

Understanding the Impact of Individual Perceived Image Quality Features on Visual Performance

Julia S. Benoit^{1,2}, Ayeswarya Ravikumar¹, Jason D. Marsack¹, and Heather A. Anderson³

¹ University of Houston, College of Optometry, Houston, TX, USA

² Texas Institute for Measurement, Evaluation, and Statistics, University of Houston, Houston, TX, USA

³ The Ohio State University College of Optometry, Columbus, OH, USA

Correspondence: Julia S. Benoit, University of Houston, College of Optometry and Texas Institute for Measurement, Evaluation, and Statistics, 4811 Calhoun Rd, Houston, TX 77204, USA. e-mail: Julia.Benoit@times.uh.edu

Received: April 16, 2019

Accepted: November 8, 2019

Published: April 15, 2020

Keywords: visual performance; image quality; visual quality; wavefront measure; Down syndrome

Citation: Benoit JS, Ravikumar A, Marsack JD, Anderson HA. Understanding the impact of individual perceived image quality features on visual performance. *Trans Vis Sci Tech.* 2020;9(5):7. <https://doi.org/10.1167/tvst.9.5.7>

Purpose: This study aimed to quantify the impact of blur, contrast, and ghosting on perceived overall image quality (IQ) as well as resultant predicted visual acuity, utilizing simulated acuity charts from objective refraction among eyes of individuals with Down syndrome (DS).

Methods: Acuity charts were produced, simulating the retinal image when applying 16 different metric-derived spherocylindrical refractions for each eye of 30 adult patients with DS. Fourteen dilated adult observers (normal vision) viewed subsets of logMAR acuity charts displayed on an LCD monitor monocularly through a unit magnification 3-mm aperture telescope. Observers rated features blur, ghosting, and contrast on 10-point scales (10 = poorest) and overall IQ on a 0- to 100-point scale (100 = best) and read each chart until five total letters were missed (logMAR technique). Mixed modeling was used to estimate feature influence on overall perceived IQ and relative acuity (compared with an unaberrated chart), separately.

Results: Perceived IQ spanned the entire scale (mean = 59 ± 22) and average reduction in relative acuity was two lines (0.2 ± 0.14 logMAR). Perceived blur, ghosting, and contrast were individually correlated with overall IQ and relative acuity. Blur, contrast, and ghosting exert unique effects on overall perceived IQ ($P < 0.05$). Blur ($b = -.009$, $P < 0.001$) and ghosting ($b = -.003$, $P < 0.001$) influence relative acuity over and beyond their effects on overall IQ ($b = .001$, $P < 0.0001$) and contrast.

Conclusions: Objectively identified refractions would ideally provide high contrast, low blur, and low ghosting. These data suggest that blur and ghosting may be given priority over contrast when improving acuity is the goal.

Translational Relevance: Findings may guide objective refraction in clinical care.

Introduction/Purpose

Visual image quality is a function of both the optical and neural components associated with seeing.¹ Visual acuity (VA) and contrast sensitivity serve as performance-based measurements of an eye's visual image quality. VA is the most commonly used measure of visual performance during clinical assessment and identification of an optical correction, with the goal of refracting a patient to provide the best possible VA. However, it has been reported that VA may not be the best indicator of visual quality in the presence

of higher-order aberrations (e.g., patients with keratoconus) or the best method to assess performance of ocular improvement procedures (monovision contact lenses and corrective surgeries, among others).²⁻⁶

The influence of blur on VA has typically been studied under two different paradigms, either by placing spherical/cylindrical lenses in front of the eye⁷⁻¹⁵ or by simulating blur through computerized images.¹⁶⁻²² The advantage of the latter is the ability to study predefined interactions between lower- and higher-order aberrations in a controlled manner. Spherical blur primarily causes reduction in contrast. Asymmetrical blur, a consequence of residual

astigmatism or irregularities in the corneal surface (e.g., keratoconus or refractive surgery patients) in addition to reduction in contrast, also leads to doubling/tripling of images (i.e., ghosting).²³ Thus, ocular aberrations can reduce image quality by causing blur, doubling of images (ghosting), or reduced contrast (described further in the [Methods section](#)). While many studies have reported relationships between these features and visual performance either in isolation or under varying designs (experimental, simulation) and for various reasons (i.e., postsurgical), and while one study reported on the additive model of blur and contrast,²⁴ none have analyzed how these disparate subjective image quality features simultaneously conspire to influence visual performance.

Lately, prescriptions identified through optimization (an objective metric-based method that removes the subjective component from the exam) have been used to investigate the impact of higher-order aberrations on retinal image quality.^{25–27} In these studies, an objective refraction based on wavefront error measurement is determined through the optimization of a specific metric (a mathematical model meant to approximate some aspect of optical or visual performance). Each metric may emphasize different features of image quality (e.g., blur, contrast, and ghosting) over others.

Improving visual performance in the presence of objectively defined refractions requires an understanding of how the optimization affects the visual percept in patients. Failure to consider this relationship may leave a patient struggling with features that hinder the percept, even as visual performance seems to be improved (“I can read the letters, but they just aren’t clear.”). This is especially important in vulnerable populations such as individuals with Down syndrome (DS). Individuals with DS may have more difficulty communicating their perceived image quality, which may result in a refraction that is less than optimal. The purpose of this study was to understand the relative impact of perceived image quality features (blur, contrast, ghosting), both separately and simultaneously, on overall perceived image quality and visual acuity.

Methods

This study was approved by the University of Houston Committee for the Protection of Human Subjects and adhered to the tenets of the Declaration of Helsinki. Adult participants without DS provided informed consent. Parental or guardian permission was obtained for individuals with DS, as well as participant assent.

