Visual function assessment and metamorphopsia after macular hole surgery*,†

Walter Wittich1, Olga Overbury1,2,3, Michael A. Kapusta2, Jocelyn Faubert3

1Psychology Department, Concordia University, Montreal, 2Department of Ophthalmology, Sir Mortimer B. Davis Jewish General Hospital, McGill University, Montreal, Quebec, and 3École d’Optométrie, Université de Montréal, Québec, Canada

Abstract
The purpose of this study was to develop a new resolution acuity measure for patients after macular hole surgery. Fifty eyes of 44 patients who had undergone successful treatment were tested. Visual acuity was measured with the Snellen, Early Treatment of Diabetic Retinopathy Study and Landolt-C charts. A Line Resolution Test was performed as part of their follow-up exam where a vertical line was presented. Participants were categorized by their perception of the line as solid, bent or broken. The line could be distorted into a sine-wave pattern in order to determine the participants’ detection threshold for the distortion. Chart acuities did not differ among the three groups, as categorized by their line perception. Only the distortion measure was sensitive enough to differentiate the solid- from the broken-line group. The distortion measure assesses resolution power of the macula in smaller increments than acuity charts. This hyperacuity approach is more appropriate in the assessment of functional outcome after microsurgery.

Keywords: hyperacuity, macular hole, metamorphopsia, psychophysical assessment, recognition acuity, resolution acuity, visual acuity

Introduction
An idiopathic macular hole (MH) is a localized circular displacement of photoreceptors in the central retina (see Figure 1a) that occurs spontaneously and predominantly affects women above the age of 60 (la Cour and Friis, 2002). The aetiology of this condition is still not completely understood (Altaweel and Ip, 2003; Smiddy and Flynn, 2004), but retinal tissue can either be torn laterally because of tension or separated from the underlying membrane because of shrinkage of the vitreous (Kang et al., 2003). Symptoms of a MH include a sudden drop in visual acuity, caused by a central scotoma (blind spot) of generally 0.5–2.5 degrees of visual arc in diameter (approximately 150–800 microns), which is usually the initial cause for contacting an optometrist or ophthalmologist. Treatment includes a surgical procedure (vitrectomy) which facilitates closure of this type of retinal defect. Since Gass (1988, 1995) proposed a theoretical model of the evolution of MHs, and established MH stage criteria, research into treatment as well as vitreo-retinal microsurgery have advanced rapidly, resulting in high surgical success rates over the last 10 years (Benson et al., 2001). Anatomically successful treatment is defined by reattaching the edges of the elevated retina and closing the MH, which aims to restore the fovea (Tornambe et al., 1998; Mori et al., 2003; Kitaya et al., 2004). Functional improvement of vision is not part of the definition of successful surgical outcome, although it is the ultimate goal of the surgeon as well as the patient.

In the attempt to predict the degree to which vision will improve by surgical intervention, several preoperative...
factors have been identified, such as Snellen acuity (Kokame and de Bustros, 1995; Scott et al., 2000), size and duration of the MH (Benson et al., 2001; la Cour and Friis, 2002), as well as patient compliance with face-down positioning after the surgery (Thompson et al., 1996). A meta-analysis by Benson et al. (2001) has shown that approximately 80% of patients have improved function after surgery and 22–49% reach acuities of 20/40 (6/12) or better, although these results do not always easily replicate (Tranos et al., 2004). When attempting to assess this functional improvement, the ophthalmology community has, so far, predominantly relied on the most common measure of acuity, the Snellen eye chart. However, detailed evaluation of this measure has revealed several flaws and problems, such as high test-retest variability (Brown and Lovie-Kitchin, 1993), unequal number of letters per line and non-uniform progression in letter size and spacing (Ferris et al., 1982), as well as unequal letter difficulty (Sloan et al., 1952).

The need for more sensitive assessment of macular functioning led to the development of eye charts specifically for MH patients (Horiguchi et al., 2001). These charts each display multiple versions of the same Landolt-C target. The concept behind these charts is that, by looking anywhere on the chart, at least one of the targets will fall on the most sensitive area of the retina and the patient will be able to identify it. The authors did question if this measure reflects practical visual ability or assesses optimal visual acuity. Practical acuity would be the conscious ability of the patient to use the most sensitive part of the retina whereas optimal acuity is the best possible resolution when a target is passively presented to the same location, without conscious effort by the patient. Considering the array of highly sensitive psychophysical tests available today, their application is the next logical step in the development of more sensitive assessment of microsurgical results.

