henson, Chaudry, Artes, Faragher, and Ansons (2000):

$$\ln(\mathrm{SD}) = 3.27 + 0.81 * \log(\alpha), \tag{3}$$

where SD is the standard deviation of noise in the perimetic signal in dB units and $1/\alpha$ is the measured sensitivity. We the computed the slope parameter, β, from

$$\beta = 10/(\mathrm{SD} * \mathrm{sqrt}(2)). \tag{4}$$

Typical slopes from Eq. (4) were 5.0 for a sensitivity of 36 dB (3981 L^{-1}) and 1.4 for a sensitivity of 20 dB (100 L^{-1}).

Heterogeneous damage was modelled by randomly removing ganglion cells from a mosaic and computing the effect on sensitivity of psychophysical spatial mechanisms which sample the responses of the degraded ganglion cell mosaics, using the model given by Pan, Swanson, and Dul (2006). Shifts in eye position by 0.5–1.0° are not uncommon in perimetry (Henson, Evans, Chauhan, & Lane, 1996) and in damaged eyes such minor changes in stimulus location can

cause dramatic changes in measured sensitivity (Fellman, Lynn, Starita, & Swanson, 1989). For each of seven levels of ganglion cell loss between 0% and 99%, values of threshold (α) were obtained for stimuli at thirteen different locations within ±1° of putative stimulus centre. This array of 13 values for threshold represents potential effects of normal fixational eye movements, and on each stimulus presentation the value for sensitivity was drawn randomly from this array of 13 values. Values for threshold were computed using spatial filters with a peak spatial frequency of 1.0 cycle per degree (cpd) sampling a sparse mosaic of ganglion cells with large receptive fields, as these yielded values for standard deviation versus sensitivity which were consistent with the equation of Henson et al. (2000) for the slope parameter, and for which perimetric loss remained a linear function of ganglion cell loss.

FP and FN were fixed at either 0.0 or 0.2, to mimic the responses of a reliable subject versus responses of an unreliable subject.

2.1.3. Perimetric strategies—Three different perimetric strategies were simulated: the standard full-threshold (FT) strategy and two strategies, termed 'Linear strategies,' which utilised linear sensitivity steps (Table 1). The total number of reversals and the step size at each reversal defined each strategy. For all strategies, the sensitivity value was taken as reciprocal of the 'last seen' stimulus luminance, or as zero when the maximum stimulus was not seen.

2.1.3.1. Full-Threshold strategy: The Full-Threshold (FT) strategy of the Humphrey Field Analyzer was replicated. This strategy terminated after two reversals. The step size was 4 dB prior to the first reversal, then 2 dB until the second reversal (Allergan Humphrey, 1986).

2.1.3.2. Linear strategies: The Linear strategies were designated Ln1 and Ln2; these were both two-reversal strategies.

Fig. 1 illustrates sample staircases from these four perimetric strategies with black symbols for the FT staircase and the Ln2 staircase (which had overall efficiency most like that of the FT staircase). Results are shown for an observer with minimal intrinsic and extrinsic noise, using a starting luminance corresponding to 500 L^{-1} and input sensitivities of 1 and 2000 L^{-1}. For an input sensitivity of 2000 L^{-1}, (α = 1/2000 L) the FT staircase had the smallest number of trials, while for a sensitivity of 1 L^{-1} (α = 1 L) the FT staircase had the largest number of trials.

2.1.4. Data analysis—*Bias* was defined as the difference between the input sensitivity and the mean of the (thousand) output sensitivities obtained with simulated staircases. Bias was positive when mean staircase sensitivity was higher than input sensitivity and negative when mean (output) staircase sensitivity was lower than input sensitivity. *Test–retest variability* (*TRV*) was computed as the standard deviation of 1000 output sensitivities for a given condition.

The *number of trials* for each staircase to run to completion was recorded. For each input sensitivity value, *perimetric efficiency* was defined as the inverse of the mean number of trials for the 1000 staircase runs. Therefore efficiency was low if, on average, a large number of trials were needed for a staircase to terminate. Simulations were implemented using Matlab 5.2.1 (The Mathworks, Inc., Natick, MA).

2.2. Collection of clinical data

To verify the clinical validity of predictions from the simulations, pilot data were gathered by testing 41 subjects with the FT and Ln2 perimetric strategies on two separate occasions. Prior to these two sessions, all subjects underwent a 'learning visit' to familiarise themselves with the nature of the test and the stimulus. The Ln2 strategy was chosen because the simulations suggested this to be more efficient than Ln1. The order of tests (FT/Ln2) was randomised and the number of tests for each strategy was counterbalanced across subjects.

2.2.1. Subjects—Twenty-one patients with glaucoma and 20 age-similar control subjects were recruited for this study. All participants were aged between 50 and 80 years, had best-corrected visual acuity 20/40 or better, spherical refraction <7 D and cylinder ≤ 2D and did not have any other eye disease affecting the posterior segment. All the glaucoma patients were regular attendants at the Glaucoma Institute of SUNY State College of Optometry. They all had evidence of reproducible visual field defects on at least two previous Humphrey Visual Field tests (average Mean Deviation, MD −6.7 dB, range −0.31 to −22.7). If both eyes were eligible, the eye with the least field damage was tested. The normal subjects all had passed a comprehensive eye examination at the University Optometric Center of SUNY and were excluded if there was a positive history of glaucoma within a first-degree relative. The study was conducted under an approval by the Institutional Review Board at SUNY and followed the tenets of the Declaration of Helsinki. After the purpose and procedures were discussed with each subject, written informed consent was obtained prior to testing.

2.2.2. Perimetric conditions—The Goldmann size III stimulus is 0.43° in diameter and is a standard stimulus used in clinical perimetry (Anderson, 1987). This stimulus was presented on a mean background luminance of 10 cd/m^2. This background luminance allowed a large dynamic range of stimulus luminances. Stimuli were presented with a Gaussian temporal profile. The time constant of the Gaussian was 100 ms. Low temporal frequencies reduce the luminance required to reach the Weber region (Graham & Hood, 1992). We used a fixed starting stimulus of 1000 L^{-1} (30 dB), at each location, to reduce test-time in normal-seeing regions of the visual field.

2.2.3. Apparatus—Stimuli were displayed on a 21″ SONY Trinitron monitor driven by a VSG 2/5 system (Cambridge Research Systems, Rochester, Kent, UK). The resolution of this monitor was set to 800 × 600 pixels. The visible portion subtended 54° × 42° at a viewing distance of 40 cm. The frame rate was 150 Hz. The VSG system provided 14-bit resolution for each phosphor. The VSG OptiCAL photometer was used to measure photometric values for each phosphor, to calibrate display gamma functions and produce a linear lookup table. A luminance meter (Minolta LS-100, Konica Minolta, Mahwah, NJ, USA) was used to ensure that the mean background luminance of the monitor was 10 cd/m^2.

2.2.4. Test reliability criteria—Blank trials were presented at pseudo-random intervals, to provide an estimate of the false positive rate. Fixation was monitored with a CCTV video camera and using the Heijl–Krakau method. Tests were excluded and subsequently repeated when false-positive rate was greater than 20% or when fixation was deemed unstable (>30% fixation losses). Patients with deep defects and variable responses may not see the maximum available stimulus (15 dB on the VSG) all of the time (Bengtsson & Heijl, 2000) and so false negative rate was not used as an indicator of test reliability.

2.2.5. Test locations—Stimuli were presented at eight locations at eccentricities of 9.5°, 15.0°, and 21.2° (Fig. 2). These locations were chosen to reflect the distribution of the retinal nerve fibre pattern. Locations were symmetrical across the horizontal midline. Since glaucomatous loss tends to be asymmetrical, this arrangement allowed sampling of retinal locations with damage ranging from mild to severe within the same eye.

2.2.6. Analysis of pilot data and comparison with simulation data—For each point the *average sensitivity* of two tests for each of the eight locations and each subject was computed as the means of the sensitivities measured by each of the two strategies, FT and Ln2.

Agreement of measured sensitivity values obtained by FT and Ln2 was evaluated using a Bland–Altman plot (Bland & Altman, 1986) and was compared to corresponding (output) sensitivity values yielded by the Monte Carlo simulation. For the purposes of comparing the

pilot and simulation data, output sensitivity values (rather than input sensitivity) were used, since in the practical situation *actual* sensitivity at a given test point is unknown and only a measured value is obtained by running a perimetric staircase.

