Change in visual acuity is well correlated with change in image-quality metrics for both normal and keratoconic wavefront errors

We determined the degree to which change in visual acuity (VA) correlates with change in optical quality using image-quality (IQ) metrics for both normal and keratoconic wavefront errors (WFEs). VA was recorded for five normal subjects reading simulated, logMAR acuity charts generated from the scaled WFEs of 15 normal and seven keratoconic eyes. We examined the correlations over a large range of acuity loss (up to 11 lines) and a smaller, more clinically relevant range (up to four lines). Nine IQ metrics were well correlated for both ranges. Over the smaller range of primary interest, eight were also accurate and precise in estimating the variations in logMAR acuity in both normal and keratoconic WFEs. The accuracy for these eight best metrics in estimating the mean change in logMAR acuity ranged between $\pm 0.0065$ to $\pm 0.017$ logMAR (all less than one letter), and the precision ranged between $\pm 0.10$ to $\pm 0.14$ logMAR (all less than seven letters).

Introduction

It is well known that the normal human eye suffers from wavefront errors (WFEs) not correctable with spectacle lenses (Helmholtz, 1896; B. Howland & Howland, 1976; H. C. Howland & Howland, 1977; Smirnov, 1961; Walsh, Charman, & Howland, 1984). For highly aberrated eyes, the higher-order (HO) WFEs are of such a magnitude that, in many instances, spectacle corrections do not provide acceptable vision. Even with rigid gas-permeable (RGP) contact lenses, the residual aberrations of highly aberrated keratoconic
eyes are well outside the normal range (Kosaki et al., 2007; Marsack, Parker, Pesudovs, Donnelly, & Applegate, 2007; Negishi, Kumanomido, Utsumi, & Tsubota, 2007), leading to the investigation of wavefront-guided corrections designed to target patient-specific levels of HO WFE (Marsack, Parker, & Applegate, 2008; Sabesan et al., 2007). Intuitively, one may think the larger the amount of residual RMS WFE, the poorer the visual performance. However, RMS WFE is not necessarily a good predictor of visual performance as measured by visual acuity (VA) because (a) individual aberrations do not impact visual performance equally (Applegate, Ballentine, Gross, Sarver, & Sarver, 2003; Applegate, Sarver, & Khemsara, 2002; Chen, Singer, Guirao, Porter, & Williams, 2005) and (b) individual aberrations interact to increase or decrease visual performance (Applegate, Marsack, Ramos, & Sarver, 2003; McLellan, Prieto, Marcos, & Burns, 2006). Several other metrics for quantification of optical quality have been proposed (Thibos, Hong, Bradley, & Applegate, 2004) and will be examined here with respect to VA.

Initially, studies examined several image-quality (IQ) metrics to understand the impact of individual aberrations and combinations of aberrations on VA (Cheng, Bradley, Ravikumar, & Thibos, 2010; Cheng, Bradley, & Thibos, 2004; Legras & Rouger, 2008; Marsack, Thibos, & Applegate, 2004; Rouger, Benard, & Legras, 2010). More recently, studies have explored the impact of whole-eye WFE on VA and found certain IQ metrics accounted for greater variance in VA than others (Ravikumar, Applegate, Shi, & Bedell, 2011; Ravikumar, Sarver, & Applegate, 2012; Schoneveld, Pesudovs, & Coster, 2009; Shi, 2012; Yoon, 2008). All these studies used regression analysis and reported the best metrics as the metric with the largest coefficient of determination ($r^2$).

Table 1 lists the coefficients of determination ($r^2$) previously reported between VA and IQ metrics separately in normal eyes (Chen et al., 2005; Cheng et al., 2010; Cheng, Bradley, et al., 2004; Legras & Rouger, 2008; Marsack et al., 2004; Ravikumar et al., 2011; Ravikumar et al., 2012; Rouger et al., 2010) and in highly aberrated eyes (Schoneveld et al., 2009; Shi, 2012; Yoon, 2008).

Most of the literature cited in Table 1 aims to correlate and/or predict the absolute VA from the absolute IQ metrics. However, here we emphasize predicting the change ($\Delta$) in VA from $\Delta$ in IQ metrics. The reasons are (a) describing $\Delta$ facilitates comparison across subjects, regardless of their baseline acuity, and (b) $\Delta$ describes how the new prescription or correction alters VA from the present habitual correction.

To understand how WFE impacts visual performance of an individual as reflected by VA, it is important to have IQ metrics that are well correlated with $\Delta$ VA regardless of the nature of the underlying WFE (whether the eye is normal or highly aberrated). Such a metric defines the optical properties of the eye-correction system that are consistent with good acuity and provides objective criteria for the design of optical corrections. Here we examine the correlation, the confidence interval (accuracy), and the prediction interval (precision) between $\Delta$ logMAR acuity and $\Delta$ in IQ metrics, both in normal and keratoconic WFEs. The confidence interval is an estimate of reliability of the mean data whereas the prediction interval is an estimate of an interval within which future observations will fall.