Research examining different measures of acuity has produced tests beyond the familiar Snellen chart for recognition acuity. The Bailey–Lovie chart addressed issues of letter size, spacing and difficulty (Bailey and Lovie, 1976). Later on, these improvements were incorporated in a chart used for the Early Treatment of Diabetic Retinopathy Study (ETDRS). The Landolt-C chart, similar in structure to the ETDRS chart, requires the detection of a gap in a ring, thereby assessing resolution acuity (Pointer et al., 1980). As MHs are generally round in appearance, scar tissue or residual scotomas after surgery can be expected to be circular as well, thereby specifically impairing performance on the Landolt-C chart more than on other types of visual acuity measure. Furthermore, letter optotypes contain additional information, such as diagonals and corners, which may facilitate recognition of a letter. This possibly allows for cognitive top-down processes to help a patient recognize a target, even if not all parts were resolved at the retinal level. Conversely, the detection of a gap in a ring is simply limited by resolution ability. It was, therefore, hypothesized that resolution acuity would be preferentially impaired in patients who reported distortions in their central visual field.

A type of acuity that is of particular interest in the assessment of macular function is hyperacuity. The visual system has the ability to integrate a line stimulus across the macular area, thereby detecting gaps, misalignments or curvature (McKee, 1991). As this ability has been shown to be independent of the aging process (Lakshminarayanan and Enoch, 1995), it holds promise in the assessment of age-related macular pathology. The macula is specialized for the resolution of high spatial frequency information. The distortion measure proposed in the present study creates a high spatial-frequency pattern and assesses the resolution power of the macula with such a hyperacuity approach.

In addition to acuity measures, a perceptual test has been used by ophthalmologists to diagnose full-thickness MHs and to evaluate their anatomical closure. In
1969, Gerrits and Timmerman (1969) reported on the use of a narrow slit of light in the assessment of patients with central retinal scotomas. In the same year, Watzke and Allen (1969) published their article on the development of a simple test which now carries their names. The Watzke-Allen (WA) test is performed on dilated eyes, using a slit lamp to place a narrow vertical beam of white light across the fovea. Patients are then asked what they see. The strip should be perceived as a continuous line if the macula is intact (negative WA). In the presence of macular pathology, other categorizations are possible: the line may be reported as broken (positive WA), distorted (metamorphopsia) or the patient is non-responsive or unsure (equivocal).

Its elegance and simplicity has made it most popular as a diagnostic tool in the assessment of MHs (Hikichi and Trempe, 1993; Ho et al., 1998; von Ruckmann et al., 1998; Kang et al., 2000; Scott et al., 2000; Gurwood and Jones, 2001; Krasnik et al., 2001; Hui and Guan, 2002; Wolf et al., 2003). Still, recurring inconsistencies in the responses to this test have questioned its reliability and sensitivity. In part, this weakness may be based in the phenomenon of perceptual filling-in at the cortical level. As clear parameters for the WA do not exist, patients often have the opportunity to scan up and down the line while it is being placed and adjusted. Relatively few reports mention the possibility of filling-in during ophthalmology examinations and are more common in neuro-ophthalmology circles (Larkin, 1980; Safran and Landis, 1996, 1999; Safran, 1997).

The development of optical coherence tomography (OCT, see Figure 1) has greatly influenced the way MHs are now diagnosed (Hee et al., 1995; Altaweel and Ip, 2003). Previous work has demonstrated that a clear positive WA may only be present in approximately 23% of full-thickness MHs that were confirmed by OCT (Tanner and Williamson, 2000). Meanwhile, 77% of patients report some form of metamorphopsia or orientation-dependent WA. In order to establish parameters that would allow a more objective interpretation of WA responses, Martinez et al. (1994) asked patients to choose between three diagrams of lines (solid, central thinning, broken). If the patients’ perception did not fit these criteria, they were encouraged to draw what they saw. Tanner and Williamson (2000) proposed a categorical scoring method. Patients were able to choose among five line representations, including three versions of metamorphopsia (central thinning, bent right, bent left). Both these studies agreed that the perception of a break in the line confirmed the clinical diagnosis of a MH; yet, the classifications of metamorphopsia or negative WA were not able to clearly identify macular status. In order to learn more about the limitations of this test, it was decided to take a closer look at three parameters of the WA task (dimension of the line, brightness, and display time) and to develop an improved Line Resolution Test (LRT).