Test–retest difference was defined as the difference in sensitivity at each location between the two tests ($t_2 - t_1$) for each strategy. *Test–retest variability* was computed as the standard deviation of test–retest differences across all locations and all subjects for each of the two strategies. Tests where the maximum luminance stimulus was not seen on both occasions were excluded from variability analysis.

Standard deviation from the simulations (1000 output sensitivities for each condition) was multiplied by the square root of two for direct comparison with the measured test–retest variability obtained from the pilot data.

3. Results

3.1. Results of the Monte Carlo simulations

3.1.1. Results in dB units—Our goal was to evaluate the sensitivity-dependence of bias, variability, and efficiency of FT and alternative staircases, when sensitivities are expressed in linear units. To demonstrate that our choice of parameters was consistent with data in the literature which expressed sensitivities in dB units, the relationship between simulated variability and sensitivity in dB units was qualitatively compared (Fig. 3). A starting point of 500 L^{-1} (27 dB) was used for all these simulations. The solid grey line in Fig. 3 shows results of the simulation for intrinsic variability modelled by varying the slope of the psychometric function according to Henson's equation. Henson's equation is not defined for sensitivities below 10 L^{-1} (10 dB). The variability of DLS measurement in dB is typically highest around 10 dB (Artes, Iwase, Ohno, Kitazawa, & Chauhan, 2002), as measurement of variability for sensitivities below this value is limited by the maximum stimulus of the perimeter. Therefore, for these simulations in dB units and for input sensitivities below 10 L^{-1}, the psychometric function slope was fixed to 0.6 (this slope corresponded to an input sensitivity of 10 L^{-1}). Psychometric slopes below 0.6 gave rise to an artificial increase and higher estimates of variability for sensitivities in the range 10–15 dB. With this slope of 0.6, input sensitivities of <1 L^{-1} (0 dB) sometimes yielded output sensitivities ≥ 10 dB. For a reliable estimate of dB variability for output sensitivities ≥ 10 L^{-1} (10 dB), input sensitivities in the range of 0.016–10 L^{-1} (−18 dB–10 dB) at 1 dB intervals were used. In accordance with the plot of Artes et al. (2002), average variability was plotted for 5 dB sensitivity bins. The solid black line shows the variability resulting from heterogeneous ganglion cell damage interacting with minor changes in stimulus location. The two methods of simulating variability (slope of psychometric function and heterogeneous damage) exhibited similar qualitative trends: variability was less than 2 dB at high sensitivities (>30 dB) and became much higher in regions of low sensitivity. This is consistent with previous reports of high variability in damaged regions of the visual field (Artes et al., 2002;Chauhan, Tompkins, LeBlanc, & McCormick, 1993;Piltz & Starita, 1990). For comparison with published data, variability values from Artes et al. (2002) are also shown (dotted black line).

The variability predicted by the simulations using a variable psychometric function slope computed from Henson's equation was as much as 1 dB lower than the variability estimated by Artes et al., 2002. Further simulations (not shown) demonstrated that this difference could be explained by starting point effects. Our simulations used a fixed starting luminance corresponding to 500 L^{-1} (27 dB). Artes' data were obtained using the Full Threshold strategy of the Humphrey Field Analyzer, which uses a variable initial stimulus, suprathreshold to an initial estimate of sensitivity at each location.

3.1.2. Bias—Fig. 4 shows the effect of sensitivity on bias for each of the three strategies, for starting luminances of 500 L^{-1} (left panels) and 2000 L^{-1} (right panels). In the absence of extraneous noise (FP = FN = 0) bias varied systematically with sensitivity, in the range −500 to +650 L^{-1} across strategies and sensitivities. The greatest underestimates were at higher sensitivities, dropping as low as −500 L^{-1} for the FT strategy, but not lower than −300 L^{-1} for the Linear strategies. The greatest overestimates were as high as +650 L^{-1} for the Ln1 strategy at a high starting stimulus, but for the FT and Ln2 staircases there were no overestimates exceeding +400 L^{-1}. For a given strategy, bias varied by no more than 800 L^{-1} across sensitivities.

As shown in the lower panels of Fig. 4, much more extreme forms of bias were found in the presence of extraneous noise (FP = 0.2 or FN = 0.2). Generally, FP = 0.2 made bias more positive whilst FN = 0.2 made bias more negative. For a starting stimulus of 500 L^{-1}, with FP = 0.2 bias for FT increased from +100 L^{-1} at low sensitivities to +500 L^{-1} at high sensitivities, while bias for Ln1 and Ln2 showed much less variation with input sensitivity. For a starting luminance corresponding to 2000 L^{-1}, with FP = 0.2 all staircases yielded overestimates greater than 500 L^{-1}, and bias for linear staircases became highly dependent on input sensitivity, r = 0.99, slope = −0.54 for Ln1; r = 0.96; r = 0.98, slope = −0.37 for Ln2 (p values for slope < 0.0001 in both cases).

For FN = 0.20, bias for all three staircases became increasingly negative at high sensitivities, more rapidly for a starting stimulus of 2000 L^{-1}.

3.1.3. Test–retest variability—Qualitatively, results were similar for intrinsic noise and heterogeneous damage (Fig. 3), and in Fig. 5 are shown for intrinsic noise. In general, variability was less dependent on sensitivity for the linear staircases than for the logarithmic (FT) staircases. When extraneous noise was eliminated (upper panels), variability for FT and linear staircases changed by 400 L^{-1} across sensitivities, with variability for linear staircases being highest at 1 L^{-1}. For sensitivities above 1500 L^{-1} (32 dB), variability was lower for linear strategies than for FT. The Linear strategy with the smallest final step size (Ln1) had the least variability over most of the sensitivity range.

Extraneous noise (bottom two graphs, Fig. 5) increased variability for all staircases. This was most pronounced for FP = 0.2, where variability for FT often exceeded 2000 L^{-1} (shown in Fig. 5 with symbols pinned at 2000 L^{-1}). For the Linear staircases variability rarely exceeded 500 L^{-1} and never exceeded 800 L^{-1}. For Ln1 and Ln2, sensitivity-dependence with FP = 0.2 was greater for a starting point of 2000 L^{-1} (z = 7.8, p < 0.0001 for Ln1 and z = 16.5, p < 0.0001 for Ln2) while for FN = 0.2 sensitivity-dependence was greater for a starting point of 500 L^{-1} (z = 8.4, p < 0.0001 for Ln1 and z = 19.7, p < 0.0001 for Ln2).

3.1.4. Efficiency—Efficiency for FT and Linear strategies had opposite forms of sensitivity-dependence: FT tended to be more efficient at high sensitivities whilst the Linear strategies tended to be more efficient at low sensitivities (Fig. 6). For a starting stimulus of 500 L^{-1} (27 dB), linear staircases had the highest efficiency for locations with sensitivity 0–20 dB, while the FT staircase had the greatest efficiency for locations with sensitivity 27–33 dB. Change to a starting stimulus of 2000 L^{-1} (33 dB) caused a decrease in efficiency of Linear strategies at sensitivities 0–30 dB.

To compare potential efficiencies of the staircases in patients with varying amounts of visual field loss, we selected 24-2 SITA Standard field test results from one control subject and three patients with glaucoma (Fig. 7). For each field, dB sensitivities at all 54 locations were converted to L^{-1} and used as input sensitivities for the Monte Carlo simulation, with a starting stimulus of 500 L^{-1}, false positive and false negative rates of zero. Psychometric slopes were

computed from Henson's equation were computed. Each simulation generated all 54 staircases and used the total number of trials as an estimate of test duration. This was repeated for a total of 100 simulations per visual field, and mean test durations were computed. For the fields with minimal or no loss (top two panels), FT had the shortest test duration. For the fields with moderate or advanced damage (lower two panels), test duration increased for FT and decreased for the Linear staircases, so Ln1 and Ln2 had lower test durations than FT. Of the Linear strategies, Ln2 had the shortest test duration.

3.2. Results from clinical data

3.2.1. Measured sensitivities—Measured mean sensitivities obtained with FT and Ln2 were similar for the 20 control subjects at each of the three eccentricities (Table 2).