The purposes of this paper are

- To determine if $\Delta$ logMAR VA is highly correlated with $\Delta$ in metrics of IQ for both normal and typical keratoconic WFEs
- To determine the accuracy and precision of the best correlating IQ metrics in estimating the variation in VA over a clinically relevant range of acuities

### Methods

#### Study subjects

The study adhered to the Declaration of Helsinki. Before participation, all subjects signed an informed-consent form approved by the University of Houston Institutional Review Board after learning the nature and the possible consequences of the study. Five healthy, normal subjects were recruited to read a series of aberrated letter charts generated from scaled versions of WFE from both normal ($n = 15$) and keratoconic ($n = 7$) eyes from previously recorded WFE data sets. The five subjects who read the aberrated charts (test subjects) were between the ages of 24 and 29 years old and were free of systemic and ocular pathology with best-corrected manifest distance VA better than 20/20.

#### Refraction

For each test subject, the eye with better acuity was dilated and accommodation paralyzed with two drops of 1% tropicamide. Between 20 and 35 minutes after instilling the drops, subjective trial frame refraction (Table 2) was performed through a 3-mm artificial pupil while the subject viewed a logMAR acuity chart at a test distance of 12.2 feet. A 3-mm pupil was chosen as it is close to the optimal pupil diameter to balance between diffraction effects and aberration effects (Charman, 1991).
Aberrometry was performed for each subject 10 times using a COAS-HD aberrometer (AMO-Wavefront Science, Inc., Albuquerque NM) while wearing their best spectacle correction in a trial frame with the artificial pupil removed. It was assumed that the artificial pupil and the WFE measured pupil were colocated. Each measurement was fitted with a normalized Zernike expansion (Thibos, Applegate, Schwiegerling, & Webb, 2002) through the 10th radial order over a 3-mm pupil. The average of the 10

<table>
<thead>
<tr>
<th>Author</th>
<th>Aberrations included</th>
<th>Eyes</th>
<th>Pupil diameter (mm)</th>
<th>Correlation/prediction (c/p)</th>
<th>Metrics ($r^2$)</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheng (2004)</td>
<td>Through focus</td>
<td>Normal</td>
<td>5</td>
<td>c</td>
<td>VSOTF (0.67), PFSt (0.70), PSFSTD (0.67)</td>
<td>Sloan letters</td>
</tr>
<tr>
<td></td>
<td>1. Defocus + spherical aberration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. Astigmatism + coma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Primary astigmatism + secondary astigmatism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marsack (2004)</td>
<td>1. Defocus + spherical aberration</td>
<td>Normal</td>
<td>6</td>
<td>c</td>
<td>VSOTF (0.81), six metrics (PFSt, VSX, NS, AreaMTF, AreaOTF, VSMTF −0.70)</td>
<td>logMAR acuity</td>
</tr>
<tr>
<td></td>
<td>2. Primary astigmatism + secondary astigmatism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Spherical aberration + quadrafoil</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. Spherical aberration + secondary astigmatism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legras (2008)</td>
<td>Spherical aberration/ trefoil/coma</td>
<td>Normal</td>
<td>Natural</td>
<td>c</td>
<td>VMTF (0.84)</td>
<td>logMAR acuity</td>
</tr>
<tr>
<td>Cheng (2010)</td>
<td>Second through fourth order</td>
<td>Normal</td>
<td>5</td>
<td>c</td>
<td>NS (0.74), VSMTF (0.73)</td>
<td>Sloan letters</td>
</tr>
<tr>
<td>Rouger (2010)</td>
<td>Astigmatism/spherical aberration/coma/ trefoil/defocus</td>
<td>Normal</td>
<td>5.5</td>
<td>c/p</td>
<td>rmtf (0.85)</td>
<td>Contrast sensitivity (10 and 25 cpd) and VA</td>
</tr>
<tr>
<td>Ravikumar (2011)</td>
<td>Second through fifth order (scaled)</td>
<td>Normal</td>
<td>3</td>
<td>c</td>
<td>VSOTF (0.96)</td>
<td>logMAR acuity</td>
</tr>
<tr>
<td>Ravikumar (2012)</td>
<td>Three WFE - 10th radial order</td>
<td>Normal</td>
<td>2, 3, 4, 5, 6, and 7</td>
<td>c</td>
<td>NS, VSX, VSMTF, PFSc, PFSt, and RMSs (0.80)</td>
<td>logMAR acuity</td>
</tr>
<tr>
<td>Yoon (Wavefront Congress, 2008)</td>
<td>Correction of second through fifth radial orders</td>
<td>KC</td>
<td>6</td>
<td>c</td>
<td>Cross-correlation (0.76), RMS (0.72)</td>
<td>logMAR acuity</td>
</tr>
<tr>
<td>Schoneveld (2009)</td>
<td>Aberrations are corneal first surface measurement</td>
<td>KC, PK</td>
<td>3, 4, and 5</td>
<td>c</td>
<td>logPFWc (0.64), logVOTF (0.58)</td>
<td>logMAR</td>
</tr>
<tr>
<td>Shi (2012)</td>
<td>Residual aberrations following fourth-order correction</td>
<td>KC</td>
<td>3</td>
<td>c/p</td>
<td>logPFSt, logVSX, logVSMTF (0.80)</td>
<td>logMAR acuity</td>
</tr>
</tbody>
</table>

Table 1. Previous studies investigating the relationship between IQ metrics and VA. Please refer to Appendix 1 for the description of the metrics.