This simulation study utilizes a database of wavefront aberrations among eyes of adults with DS. These wavefront errors (WFEs) are the basis upon which charts are generated to be rated to answer research questions concerning visual performance and perceived image quality. Below we discuss how the WFE measurements were obtained from the individuals with DS, how the charts were then blurred using the WFE data to simulate the retinal image of the individuals with DS, and how the simulated charts were read and graded by typical observers.

Patients

As a part of a larger primary study, adult patients with DS were recruited and screened for nystagmus and ocular pathology that would prohibit obtaining wavefront measurements, resulting in 30 participants for this study. Each participant underwent a complete ophthalmic examination, including clinical refraction utilizing techniques that were best suited to each individual (e.g., retinoscopy and subjective refraction, where possible) as well as autorefractometry on both eyes.

Individuals were then dilated with 1% tropicamide and 2.5% phenylephrine, and WFE was collected using the Discovery System (Innovative Visual Systems, USA) 30 minutes after dilation. Five WFE measurements were taken for each eye (measured at least 6 mm pupil diameter) with the goal of capturing images with minimal missing WFE spots and no reflection in the spot image.

Prescription Derivation and Chart Simulation

The resulting WFE data were used to calculate normalized Zernike coefficients over each eye’s average habitual pupil size (obtained by recordings from an infrared video camera during visual acuity testing prior to dilation). A brute-force algorithm was used to apply >25,000 spherocylindrical correction combinations to the uncorrected WFE of each eye, and these data were used to calculate 30 different image quality metrics (as described by Thibos et al.²⁶). The refraction that optimized each individual metric was identified and termed the optimal refraction for that metric.

The residual WFE for each optimized refraction was then used to generate acuity charts simulating predicted retinal image quality in the presence of each optimized refraction. Methods used to generate the simulated acuity charts have been previously reported.²⁸ To reduce the number of conditions evaluated, charts from metrics that were identified as consistently yielding poor-quality acuity charts were

eliminated, as well as any charts from redundant refractions whereby multiple metrics identified the same optimized refraction for a given eye. This methodology has been described in detail elsewhere.²⁸ Experimental acuity charts were also created based on applying refractions derived from WAM-5500 Grand Seiko (RyuSyō Industrial Co., Hiroshima, Japan) autorefractometer (AutoRef) measures, each patient's habitual refraction (based on lensometry of their current glasses or plano refraction if they presented unaided), and a theoretical zeroing of all lower-order aberrations (LOAZ). Thus, for this study, a total of 669 unique experimental charts were viewed (out of a possible 1140) and graded (including unique derived prescriptions + habitual + AutoRef + LOAZ).

Measurement of Visual Performance and Chart Reading Session

Fourteen typically sighted observers without DS, with best-corrected distance visual acuity of at least 20/20 and free of ocular and systemic pathology, were recruited to read the simulated acuity charts and rate perceived image quality.

The simulated acuity charts were randomly divided into 10 chart sets, where each set was derived from a block of at least three patients with DS. Each set of charts was viewed by 5 of 14 raters during an individual session, representing an incomplete cross-classification of rater and patients with DS with partial nesting with five ratings for every simulated chart, allowing assessment of interrater measurement error. During each session, a rater underwent dilation with 1% tropicamide and 2.5% phenylephrine. Thirty minutes after dilation, each rater viewed logMAR simulated acuity charts displayed on a high-contrast gamma-corrected LCD monitor monocularly through a unit magnification telescope with a 3-mm aperture and best spectacle correction in place. During each rater session, an unconvolved “clear” chart was randomly inserted in each set to collect a baseline acuity and baseline perceived image quality.

Simulated Acuity Chart Visual Performance and Gradings

Subjective measures collected and rated for each chart under each condition included assessment of perceived image quality (overall and individual features) and visual performance (logMAR visual acuity).

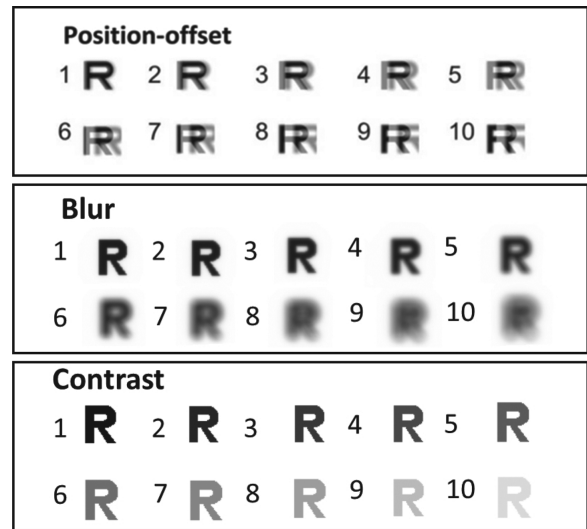


Figure 1. Scales of ghosting (quantified in this study as position offset as in the work by Kollbaum et al.²⁹), blur, and contrast used by observers to rate each chart. Observers rated their perceived level of each of these three features on 10-point scales.

Perceived Image Quality

Overall perceived image quality was rated for each chart. Overall subjective assessments were rated based on a 0- to 100-point scale, where 0 represented the worst and 100 the best.

Individual image quality features were rated for each acuity chart. Perceived ghosting (quantified as position offset), perceived blur, and perceived contrast were rated on a 10-point scale (10 = poorest) (Fig. 1). The grading scale shown in Figure 1 was displayed in between each chart presentation for reference and to encourage consistency of applying the scale by the observer. For this study, position offset, a dimension of ghosting, was primarily judged by raters to represent the perceived ghosting feature, and these two terms will be used synonymously for purposes of this study. The position offset scale was adopted from the “position offset series” of the ghosting scale found in Kollbaum et al.²⁹ displaying 10 horizontally offset ghost images from the focused “R” in steps corresponding to 3.2-arcmin increments with added blur of 0.50 D and relative intensity (contrast) of 50. The *blur* scale was derived to range from 1 to 10 representing 0.03 to 3.0 microns of Zernike defocus in increments of 0.3-micron steps for each level of blur. The *contrast* scale, ranging from 1 to 10, represents 100% to 10% in 10% decrement steps for each level of contrast (1 = 100% contrast; 10 = 10% contrast).