In light of the available information about visual acuity and perceptual abilities in MH patients, the present study addressed the following hypotheses: (1) The perception of the LRT will contradict the WA test in patients after anatomically successful MH surgery. (2) When grouping these participants by their perception of the LRT as solid, bent or broken, their visual acuities will differ, but more so on resolution tasks (Landolt-C and hyperacuity) than on recognition acuity (Snellen ETDRS).

Methods

Participants

The sample consisted of 50 eyes of 44 patients with a diagnosis of idiopathic MH and no concomitant retinal disease, currently being treated by one retinal surgeon. All had undergone conventional pars plana vitrectomy with gas tamponade between 1998 and 2004. Preoperatively, all participants had a positive WA (broken line), as tested by the surgeon, and had then undergone anatomically successful surgery at least 4 months before their testing session. The participants were 33 women and 11 men with a mean age of 70.16 years (S.D. = 5.71), ranging from 58 to 80 years. Mean follow-up time was 2.15 years (S.D. = 1.53) ranging from 4.0 months to 5.6 years. Thirteen eyes were phakic without signs of cataract formation, while 37 eyes had undergone cataract surgery and were implanted with intraocular lenses.

Apparatus

All participants were optimally refracted before testing began. Acuities were determined with a standard Snellen acuity chart, as well as with an ETDRS and a Landolt-C chart. The WA test was administered by the clinician, using slit-lamp biomicroscopy with a hand-held or contact fundus lens. A narrow white slit beam was vertically projected across the fovea. The LRT, developed by the Faubert Perception Lab of the School of Optometry at the University of Montreal, was displayed on a Toshiba 2060CDS portable computer (Toshiba, Markham, Ontario, Canada) with liquid crystal display (LCD). Participants placed their heads in a chin rest and the screen display was positioned perpendicular to their viewing direction. They were instructed to fixate on the centre of an incomplete fixation cross (see Figure 2). This task is manageable for patients with central scotomas as the components of the fixation cross fall on peripheral retina. The stimulus was a vertical line pattern (see Figure 2). Its luminance distribution was
Based on the fourth derivative of the Gaussian function, peak luminance at the centre of the line was 130 cd m\(^{-2}\) and the line display had the same mean luminance (33 cd m\(^{-2}\)) as the grey background on which it was presented. The line could be distorted into a sine-wave pattern. The amplitude of the wave was measured in pixels and ranged from 0.5 to 20 (0.8–32 min arc). The wave pattern created a spatial frequency of 16 cycles per degree of visual angle. The line was adjusted for the 65 cm testing distance to form an image of length 3 and width 1 degree (diameter of the fovea).

**Procedure**

The protocol was approved by the Institutional Ethics Review Board, in accordance with the Canadian Tri-Council Policy Statement of ethical conduct for research involving humans. Patients were contacted at home via telephone and the purpose and procedure of the study were explained. They were then invited to arrive at the Ophthalmology Department 30 min before their next scheduled appointment with the retinal surgeon. Written informed consent was obtained and the participants completed the five psychophysical tests. Participants were tested under optimal artificial lighting conditions in an examining room as part of their post-surgical follow-up examination. Acuities were determined in the eye which received surgery while the other eye was occluded.

The LRT required the participant to decide which of two lines was distorted, using a two-alternative temporal forced-choice paradigm. Forty randomly alternating line-pairs were presented on a computer screen for 500 ms consecutively, with a 500-ms delay between lines of a pair. One line was without distortion, the other line varied in distortion, depending on the accuracy of the previous choice. Participants said aloud which line of each pair was distorted. Responses were recorded by the examiner. Then the display moved on to the next line-pair. Incorrect answers resulted in amplification of the next distortion, whereas correct answers made the task more difficult by decreasing the distortion. The program used a quick estimate by sequential testing (QUEST) procedure to determine the minimum threshold of pixel distortion (Watson and Pelli, 1983). Participants reported if they saw the line as solid, bent or broken, by choosing which option in a display most resembled their perception (see Figure 3). Test duration for the LRT was approximately 3 min per patient.