Subject’s WFE

Aberrometry was performed for each subject 10 times using a COAS-HD aberrometer (AMO-Wavefront Science, Inc., Albuquerque NM) while wearing their best spectacle correction in a trial frame with the artificial pupil removed. It was assumed that the artificial pupil and the WFE measured pupil were colocated. Each measurement was fitted with a normalized Zernike expansion (Thibos, Applegate, Schwiegerling, & Webb, 2002) through the 10th radial order over a 3-mm pupil. The average of the 10
Table 2. For all five subjects: Best-corrected sphero-cylindrical refraction and associated logMAR acuity.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Best-corrected sphero-cylindrical refraction of the test eye</th>
<th>Best-corrected logMAR acuity of the test eye</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>Plano</td>
<td>−0.14</td>
</tr>
<tr>
<td>Subject 2</td>
<td>+1.25 D</td>
<td>−0.08</td>
</tr>
<tr>
<td>Subject 3</td>
<td>+0.75 D/−0.25 D × 175</td>
<td>−0.10</td>
</tr>
<tr>
<td>Subject 4</td>
<td>Plano</td>
<td>−0.12</td>
</tr>
<tr>
<td>Subject 5</td>
<td>Plano</td>
<td>−0.12</td>
</tr>
</tbody>
</table>

Normal and keratoconic WFEs used to degrade acuity charts

Fifteen right-eye normal WFEs over a 6-mm pupil were randomly selected from the best-corrected 100-subject normal WFE data set (Thibos, Hong, Bradley, & Cheng, 2002). Because WFE was measured at 633 nm, a chromatic aberration correction was applied to the defocus term to shift from 633 nm to 555 nm (Thibos, Bradley, Still, Zhang, & Howarth, 1990) for a 4-mm pupil. Residual keratoconic WFEs at 555 nm of seven keratoconic eyes over their habitual RGP lenses were obtained from Marsack et al. (2007) for a 4-mm pupil. (Note: The 3-mm pupil will allow all visually relevant spatial frequencies of interest in the simulated aberrated charts to pass. Chart generation is discussed in the section below titled final target generation).

Scaling of WFEs and calculation of metrics of IQ

The optical quality in normal eyes ranges from −0.4 to −1.8 logVSX (Ravikumar et al., 2012). In the seven keratoconic eyes, optical quality ranged from −0.8 to −2.5 logVSX. We assumed in the case of more severe keratoconus (KC) IQ would exceed this range and therefore scaled KC WFEs over a larger range. We also wanted to have significant overlap between the two populations and scaled Zernike coefficients in 12 approximately equal steps such that in normal eyes IQ ranged from −0.12 to −2.5 logVSX (intended to induce an approximately seven- to eight-line change in VA) and in KC eyes ranged from −0.45 to −3.5 logVSX (intended to induce an approximately 10- to 11-line change in VA). The resulting 264 (12 steps × 22 WFEs) WFEs were used to calculate 29 IQ metrics as previously described (Thibos et al., 2004).

Source versus observer methods

The impact of WFE on vision is often studied using two different approaches, which Chan, Smith, and Jacobs (1985) described as “observer” (blurring the eye) and “source” (blurring the stimulus) methods. In the observer method, the WFE can be induced in the eye through lenses (Bradley, Thomas, Kalaher, & Hoerres, 1991; Herse & Bedell, 1989), phase plates (Lopez-Gil, Howland, Howland, Charman, & Applegate, 1998; Navarro, Moreno-Barriuso, Bara, & Manecebo, 2000), and adaptive optics (Liang, Williams, & Miller, 1997). In contrast, the source method computationally incorporates the error into the target (Akutsu, Bedell, & Patel, 2000; Applegate, Ballentine, et al., 2003; Applegate, Marsack, et al., 2003; Chan et al., 1985; Cheng, Bradley, et al., 2004). The source method was used here.

Generation of 30% logMAR acuity charts

Thirteen-line logMAR acuity charts (0.9 to −0.3 logMAR) of 30% contrast were generated using Visual Optics Laboratory (VOL) Professional software (version 6.89, Sarver & Associates, Inc.). VOL uses equally identifiable letters to randomly generate logMAR charts such that no letter is repeated in any given line. For each of the 264 test conditions, three unique unaberrated acuity charts of 30% contrast (792 total) were generated.