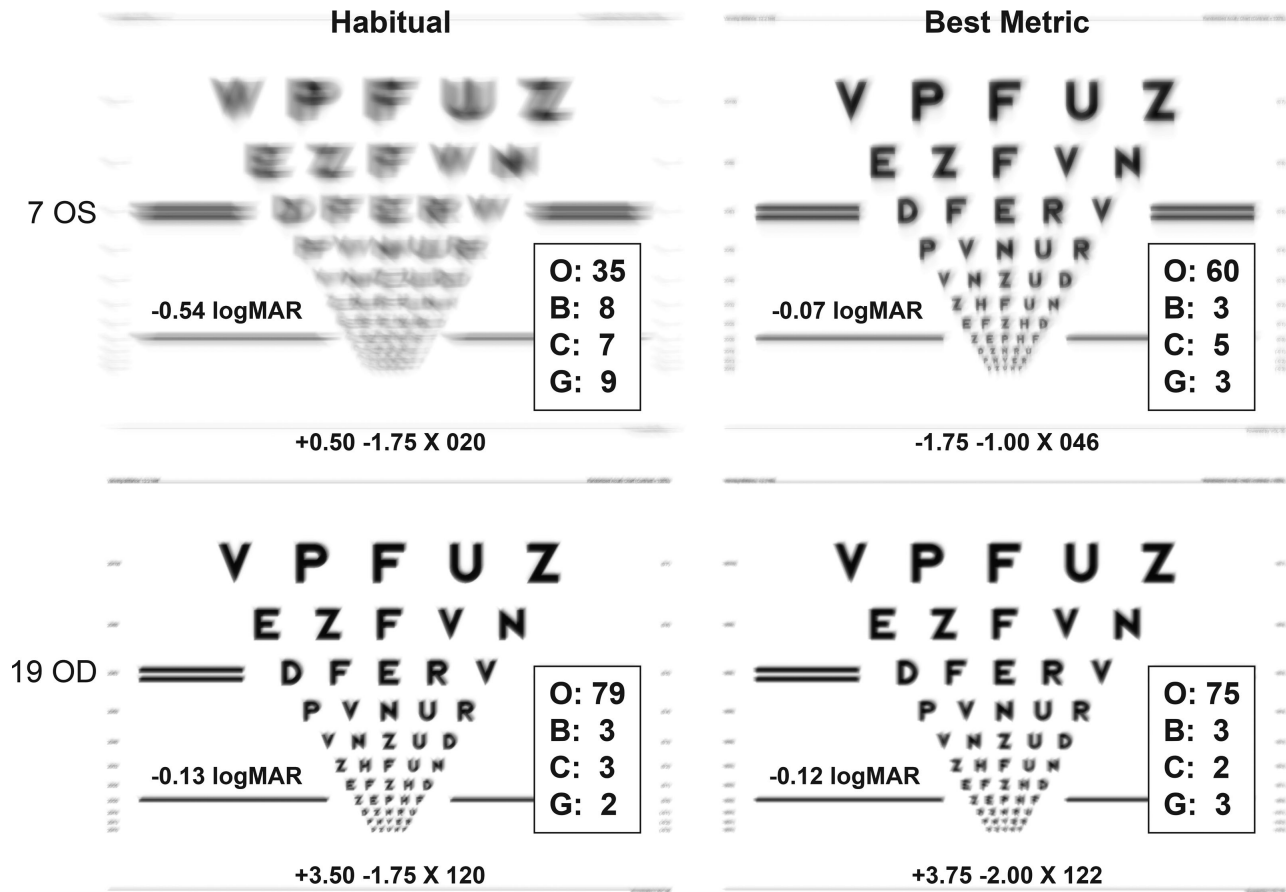


Figure 2. Example of a rater’s gradings of image quality (O, overall; B, blur; C, contrast; and G, ghosting) of four simulated charts from habitual and best metric prescriptions from eyes of two individuals with DS. These two examples show variability in the benefit of the optimized method in the simulated charts as well as the range of potential chart ratings for two conceivable cases: (1) case where habitual refraction provides poorer perceived image quality than that identified by the metric conditions based on the optimized method (7 OS) and (2) the case where the two refractions yield charts with similar acuity (19 OD).

Visual Performance Outcome

Visual acuity measured in logMAR was recorded using letter-by-letter scoring previously reported.³⁰ As the largest line presented on each chart was 0.7 logMAR, we expected that observers would read at least better than 0.7 logMAR. Relative acuity was defined as the difference between acuity achieved on the clear chart and acuity achieved on the aberrated or experimental chart, with a negative value representing a loss in acuity relative to the clear chart (i.e., relative acuity = clear chart acuity – acuity from experimental chart).

In Figure 2, we illustrate an example of simulated charts read by a control observer for eyes originating from two patients with DS under a specific condition. The purpose of this example is to show the range of potential self-reported ratings for two conceivable cases: (1) case where the habitual refraction provides

poorer perceived image quality than the refraction determined based on the optimized method (7 OS) and (2) the case where the two refractions yield acuity charts of similar perceived quality (19 OD).

A rater who read the habitual derived chart from DS-7OS with measured VA of -0.54 logMAR self-reported perceived blur = 8, ghosting = 9, contrast = 9, and image quality = 35. The same rater read the chart simulating the best metric-derived refraction on the same patient with measured VA of -0.07 logMAR self-reported perceived blur = 3, ghosting = 3, contrast = 5, and overall perceived image quality = 60. For a different eye, DS-19 OD, the rater had similar VA in reading the habitual and best metric-derived refraction, and subjective measurements varied slightly. We also illustrate an example of a rating from an unaberrated chart (Fig. 3) to demonstrate that even in these cases (no aberration convolved into the chart), the best scores were not reported by this rater.

