September 1, 2006

Evaluating Equiluminant Chromatic Stimuli As Stimuli for Assessing Glaucomatous Damage

By

Emily Ly-Schroeder, O.D.

M.S. RESEARCH PAPER

In partial satisfaction of the requirements for the degree of

Master of Science
In
Vision Science

State University of New York
State College of Optometry

September, 2006

Approved by M.S. Research Paper Committee:

_____
Dr. William H. Swanson, Ph.D., FAAO
Principal Investigator and Advisor

_____
Dr. Mitchell W. Dul, O.D., M.S., FAAO

_____
Dr. Harry J. Wyatt, Ph.D., FAAO

Abstract

Purpose: To examine the potential clinical utility of equiluminant chromatic stimuli for assessing glaucomatous damage. Pan et al. (2006) found that equiluminant red-green chromatic stimuli could have good ability to detect defects as well as low test-retest variability, but clinical utility was limited due to the small dynamic ranges for their stimuli. The current study increased the dynamic range by using larger stimuli and including tritan stimuli.

Methods: Luminance, red-green (R-G), and tritan stimuli were created by modulating a large square (3 degrees per side) from an equal energy white (20 cd/m^2) along three cardinal directions in color space. Contrast sensitivity was measured at four locations with an eccentricity of 12 deg, along the 45^o, 135^o, 225^o and 315^o meridia. Twenty-five patients with glaucoma and twenty-six control subjects free of eye disease were tested monocularly at two separate sessions within a two-week time period. Sensitivities were reported in decibel (dB) units, where 1 dB = -1 (log contrast threshold) x 10.

Results: The dynamic ranges were 11 and 13 dB for the tritan and red-green stimuli. Test-retest variability was dependent on depth of defect for the two chromatic stimuli ($r > 0.2$, $p < 0.25$) but not for the achromatic stimuli ($r=0.01$, $p =0.46$). Matched t-tests found that, on average, defect depths were similar for the red-green and luminance stimuli ($t=0.5$, $p = 0.30$), and were slightly deeper for the luminance stimulus than for the tritan stimulus ($t=4.2$, $p < 0.0001$). The relationship between defect depths for the luminance and tritan stimuli was dependent on mean defect ($r = -0.42$, $p < 0.0001$).

Discussion: The effort to increase the dynamic range for the chromatic stimuli by increasing the size of the stimulus and using tritan modulation were successful. However, this came at the cost of increased test-retest variability and decreased ability to detect glaucomatous defects.

Conclusion: Equiluminant chromatic stimuli in CRT-based tests may not be clinically useful as perimetric stimuli, since increased dynamic range comes at the expense of increased test-retest variability and decreased ability to detect visual loss. These findings further support the works of Hart (1988) and Sample et al. (2006).

Glaucoma is one of the world's leading causes of irreversible blindness, with an estimated 3 million people affected in the United States and approximately 70 million people affected world wide (Thomas and Melton, 2004; Quigley, 2005). Glaucoma is a heterogeneous group of chronic optic neuropathies which cause progressive visual field loss that is routinely detected and followed by automated perimetry (Delgado et al., 2002; Tan et al., 2002). Conventional automated static perimetry, which uses an achromatic stimulus of size III (0.43 deg) on a white background, shows high test-retest variability in areas with glaucomatous defects (Heijl et al., 1989). This high variability makes it difficult for clinicians to determine if a change in depth of a scotoma constitutes progression or just fluctuation. If variability could be reduced, then clinicians could more readily determine whether a change in scotoma depth constitutes progression that requires an adjustment in a patient's treatment.

Variability can be reduced in moderately damaged and normal areas of the visual field by increasing the stimulus size from Goldman III (0.43 deg) to V (1.72 deg) (Wall et al., 1997). However, the use of a larger stimulus can result in decreased measured depth of defect (Wall et al., 1997). Decreased defect depth can be averted to some extent by using chromatic stimuli instead of achromatic stimuli, as demonstrated by Pearson et al. (2001), who found that on average contrast sensitivities were lower for large chromatic stimuli than for large achromatic stimuli. However, Pearson et al. used a large and complex xenon-arc-based optical system and a lengthy forced-choice protocol which made it unsuitable for clinical use. Pan et al. (2006) developed a more clinically suitable computer-based testing station in which the chromatic stimuli were created by using equiluminant chromatic modulation. Their circular chromatic stimuli, which were approximately the same size as the chromatic spatial summation area, were designed to tap the red-green chromatic pathway. They obtained results consistent with those of Pearson et al.: low test-retest variability with good ability to detect defect. However, Pan et al.'s chromatic stimuli had a limited dynamic range; the maximum stimulus was only 0.4 to 0.7 log unit greater than mean normal contrast threshold. Chromatic stimuli have been used in perimetry for some time in SWAP (Short-Wavelength Automated Perimetry), which uses chromatic increments on a uniform background rather than equiluminant chromatic modulation. However, SWAP has been found to have even higher test-retest variability than conventional achromatic perimetry (Wild et al., 1998), which may be due to the high luminance adapting field required to isolate chromatic sensitivity with an increment threshold paradigm (Felius & Swanson, 2003).

In this study, we used Pan et al.'s approach of equiluminant modulation, and increased the dynamic range by increasing the stimulus size and by using tritan as well as red-green

chromatic stimuli. We evaluated whether with these stimuli we could replicate the results of Pan et al. and Pearson et al., namely low variability with good ability to detect glaucomatous defects. Successful replication would indicate the potential for large equiluminant chromatic stimuli to be useful as perimetric stimuli. However, failure to replicate the results of Pan et al. and Pearson et al. would instead support Hart's conclusion that color contrast stimuli may have limited clinical utility (1988).

MATERIALS AND METHODS

Apparatus

A customized program was written in C, using the Psychophysics Toolbox, and run on a CRT-based iMac computer (Apple Computers, Cupertino, CA). The screen was filled with a uniform equal-energy white background having a mean luminance of 20 cd/m^2, chosen to reduce the effects of lenticular density on tritan sensitivity (Pearson et al., 2006; Felius & Swanson, 2003). Chromaticity of the stimuli was modulated along three cardinal directions in color space: equiluminant red-green (L-M) and tritan axes (S-(L+M)), and the luminance axis (L+M), where L, S, and M represent the long-wavelength-sensitive, short-wavelength-sensitive, and middle-wavelength-sensitive cone types, respectively. The stimuli were perceived by observers as varying degrees of red, violet, and white, respectively. Details on calibrations have been described previously (Pearson et al., 2006). Sensitivities are expressed in terms of the logarithm of reciprocal contrast at threshold, in decibel (dB) units, where 1 dB = 0.1 log unit (in other words, sensitivity was expressed as -1 (log contrast threshold) x 10).