After completion of all tests, the participants’ pupils were dilated and they were seen by the retinal surgeon. During his examination, he confirmed anatomically successful MH closure and administered the WA test. A vertical strip of white light was projected across the macula and the participants were asked if they perceived a solid line or if a break or bend in the line was present (see Figure 4). All test results were recorded on a score sheet and a copy was made available to the participant upon request. Total duration of the session, including

---

**Figure 2.** The Line Resolution Test stimulus at the centre of the incomplete fixation cross forms an image that is 3 degrees long (diameter of the macula) and 1 degree wide (diameter of the fovea).

**Figure 3.** Categorical scoring options presented to participants for perception of the Line Resolution Test (LRT).

**Figure 4.** Categorical scoring options presented to participants for perception of the Watzke-Allen test (WA).
pre-experimental preparation time, did not exceed 30 min.

Results

For the WA and the LRT, participants were categorized by their perception of either line as solid, distorted/bent or broken. Chart acuities were expressed in the traditional Snellen fraction and then transformed into the logarithm of the minimum angle of resolution (logMAR) by calculating the inverse of the fraction value and taking the logarithm to base 10. This transformation expressed acuity in a form that fulfils the statistical requirements of linearity in the data (Holladay, 1997). For the distortion measure, the required amplitude for detection was transformed into logMAR as well by taking the log of $\alpha$, using the formula: $\alpha = \arctan(\text{stimulus size}/\text{stimulus distance})$, where $\alpha$ is the minimum angle of resolution. Preoperative data on Snellen acuity as well as the time since the surgical intervention were available. Statistical analyses of group differences in visual acuity were done using one-way analyses of covariance (ANCOVAs) by removing the possible effects of preoperative acuity as well as time since surgery. All data were normally distributed and Levene’s test confirmed that the assumption of homogeneity of variance had not been violated.

Hypothesis 1: Comparison of the two line tests

The distribution of patients according to their line perceptions is displayed in Figure 5. Postoperatively, 47 eyes (94%) saw a negative WA (solid line). Three eyes (6%) saw some type of distortion but were still able to see a continuous line. Overall, 34 eyes (68%) were able to see the LRT as a solid line, whereas 11 eyes (22%) saw a distorted line and five eyes (10%) saw a gap in the stimulus. Two of the three eyes with metamorphopsia on the WA perceived a break in the LRT. Thus, the perception of the LRT contradicted the WA in 15 eyes (30%).

Hypothesis 2: Differences in visual acuities

Overall mean visual acuities were 20/60 (range 20/25 to 20/400) for Snellen, 20/50 (range 20/25 to 20/130) for ETDRS and 20/70 (range 20/25 to 20/160) for Landolt-C. Pearson’s correlation coefficients for all acuity measures are displayed in Table 1. When the sample was divided into three groups based on the participants’ perception of the LRT, ANCOVAs for the mean visual acuities revealed that none of these chart acuities differed statistically across groups after accounting for preoperative Snellen acuity and time since surgery. Only the distortion measure was able to demonstrate statistically significant differences between the solid- and the broken-line group, $F(44,2) = 3.70, p < 0.03, \eta^2 = 0.14$, observed power = 0.65. The comparison of group differences in visual acuities is displayed in Figure 6. The height of the bars represents impairment in acuity; the error bars represent one standard error from the mean.

Discussion

The first component of the study confirmed the hypothesis that the LRT would contradict the WA in patients’ perception of the line stimulus. The number of contradicting results between the WA and the LRT was surprisingly high; 15 of 50 eyes (30%) reported a break or bend in the line display of the LRT that was undetected by the WA. These results indicate that the combined controlled components of line size, brightness and display time are of crucial importance for the perception of the LRT.

Line size

The importance of the line width initially became apparent in a detailed psychophysical examination by Burke (1999) who described his own experiences of
perceptually filling in lines that were wider than 0.5 degrees of visual angle (approximately 150 μm). Martinez et al. (1994), as well as Tanner and Williamson (2000), proposed a line width of 100 μm. Considering that the fovea has an approximate diameter of 330 μm, a line stimulus of approximately the same width seemed appropriate, in order to maximize the sensitivity of the LRT (see Figure 2). The issue of the line length has so far been ignored but may well be of crucial importance. It was decided to keep the length of the line at three degrees of visual angle (approximately 900 μm) which approximates the diameter of the area in which MHs are usually centred.