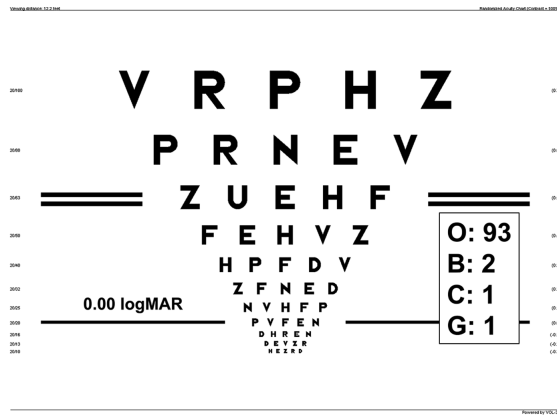


Figure 3. Example of a single rater's subjective grading of a clear (unaberrated) chart in terms of overall perceived image quality (maximum or best = 100), blur (minimum or least = 1), contrast (maximum or highest = 1), and ghosting (minimum or least = 1).

Statistical Analysis

Relative acuity and overall rating each represent an overall measure of image quality. Both are treated as continuous variables. The distribution of the overall rating scale (0–100) demonstrates a continuous form, covering a range of perceived good and poor quality. Given the number of response categories for the overall rating of perceived image quality as well as the study design, treating the 100-point ratings as continuous preserves the order of the scale and is appropriate given the psychometric literature around ordinal responses with many categories. Based on psychometric work that has been done fitting latent variable models to polytomous data with many categories, measures with more than seven categories are best treated as continuous rather than polytomous.^{31–33} As mentioned prior, we have an incomplete (or partially crossed) cross-classified data structure, meaning that observations do not occur in all combinations of raters and conditions (metrics used to generate the charts). Mixed models with random effects were used in SAS software to model image quality while accounting for the within-subject correlation of the repeated measurements and cross-classification of the data. Multiple models were evaluated, consisting of two sets: unconditional variance component models and conditional models building upon the unconditional models, including predictors of individual perceived image quality features. Residual analyses were performed to assess the assumptions of our primary model.

Reliability

The unconditional models were evaluated for each image quality measure (both visual acuity and perceived image quality) to estimate the proportion of variation attributable to raters or reliability. Due to the data structure, potential variability for each outcome may be contributed by eye, rater, and condition in addition to unobservable or unexplainable variation. Since the purpose of this analysis is to explain how subfeatures of image quality explain both overall image quality and the method by which this occurred in random raters, we wanted to ensure that the judgments of the raters were reliable or that the proportion of total variation was not highly attributable to the method. To determine the proportion of variance explained by rater, the proportion of variance attributable to rater from the decomposed variance estimates relative to the total variance was calculated and reported along with reliability measurements.

Bivariate Correlations of Residuals

Bivariate correlations were computed and scatterplots constructed between overall perceived image quality, image quality features (blur, contrast, ghosting), and relative acuity after removing the correlation due to the rater, subject, and eye within subject. Standardized scores for each perceived measure were calculated and stacked such that all observations were on the same scale. We modeled the standardized scores using a mixed modeling approach. Fixed effects included the indicator of the outcome and eye while the random effects were specified at the rater, subject, and eye within subject levels.

Influence of Perceived IQ on VA

To determine the association between each feature (ghosting, blur, and contrast) and overall perceived image quality, we used a linear mixed regression controlling for fixed effects of condition with random and repeated components for DS patient, eye, condition, and rater. We used a similar approach for outcome visual acuity.

All statistical analyses were performed using SAS 9.4 software (SAS Institute, Inc., Cary, NC, USA) and R (v. 3.6.2).³⁴

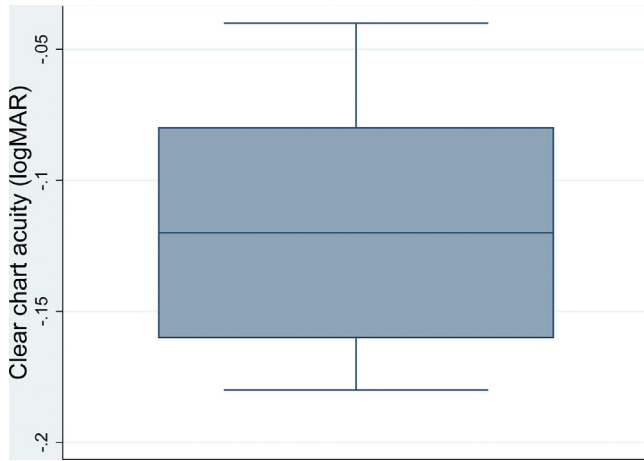


Figure 4. Distribution of clear chart acuity (logMAR) among all raters.

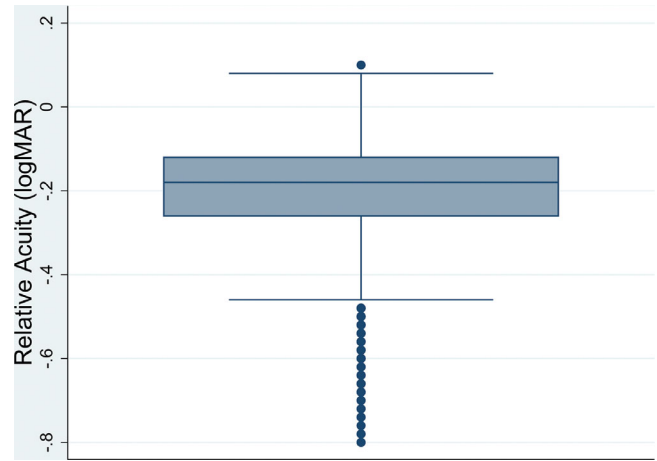


Figure 5. Distribution of relative acuity (logMAR) among all charts read. Relative acuity = clear chart acuity – acuity from experimental charts.