Final target generation

These 792 charts were used to create the aberrated visual simulations that were precompensated for the 3-mm WFE of each individual subject, using techniques first described by Burton and Haig (1984) and later used by others (Alonso & Barreto, 2003; Cheng et al., 2010; Ravikumar et al., 2012; Shi et al., 2013). The precompensation procedure modifies the image to be displayed to compensate for the impact of residual aberration present in each subject’s eye. To illustrate the principles of our approach, the precompensation procedure is explained mathematically in Equation 1 and shown diagrammatically in Figure 1.

$$I_{pre}(x, y) = I_{unaberrated}^{-1}(x, y) \times PSF_o(x, y) \times PSF_o^{-1}(x, y)$$ (1)

where

- $I_{pre}$ = precompensated image
- $I_{unaberrated}$ = unaberrated image
- $PSF_o$ = point spread function (PSF) of the scaled WFE
- $PSF_o^{-1}$ = inverse of PSF of the observer’s WFE
Apparatus

The 792 aberrated and precompensated acuity charts for each subject were displayed on a gamma-corrected, black-and-white, high-resolution (Totoku M253i2, 1,200 × 1,600 pixels) LCD monitor at a luminance of 410 cd/m². The subjects viewed the charts through a unit magnification telescope, which imaged the 3-mm artificial pupil and the best-corrected sphero-cylindrical prescription onto the geometric center of the subject’s dilated pupil. Alignment was maintained using a dental bite bar and a three-dimensional translational stage and monitoring system (see Figure 2). The monitoring system consisted of an alignment target, which was conjugate and centered with the artificial pupil and the eye’s pupil. The display was placed at 12.2 feet and subtended 4.98° from the position of the unit magnification telescope.

Measurement of acuity

The 792 charts for each subject were displayed in random order (KC and normal charts) using a custom Matlab program that utilized Psychtoolbox (Brainard, 1997). A letter-by-letter scoring system was used, in which subjects were given credit for each letter read correctly up to the fifth miss (Carkeet, 2001). LogMAR

Figure 1. The clear 30% contrast chart is convolved with the PSF of the test WFE and deconvolved with the subjects best-corrected WFE, resulting in the precompensated chart displayed to the subject. Thirty percent contrast clear charts are used to allow for the contrast enhancement of the precompensation to occur without exceeding the dynamic range of the monitor.

Figure 2. Experimental apparatus. The artificial pupil and correcting lens (sphere and cylinder) are imaged into the eye’s pupil with unit magnification (dashed lines). The display monitor is conjugate with the retina. The heavy black lines are marginal rays to illustrate that the display monitor can be fully imaged onto the retina. The pellicle beam splitter and the mirror allow the pupil-monitoring camera to view the artificial pupil and the eye’s pupil simultaneously.
acuity for each chart was calculated using Equation 2:

\[ \text{logMAR acuity} = x - (LC \times 0.02) \]  

where \( x \) is logMAR acuity for the line above the largest test line on the chart, \( LC \) is the total number of letters read correctly up to the fifth missed letter, and 0.02 is the logMAR equivalent of one letter.

A short training session was given to each subject, enabling the subjects to get familiar with the apparatus and procedure.

**Normalization of acuity data**

To compare data across subjects, the data for each subject were first normalized to the subject’s mean best-corrected logMAR acuity determined by reading an additional three unaberrated pre-emphasized logMAR charts.

\[ \Delta L(\text{change in acuity}) = L(A) - L(UA) \]

where

\[ \Delta L = \text{logMAR acuity gained or lost} \]
\[ L(A) = \text{logMAR acuity on an aberrated chart} \]
\[ L(UA) = \text{average logMAR acuity on an unaberrated chart} \]

In this paradigm, negative values indicate a gain in acuity and positive numbers a loss of acuity. The average gain in acuity for the five subjects reading unaberrated pre-emphasized charts was 0.05 logMAR (2.5 letters) and was in the same direction as that reported by Cheng et al. (2010). Figure 3 diagrammatically represents the experimental procedure and logMAR acuity calculation.

**Data analysis**

**Correlations over full test range**

To determine the degree to which \( \Delta \) logMAR acuity is accounted for by \( \Delta \) IQ, the coefficient of determi-

---

**Figure 3.** (A) Diagrammatic representation of the experimental procedure in words and (B) pictorially. The five red and blue circles indicate five letters read incorrectly in the aberrated and unaberrated chart, respectively.
nation ($r^2$) was calculated by regressing the $\Delta$ logMAR acuity against $D$ metric value for each of 29 IQ metrics. These were calculated separately for the simulated WFEs of normal and keratoconic eyes and for both sets of WFEs in the aggregate. A $t$ test was performed to compare the difference between the correlation coefficients ($r$) for normal and KC simulated WFEs.