Stimulus Design

Due to the limited number of color and luminance intensities available with 8-bit resolution of the digital-to-analog converters (DACs) controlling the phosphors on the CRT, a dithering technique was applied to obtain finer control over stimulus chromaticity. The stimulus was divided into squares 10 pixels across, and the 100 pixels within each square were assigned to two adjacent DAC values (e.g. 60 pixels at DAC value 217 and 40 pixels at DAC value 218 resulted in a mean phosphor luminance corresponding to a DAC value of 217.4). This dithering produced an effective resolution of 14 bits per phosphor (Pearson et al., 2006). As a result of the dithering technique, it was more practical to create a square-shaped stimulus rather than the

conventional circular-shaped stimulus due to the serrated appearance of the circle produced by the pixel arrays.

Achromatic (L+M) and chromatic ((L-M) and (S-(L+M))) sensitivity were assessed at 12 degrees eccentricity using a 3 degree target size (see Figure 1 for stimulus layout). In a separate study we found that this size of chromatic stimuli yielded a much larger dynamic range than obtained with smaller stimuli (Pearson et al., 2006). A 0.6 degree cross was placed at the center of the screen as a fixation target, where a 0.6 degree square stimulus was randomly presented to monitor the subject's attention. The small size of the central stimulus encouraged subjects to fixate carefully, since any eye movement away from the central target would decrease their sensitivity and reaction time to the central stimulus (Schiefer et al., 2001). The Heijl-Krakau blind spot monitoring technique was also used to assess fixation by periodically presenting stimuli in the physiological blind spot.

Stimuli were presented for 500 ms with a one second interval between stimulus presentations. Schiefer et al. showed that the reaction time within the central 30 degree radius visual field in healthy young individuals was approximately 400 ms (Schiefer et al., 2001); thus, a one second interval between stimuli presentations should be sufficient for older individuals with visual defects to respond. Swanson (1987) demonstrated that detection of chromatic pulses showed temporal integration up to 300 ms at 9 Td (trolands), so 500 ms stimulus duration should be sufficient for stimulating the chromatic pathway. To determine whether the chromatic pathways had been indeed stimulated, control experiment sensitivities were compared for durations of 50 vs. 500 msec. If detection were mediated by the luminance channel, then sensitivity should show little increase with stimulus duration, while if a chromatic channel mediated detection then sensitivity should increase with stimulus duration (Swanson, 1987).

Thresholds were measured using a one-up one-down staircase procedure with contrast starting at the maximum contrast available: 3.8 dB, -7.0 dB, and 4.8 dB for the tritan, red-green, and luminance stimulus, respectively. The stimulus contrast was decreased in steps of 3 dB prior to the second reversal and varied in 1.5 dB steps thereafter up to a total of eight reversals. Contrast sensitivity was computed from the mean of the last 6 reversals. In addition to stimuli presented in the staircases, "free trials" (stimuli at 6 dB greater contrast than the ongoing mean of reversals) were interspersed to improve the subject's attention to the stimuli and to better estimate false negative rates. A maximum likelihood estimation (MLE) technique was used to estimate reliability of the responses (Swanson & Birch 1992). Data were considered unreliable and were rejected when false negative rate exceeded 14%, false positive rate exceeded 20%, rate

of fixation losses exceeded 33 percent, and/or the mean of reversals was more than 2 dB different from the MLE sensitivity estimate. Some participants had up to three measurements for the same stimulus test if more than one location was unreliable on the first test. The test with the least number of unreliable peripheral points from each session was used for data analysis.

Subjects

A total of fifty-one subjects were recruited for the study: twelve younger (<30 years old) control subjects recruited from students and staff at SUNY State College of Optometry; fourteen older (>50 years old) control subjects recruited from College staff and the University Optometric Center Primary Care Clinic; and twenty-five patients with primary open angle glaucoma recruited from the SUNY Glaucoma Institute. Table 1 lists the gender and ages of the groups. The inclusion and exclusion criteria for the control subjects are listed in Table 2. The inclusion and exclusion criteria for the patients with glaucoma were the same as for the controls with the following exceptions: allowed to have best-corrected visual acuity as low as 20/40 in the tested eye, allowed to have positive family history of glaucoma, required to have diagnosis of primary open angle glaucoma, and allowed to have IOP greater than 20mmHg.

Procedure

This study was conducted in accordance with the Declaration of Helsinki and the protocols were approved by the Institutional Review Board (IRB) of SUNY State College of Optometry. Prior to testing, an explanation of the tests was given and informed consent was obtained from each participant. All subjects were tested for visual acuity using the "ETDRS" chart (Logarithmic VA chart by Precision Vision, Illinois); monocular color vision with the Ishihara and SPPII plates under Standard Illuminant C; and contrast sensitivity with the Pelli-Robson Contrast Sensitivity chart (by HS International, UK). Since the patients with glaucoma were seen in our Primary and Ocular Disease Clinics, information on the status of their eyes was obtained from their charts and reviewed by the authors (ELS and MWD).

One eye was tested for each participant. To avoid selection bias, the right eye was always selected unless it did not meet the inclusion criteria (i.e. BCVA 20/20 for controls, 20/40 for patients). The testing distance (from the monitor to the subject's outer canthus) was maintained at 33 cm by using a headrest and by adjusting the CRT monitor, which was mounted on a retractable arm of an examining chair. A 33 cm string above the monitor was used as a measuring tape.

Participants over the age of 45 with no significant distance refractive error were given a pair of +3.00D corrective lenses and those with significant refractive errors had a pair of +3.00D clip-on lenses placed in front of their distance prescription to compensate for their inability to accommodate. A black plastic adhesive was placed over the corrective lens of the non-participating eye. A participant's blind-spot was determined at the beginning of each practice run, which took less than a minute. The practice run also allowed participants to familiarize themselves with the testing conditions. Participants were instructed to click on a computer mouse whenever they saw a stimulus in the central or peripheral locations, while maintaining fixation on a central cross. Total test time for one measurement was approximately five minutes. Each subject was tested twice within a two-week period.