Brightness

Exact parameters for the WA do not exist and are difficult to establish as the administration of this test depends heavily on the slit-beam lamp available and the person operating it. Depending on the make of the lamp, brightness can be adjusted gradually or in discrete units and can range from approximately 300 to over 900 cd m⁻². In ophthalmology departments, where several slit-beam lamps are available, it becomes difficult to assure that a patient will always be examined with the same machine. Furthermore, over time, light bulbs need replacing and the original luminance parameters may change. In order to reduce light scatter and to maintain constant luminance levels, it was proposed to display the LRT on a computer screen at low brightness (peak luminance 130 cd m⁻², mean luminance 33 cd m⁻²).

Display time

In a clinical setting, the patient continuously looks at the slit beam as it is being adjusted and placed across the fovea by the ophthalmologist. Therefore, it is possible for the patient to scan up and down the line. At the same time, the pupil is dilated and an extremely large amount of light reaches the retina. Photoreceptors may briefly bleach and the patient’s perception of the line may become altered. In order to avoid any possible integration of the stimulus over time, the LRT display appeared only briefly on the screen (500 ms). It was not necessary to dilate the pupil for this test, thereby creating a more realistic and practical way to assess the actual visual capabilities of the patients. The ability to fill in missing information after central field loss has previously been described (Craik, 1966; Gerrits and Timmerman, 1969; Burke, 1999, 2002). Gerrits and
Timmerman (1969) pointed out that the filled-in components did not appear and disappear instantly but faded in and out slowly. By presenting the LRT at low luminance and limiting the exposure to 500 ms, any residual damage to the macula that required filling-in may have become detectable.

The second hypothesis found partial support as only one of the resolution acuity measures was sensitive enough to detect group differences. The new distortion measure outperformed all three chart acuities in its ability to separate two of the groups based on line perception. Although acuities on the Snellen, ETDRS, and the Landolt-C charts decreased as perception of the LRT worsened, the within-group variances were too large to be statistically significant. As perception of the line degraded, only the distortion threshold declined in a manner that allowed for statistical separation of the groups. It is speculated that this is the case because of the increased sensitivity of this new measure. The distortion measure assesses the resolution power of the retina in minutes of visual arc (min arc). Specifically, the smallest possible target on the Snellen chart has a diameter of five min arc and the parts of each letter are one min arc wide, whereas the distortion measure has a minimum of 0.8 min arc and can easily be adjusted to an even smaller scale.

As an assessment tool for postoperative visual impairment, a line perception task may be more sensitive than reading letters; yet, this type of test encounters a perceptual problem. The WA operates on the assumption that no perceptual filling-in occurs if retinal damage is present. The research literature has repeatedly demonstrated that this assumption is not true (Gerrits and Timmerman, 1969; Burke, 1999; Sakaguchi, 2001, 2003; Zur and Ullman, 2003). Therefore, the use of the WA in the follow-up assessment after MH surgery must be considered with caution. If this test is able to detect perceptual loss, the defect is probably substantial, yet, if no such loss is found, this may not be sufficient evidence for complete restoration of retinal function. Research has shown that after loss of photoreceptors in age-related macular degeneration, central scotomas of up to seven degrees of visual angle can still be successfully filled-in on the perceptual level (Zur and Ullman, 2003).

The interpretation of the Snellen chart acuities encountered an important problem. The results for the ‘broken’ group created a large error bar, based on the amount of variability and small sample size of this group. This error bar was, in fact, so large that it contained the means of both the ‘solid’ and the ‘bent’ group. This is further evidence that the use of the Snellen chart as a scientific measure is not appropriate when working with patients with visual impairment. Overall, the three chart acuities were significantly correlated with each other, more so than their correlations with the distortion measure. Mostly, this is because of the similar format and task in reading from an eye chart. Furthermore, the distortion measure assesses the resolution power of the retina in smaller increments of min arc, a scale that is usually only applied in psychological research on the fully functioning visual system.

In conclusion, the assessment of closure of a MH with the traditional WA test could be improved upon by the suggested parameters for a LRT. Displaying the line only briefly at a low brightness setting may be the key factor in detecting retinal damage that otherwise may be undetectable because of perceptual filling-in. Future work in this area should incorporate preoperative measures of all acuity types as well as the LRT measures. Considering that the surgical tools used in ophthalmology get updated and improved on a monthly basis and are rarely older than a few years, it is time to introduce assessment tools for visual functioning that are more recent and sophisticated than a Snellen chart, developed in 1854.

Acknowledgements

This study was partially supported by the Canadian Institutes of Health Research (Grant no. MOP-5353291) and the Reseau de Recherche en Santé de la Vision.