When data from both patient and control groups were combined, analysis of agreement (Fig. 8) found no significant mean difference between sensitivities obtained with the FT and Ln2 staircases (mean = 16 L^{-1}, 95% limits for agreement −435 to +467 L^{-1}). There was a tendency for the difference between FT and Ln2 sensitivities to increase with average sensitivity (r^2 = 0.01, slope = 0.06, p = 0.03), with Ln2 giving a higher value than FT at higher average sensitivities.

3.2.2. Test–retest variability—Test–retest variability was calculated as the standard deviation of test–retest difference at each location separately for each subject. For the clinical data (solid symbols), sensitivity was grouped into eight bins and the standard deviation (SD) of test–retest difference for each bin was computed (Fig. 9). The solid lines show the regression lines for these points. The dotted lines show the predicted standard deviations from the simulations. For both predicted and measured values of FT and Ln2, SD was well described by a linear fit in each case (r^2 = 0.79 FT and r^2 = 0.68 for Ln2). As expected, the sensitivity-dependence of the variability was greater for FT than for Ln2. For the clinical data, the slope of the regression line was steeper (z = 1.7, p = 0.04) for FT (slope 0.4, SE 0.085) than for Ln2 (slope 0.26, SE 0.072). In each case, the slopes of the regression lines for the predicted values and measured values were similar (z = 0.25, p = 0.4 for FT; z = 0.97, p = 0.17 for Ln2).

Pointwise test–retest differences were divided into three discrete sensitivity groups (1–500, 501–999 and ≥ 1000 L^{-1}), and comparisons of variability for FT and Ln2 was computed in L^{-1} (Fig. 10A) This revealed significantly higher variability for FT in the group with high sensitivities (F = 2.1, p = 0.002) but not for the groups with intermediate or low sensitivity (F < 1.3, p > 0.35).

Although the primary purpose of the study was to assess the effect of the type of staircase step (log versus linear) on variability, for comparison, we also analysed pointwise test–retest differences in dB units (Fig. 10B). As expected, this showed that for dB units, variability increased for FT and Ln2 as sensitivity decreased. Variability was still higher for FT compared to Ln2 at higher (≥ 1000 L^{-1}) sensitivities (F = 2.32, p = 0.001) but no different at intermediate sensitivities (501–999 L^{-1}), F = 1.34, p = 0.18. At low sensitivities, variability (in dB) for FT was much lower than for Ln2 (F = 3.15, p < 0.001).

3.2.3. Efficiency—The mean sensitivities for the control group for the different test locations are given in Table 2. At all locations tested the mean sensitivities were in the range 700– 1300 L^{-1}, for which both FT and Ln2 staircases are predicted to have similar efficiency (upper left of Fig. 6). Across all subjects, the mean number of trials required for a staircase to terminate is shown in Fig. 11 for each sensitivity value. The mean number of trials was indeed quite similar for sensitivities of 700–1300 L^{-1} (t < 1.5, p > 0.15 for all average measured sensitivities ≥ 700 and <1300 L^{-1}), and as predicted (dotted lines) the Linear staircase was more efficient

at low sensitivities. The total number of trials across all locations and all subjects was 4580 for FT and 4353 for Ln2, respectively.

4. Discussion

4.1. Summary of findings

Conventional perimetric algorithms return sensitivities in equal logarithmic steps, while structural and electrophysiological measures are typically returned in linear steps. The conversion of perimetric data to linear sensitivity could produce substantial statistical artefacts due to sensitivity-dependence for bias, variability and/or efficiency. We used Monte Carlo simulations to evaluate sensitivity- dependence of these three factors for conventional strategies, as well as for two strategies returning sensitivity in equal linear steps. The simulations identified a Linear staircase, Ln2, which was expected to have similar efficiency to the conventional logarithmic full-threshold (FT) strategy when averaged across locations with a range of degrees of loss (Fig. 7), and to have weaker sensitivity-dependence than conventional staircases. These predictions were tested by making repeated measures of perimetric sensitivity of patients with glaucoma and control subjects, using both logarithmic and Linear staircases.

As predicted, the Ln2 staircase reduced sensitivity-dependence for both efficiency and precision (Figs. 9 and 11). Sensitivity-dependence of bias cannot be directly determined, as the true sensitivity for a given location is unknown, but the data gathered with FT and Ln2 staircases do conform to the prediction that the difference between sensitivities measured with Ln2 and FT increases with sensitivity (Fig. 8).

4.2. Units of DLS measurement

In conventional perimetry sensitivity is measured with dB steps, for which a 3 dB decrease in DLS is equivalent to a doubling of luminance. The use of logarithmic steps has some advantages: the variation of normal thresholds across eccentricity (Heijl et al., 1987) is smaller with logarithmic units, and use of dB units compresses the normal range of sensitivities while expanding the range of abnormal sensitivities, thereby facilitating the identification of abnormal levels of sensitivity. However, DLS in dB units may not be linearly related to the number of functioning ganglion cells, particularly in the early stages of glaucomatous disease (Harwerth, Carter-Dawson, Shen, Smith, & Crawford, 1999; Swanson et al., 2004).

The complex relationship between DLS (in dB) and structural measures makes it difficult to accurately grade the severity of glaucoma, particularly in the early stages of disease. The impression of a ‘functional reserve’ may result from the logarithmic nature of the dB scale (Garway-Heath, Caprioli, Fitzke, & Hitchings, 2000; Garway-Heath et al., 2002). Linear units of DLS may provide a more precise indication of underlying ganglion cell number. Both theoretical (Swanson et al., 2004) and clinical evidence (Garway-Heath et al., 2002; Reus & Lemij, 2004; Schlottmann, De Cilla, Greenfield, Caprioli, & Garway- Heath, 2004) suggests that a continuous (linear) structure –function relationship may be obtained if DLS is computed in a linear metric of sensitivity.

Studies of electrophysiology have also provided support for the linear relationship between ganglion cell damage and a linear sensitivity. A linear relationship exists between DLS in linear units and Pattern ERG (PERG) amplitude (Garway-Heath et al., 2002). PERG amplitude is thought to reflect the number of functioning ganglion cells. Hood et al. (2002) reported a linear relationship between the log ratio (right versus left eye) of multifocal visual evoked potential amplitude and log ratio visual field losses and postulated that both these parameters may be directly related to the local loss of ganglion cells.

The relationship between structural and functional measures can be linearised by plotting both parameters in logarithmic units (Harwerth et al., 2004; Harwerth et al., 2005) but there is currently no agreed method of measuring optic disc and retinal nerve fibre layer parameters in logarithmic metrics. Several investigators have converted dB values to linear sensitivity to aid structure/function comparisons (Garway-Heath et al., 2002; Harwerth et al., 2005; Schlottmann et al., 2004). Inspection of data from these studies suggests that dB perimetric sensitivity, when scaled in linear units, exhibits large variability particularly at higher sensitivities. The aim of the present study was to evaluate the effects of measuring linear sensitivity with dB steps and test the hypothesis that measurement of linear sensitivity with linear steps would reduce the effect of sensitivity on variability. The results of our study, comparing log and linear steps, are specific for linear measures of DLS and the variability findings are intended to be interpreted in the context of the linear model of structure/function relationship. The interpretation of variability in dB units, outside the linear structure function model, may be misleading (Fig. 10), as apparently large variability in dB units at low sensitivity may actually be low variability in terms of structural units.

4.3. Logarithmic versus linear steps for measuring linear DLS

By Fechner's law, equal steps in the change of the apparent brightness of a stimulus is related to the logarithm of the luminance increment and therefore the use of logarithmic steps may provide a representation of the corresponding changes in sensory magnitude.