To assess potential effects of blur (Hersh 1992, Anderson et al., 2001), a control experiment was conducted with seven young controls. This experiment simulated six diopters of blur by placing a +9.00 diopter lens in front of the tested eye. Contrast sensitivity measurements were obtained for the tritan, red-green, and luminance stimuli and the results were compared to the standard (no blur) data.

Statistical Design

Depth of defect was computed for each peripheral location (each patient contributed about four locations) by subtracting (in dB units) the patient's sensitivity at that location from the mean sensitivity of the older control group. Test-retest variability was computed for each patient at each of the peripheral locations, by subtracting (in dB units) the contrast sensitivity at the first visit from the contrast sensitivity at the second visit.

The two main questions of interest were:

1. Could large chromatic stimuli yield test-retest variability as low as that for large achromatic stimuli in areas of glaucomatous defect?
2. Could large chromatic stimuli be more sensitive than large achromatic stimuli in detecting ganglion cell loss? In other words, do large chromatic stimuli reveal deeper defects than large achromatic stimuli?

For the first question, four statistical tests were used. First, F-tests were used to compare the standard deviations (SD) of the test-retest differences in contrast sensitivity (CS) for chromatic and achromatic stimuli. The SD of the test-retest differences gave an approximation

of the amount of intrinsic variability, or the SD across repeated tests. Second, to determine if there is a learning effect besides intrinsic variability contributing to test-retest (t-rt) variability, mean t-rt variability was computed for each of the three stimuli across all patients by locations (each patient contributed about four points). Third, linear regression was used to assess whether the overall average of the difference in CS (test1 minus test2) varied with sensitivity, in which case is the SD of test-retest would include be affected by the learning-effect as well as intrinsic variability. Fourth, the absolute value of test-retest difference was correlated with the depth of defect, using a one-tailed test to determine whether variability increased with defect depth (Bland-Altman, 1991). The raw value for test-retest difference was correlated with depth of defect to assess whether learning effect varied with depth of defect.

For the second question, two statistical tests were used. For each patient and eccentric location, depth of defect was compared across stimulus types. First, a matched t-test was used to compare the average depths of defect for the different stimuli. Second, the method of Bland and Altman (1986) was used to compare depths of defect on a pointwise basis. Here, for each patient and location the difference in depth of defect for two stimuli was plotted versus the mean of the two defect depths. Linear regression was used to determine whether the difference in depth of defect varied systematically with the mean depth of defect.

Two secondary questions were:

1. Were the chromatic pathways successfully isolated?
2. Could this psychophysical test tolerate up to six diopters of blur?

For the first question, a matched t-test was used to compare the difference in contrast sensitivities of the long- and short-duration tests for the chromatic versus achromatic stimuli in the younger control group. For the second question, the effect of blur was evaluated by performing a matched t-test of the difference in contrast sensitivity between the standard tests of no blur with six diopters of blur for the chromatic versus achromatic stimuli in the younger control group. A p-value of less than 0.025 was considered significant for all the above calculations.

RESULTS

Norms and control experiments

Mean sensitivities for the young and old control groups (Table 3) showed age effects: the decline with age was approximately 1 dB, 4 dB, and 2 dB for the luminance, tritan and red-green, stimuli, respectively. The dynamic ranges for the older controls were approximately 11, 13, and 18 dB for the tritan, red-green, and luminance stimuli, respectively. For both patient and control groups the mean test-retest differences (Table 3) revealed a small learning-effect for all stimuli, on the order of 1 dB (-0.3 to -1.3). The standard deviations for the older control group yielded 95% confidence lower limits for normal that were 1.5, 2.8, and 2.5 dB below mean normal for the luminance, tritan, and red-green, stimuli, respectively.

Isolation of the chromatic pathways was confirmed: increase in stimulus duration caused a 6 dB increase in sensitivity for the chromatic stimuli, which was significantly larger than the 1 dB increase for the achromatic stimulus ($t > 12$, $p < 0.0001$). Six diopters of optical defocus resulted in approximately 2 dB loss in sensitivity for all three stimuli, a small but significant decline ($t > 3.9$, $p < 0.01$).

Test-Retest Variability of Patients

The chromatic stimuli did not yield as low test-retest variability as the achromatic stimulus. The standard deviation of mean contrast sensitivity was on average 0.4 dB greater for the chromatic than the achromatic stimuli (Table 3), a small but significant increase in intrinsic variability ($F= 1.67$, $p < 0.025$). Figure 2 shows the absolute value of test-retest difference versus depth of defect; variability was dependent on defect depth for the chromatic stimuli (Tritan, $r = 0.22$, $p<0.025$); R-G, $r = 0.32$, $p<0.001$) but not for the achromatic stimuli ($r= 0.01$, $p=0.46$). As depth of defect increased so did variability for the chromatic stimuli. Table 5 summarizes the slope, intercept, r, and r^2 values for the regressions in Figure 2. The slopes for chromatic stimuli were significantly steeper than for the luminance stimulus (Tritan, $z = 2.12$, $p = 0.017$; R-G, $z = 3.37$, $p < 0.001$) and the intercept for the luminance stimulus was no greater than for the chromatic stimuli. These findings demonstrate that the chromatic stimuli were not as successful as the achromatic stimulus in producing low test-retest variability in glaucomatous defects.

Figure 3 shows a scatter-plot of contrast sensitivities from the first and second visits. In general, there was a small learning effect for all three stimuli as demonstrated by the low mean test-retest variability for patients shown on Table 3 (-0.3 to -0.7). However, when comparing the

difference in contrast sensitivity (from week 1 to week 2) with the average contrast sensitivity as illustrated in Figure 4, linear regression found no dependence of test-retest difference on mean sensitivity for the achromatic stimulus ($r = 0.005$, $p = 0.50$), but a significant dependence for the chromatic stimuli (tritan, $r = -0.25$, $p = 0.017$; R-G, $r = 0.35$, $p < 0.001$). For the tritan stimulus patients with low sensitivity tended to do worse on the retest (evidenced by the negative slope), whereas for the red-green stimulus patients with low sensitivity tended to do better on the retest (evidenced by the positive slope).