Correlations over a clinically relevant range

To estimate the range of acuity anticipated in keratoconic eyes, we turn to the CLEK study (Zadnik, Barr, Gordon, & Edrington, 1996), in which 77.9% had best-corrected VA better than or equal to 0.3 logMAR (20/40) in both eyes. We therefore elected to limit the clinically relevant range to 20/40 and better. The data that resulted in an absolute VA of 0.3 logMAR were identified, and then the change in logMAR acuity was calculated and plotted against the change in IQ metrics. The accuracy (95% confidence interval) and precision (95% prediction interval) from the resulting regression lines were calculated.

### Results

Correlations over full test range

Table 3 reports the metrics in which the slope of the correlation is statistically indistinguishable ($p > 0.05$) for both normal and KC WFEs. The first three metrics (logNS, logVSX, and logLIB) accounted for >80% of the variance in the $\Delta$ logMAR acuity. The next set of six metrics (logSRX, logEW, logVSMTF, logAreaMTF, logSTD, and logVSOTF) accounted for >70% of the variance in the $\Delta$ logMAR acuity. The remaining five metrics (logPFWt, logENT, logSM, logSFCMTF, and logSFCOTF) were correlated less with the $\Delta$ logMAR acuity ($r^2 < 0.6$).

The metrics logD50, logPFS, logHWHH, logAreaOTF, and logSRMTF had a significantly greater correlation value in normal eyes compared to highly aberrated eyes. In contrast, logRMSw, logPV, logRMSs, logPFWc, logPFSc, logBave, logCW, logVOTF, and logVNOTF had a significantly greater correlation in KC eyes compared to normal eyes.

### Full data sets for top three metrics

Figure 4 shows the individual data for $\Delta$ logMAR acuity for all five subjects for all conditions as a function of the $\Delta$ in the three best metrics: log neural sharpness, log visual Strehl ratio, and log light in the bucket. The left column displays the data for all 22 WFEs (15 normal and seven KC), and the right column displays the data set excluding two of the normal WFEs that behaved differently (D64 and D37). The black/grey open circles represent the data points for charts with normal/KC WFE, respectively, and the thick solid/dashed lines represent the best fitting lines for the normal/KC WFEs. The dark and light grey filled symbols represent the two normal WFEs (D64 and D37). Because these two data sets are nonlinear, they are fitted with a polynomial.
Figure 4. Δ logMAR acuity as a function of Δ in three best correlating metrics: log visual Strehl ratio (VSX), log neural sharpness (NS), and log light in the bucket (LIB). Each data point represents the mean acuity obtained on three chart readings for each subject. The black/grey open circles represent the data points for charts convolved with normal/KC WFE, respectively. The thick solid/dashed lines represent the best-fitting lines to the logMAR acuities obtained with the normal/KC WFE, respectively. The dark and light grey filled symbols represent the data for two normal WFEs (D64 and D37) that behaved differently. These data are fitted with polynomial equations. The left column shows the data and fits for all 22 WFEs (15 normal and seven KC), and the right column represents the same data set, excluding two WFEs (D64 and D37).
Figure 5. The Δ logMAR acuity as a function of Δ in metric value for nine different metrics over the range when the logMAR acuity is better than 0.3 (20/40) for both normal and KC WFEs. Each data point represents the mean of three trials for each subject. Red is used for normal WFEs and blue for KC WFEs. The solid lines represent the best-fitting lines. The dashed/dashed-dotted lines represent the 95% confidence interval and 95% prediction interval, respectively.

**Accuracy and precision of regression fits for the clinically relevant range of acuity**

Figure 5 displays the regression analysis, the 95% confidence interval, and the 95% prediction interval for the top nine metrics that had statistically indistinguishable correlations over the smaller clinically relevant acuity range. For all of the metrics except for log VSOTF, there is no statistically significant (p > 0.05) difference between the slopes of the regressions for normal and KC eyes, and there is a good agreement in the 95% confidence and prediction intervals. For logVSOTF, a statistically significant (p < 0.05) difference exists; therefore, the accuracy and precision was reported only for the top eight metrics.

The accuracy for the top eight metrics ranged between ±0.0065 and ±0.017 logMAR (all less than one letter), and the precision ranged between ±0.10 and ±0.14 logMAR (all less than seven letters). The average standard deviation of the three measures of acuity for each condition across all subjects is a measure of the test-retest variability and is 0.033 ± 0.021 logMAR (approximately two, plus or minus one, letters). The accuracy (95% confidence interval) of all the metrics listed in Table 4 lies within the test-retest variability for the five subjects.

**Discussion**

When the acuity is better than 20/40, the agreement between the confidence intervals, the prediction intervals, and a negligible difference between the fitted slopes and intercepts (Figure 5) for logNS, logVSX, logLIB, logSRX, logEW, logVSMTF, logAreaMTF, and logSTD suggest that these metrics can be used to
predict Δ VA regardless of whether the eye is normal or keratoconic. The metrics that are best correlated, accurate, and precise in estimating Δ logMAR acuity are all image plane metrics.