Results

Description of Study Data and Sample Statistics

A total of 14 observers read the charts (median 71 charts read per observer, 25th percentile = 60; 75th percentile = 343). The median number of conditions read for each eye among each DS patient was 11 (unique metric-derived prescriptions + habitual + AutoRef + LOAZ). Among all acuities obtained from viewed charts, a total of 61 (1.8%) were excluded from analysis due to measurements worse than 0.7 logMAR. On average, performance of clear chart logMAR acuity was -0.1 (0.06) logMAR (mean [SD]) (Fig. 4). For the entire sample of refractions that were evaluated (including metric based, autorefraction, habitual, and LOAZ), on average, two lines of acuity were lost (-0.2 [0.14] logMAR) (Fig. 5). In other words, the residual aberrations associated with individuals with DS resulted in a loss of two lines on average. Among all clear charts observed, perceived image quality features of blur, contrast, and ghosting were rated the optimal score most frequently (median/mode = 1) (Table 1)

with a median overall rating of 97. This sample yielded a mixture of perceived good and bad charts. Among all experimental charts read, median (25th percentile, 75th percentile) overall perceived image quality ratings were 65 (45, 78) with a mean (SD) of 59.9 (21.9) (Table 1). Median ratings for the perceived image quality features for blur, ghosting, and contrast, among experimental charts, were 4, 2, and 4, respectively (Table 1).

Reliability of Rater Grading Method

To evaluate the between-rater method of chart grading, we estimated from variance component models the proportion of total variance attributable between raters, using each measure of image quality as the dependent outcome separately (Table 2). Generally, we found that the judgments made by raters across all measures demonstrated high reliability (intraclass correlation coefficients all above 80%). Judgments of contrast and overall measures were found to be slightly less consistent across raters relative to other measures.

Table 1. Summary of Rated Perceived Image Quality Performance Among Clear and Experimental Charts

Characteristic	Clear Charts	Experimental Charts
	Median (25th, 75th Percentile)	Median (25th, 75th Percentile)
Overall perceived image quality	97 (93, 100)	65 (45, 78)
Perceived blur	1 (1, 1)	4 (3, 5)
Perceived ghosting*	1 (1, 1)	2 (2, 4)
Perceived contrast	1 (1, 1)	4 (2, 5)

* Measured by position offset scale.

Table 2. Variance Component and Reliability of Judgments on Measures Calculated From Results of Multilevel Models Using Each Measure of Image Quality as Dependent Outcome, Controlling for Metric Type

Measure of Image Quality	Total Variance	Variance Due to Rater	Proportion of Variance Due to Rater	ICC
Relative acuity	0.0198	0.001	0.044	0.956
Overall	453.62	69.15	0.152	0.848
Blur	3.45	0.18	0.052	0.947
Contrast	3.78	0.67	0.178	0.822
Ghosting*	4.59	0.25	0.054	0.946

The models divided the variance of each outcome into components contributed at the rater, DS patient, eye within subject level, and residual level. Displayed is the total variance, variance (and proportion) explained by rater, and reliability measurements. ICC, intraclass correlation coefficient.

*Measured by position offset scale.