B. Ability to detect glaucomatous defects

Figures 5 and 6 show a Bland-Altman plot comparing depth of defect for chromatic and achromatic stimuli. The average defect depth was 1.5 dB deeper for the luminance than the tritan stimuli, a small but significant difference ($t(87, 87) = -4.20$, $p<0.0001$). The average defect depth was 0.6 dB deeper for the red-green than luminance stimuli, which was too small to be significant ($t(93, 93) = -0.52$, $p = 0.30$). The difference in defects for the two stimuli was dependent on depth of defect for the tritan versus luminance stimuli ($r = 0.42$, $p < 0.0001$), but not for the luminance versus red-green stimuli ($r = 0.09$, $p = 0.2$). Overall, defects for the red-green stimulus were always within 4 dB of defects for the achromatic stimulus, but for the tritan stimulus more than 10% of defects were more than 4 dB different that the defects for luminance stimulus.

DISCUSSION

The purpose of this study was to evaluate the clinical usefulness of large equiluminant chromatic stimuli for monitoring progression of glaucomatous defects. It is known that use of large luminance increments allow reduction in test-retest variability compared to the small luminance increments used in conventional perimetry, but this comes at the expense of weakened ability to detect defects (Wall et al., 1997). Pearson et al. (2001) found that large chromatic increments provided test-retest variability as low as for large luminance increments while yielding defects intermediate between those for large and small luminance. Pearson et al. hypothesized that this result was due to larger spatial summation areas for the chromatic pathways than the achromatic pathway. For example, if stimulus size equaled the chromatic spatial summation area then loss of ganglion cells in a given retinal region would reduce sensitivity of all the chromatic cortical processes mediating detection of the stimulus, resulting in

an overall decrease in sensitivity. By comparison, achromatic cortical processes would sample only a portion of the stimulated ganglion cells, and if an achromatic process sampled a subset of ganglion cells which remained relatively intact then this could support near-normal sensitivity for the large achromatic stimuli.

Pan et al. (2006) showed that equiluminant chromatic stimuli could yield low variability and defects equally as deep as those for the small luminance increments of conventional perimetry. Pan et al. limited the size of their chromatic stimuli to their estimates of chromatic critical diameters, giving a restricted dynamic range which makes these stimuli unsuitable for clinical use. Our intent was to increase dynamic range by using large equiluminant stimuli, and to see whether such stimuli were superior to large luminance increments in terms of test-retest variability and depth of defect.

We found that dynamic range was indeed increased substantially by increasing the size of the equiluminant stimuli, but that the large chromatic stimuli were not superior to the large luminance stimuli in either test-retest variability or depth of defect. In fact, test-retest variability was slightly higher for the chromatic stimuli, and increased with depth of defect. This was the opposite of the finding by Pearson et al. (2001) with chromatic increments, and we found a possible explanation by evaluating learning effects. Unlike Pan et al.'s and Pearson et al.'s study, where the mean of (test1-test2) was very close to zero, ours was on average -0.5 dB (across all subjects and stimuli). The negative sign indicates that on average participants performed better on the second test. Although the intrinsic variability and learning effect were computed from the same set of data (intrinsic variability was obtained from standard deviation of contrast sensitivity while learning effect was obtained from variance of test-retest), according to Table 3, learning effect contributed only one fifth of the overall maximum test-retest variability (maximum test-retest variability contributed by both learning effect and intrinsic variability would be 2.5 dB). Thus, most of test-retest variability was contributed by intrinsic variability (or standard deviation across repeated tests).

Although on average learning effect was small, we wanted to see who tended to have a learning effect. For the red-green equiluminant stimuli, patients with low sensitivity showed higher sensitivity on retest, consistent with a learning effect. For the tritan equiluminant stimuli, patients with low sensitivity tended to do worse on retest, the opposite direction of a learning effect. Many participants complained of a yellowish after-image following presentations of the tritan stimulus. If they waited to respond until they saw the afterimage, their response could have been too late to be recorded as "seen". Pearson et al. used two-alternative forced-choice

procedures, which should be less affected by learning effects and by distractions caused by afterimages.

The red-green equiluminant stimuli yielded an insignificant 0.6 dB increase in depth of defect as compared with the luminance stimuli, while Pearson et al. (2001) found a more substantial 3.6 dB increase. This may be due to our choice of stimulus size and eccentricity. We used a similar mean luminance (20 cd/m^2 vs. 21 cd/m^2) and temporal presentation (500 msec) as used by Pearson et al., and our stimulus had a slightly larger area than theirs (9 deg^2 vs. 7 deg^2). However, our measurements were exclusively at an eccentricity of 12 deg, while Pearson et al. used eccentricities of 12 to 20 deg. Subsequent to the time we gathered our data, Pan et al. (2006) measured Ricco's area and obtained values of 1.0 deg^2 at 10°, 2.5 deg^2 at 15° and 4.5 deg^2 at 21°. Our stimulus would then be in the range 4 to 9 times larger than Ricco's area at our eccentricity, while Pearson's stimuli would have been 2 to 3 times larger for most of their eccentricities.

Pearson et al. (2001) found that the defect depth for their red-green stimulus averaged 5.4 dB less than for the standard perimetric stimulus, while Pan et al. (2006) found that the defect depth for their chromatic stimuli averaged within 0.4 dB of defect depths for the standard perimetric stimulus. Pan et al. (2006) presented an analysis of potential effects of stimulus size for circular stimuli, and found that defects should be similar for different stimuli as long as they are smaller than Ricco's area, and that as the stimulus area exceeds Ricco's area defects will become less deep. The 3°-diameter luminance increments that Pearson et al. used would have been at least 9 times Ricco's area for luminance increments, compared to 2-3 times larger than Ricco's area for chromatic increments, consistent with an average of 3.4 dB deeper defects for their chromatic than luminance stimuli. Since our chromatic stimuli were probably 4-9 times larger than Ricco's area, our stimuli would be expected to give defects more similar to those for the large achromatic stimuli (since both were considerably larger than Ricco's area). If this analysis is correct, then equiluminant red-green stimuli cannot be made suitable for clinical use: either they are limited to the size of Ricco's area, in which case dynamic range is too small for clinical use, or else they are too large to yield advantages over large luminance stimuli in terms of variability and defect depth. Large luminance stimuli can already be produced by standard perimeters, and red-green stimuli are not appropriate for patients with hereditary red-green color deficiencies, so if there is no advantage to red-green stimuli over large luminance stimuli then there would be little motivation to use them clinically.