## Two uniquely behaving WFEs

Figure 4 shows the correlation between Δ VA as a function of Δ in the three best metrics. The data points from two of the WFEs followed a different trend than all other WFEs. Ideally, the diffraction-limited PSF takes the shape of an Airy’s disc, which has a clearly defined circular hot spot at which maximum light intensity is pooled. In the presence of residual WFE, the PSF takes an irregular shape. Twenty WFEs had a well-defined hot spot in their PSF. As seen in Figure 6, the WFEs D37 and D64 have an elongated PSF in which the intensity is distributed approximately equally along its length possibly due to astigmatism (D64) and trefoil (D37). This effect exaggerated as the WFE was scaled up. To illustrate, Figure 7 compares the simulated logMAR acuity charts all having a logNS value of −0.9. Further, in experiments in which scaling

<table>
<thead>
<tr>
<th>Rank</th>
<th>Metrics</th>
<th>Normal</th>
<th>Metrics</th>
<th>KC</th>
<th>Metrics</th>
<th>Normal</th>
<th>Metrics</th>
<th>KC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>logVSX</td>
<td>±0.0065</td>
<td>logVSX</td>
<td>±0.0125</td>
<td>logVSX</td>
<td>±0.1000</td>
<td>logVSX</td>
<td>±0.1037</td>
</tr>
<tr>
<td>2</td>
<td>logNS</td>
<td>±0.0066</td>
<td>logNS</td>
<td>±0.0129</td>
<td>logNS</td>
<td>±0.1021</td>
<td>logNS</td>
<td>±0.1059</td>
</tr>
<tr>
<td>3</td>
<td>logLIB</td>
<td>±0.0066</td>
<td>logLIB</td>
<td>±0.0137</td>
<td>logLIB</td>
<td>±0.1046</td>
<td>logLIB</td>
<td>±0.1071</td>
</tr>
<tr>
<td>4</td>
<td>logEW</td>
<td>±0.0071</td>
<td>logVSMTF</td>
<td>±0.0150</td>
<td>logEW</td>
<td>±0.1138</td>
<td>logSRX</td>
<td>±0.1264</td>
</tr>
<tr>
<td>5</td>
<td>logSRX</td>
<td>±0.0071</td>
<td>logSRX</td>
<td>±0.0152</td>
<td>logSRX</td>
<td>±0.1138</td>
<td>logSRX</td>
<td>±0.1266</td>
</tr>
<tr>
<td>6</td>
<td>logSTD</td>
<td>±0.0081</td>
<td>logSTD</td>
<td>±0.0161</td>
<td>logSTD</td>
<td>±0.1200</td>
<td>logSTD</td>
<td>±0.1377</td>
</tr>
<tr>
<td>7</td>
<td>logVSMTF</td>
<td>±0.0082</td>
<td>logAreaMTF</td>
<td>±0.0168</td>
<td>logVSMTF</td>
<td>±0.1295</td>
<td>logAreaMTF</td>
<td>±0.1385</td>
</tr>
<tr>
<td>8</td>
<td>logAreaMTF</td>
<td>±0.0084</td>
<td>logEW</td>
<td>±0.0174</td>
<td>logAreaMTF</td>
<td>±0.1369</td>
<td>logEW</td>
<td>±0.1407</td>
</tr>
</tbody>
</table>

Table 4. The accuracy and precision when estimating the Δ logMAR acuity as determined by the 95% confidence and prediction intervals for the eight most highly correlating IQ metrics.

Figure 6. Panels A through G display PSFs for seven WFEs all having a logNS value of −0.9. The top row three KC WFEs and the middle row two normal WFEs, which behaved typically, and the bottom row two normal WFEs that behaved atypically (D37 and D64 show an elongated region of high intensity).
was not used (Ravikumar, Marsack, Shi, & Applegate, 2013), no such nonlinear behavior was observed.

Comparison with other studies

The nine metrics that are well correlated ($r^2 > 0.7$) with Δ logMAR VA have a potential application beyond estimating Δ VA given a Δ metric value. The same metrics were among the top 15 metrics that were shown by Thibos et al. (2004) and Guirao and Williams (2003) to accurately and precisely predict observers’ sphero-cylindrical refractive corrections from their WFE. Tarrant, Roorda, and Wildsoet (2010) evaluated the top five metrics identified by Thibos et al. (2004) to accurately and precisely predict observers’ wavefront refraction and found that logNS and logVSMTF also are good predictors of accommodative response.

We found the metrics logNS and logVSMTF account for $>70\%$ of the variance in both normal and keratoconic WFEs, which is in accordance with previous studies (Cheng et al., 2010; Cheng, Bradley, et al., 2004; Marsack et al., 2004; Ravikumar et al., 2012; Shi, 2012). Marsack et al. (2004), Ravikumar et al. (2012), and Shi et al. (2013) have also shown that logVSX accounts for $>80\%$ of the variance in predicting logMAR acuity. Our finding that logVSOTF accounts for $>70\%$ of the variance in the logMAR acuity produced by normal and KC WFEs is in agreement with results reported previously (Cheng, Bradley, et al., 2004; Marsack et al., 2004; Ravikumar et al., 2011). However, even though logVSOTF is well correlated with logMAR acuity, the correlations for normal and KC eyes were not equivalent. In addition, like Cheng et al., 2004, and Marsack et al., 2004, the results of the current study show that the metrics VOTF and VNOTF correlate poorly with Δ logMAR acuity.