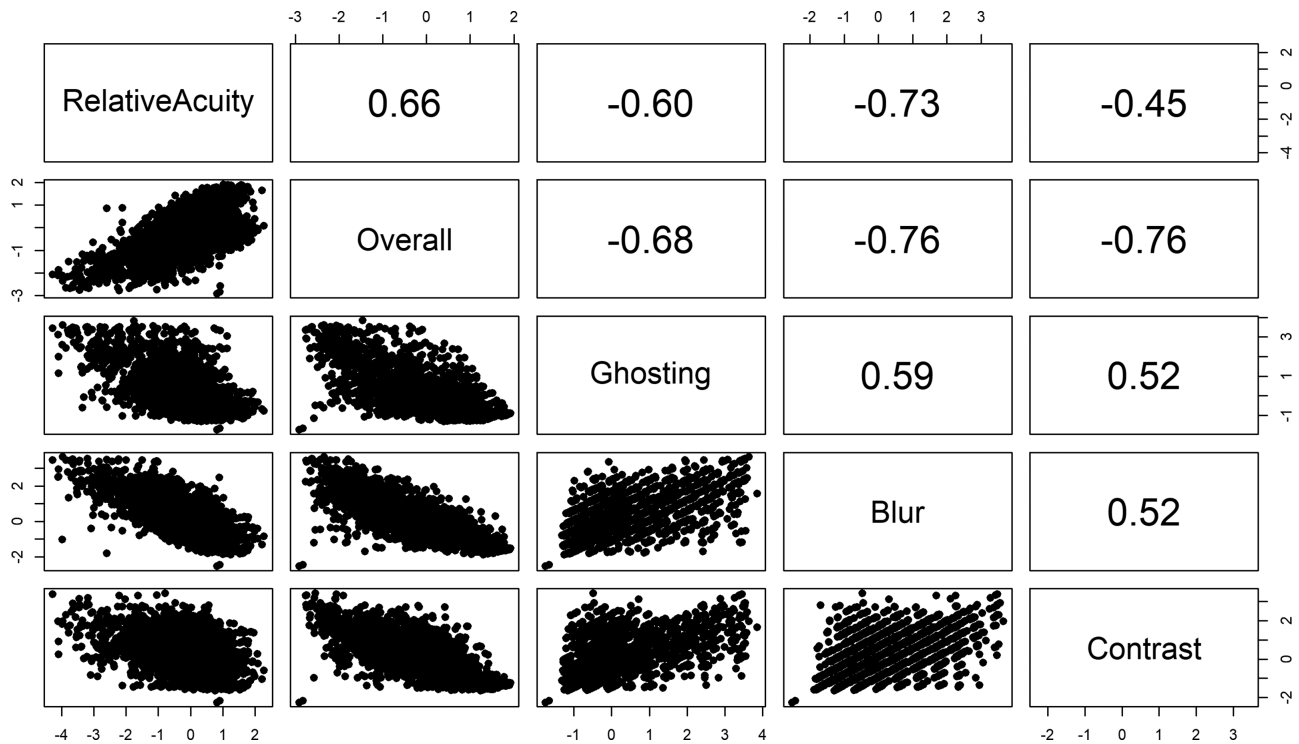


Figure 6. Bivariate scatterplots (lower) with associated correlations (upper) between overall perceived image quality, image quality features (blur, contrast, ghosting), and relative acuity after accounting for rater, metric, subject, and eye within subject. Scores were standardized to achieve the same scale and modeled via a mixed modeling approach. For example, relative acuity was positively correlated with the overall perceived score (0.66) (upper panel) and the actual data points for this correlation are reflected in row 2, column 1 (lower panel).

Specifically, we found that reliability of ratings of contrast due to raters resulted in 82% reliability and that the overall measure resulted in a reliability measure of 85%. Blur (97%) and ghosting (94%) yielded more reliable measures across raters, having had the least amount of variance explained due to rater of the total variation.

Bivariate Relationships

Figure 6 describes the bivariate relationships among residuals of visual performance and individual percept calculated from mixed models upon normalization of measures. Relative acuity was significantly correlated with all rated image quality measures: blur

Table 3. Maximum Likelihood Parameter Estimates of Fixed Effects from Mixed Regression Models that Quantify (1) the Influence of Perceived Image Quality Ratings on Relative Acuity and (2) the Influence of Perceived Contrast, Blur, and Ghosting (Measured as Position Offset) on Overall Perceived Image Quality After Controlling for the Mean Effects of Metric Type

Characteristic	Estimate	Standard Error	95% Confidence Intervals	P Value
(1) Relative acuity*				
Overall perceived image quality	0.001	0.00014	0.0006 to 0.0012	<0.0001
Perceived blur	-0.012	0.001	-0.014 to -0.005	<0.0001
Perceived ghosting	-0.003	0.001	-0.001 to -0.009	0.0053
Perceived contrast	0.002	0.001	-0.000 to 0.004	0.1285
(2) Overall perceived image quality**				
Perceived blur	-4.37	0.13	-4.64 to -4.11	<0.0001
Perceived ghosting	-2.17	0.18	-2.40 to -1.94	<0.0001
Perceived contrast	-2.56	0.14	-2.84 to -2.28	<0.0001

*Difference between clear chart acuity and experimental derived chart acuity.

**Sensitivity analysis for model 2: bootstrap means and confidence intervals: blur (-4.72, -3.78), ghosting (-2.50, -1.77), and contrast (-2.93, -1.89).

Note: Additionally, deviance χ^2 statistics for the overall effects of blur, ghosting, and contrast were assessed in unconstrained models, with predictors entered categorically, comparing reduced and full models (i.e., $\Delta_{-2LL} = [-2LL(\text{Reduced}) - 2LL(\text{Full})] \sim \chi^2$ (9)). Estimated overall effects for model 1 were as follows: blur ($\Delta_{-2LL} = 92.9$), ghosting ($\Delta_{-2LL} = 25.4$), and contrast ($\Delta_{-2LL} = 10.1$). Estimated overall effects for model 2 were as follows: blur ($\Delta_{-2LL} = 880.4$), ghosting ($\Delta_{-2LL} = 318.2$), and contrast ($\Delta_{-2LL} = 357.7$). All except the effect of contrast in model 1 surpass the .05 critical value, $\chi^2(9) = 16.92$ ($P < 0.05$) (i.e., $P < 0.05$ when $\Delta_{-2LL} > 16.92$).

(-0.73), contrast (-0.45), ghosting (-0.60), and overall (0.66). Overall score was significantly correlated with characteristics of image quality: blur (-0.76), ghosting (-0.68), and contrast (-0.76).

Multiple Regression-Primary Analyses

In this section, we report the impact of perceived image quality features (blur, contrast, and ghosting) simultaneously on overall perceived image quality and visual performance (relative acuity), as measured by visual acuity. Our results from mixed multiple regression models, after statistically controlling for all features and metric type, blur, contrast, and ghosting, were found to have unique effects on overall perceived image quality: blur ($b = -4.37$; 95% confidence interval [CI]: -4.64 to -4.11), contrast ($b = -2.56$; 95% CI: -2.84 to -2.28), and ghosting ($b = -2.17$; 95% CI: -2.40 to -1.94) ($P < 0.05$) (Table 3). The estimated effect size for the influence of perceived blur on overall perceived image quality was approximately twice that of ghosting and contrast. For visual performance (relative acuity), contrast did not exert an effect over and above the other two ($P = 0.1285$). Blur ($b = -0.012$, $P < 0.001$) and ghosting ($b = -0.003$, $P = 0.0053$) influenced relative acuity over and beyond their effects on image quality ($b = 0.001$, $P < 0.0001$).

As a sensitivity analysis to model 2, bootstrapped 95% confidence intervals for blur (-4.72 to -3.78), ghosting (-2.50 to -1.77), and contrast (-2.93 to -1.89) indicate that our estimates do not depend on the validity of normal theory for the data. We obtained similar results when assessing overall effects of blur, contrast, and ghosting entered as categorical predictors in an unconstrained model using χ^2 deviance statistics to compare full and reduced models (Table 3).

Discussion

The image quality metrics utilized in objective refraction techniques have been correlated with visual performance.²¹ It was unclear whether the optimized factor corresponds to features most influential on the patients' perceived or preferred viewing visual quality or level of visual acuity and the extent to which visual acuity associates with perceived image quality. As would be expected, visual acuity decreased independently with reported ratings of increased blur, reduced contrast, and increased ghosting. This research suggests that perceived image quality is moderately correlated with VA and not necessarily in agreement.