Our results suggest that the precision, bias and efficiency of linear DLS measurement is, to some extent, determined by the size of the stimulus steps in relation to the underlying sensitivity. The precision of a sensitivity estimate is largely determined by the staircase step size in relation to the slope of the psychometric function. In general, the psychometric function slope is shallow at low sensitivities and steep at high sensitivities. In linear units, the number of possible sensitivity outcomes obtained with a dB-increment staircase is relatively high at low sensitivities and low at high sensitivities. In linear units, this results in lower precision at higher sensitivities and higher precision at lower sensitivities. Precision improves at higher sensitivities when sensitivity is measured with a linear-increment staircase (Fig. 9). One might argue, therefore, that whilst a linear strategy may be suited to detecting loss at higher sensitivities (for the detection of glaucoma), a dB staircase may be more suitable for detecting progression of established disease. Therefore, the identification of which algorithm is 'best' depends on the question or task in question. We have selected a strategy that we believe is more appropriate for structure/function analyses, where the step size is selected in the context of a linear model of the structure/function relationship, for which there is growing evidence. The Ln2 algorithm results in greater precision at higher sensitivities and lesser precision at lower sensitivities, when compared with the dB algorithm. Assuming a linear structure/function relationship, a 2 dB loss from 0 to −2 dB represents a 37% loss in ganglion cells, yet a 2 dB loss from −10 dB to −12 dB is equivalent to only 3.7% loss (from 90% to 93.7% loss). In this scenario, when trying to quantify structural damage from a functional test, it makes little sense to have high precision at low sensitivity and low precision at high sensitivity. In alternative scenarios, such as the evaluation a subject's ability to manage in his/her visual environment, measuring small amounts of remaining visual function with greater precision may be more important. There is a reciprocal relationship between precision and efficiency, so that an increase in precision occurs at the cost of reduced efficiency.

Various algorithms are possible, depending on the purpose for field testing. These include a 'hybrid' log-linear scale, with linear increments for higher sensitivities, switching to log increments for lower sensitivities, an algorithm where the step size is a percentage of the previous luminance value (this would fall somewhere between the linear and logarithmic staircase), or an algorithm which matches the step size to the estimated threshold.

4.4. Validating the variability predictions of the Monte Carlo model

The usefulness of conventional perimetry for monitoring glaucoma has been limited by high test–retest variability in damaged regions of the visual field: more than seven visual fields may be required to accurately determine progression of disease (Johnson, 2001). The simulations replicated this finding (Fig. 3), in that the magnitude of test–retest variability in simulations of the standard FT strategy was similar to previous clinical studies.

Heijl et al. (1987) reported a standard deviation for intertest variability of around 2 dB at sensitivities between 30 and 35 dB (in normal eyes). Piltz & Starita's data (1990) shows SDs of about 1 – 2 dB at a mean sensitivity of 30 dB, and SDs ranging from 4 to 10 dB at a mean sensitivity of 15 dB. Artes et al. (2002) found a mean root-mean-square (RMS)-error of between 5 and 6 dB at a mean sensitivity of 15 dB, and a mean RMS error of less than 2 dB at sensitivities in the range 30–35 dB. These findings are consistent with values obtained from the simulations (SD near 6 dB for an input sensitivity of 15 dB, and near 1 dB for an input sensitivity of 30 dB, Fig. 3).

4.5. Importance of findings

The simulations and clinical data both found that the use of a conventional dB-increment staircase results in lower number of trials and higher linear-unit variability in areas of the visual field with higher sensitivities. The simulations demonstrated that, in the absence of extraneous noise, the FT staircase should have a tendency to underestimate sensitivity in regions of high sensitivity, and that this bias will vary with starting luminance for subjects with substantial extraneous noise. These forms of sensitivity-dependence for the FT staircase could produce artefacts when FT sensitivities are converted to linear units for comparisons with structural indices, and could potentially produce a curvilinear relation even when the underlying relation is linear. The increase in variability with sensitivity for FT staircases is likely to decrease the strength of correlations between structural and visual field measurements, especially if many locations tested have near-normal sensitivities. The use of Linear staircases to measure perimetric sensitivity could potentially reduce these sources of artefact in structure–function comparisons.

Greater accuracy and improved precision at normal sensitivities has important implications for the early detection of glaucoma, particularly for inexperienced perimetric subjects with a high false positive rate. The normal sensitivity for an average 50-year-old at a peripheral nasal location is 29 dB (800 L^{-1}) (Heijl et al., 1987). The simulations suggest that the SD of variability for a subject with a FP rate of 20% at this sensitivity is more than halved with the use of Ln2 rather than FT (Fig. 5). Hence, it is expected that the ability to detect early defects at such locations would be greatly improved with the use of linear steps. This has important implications for glaucoma screening and diagnosis.

This study has shown that, when DLS is measured in linear units, the relation between variability and sensitivity is inverted from that found with dB units: variability of sensitivities derived from the FT staircase (dB staircase algorithm) is greater in areas with higher sensitivity (Figs. 5 and 10). The use of linear staircases made variability less dependent on sensitivity, so that sensitivity estimates would have nearly equal variability at all levels of sensitivity. It was also found that FT tends to be less efficient in regions with reduced sensitivity, which means that patients with more profound defects will tend to require a larger number of trials, increasing the likelihood that prolonged test time will produce fatigue effects.

4.6. Comments on study methodology

For this study, fixed levels of extraneous noise were employed. For a given individual, the lapse (FN) or guess (FP) rate may vary with factors such as the length of the test, the location

being tested and the time of day. Some investigators have tried to overcome this problem by using computer simulations that are based on stimulus-response data from actual subjects (e.g., Chauhan & Johnson, 1994; Johnson et al., 1992). The present simulations provided a method of introducing extraneous noise that did not require human subjects. With this approach, a large number of noise conditions could be modelled in a short time.

For the simulations, the relationship between sensitivity and the slope of the psychometric function, as given by Henson et al. (2000) has been utilised. In general, slopes are steep in normal sensitivity regions and shallow in defective regions of the visual field. However, some patients with glaucoma can have shallow slopes in normal sensitivity regions (Chauhan et al., 1993). The simulations allowed the effects of slope and sensitivity to be considered separately. Although, for the purpose of the simulations reported here, a variable slope (dependent on the underlying sensitivity value) was employed, the simulation also allowed the use of fixed levels of slope at any sensitivity value. These simulations (not shown) found that intrinsic noise increased variability and that the sensitivity-dependence of variability increased as the slope (β) was reduced.

It should be noted that Henson's equation has not been validated for sensitivities below 10 dB (10 L^{-1}) (Henson et al., 2000), so simulated values of bias, variability and efficiency computed for an input sensitivity of 1 L^{-1} (Figs. 4–6) should be interpreted with caution.

5. Conclusion

Monte Carlo simulations and clinical pilot data support the hypothesis that the dependence of bias, variability and efficiency on sensitivity can be reduced with the use of linear DLS steps.

The use of perimetric strategies which utilise linear steps for perimetric sensitivity can reduce extremes of variability in normal regions of the visual field, giving rise to relatively uniform variability characteristics across the sensitivity range. It is likely that this would improve the precision of measurements in normal regions of the field, when DLS is recorded in linear units, potentially narrowing the normal range of linear sensitivity.

It remains to be ascertained clinically whether the use of algorithms which measure sensitivity with linear steps lead to better agreement between perimetric and structural measures of ganglion cell damage in the initial stages of disease and earlier identification of glaucoma.

Acknowledgements

The authors thank Alisa Surkis, PhD for writing the Monte Carlo Simulation programme and May Ellis, Mitchell W. Dul, OD, MS and Richard J. Madonna, OD for their assistance with patient recruitment.

References

Allergan Humphrey. The field analyzer primer. Allergan Humphrey; San Leandro, California: 1986.

Anderson, DR. Perimetry —with and without automation. 2. The C.V. Mosby Company; 1987. Instruments and strategies (programs).

Anderson AJ. Spatial resolution of the tendency-orientated perimetry algorithm. Investigative Ophthalmology and Visual Science 2003;44:1962–1968. [PubMed: 12714630]

Anderson AJ, Johnson CA. Anatomy of a supergroup: does a criterion of normal perimetric performance generate a supernormal population? Investigative Ophthalmology and Visual Science 2003;44:5043–5048. [PubMed: 14578433]

Artes PH, Iwase A, Ohno Y, Kitazawa Y, Chauhan BC. Properties of perimetric threshold estimates from full threshold, SITA standard and SITA fast strategies. Investigative Ophthalmology and Visual Science 2002;43:2654–2659. [PubMed: 12147599]

Bengtsson B, Heijl A. False-negative responses in glaucoma perimetry: indicators of patient performance or test reliability? Investigative Ophthalmology and Visual Science 2000;41:2201–2204. [PubMed: 10892863]

Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–310. [PubMed: 2868172]

Chauhan BC, Tompkins JD, LeBlanc RP, McCormick TA. Characteristics of frequency-of-seeing curves in normal subjects, patients with suspected glaucoma and patients with glaucoma. Investigative Ophthalmology and Visual Science 1993;34:3534–3540. [PubMed: 8258511]

Chauhan BC, Johnson C. Evaluating and optimizing test strategies in automated perimetry. Journal of Glaucoma 1994;3(Supplement):S75–S81.