At the 2008 Wavefront Congress, Yoon (2008) reported the correlation between logMAR acuity and several IQ metrics for KC eyes and found a cross-correlation metric to be best, accounting for 76% of the variance in logMAR acuity. The correlation was similar to correlations reported in the current study over a similar range of acuity loss (Figure 5). (For a description of the cross-correlation metric, refer to Lewis, 1995; Ravikumar, Bradley, & Thibos, 2010.) Schoneveld et al. (2009) correlated logMAR acuity and contrast sensitivity with 30 IQ metrics derived from corneal topography in 26 KC, eight penetrating keratoplasty, and 18 normal eyes. They concluded that logPFWc (pupil fraction metrics) and logVOTF (volume under optical transfer function) accounted for 60% and 58% of the variance in logMAR acuity, respectively. In the current study, logPFWc and logVOTF showed significantly higher correlations ($p < 0.05$) with logMAR acuity in KC eyes compared to normal eyes. Both Yoon and Schoneveld reported absolute acuity values whereas here we report changes in acuity. Shi et
al. (2013) showed that logVSX accurately predicts the Δ logMAR acuity in highly aberrated eyes, which is in agreement with the current results.

Potential limitation and considerations

Eight metrics provide statistically indistinguishable correlations between normal and KC WFEs. For these metrics, no change in metric value (Figure 4) predicts a significant gain in acuity ($p < 0.05$), suggesting there are variations in IQ due to other factors not captured by wavefront sensing (scatter, photoreceptor sampling limits), limiting gains in acuity. For example, Ravikumar et al. (2011) showed that, on average, there are six (range five to eight) just noticeable differences in metric value before a one-line loss in acuity.

Lack of agreement for worse acuities between normal and KC WFEs over the full data set may impose a limitation on the use of metrics when the WFEs are very large, for example, in eyes with very poor refractive surgery outcomes or advanced KC. Caution is warranted when applying the model in a predictive manner, outside the range used to generate the model.

The maximal benefit for any correction strategy may not be experienced immediately due to long-term adaptation to the habitual retinal IQ in highly aberrated eyes (Pesudovs, 2005; Rouger, Benard, Gatinel, & Legras, 2010; Sabesan & Yoon, 2009, 2010) and in normal eyes (Artal et al., 2004; George & Rosenfield, 2004; Mon-Williams, Tresillian, Strang, Kochhar, & Wann, 1998; Sawides, de Gracia, Dorrorsoro, Webster, & Marcos, 2011; Sawides et al., 2010; Sawides et al., 2012; Webster, Georgeson, & Webster, 2002). Here the emphasis is on visual performance following an abrupt change in WFE structure given the prior WFE and/or VA. We do not consider changes that are dynamic or gradual in nature, such as increase in aberration with age (Applegate, Donnelly, Marsack, Koenig, & Pesudovs, 2007; Artal, Ferro, Miranda, & Navarro, 1993; Guirao, Redondo, & Artal, 2000; McLellan, Marcos, & Burns, 2001), accommodation (Cheng, Barnett, et al., 2004), tear-film dynamics (Li & Yoon, 2006; Montes-Mico, Alio, & Charman, 2005; Wang et al., 2009), or pupil size (Winn, Whitaker, Elliott, & Phillips, 1994).

Conclusions

In summary, our findings suggest that changes in eight IQ metrics (logNS, logVSX, logLIB, logSRX, logEW, logVSMTF, logAreaMTF, and logSTD) are highly correlated, accurate, and precise in estimating the changes in VA that are produced by WFEs of either normal or keratoconic eyes. Further investigation of the predictive ability of these metrics for a given individual will establish whether these metrics can be used as an objective surrogate to predict change in logMAR VA resulting from therapy (e.g., wavefront-guided customized refractive corrections, intraocular lens designs, and novel contact lens designs for the highly aberrated eye).

Keywords: image quality, image-quality metrics, visual acuity, normal and keratoconic wavefront errors

Acknowledgments

This work is supported by NIH/NEI R01 EY08520 (RAA), NIH/NEI R01 EY019105 (RAA), NIH/NEI P30 EY 07551 (Core Grant), Navy contract N0025910 P1354 (RAA), and Borish Endowment (RAA). In addition, the authors thank Dr. Larry N. Thibos, Hope Queener and her team of programmers for developing the GUI for metric calculations, and Dr. Scott B. Stevenson for help implementing gamma correction and other programming for the image display. The authors also thank Dr. Edwin J. Sarver for programming the precompensation of the VA charts and Dr. Jason Porter for assisting in the construction of the apparatus.