Prior to this study, it was expected that due to the subjective nature of how one perceives the quality of an image, visual acuity performance may not necessarily reflect individual perceived image quality. However, blur tends to impact both subjective judgment about an image and VA. Perceived blur was most highly correlated with VA and had the most impact on VA over and beyond contrast and ghosting. In the presence of two ghosted images, the natural tendency is for the brain to choose a higher-quality image and ignore the lower-quality image to achieve better performance. Consequently, one can learn to ignore the ghosting component, but as our study shows, this may not be the case with blur. This finding seemed to be corroborated anecdotally in follow-up interviews with the raters after viewing numerous combinations of charts. Our results also indicate that perceived blur may play a larger role in overall perceived image quality than other image quality features. This finding can be supported by the literature³⁵ and explained by the asymmetric blur as a consequence of residual astigmatism or irregularities in the corneal surface caused by ocular aberrations among individuals with higher-order aberrations (e.g., keratoconus or refractive surgery patients).²³

The finding here that blur and ghosting may be given priority over contrast when the goal is to improve acuity is contingent upon our use of high-contrast letters (as is ubiquitous in clinical testing). If the level of blur and ghosting had been held constant, and low-contrast letters had replaced the high-contrast letters used in the current work, the relative importance of blur, ghosting, and contrast may have been altered.

Strengths

The findings in this study can be generalized to the normal population in studies where acuity is impacted by contrast,^{24,36} blur,^{24,37} and/or ghosting.³⁸ For example, the relationship between visual quality and visual acuity fails when ghosting is introduced (based on simulation).²¹

It should be noted that the sample from which the wavefront aberrations were obtained and used to answer our research question relating perceived image quality and performance was a sample of convenience from an ongoing project and not critical to this simulation study. However, a strength of this study was that this sample offered a range of aberration profiles, as individuals with DS are commonly found to have elevated aberrations when compared with the typical population but not so severely elevated as might be seen in a population of individuals with overt keratoconus. Resultant retinal image quality was expansive while still

resulting in charts that yielded observer acuity of 0.8 logMAR or better.

These three factors have not been simultaneously studied and analyzed when the levels of each factor are derived from real eyes (in a nonexperimental setting). This study demonstrates the necessity of a multivariate approach in considering all features working together, even if the research may have emphasis only on one isolated percept. We have modified previously reported rating scales (e.g., Kollbaum et al.²⁹) with an analytical strategy and design for assessing image quality features simultaneously in a real-world setting that is clinically beneficial to not only studying questions among specific populations but also understanding what may contribute to disagreement between objective and perceived image quality among typical eyes. This strategy can be adopted and modified by others for future work when the goal is to assess features of image quality in visual performance.

This study provides new insights to the objective refraction literature, particularly in residual error of the simulated charts from individuals who may most benefit from this type of prescription. Our study shows that although different metrics may optimize specific factors associated with retinal image quality, that those where residual blur remains may be most unfavorable to improving image quality.

Given that objective refraction is designed to optimize aspects of image quality, to ultimately improve VA, we wanted to quantify the extent of patients' perceived optimal visual quality, despite improved acuity. We have shown that VA is moderately correlated with overall perceived image quality, suggesting some residual error in optimized refraction. Although we did not stratify by levels of aberration, it may be informative to study this relationship among those with low versus high levels of aberration to determine capabilities of objective refraction to predict VA.

After controlling for metric type, blur, contrast, and ghosting all have unique effects on overall perceived image quality, with blur seemingly to be more influential on the overall perception of the quality of the image. Past research has focused on reporting relationships between these features and visual perception either in isolation or under varying designs (experimental, simulation). These results highlight the value in studying the multivariate simultaneous impact of these features on overall perceived image quality.

Limitations

One potential limitation of this study is that we measured acute image quality without any adaptation to the chart during assessment. Thus, there was no

opportunity for neural adaptation to blur, as would likely occur in an individual with prolonged exposure to a given set of aberrations. This limitation could have impacted or overestimated blur.³⁵

While, on the one hand, our results are generalizable in the sense of the measurement evaluation being used, we were limited to objective refractions of eyes from patients with DS who may not represent the typical population. A study among a typical seeing population may provide a more general model for prediction of visual performance from ratings of percept.

Note that these results are based on monocular vision. Monocular vision may have provided somewhat reduced estimates of VA and perceived image quality given that when both eyes are combined, the VA would have likely improved due to binocular summation.

Conclusions

Refractions that are objectively identified would ideally have high contrast and low blur and ghosting, but in individuals with elevated aberrations (e.g., individuals with DS), compromises may be needed. This study provides a better understanding of the percept of image quality as it relates to visual performance and may be useful in the future pursuit of a personalized objective refraction approach. Finally, these data suggest that blur and ghosting may be given priority over contrast when the goal is to improve acuity.

Acknowledgments

The authors thank Hope Queener for developing the Spectacle Sweep software used in this study.

Supported by National Institutes of Health R01 EY024590 and P30 EY07551.