Fellman, RL.; Lynn, JR.; Starita, RJ.; Swanson, WH. Clinical importance of spatial summation in glaucoma. In: Heijl, A., editor. Perimetry update 1988/1989. Berkeley: Kugler& Ghedini; 1989. p. 313-324.

Garway-Heath DF, Caprioli J, Fitzke FW, Hitchings RA. Scaling the hill of vision: the physiological relationship between light sensitivity and ganglion cell numbers. Investigative Ophthalmolmology and Visual Science 2000;41:1774–1782.

Garway-Heath DF, Holder GE, Fitzke FW, Hitchings RA. Relationship between electrophysiological, psychophysiological and anatomical measurements in glaucoma. Investigative Ophthalmology and Visual Science 2002;43:2213–2220. [PubMed: 12091419]

Glass E, Schaumberger M, Lachenmayr BJ. Simulations for FASTPAC and the standard 4–2 dB full-threshold strategy of the Humphrey field analyser. Investigative Ophthalmology and Visual Science 1995;36:1845–1847.

Graham N, Hood DC. Modelling the dynamics of light adaptation: the merging of two traditions. Vision Research 1992;321:1373–1393. [PubMed: 1455710]

Harwerth RS, Carter-Dawson L, Shen F, Smith EL III, Crawford ML. Ganglion cell losses underlying visual field defects from experimental glaucoma. Investigative Ophthalmology and Visual Science 1999;40:2242–2250. [PubMed: 10476789]

Harwerth RS, Carter-Dawson L, Smith EL III, Barnes G, Holt WF, Crawford ML. Neural losses correlated with visual losses in clinical perimetry. Investigative Ophthalmology and Visual Science 2004;45:3152–3160. [PubMed: 15326134]

Harwerth RS, Carter-Dawson L, Smith EL III, Crawford ML. Scaling the structure–function relationship for clinical perimetry. Acta Ophthalmologica Scandanavica 2005;83:448–455.

Heijl A, Lindgren G, Olsson J. Normal variability of static perimetric threshold values across the central visual field. Archives of Ophthalmology 1987;105:1544–1549. [PubMed: 3675288]

Henson DB, Evans J, Chauhan BC, Lane C. Influence of fixation accuracy on threshold variability in patients with open angle glaucoma. Investigative Ophthalmology and Visual Science 1996;37:444–450. [PubMed: 8603850]

Henson DB, Chaudry S, Artes PH, Faragher EB, Ansons A. Response variability in the visual field: comparison of optic neuritis, glaucoma, ocular hypertension, and normal eyes. Investigative Ophthalmology and Visual Science 2000;41:417–420. [PubMed: 10670471]

Hood DC, Greenstein VC, Odel JG, Zhang X, Ritch R, Liebmann JM, et al. Visual field defects and multifocal visual evoked potentials: evidence of a linear relationship. Archives of Ophthalmology 2002;120:1672–1681. [PubMed: 12470141]

Johnson CA, Chauhan BC, Shapiro LR. Properties of staircase procedures for estimating thresholds in automated perimetry. Investigative Ophthalmology and Visual Science 1992;33:2966–2974. [PubMed: 1526745]

Johnson CA. Psychophysical measurement of glaucomatous damage. Survey of Ophthalmology 2001;45 (Suppl 3l):S313–S318. [PubMed: 11377455]discussion S322–S314

Maloney LT. Confidence intervals for the parameters of psychometric functions. Perception & Psychophysics 1990;47:127–134. [PubMed: 2304811]

Pan F, Swanson WH, Dul MW. Evaluation of the two-stage neural model of glaucomatous defect: An approach to reduce test-retest variability. Optometry and Vision Science. 2006accepted for publication

Piltz JR, Starita RJ. Test–retest variability in glaucomatous visual fields [letter]. American Journal of Ophthalmology 1990;109:109–111. [PubMed: 2297025]

Reus NJ, Lemij HG. The relationship between standard automated perimetry and GDx VCC measurements. Investigative Ophthalmology and Visual Science 2004;45:840–845. [PubMed: 14985299]

Schlottmann PG, De Cilla S, Greenfield DS, Caprioli J, Garway-Heath DF. Relationship between visual field sensitivity and retinal nerve fibre layer thickness as measured by scanning laser polarimetry. Investigative Ophthalmology and Visual Science 2004;45:1823–1829. [PubMed: 15161846]

Spry PGD, Johnson CA, McKendrick AM, Turpin A. Measurement of visual field tests in glaucoma. British Journal of Ophthalmology 2003;87:107–112. [PubMed: 12488273]

Swanson WH, Felius J, Pan F. Perimetric defects and ganglion cell damage: interpreting linear relations using a two-stage neural model. Investigative Ophthalmology and Visual Science 2004;45:466–472. [PubMed: 14744886]

Turpin A, McKendrick AM, Johnson CA, Vingys A. Properties of perimetric threshold estimates from full threshold, ZEST, and SITA-like strategies, as determined by computer simulation. Investigative Ophthalmology and Visual Science 2003;44:4787–4795. [PubMed: 14578400]

Weinreb RN, Khaw PT. Glaucoma. Lancet 2004;363:1711–1720. [PubMed: 15158634]

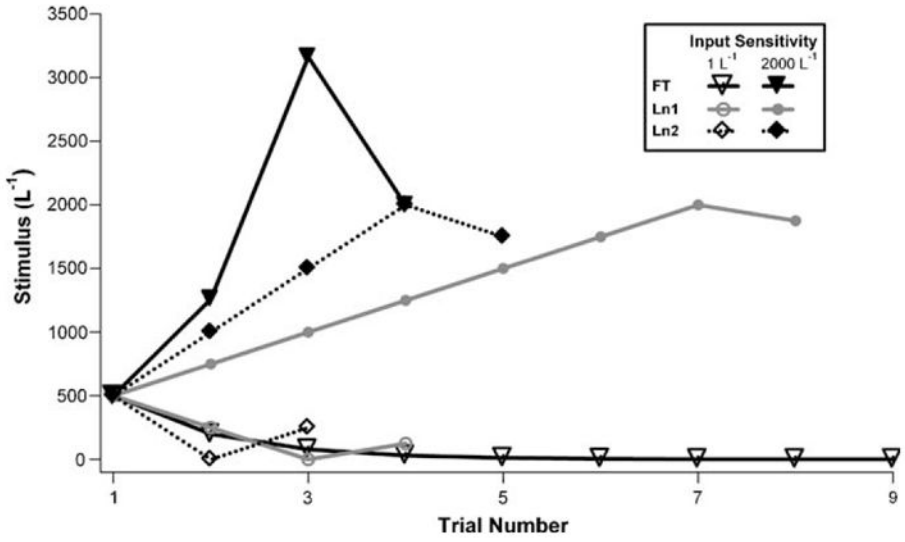

Fig. 1.

Sample staircase runs for FT, Ln1 and Ln2 with a starting stimulus of 500 L^{-1} and an input sensitivity of 1 and 2000 L^{-1}.

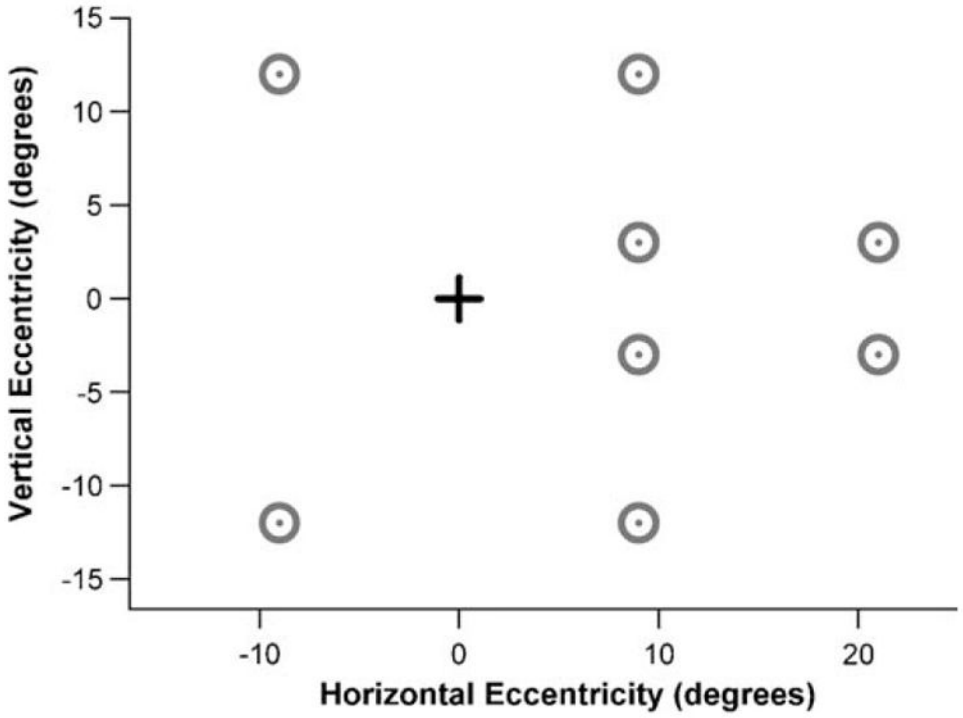

Fig. 2.
Test locations (left eye). Test locations (open grey symbols) are shown relative to the fixation point (solid black cross).