The tritan stimuli yielded defects that on average were 1.5 dB less deep than for the luminance stimulus, while for the tritan stimuli of Pearson et al. (2001) the defects averaged 3.1 dB deeper than for the luminance stimuli. In fact, some of our patients had tritan sensitivities that were 3-5 dB greater than mean normal while none of them had luminance or red-green sensitivities that were 2 dB above mean normal. A possible explanation for this odd result is our recent finding that the tritan stimuli were more strongly affected by lens yellowing than we had expected (Pearson et al., 2006). Pearson et al. (2001) found that 21 cd/m^2 for their white background was sufficient to provide Weber behavior, so that tritan thresholds were minimally affected by reductions in retinal illuminance, while Pearson et al. (2006) found that our white background was not sufficient to produce Weber behavior and that half of the effect of age on our stimuli was due to aging of the crystalline lens. The patients with high tritan sensitivities in our study may have had lenses that were less dense than the mean for our older control group. As with the red-green stimuli, if there is no advantage for the tritan stimuli over luminance stimuli, then given the availability of large luminance stimuli and the problems of tritan stimuli, there is little motivation for clinical use of tritan stimuli. In fact, recent work has cast doubt on the supposed clinical advantages of short-wavelength automated perimetry (Sample, 2006).

Despite our findings that large chromatic stimuli do not provide advantages over large luminance stimuli, our results do confirm certain advantages of using large stimuli. Test-retest variability of patients was on the average less than 2 dB for all three stimuli, and regression lines were consistent with an average variability less than 3 dB at a -10 dB defect depth. This is less than half the variability for conventional perimetry at a defect depth of -10 dB (Heijl et al., 1989). Sensitivities for these stimuli were relatively immune to blur, with 6 diopters of defocus causing only a 2 dB decrease in sensitivity, compared to an 8 dB loss in sensitivity for conventional stimuli with 4 diopters of blur (Anderson et al., 2001). The primary problem with large stimuli is that on average they yield defects that are considerably smaller than found with conventional perimetry. Our analysis suggests that this problem is due to the stimuli being larger than the critical area of the cortical mechanisms mediating their detection.

Given our assessment of the weaknesses of chromatic stimuli, the low test-retest variability for large stimuli, and the problems when stimuli are potentially large relative to the cortical mechanisms mediating detection, an alternative luminance stimulus could be low-spatial-frequency sinusoids. Detection for patches of sinusoidal stimuli should be mediated by cortical mechanisms tuned to similar spatial frequencies, whose receptive fields should be similar in size to the stimuli. Pan et al. (2006) and Sun et al. (2006) found that large 1.0 c/deg

Gabor patches yielded low variability while having similar average defect depth at conventional perimetric stimuli.

In conclusion, based on the results of this study and an assessment of the literature, chromatic stimuli in CRT-based tests may not be useful for detecting or following the progression of glaucomatous defect. Sinusoidal luminance stimuli such as Gabor patterns should be considered and evaluated further.

ACKNOWLEDGEMENTS

This study was supported by T-35 and NIH grant EY007716 (WH Swanson).

REFERENCES

Anderson RS, McDowell DR, Ennis FA. Effect of localized defocus on detection threshold for different sized targets in the fovea and periphery. Acta Ophthalmol Scand. 2001;79:60-3.

Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1:307-310.

Delgado MF, Nguyen NTA., Cox TA, Singh K, Lee DA, Dueker DK, Fechtner RD, Juzych MS, Lin SC, Netland PA, Pastor SA, Schuman JS, Samples JR. Automated perimetry: A report by the American academy of ophthalmology. The American Academy of Ophthalmology. 2002;109:2362-2374.

Felius J, Swanson WH. Effects of cone adaptation on variability in S-cone increment thresholds. Invest Ophthalmol Vis Sci. 2003;44(9):4140-6.

Hart WM. Blue/yellow contrast perimetry compared to conventional kinetic perimetry in patients with established glaucomatous visual field defects. Proceedings of the VIIIth International Perimetric Society Meeting. 1988/89:23-30.

Heijl A, Lindgren A, Lindgren G. Test-retest variability in glaucomatous visual fields. Am Journal of Ophthalmology. 1989;108:130-135.

Pan F, Swanson WH. Evaluation of the two-stage neural model of glaucomatous defect: An approach to reduce test-retest variability. Optometry and Vision Science. 2006;83:499-511.

Pearson PM, Schmidt LA, Ly-Schroeder E, Swanson WH. Ganglion cell loss and age-related visual loss: A cortical pooling analysis. Optometry and Vision Science. 2006;83:444-454.

Pearson P, Swanson WH, Fellman RL. Chromatic and achromatic defects in patients with progressing glaucoma. Vision Research. 2001;41:1215-1227.

Quigley HA. Glaucoma: macrocosm to microcosm the Friedenwald lecture. Invest Ophthalmol Vis Sci. 2005;46(8):2662-70.

Sample PA, Medeiros FA, Racette L, Pascual JP, Boden C, Zangwill LM, Bowd C, Weinreb RN. Identifying glaucomatous vision loss with visual-function-specific perimetry in the diagnostic innovations in glaucoma study. Invest Ophthalmol Vis Sci. 2006;47:3381-3389.

Schiefer U, Strasburger H, Becker ST, Vonthein R, Schiller J, Dietrich TJ, Hart W. Reaction time in automated kinetic perimetry: Effects of stimulus luminance, eccentricity, and movement direction. Vision Research. 2001;41:2157-2164.

Sun H, Dul MW, Swanson WH. Linearity can account for the similarity among conventional, frequency-doubling, and gabor-based perimetric tests in the glaucomatous macula. Optom Vis Sci. 2006;83:455-65.

Swanson WH. Temporal modulation sensitivity and pulse-detection thresholds for chromatic and luminance perturbations. Optical Society of America. 1987;4:1992-2005.

Swanson WH, Birch EE. Extracting thresholds from noisy psychophysical data. Perception and Psychophysics. 1992;51:409-422.

Tan JC, Franks WA, Hitchings RA. Interpreting glaucoma progression by white-on-white perimetry. Graefe's Arch Clin Exp Ophthalmol. 2002;240:585-592.

Thomas R, Melton R. Contemporary glaucoma therapy: A world transformed. Continuing education. Optometric study center: July 2004.

Wall M, Kutzko KE, Chauhan BC. Variability in patients with glaucomatous visual field damage is reduced using size V stimuli. Inv Ophthalmology and Visual Science. 1997;38:426-435.