Commercial relationships: RAA has patent interest in retinal image quality metrics through the University of Houston. No other author has a proprietary interest in any material or method mentioned.

Corresponding author: Ayeswarya Ravikumar.
Email: ayeswarya22@gmail.com.
Address: College of Optometry, University of Houston, Houston, TX, USA.

References


Sabesan, R., Jeong, T. M., Carvalho, L., Cox, I. G., Williams, D. R., & Yoon, G. (2007). Vision improvement by correcting higher-order aberra-
tions with customized soft contact lenses in keratoconic eyes. Optics Letter, 32(8), 1000–1002.


### Appendix

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSw</td>
<td>RMS WFE computed over the whole pupil (microns)</td>
</tr>
<tr>
<td>PV</td>
<td>Peak-to-valley difference (microns)</td>
</tr>
<tr>
<td>RMSs</td>
<td>RMS wavefront slope computed over the whole pupil (arcmin)</td>
</tr>
<tr>
<td>PFWc</td>
<td>Pupil fraction when critical pupil is defined as the concentric area for which RMSw &lt; criterion (e.g., wavelength/4)</td>
</tr>
<tr>
<td>PFWt</td>
<td>Pupil fraction when a “good subaperture satisfies the criterion PV &lt; criterion (e.g., wavelength/4)</td>
</tr>
<tr>
<td>PFS</td>
<td>Pupil fraction when a “good subaperture satisfies the criterion horizontal slope and vertical slope are both &lt; criterion (e.g., 1 arcmin)</td>
</tr>
<tr>
<td>Bave</td>
<td>Average blur strength (diopters)</td>
</tr>
<tr>
<td>PFCt</td>
<td>Pupil fraction when a “good” subaperture satisfies the criterion Bave &lt; criterion (e.g., 0.25 D)</td>
</tr>
<tr>
<td>PFCc</td>
<td>Pupil fraction when critical pupil is defined as the concentric area for which Bave &lt; criterion (e.g., 0.25 D)</td>
</tr>
<tr>
<td>D50</td>
<td>Diameter of a circular area centered on peak that captures 50% of the light energy (arcmin)</td>
</tr>
<tr>
<td>EW</td>
<td>Equivalent width of centered PSF (arcmin)</td>
</tr>
<tr>
<td>SM</td>
<td>Square root of second moment of light distribution (arcmin)</td>
</tr>
<tr>
<td>HWHH</td>
<td>Half width at half height (arcmin)</td>
</tr>
<tr>
<td>CW</td>
<td>Correlation width of light distribution (arcmin)</td>
</tr>
<tr>
<td>SRX</td>
<td>Strehl ratio computed in spatial domain</td>
</tr>
<tr>
<td>LIB</td>
<td>Light in the bucket: the percentage of total energy falling in an area defined by the core of a diffraction-limited PSF</td>
</tr>
<tr>
<td>STD</td>
<td>Standard deviation of intensity values in the PSF, normalized to diffraction-limited value</td>
</tr>
<tr>
<td>ENT</td>
<td>Entropy of the PSF</td>
</tr>
<tr>
<td>NS</td>
<td>Neural sharpness: weighting the PSF with a bivariate Gaussian weighting function normalized to the diffraction-limited case</td>
</tr>
<tr>
<td>VSX</td>
<td>Visual Strehl ratio computed in the spatial domain is an inner product of the PSF with a neural weighting function normalized to the diffraction-limited case</td>
</tr>
<tr>
<td>SFCMTF</td>
<td>Spatial frequency cutoff of radially averaged modulation-transfer function (rMTF)</td>
</tr>
<tr>
<td>AREA MTF</td>
<td>Area of visibility for rMTF (normalized to diffraction-limited case)</td>
</tr>
<tr>
<td>SFcOTF</td>
<td>Spatial frequency cutoff of radially averaged optical-transfer function (rOTF)</td>
</tr>
<tr>
<td>AREAO TF</td>
<td>Area of visibility for rOTF (normalized to diffraction-limited case)</td>
</tr>
<tr>
<td>SROT F</td>
<td>Strehl ratio computed in frequency domain (OTF method)</td>
</tr>
<tr>
<td>VOTF</td>
<td>Volume under OTF normalized by the volume under MTF</td>
</tr>
<tr>
<td>VNOTF</td>
<td>Visual Strehl ratio computed in frequency domain (OTF method)</td>
</tr>
<tr>
<td>VNMT F</td>
<td>Volume under neurally weighted OTF, normalized by the volume under neurally weighted MTF</td>
</tr>
<tr>
<td>SRMT F</td>
<td>Strehl ratio computed in frequency domain (MTF method)</td>
</tr>
<tr>
<td>VSMT F</td>
<td>Visual Strehl ratio computed in frequency domain (MTF method)</td>
</tr>
</tbody>
</table>