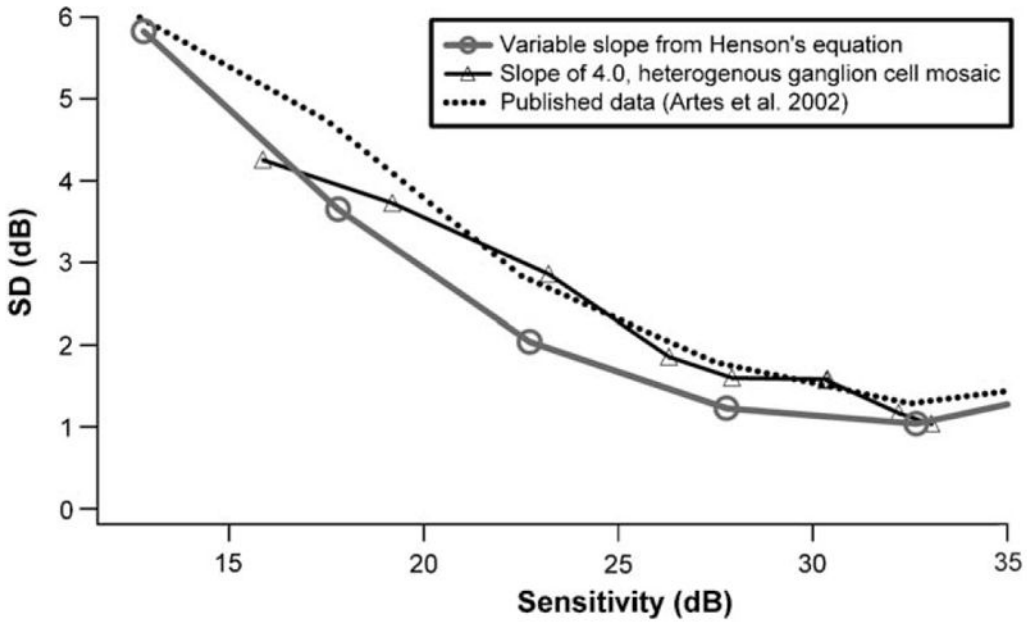

Fig. 3.

Variability of DLS (dB) versus sensitivity (dB) as predicted by the Monte Carlo simulation for two conditions: a psychometric slope which varies with sensitivity according to Henson's equation ('variable slope') and for a fixed psychometric slope of 4 with sensitivity values obtained from a model of sparse ganglion cell loss sampled by a 1.0 cpd spatial filter to represent effects of normal fixational eye movements.

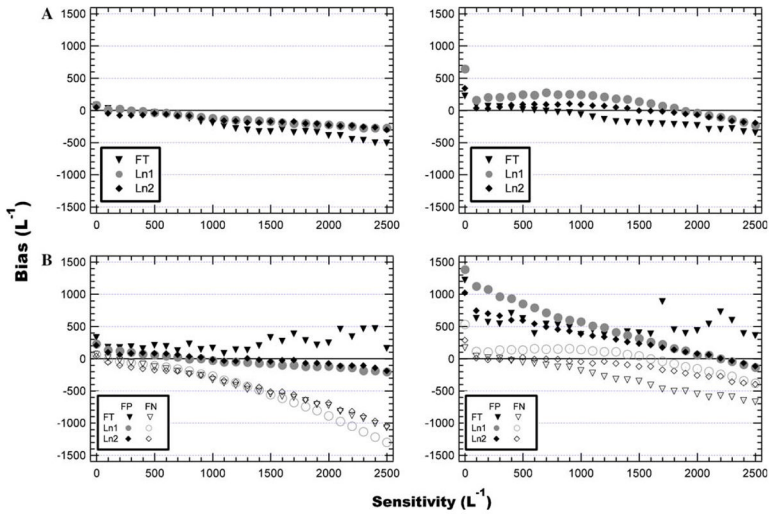

Fig. 4.

The effect of sensitivity on bias for FT and the two linear strategies Ln1 and Ln2. Simulations for two different starting stimuli are shown. The top graphs show results for minimal extraneous noise (FP = FN = 0). The bottom graphs illustrate the effect of introducing FP or FN rate of 20% (FP, false positive; FN, false negative; FT, full threshold strategy; Ln1 and Ln2, linear strategies; and L^{-1}, 1/Lambert). (A) Starting 500 L^{-1}. (B) Starting 2000 L^{-1}.

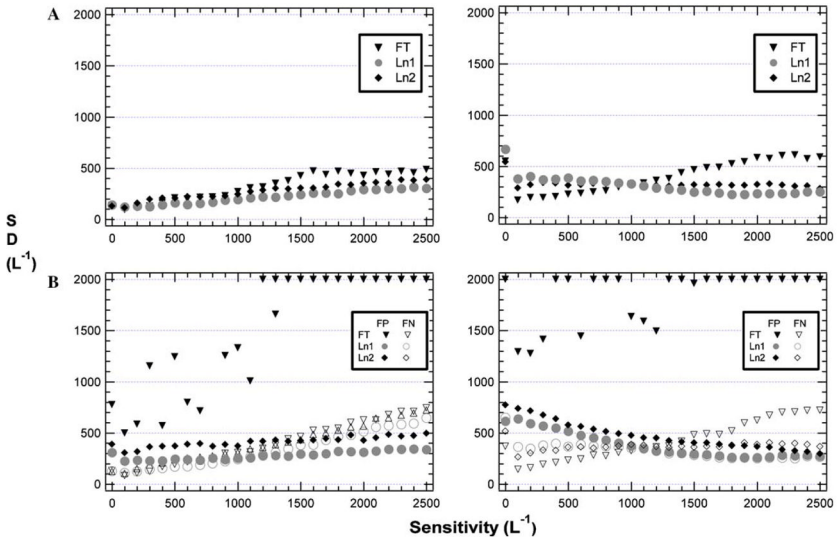

Fig. 5.

Test–retest variability (measured as the standard deviation of 1000 staircase runs) for minimal extraneous noise, FP = FN = 0 (top graphs) and FP/FN rates of 20% (bottom graphs). The graphs on the left show simulations for a starting stimulus of 500 L^{-1}. Variability for runs with a starting stimulus of 2000 L^{-1} are shown on the right. With 20% FP, SD values for FT exceeded 2000 L^{-1} at sensitivities above 1500 L^{-1} and are shown pinned at 2000 L^{-1}. (FP, false positive; FN, false negative; FT, full threshold strategy; Ln1 and Ln2, linear strategies; and L^{-1}, 1/Lambert). (A) Starting 500 L^{-1}. (B) Starting 2000 L^{-1}.