Wild JM, Cubbidge RP, Pacey IE, Robinson R. Statistical aspects of the normal visual field in short-wavelength automated perimetry. Inv Ophthalmology and Visual Science. 1998;39:54-63.

Table 1
Characteristics of Participants in the study

	Younger Controls	Older Controls	Patients
Age	25 ± 2.5	61 ± 4.3	59 ± 6.3
Age Range	20-28	54-69	50-69
Gender	3M, 9F	7M, 7F	8M, 17F

Table 2. **Inclusion and Exclusion Criteria for Control Subjects**

BVA ≥ 20/20 (in the tested eye)	Spherical ametropia ≤ ± 5.00 D
Cylindrical ametropia ≤ 3.00 D	Normal color vision
Clear ocular media	Open angles (Van Herrick)
(-) FHx glaucoma and other ocular disease	IOP ≤ 20 (applanation tonometry)
No systemic diseases that are known to affect vision	
Not taking meds known to affect vision	
Optic nerve head assessed by indirect ophthalmoscope	

Table 3.
Mean and standard deviation (SD) of the contrast sensitivity and test-retest (t-rt) variability of the controls and patients for each of the luminance, tritan, and red-green contrast stimuli.

	Younger Controls	**Older Controls**	**Patients**
Luminance	Mean CS 14.2 dB SD 0.9 Mean t-rt -0.4 SD 1.0	Mean CS 12.9 dB SD 1.2 Mean t-rt -0.4 SD 0.9	Mean CS 11.6 dB SD 2.4 Mean t-rt -0.7 SD 1.3
Tritan	Mean CS 10.4 dB SD 1.5 Mean t-rt -0.7 SD 1.4	Mean CS 6.4 dB SD 2.7 Mean t-rt -1.3 SD 1.7	Mean CS 6.6 dB SD 3.0 Mean t-rt -0.3 SD 1.7
Red-Green	Mean CS 21.7 dB SD 1.6 Mean t-rt -0.2 SD 1.3	Mean CS 19.6 dB SD 1.8 Mean t-rt -0.3 SD 1.5	Mean CS 18.4 dB SD 2.7 Mean t-rt -0.7 SD 1.7

Table 4. The effect of optical defocus on each stimulus type in the younger control group. All values expressed in dB.

	Red-Green	Tritan	Luminance
Mean	2.3	2.3	1.9
SD	1.6	2.0	1.0
N	28	28	28
p-value	<0.001	<0.001	<0.001

Table 5. Slope, intercept, r^2 and r values for Figure 2.

Stimulus type	Slope	Intercept	r^2	r
Luminance	+0.005	+1.12	0.0001	0.01
Red-Green	-0.15	+1.14	0.10	0.32
Violet	-0.08	+1.40	0.05	0.22

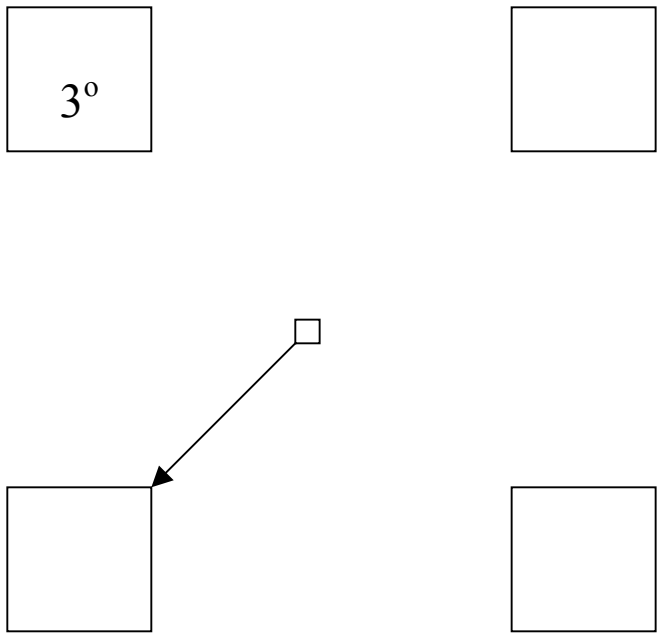


Figure 1. Lay-out of stimulus presentations. The central fixation target subtends an angle of 0.6 degree and the peripheral targets subtend an angle of 3 degrees. They are located 12 degrees away from the central fixation target at 45^{o}, 135^{o}, 225^{o}, and 315^{o} meridia.