Disclosure: **J.S. Benoit**, None; **A. Ravikumar**, None; **J.D. Marsack**, None; **H.A. Anderson**, None

References

1. Applegate RA, Thibos LN, Hilmantel G. Optics of aberroscopy and super vision. *J Cataract Refract Surg.* 2001;27:1093–1107.
2. Williams D, Yoon G, Porter J, Guirao A, Hofer H, Cox I. Visual benefit of correcting higher order aberrations of the eye. *J Refract Surg.* 2000;16:S554–S559.
3. Green DG. Regional variations in the visual acuity for interference fringes on the retina. *J Physiol (Lond).* 1970;207:351–356.
4. Kohonen T. Measuring vision in refractive surgery. *J Cataract Refract Surg.* 2001;27:1897–1898.
5. Lawless MA. Refining visual quality assessment in refractive surgery. *J Cataract Refract Surg.* 1999;25:1031–1032.
6. McLeod SD. Beyond Snellen acuity: the assessment of visual function after refractive surgery. *Arch Ophthalmol.* 2001;119:1371–1373.
7. Bedell HE, Patel S, Chung ST. Comparison of letter and vernier acuities with dioptric and diffusive blur. *Optom Vis Sci.* 1999;76:115–120.
8. Campbell F, Green D. Optical and retinal factors affecting visual resolution. *J Physiol (Lond).* 1965;181:576–593.
9. Charman WN. Effect of refractive error in visual tests with sinusoidal gratings. *Br J Physiol Opt.* 1979;33:10–20.
10. Herse PR, Bedell HE. Contrast sensitivity for letter and grating targets under various stimulus conditions. *Optom Vis Sci.* 1989;66:774–781.
11. Ho A, Bilton SM. Low contrast charts effectively differentiate between types of blur. *Am J Optom Physiol Opt.* 1986;63:202–208.
12. Smith G, Jacobs RJ, Chan CD. Effect of defocus on visual acuity as measured by source and observer methods. *Optom Vis Sci.* 1989;66:430–435.
13. Thorn F, Schwartz F. Effects of dioptric blur on Snellen and grating acuity. *Optom Vis Sci.* 1990;67:3–7.
14. Tucker J, Charman WN. The depth-of-focus of the human eye for Snellen letters. *Am J Optom Physiol Opt.* 1975;52:3–21.
15. Walsh G, Charman W. The effect of defocus on the contrast and phase of the retinal image of a sinusoidal grating. *Ophthalmic Physiol Optics.* 1989;9:398–404.
16. Burton G, Haig N. Effects of the seidel aberrations on visual target discrimination. *J Opt Soc Am A.* 1984;1:373–385.
17. Chen L, Singer B, Guirao A, Porter J, Williams DR. Image metrics for predicting subjective image quality. *Optom Vision Sci.* 2005;82:358–369.
18. Cheng X, Bradley A, Thibos LN. Predicting subjective judgment of best focus with objective image quality metrics. *J Vision.* 2004;4:7.
19. Marsack JD, Ducros MG, Parekh SH, Thomsen SL, Rylander HG, Milner TE. Imaging the primate retina using polarization-sensitive optical

- coherence tomography. *Biomed Eng.* 2001;4245:158–164.
20. Ravikumar A, Applegate RA, Shi Y, Bedell HE. Six just-noticeable differences in retinal image quality in 1 line of visual acuity: toward quantification of happy versus unhappy patients with 20/20 acuity. *J Cataract Refract Surg.* 2011;37:1523–1529.
 21. Ravikumar A, Sarver EJ, Applegate RA. Change in visual acuity is highly correlated with change in six image quality metrics independent of wavefront error and/or pupil diameter. *J Vision.* 2012;12:11.
 22. Applegate RA, Ballentine C, Gross H, Sarver EJ, Sarver CA. Visual acuity as a function of Zernike mode and level of root mean square error. *Optometry Vision Sci.* 2003;80:97–105.
 23. Atchison DA, Guo H, Charman WN, Fisher S. Blur limits for defocus, astigmatism and trefoil. *Vision Res.* 2009;49:2393–2403.
 24. Johnson CA, Casson EJ. Effects of luminance, contrast, and blur on visual acuity. *Optometry Vision Sci.* 1995;72:864–869.
 25. Iskander DR. A subjective refraction-based assessment of image quality metric. *Photonics Lett Poland.* 2011;3:150–152.
 26. Thibos LN, Hong X, Bradley A, Applegate RA. Accuracy and precision of objective refraction from wavefront aberrations. *J Vision.* 2004;4:9.
 27. Hastings GD, Marsack JD, Nguyen LC, Cheng H, Applegate RA. Is an objective refraction optimised using the visual Strehl ratio better than a subjective refraction? *Ophthalmic Physiol Optics.* 2017;37:317–325.
 28. Ravikumar A, Benoit JS, Marsack JD, Anderson HA. Image Quality Metric Derived Refractions Predicted to Improve Visual Acuity Beyond Habitual Refraction for Patients With Down Syndrome. *Transl Vis Sci Techn.* 2019;8:20.
 29. Kollbaum PS, Dietmeier BM, Jansen ME, Rickert ME. Quantification of ghosting produced with presbyopic contact lens correction. *Eye Contact Lens.* 2012;38:252–259.
 30. Ravikumar A, Benoit JS, Morrison KB, Marsack JD, Anderson HA. Repeatability of monocular acuity testing in adults with and without down syndrome. *Optometry Vision Sci.* 2018;95:202–211.
 31. Babakus E, Ferguson CE, Jöreskog KG. The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *J Market Res.* 1987;24:222–228.
 32. Dolan CV. Factor analysis of variables with 2, 3, 5 and 7 response categories: a comparison of categorical variable estimators using simulated data. *Br J Math Stat Psychol.* 1994;47:309–326.
 33. Johnson DR, Creech JC. Ordinal measures in multiple indicator models: a simulation study of categorization error. *Am Sociol Rev.* 1983;48:398–407.
 34. R Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. 2019. <https://www.R-project.org/>.
 35. Webster MA, Georgeson MA, Webster SM. Neural adjustments to image blur. *Nat Neurosci.* 2002;5:839.
 36. Barteselli G, Gomez M, Doede A, et al. Visual function assessment in simulated real-life situations in patients with age-related macular degeneration compared to normal subjects. *Eye.* 2014;28:1231.
 37. Chung ST, Jarvis SH, Cheung S. The effect of dioptric blur on reading performance. *Vision Res.* 2007;47:1584–1594.
 38. Carkeet A, Chang YA, Chang TY, et al. Simulated image doubling and visual acuity: effects of doubling magnitude, orientation, and ghost image intensity. *Ophthalmic Physiol Opt.* 2019;39:86–93.