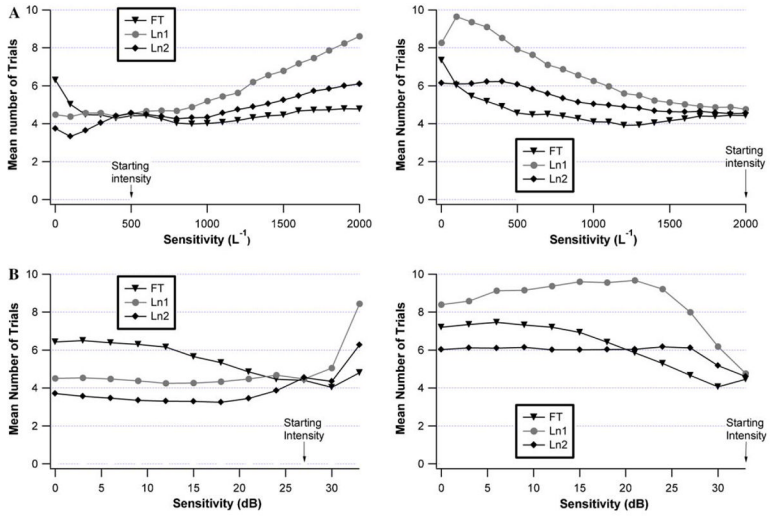

Fig. 6.

Efficiency predictions from the Monte Carlo simulation. Efficiencies for linear and dB algorithms are shown for two different starting points: 500 L^{-1} (left graphs) and 2000 L^{-1} (right graphs) and for sensitivity scaled in both in linear metrics (top graphs) and dB metrics (bottom graphs). (SD, standard deviation; FT, full threshold; Ln1 and Ln2, linear strategies, and L^{-1}, 1/Lambert). (A) Starting 500 L^{-1}.

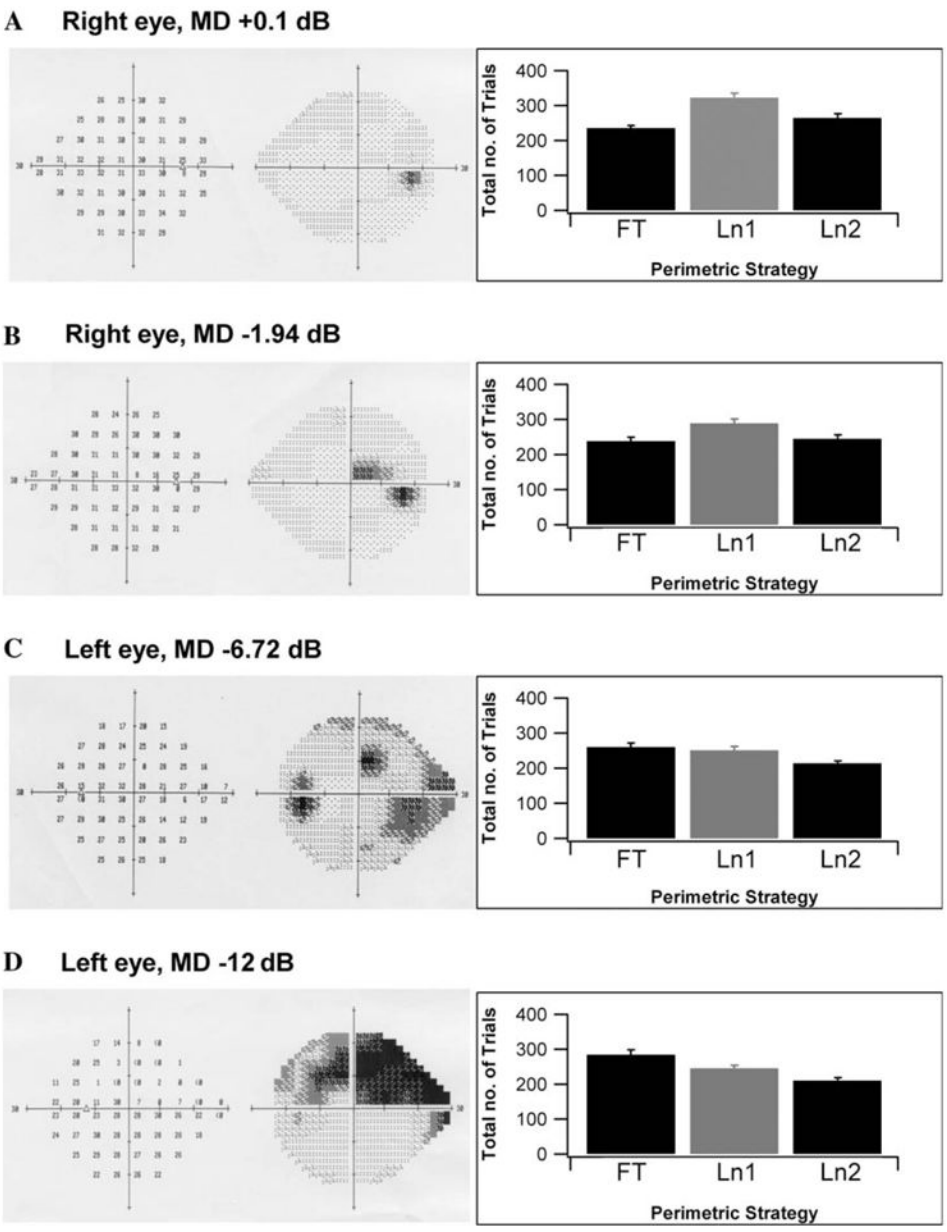

Fig. 7.

FT, Ln1, and Ln2 estimated test times for a full 24-2 visual field test. Sensitivities were obtained from 24-2 SITA Standard Visual Field tests for one normal subject (A) and three glaucoma patients with varying amounts of field loss (B–D). The sensitivities obtained in dB units were converted to linear (L^{-1}) units and used as input sensitivities for our Monte Carlo simulation to estimate the total test time (total number of trials) for each of the perimetric strategies for a 54-location test. The Ln2 and FT strategies (both black solid bars) were comparable in efficiency.

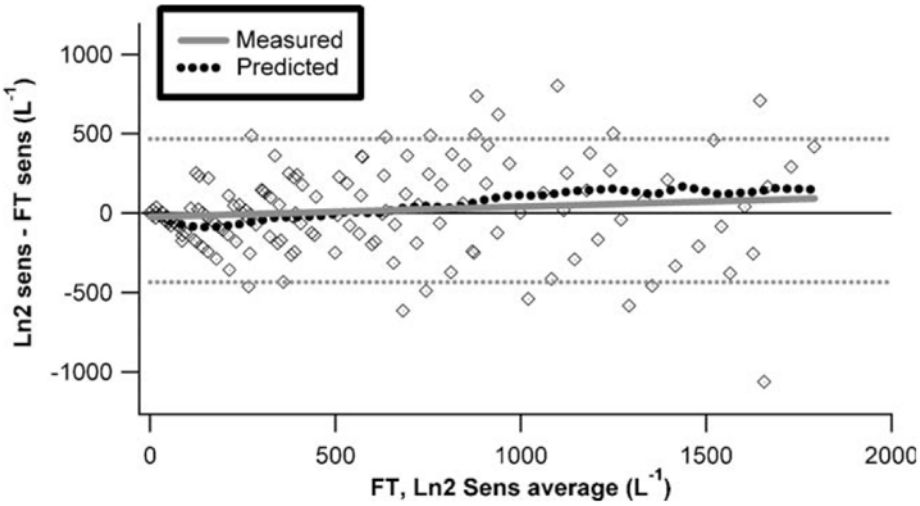

Fig. 8.

Agreement of sensitivity values as measured by FT and Ln2 for the clinical data (open symbols), 95% limits for agreement 467–435 L^{-1}, mean difference = 16 L^{-1}. The solid grey line shows the regression line for these points (r^2 = 0.01, slope = 0.06, p = 0.03). The dotted line shows the prediction obtained by the Monte Carlo simulation. (FT, full threshold; Ln2, linear 2 strategy; and L^{-1}, 1/Lambert).