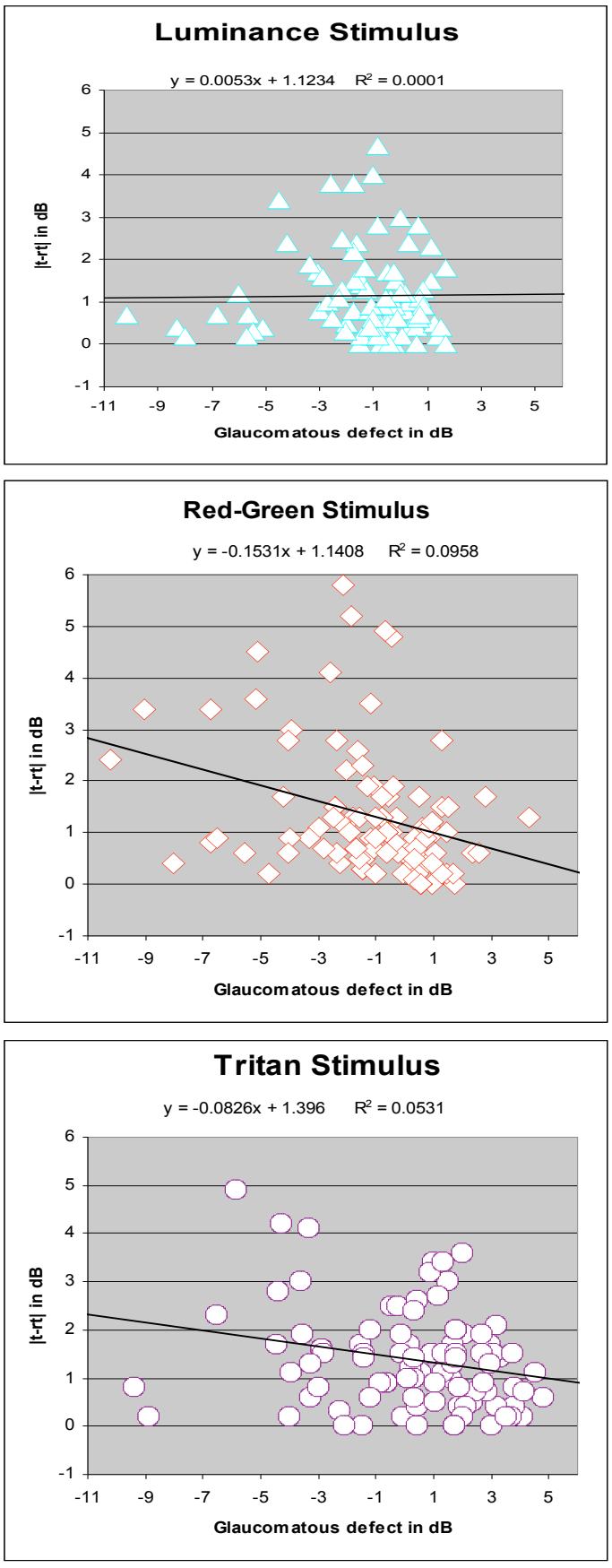
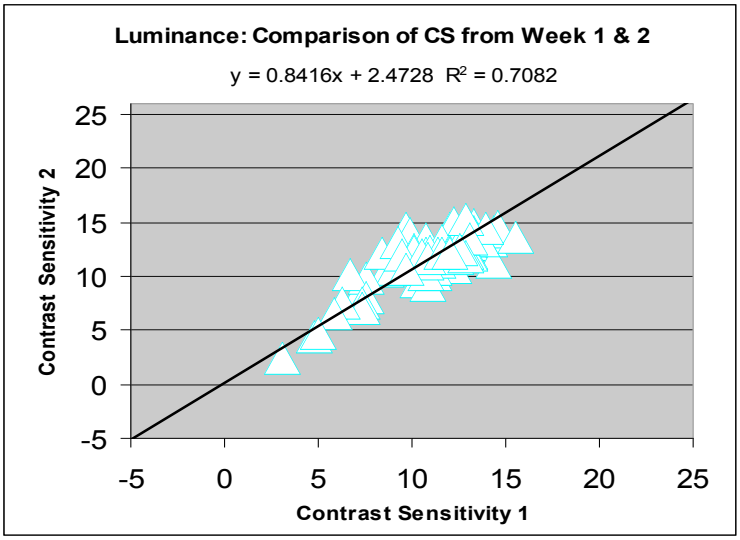
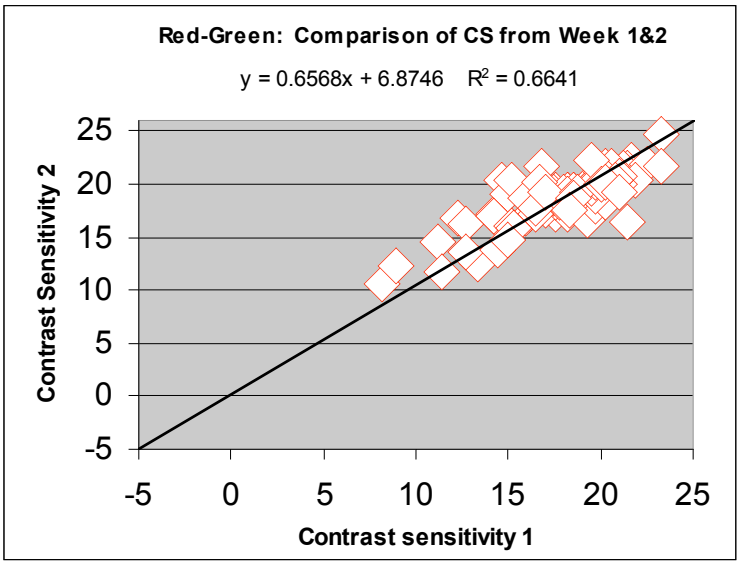


Figure 2. The absolute value of test-retest variability as a function of glaucomatous visual field defect for the luminance (top panel), red-green (middle panel), and tritan (bottom panel) stimuli.





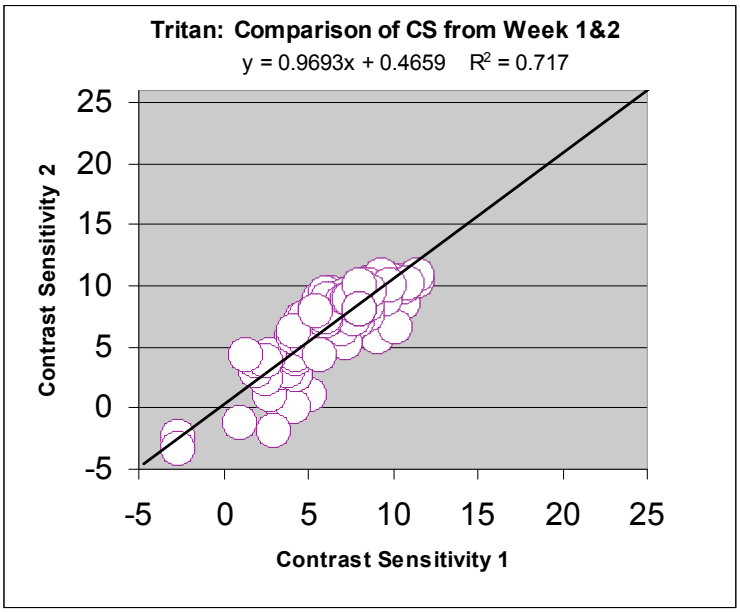
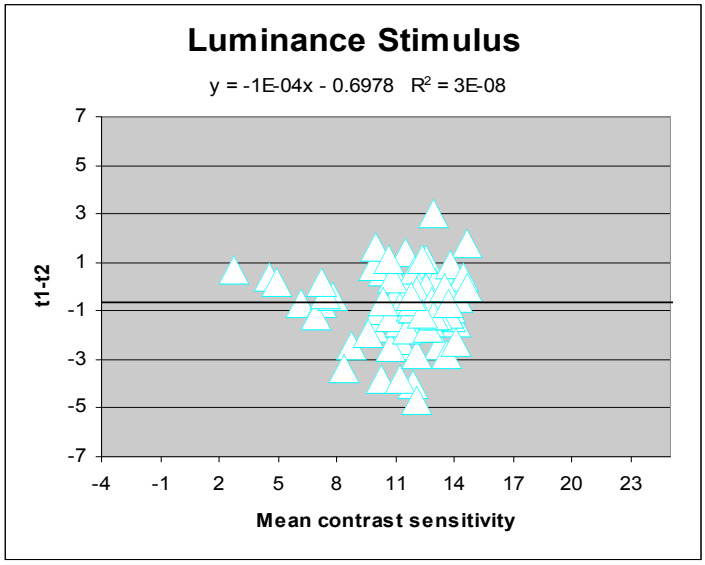
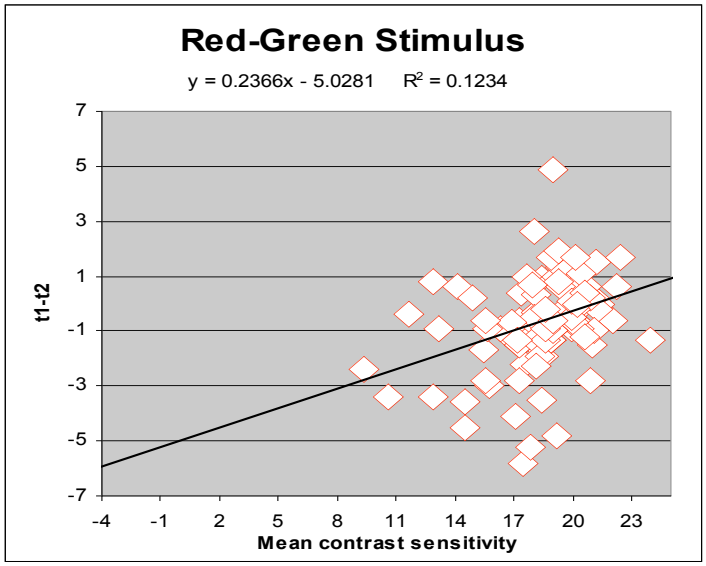