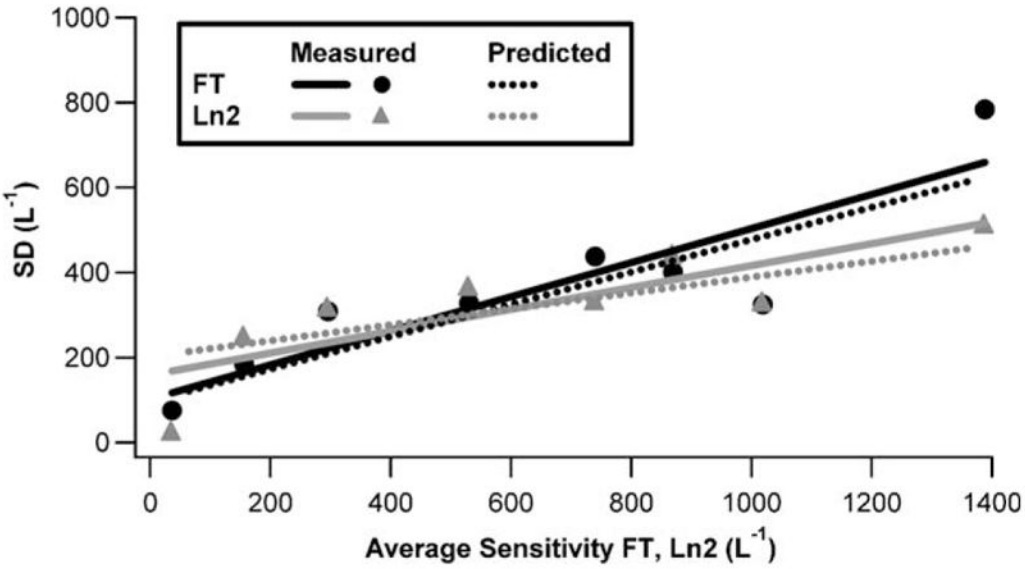

Fig. 9.

Variability (measured as the standard deviation of test–retest difference across locations with similar sensitivity) for FT and Ln2 (solid symbols). The regression lines for these points are shown (solid lines). The dotted lines show predictions from the Monte Carlo simulation. (SD, standard deviation; FT, full threshold; Ln2, linear two strategy; and L^{-1}, 1/Lambert).

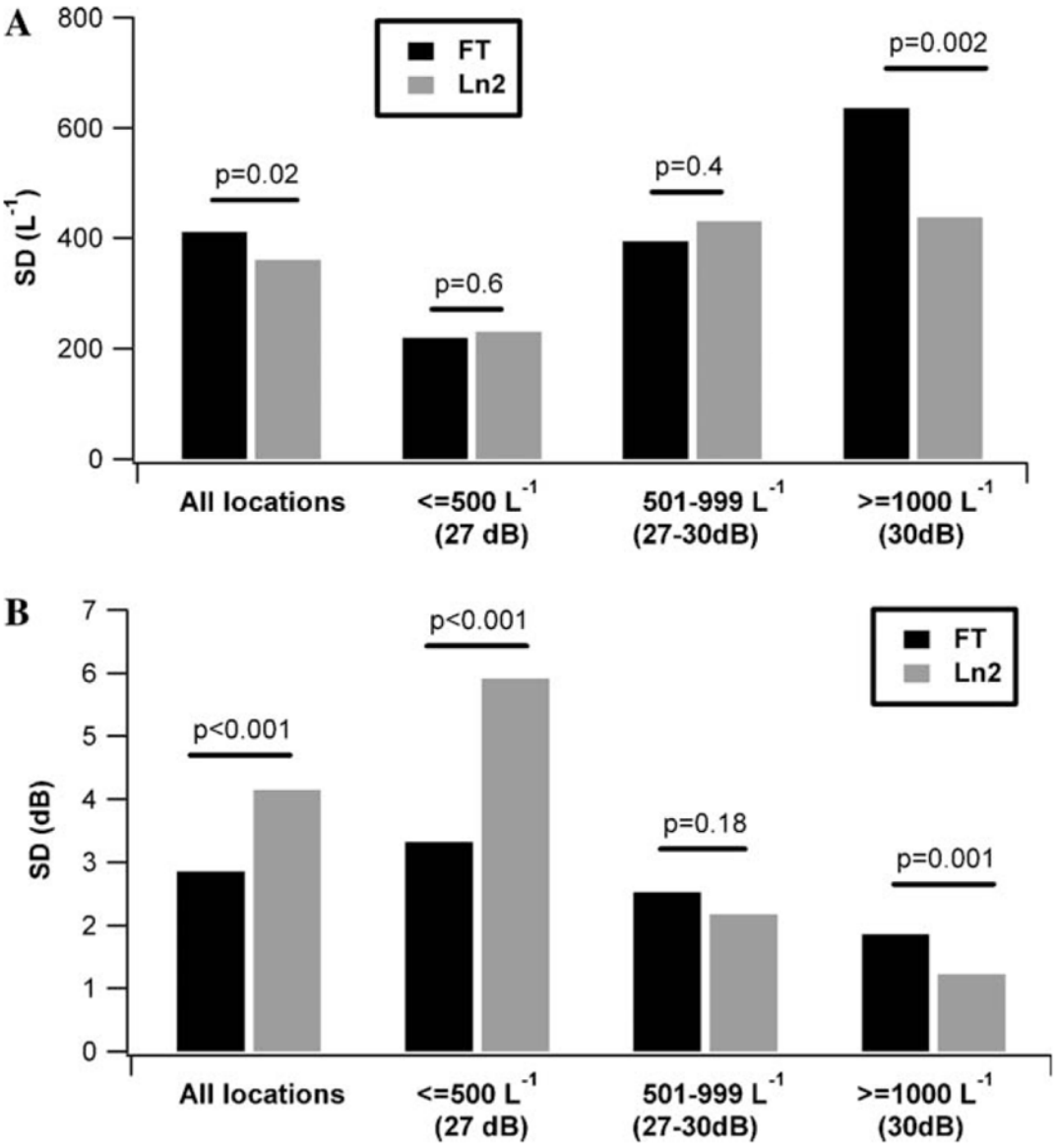

Fig. 10.

Variability by sensitivity group for FT and Ln2 for the pilot data. The first set of bars show the standard deviation across all locations. The next three sets of bars show the standard deviation for locations with sensitivity ≤500 L^{-1} (27 dB) 501–999 L^{-1} (27–30 dB) and ≥ 1000 L^{-1} (30 dB) respectively. *P* values (*F* test) for comparison of variances between FT and Ln2 are shown for each group. The top graph, (A) shows variability (vertical axis) computed in linear units. Variability computed in dB is shown in the bottom graph, (B) for comparison. (SD, standard deviation; FT, full threshold; Ln2, linear three strategy; and L^{-1}, 1/Lambert).

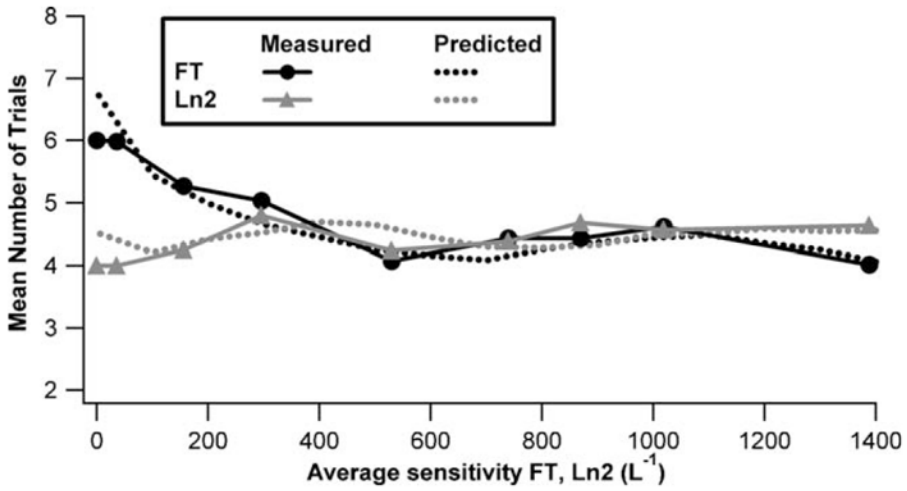

Fig. 11.

Mean number of trials taken for FT and Ln2 staircases to terminate across the sensitivity range (mean of two tests). The dotted lines show predictions from the Monte Carlo simulation. (FT, full threshold, Ln2, linear two strategy, and L^{-1}, 1/Lambert).

Table 1

Perimetric strategies simulated

Strategy	Smallest increment	Algorithm	Number of reversals
FT	2 dB	4–2 dB	2
Ln1	125 L^{-1}	250–125 L^{-1}	2
Ln2	250 L^{-1}	500–250 L^{-1}	2

Table 2

Mean sensitivity at each eccentricity for normal subjects

Eccentricity (degrees)	**9.5**	**15.0**	**21.2**
FT Sensitivity (L^{-1}) (M ± SD)	1140 ± 403	797 ± 355	707 ± 228
Ln2 Sensitivity (L^{-1}) (M ± SD)	1250 ± 387	814 ± 304	735 ± 265