Figure 3. A scatter-plot comparing contrast sensitivities of week 1 and 2 for the luminance (top panel), red-green (middle panel), and tritan (bottom panel) stimuli. All values presented in decibels.





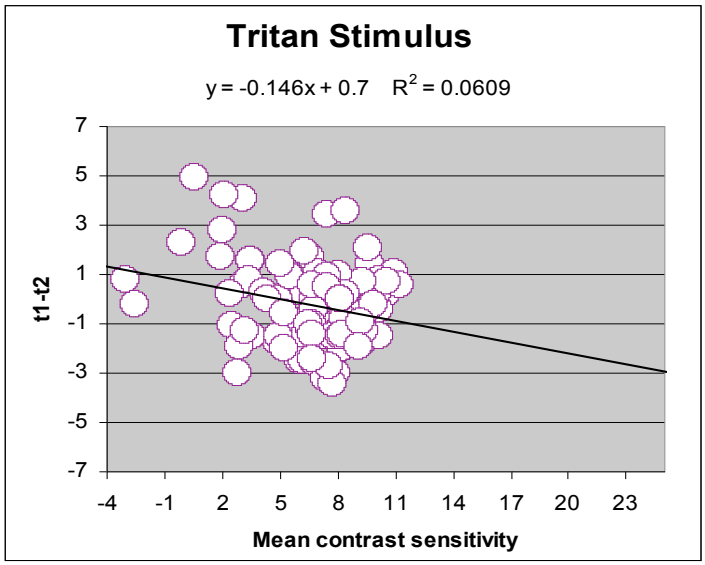


Figure 4. Comparing contrast sensitivity obtained from week 1 and 2 using the Bland-Altman method. All values presented in decibels (dB, where 1 dB = 0.1 log unit)

Figure 5. Comparison of depth of defect for the red-green and luminance stimuli using the Bland-Altman Method. All values expressed in dB. ($(R\text{-}G)_{dd}$ = depth of defect for red-green stimulus, L_{dd} = depth of defect for luminance stimulus).

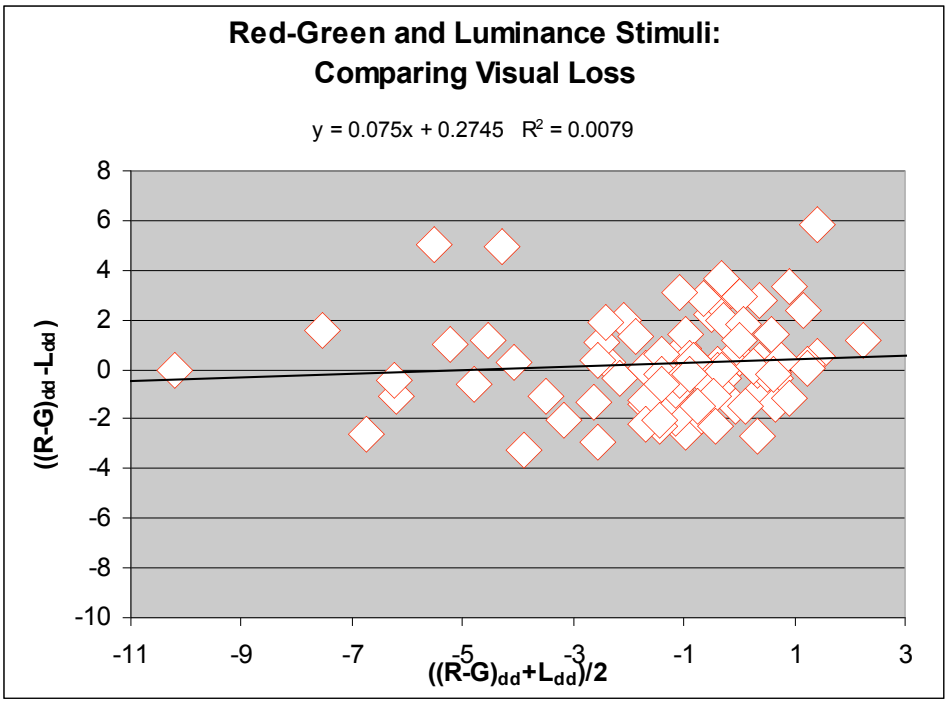


Figure 6. Comparison of glaucomatous defect for the tritan and luminance stimuli using the Bland-Altman Method. All values expressed in dB. (T_{dd} = depth of defect for tritan stimulus, L_{dd} = depth of defect for luminance stimulus).