

Different aberrations raise contrast thresholds for single-letter identification in line with their effect on cross-correlation-based confusability

Laura K. Young

Department of Physics, Durham University, Durham, UK
Department of Experimental Psychology,
University of Oxford, Oxford, UK



Gordon D. Love

Department of Physics, Durham University, Durham, UK



Hannah E. Smithson

Department of Experimental Psychology,
University of Oxford, Oxford, UK
Department of Psychology, Durham University,
Durham, UK



We previously showed that different types of aberration defocus, coma, and secondary astigmatism affect reading performance via different mechanisms. In this paper, we show the contrary result that, for identification of isolated letters, the effects of rendering different types of aberration can be described by a single cross-correlation-based metric. Aberrations reduce the effective resolution of an optical system, quantified by the high-frequency fall-off of the modulation transfer function. They additionally cause spatial-frequency-dependent phase and contrast changes, which have a size-dependent effect on letter forms. We used contrast threshold as our performance measure, instead of distance acuity, to separate the effects of form alterations from those of resolution limits. This measure is additionally appropriate in comparing single-letter-based performance to reading at a fixed distance. The relationship between a cross-correlation-based measure of letter confusability and performance was the same for all three types of aberration. For reading, we had found a different relationship for coma than for defocus and secondary astigmatism. We conclude that even when two tasks—letter identification and reading—use the same component stimulus set, the combination of multiple letters in a reading task produces functional differences between the effects of these aberrations that are not present for isolated letters.

Introduction

The distortions to the retinal image caused by ocular aberrations impair letter identification (see Applegate, Sarver, & Khemsara, 2002, for example) and associated tasks, such as reading (Young, Liversedge, Myers, Love, & Smithson, 2011). Our work aims to tease apart the mechanisms of impairment by measuring the effects of different types of aberration on both letter identification and reading. In the present paper, we report the letter identification data. The differences we measure are important because a description of image quality cannot be derived without considering both the nature of the distortion and the nature of the task.

Imperfections in the optical components of the eye distort the wavefront of incident light, causing blurring of the retinal image. In ophthalmology, the distorted two-dimensional wavefront is typically quantified by decomposing it into a series of orthogonal components called Zernike modes, which are shown in Figure 1. The low-order modes, namely defocus and astigmatism, can be corrected with spectacles, contact lenses, or laser eye surgery, but there remain higher-order aberrations that affect vision (Liang, Grimm, Goelz, & Bille, 1994). These higher-order aberrations can be corrected using a more complicated optical surface, such as a phase plate (Yoon, Jeong, Cox, & Williams, 2004) or a deformable mirror (Liang & Williams, 1997; Liang, Williams, &

Citation: Young, L. K., Love, G. D., & Smithson, H. E. (2013). Different aberrations raise contrast thresholds for single-letter identification in line with their effect on cross-correlation-based confusability. *Journal of Vision*, 13(7):12, 1–14, <http://www.journalofvision.org/content/13/7/12>, doi:10.1167/13.7.12.

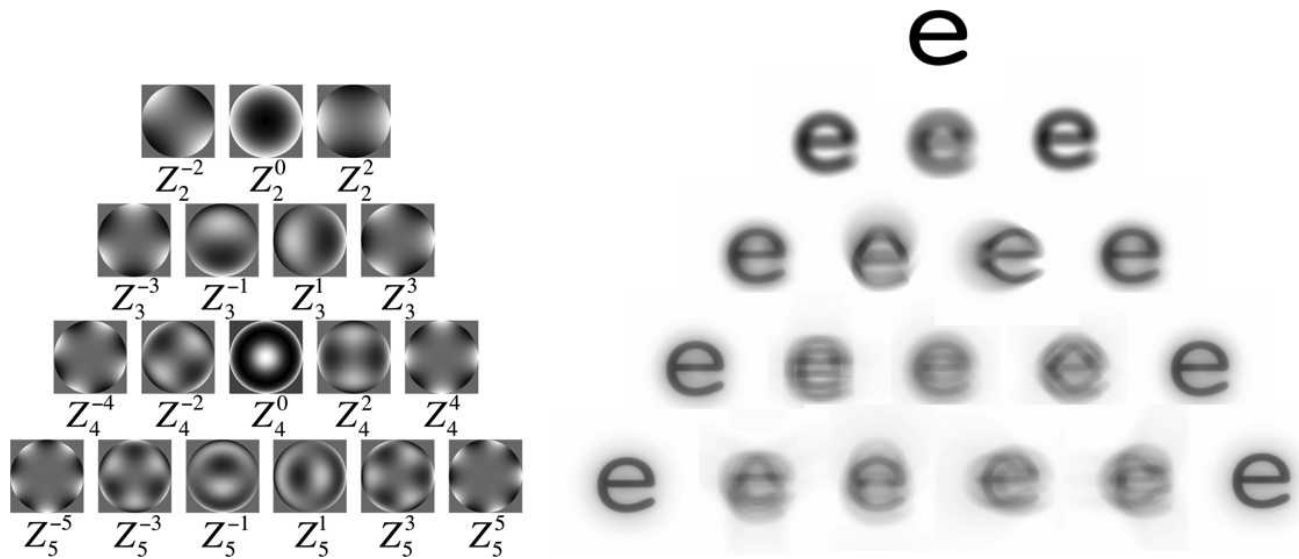


Figure 1. The Zernike pyramid showing wavefront modes (left image) from second radial order to fifth radial order. The greyscale corresponds to the wavefront deviations from a plane wave. The orders are given as Z_n^m (Z_n^m where n is the radial order and m is the angular order), and the aberrations we have tested are represented by Z_2^0 (defocus), Z_3^1 (coma), and Z_4^2 (secondary astigmatism). Spectacles can be used to correct for defocus and primary astigmatism, $Z_2^{\pm 2}$, and they also change the amount of spherical aberration, Z_4^0 , whereas higher-order aberrations (those with radial order greater than two) can only be corrected with complicated optical surfaces, such as the deformable mirror used in an adaptive optics system. The image on the right shows the effect of $0.6 \mu\text{m}$ rms of each Zernike mode on a letter “e” spanning 1° .

Miller, 1997). It is important to understand how these higher-order aberrations affect visual performance so that we can better interpret wavefront measurements and decide if a correction would be worthwhile. Higher-order aberrations may be modified during conventional laser eye surgery (see Yamane et al., 2004, for example), and it is valuable to know which aberrations are particularly disruptive to visual tasks so that efforts can be made to minimize them.

Previous studies on the effect of higher-order aberrations on letter-based visual acuity have tested the size threshold for recognition. Using logMAR acuity charts rendered with a constant root-mean-square (rms) amplitude of the first 12 Zernike modes (second to fourth radial order), Applegate et al. (2002) showed that Zernike modes concentrated near the center of the pyramid (see Figure 1) caused more disruption to visual acuity than those at the edges. Furthermore, they showed that the relationship between the amplitude of the aberration and the decrease in logMAR acuity was linear with slopes that were greater for modes near the center of the pyramid (Applegate, Ballentine, Gross, Sarver, & Sarver, 2003). This result had been confirmed by other authors using similar methods for stimuli generated computationally (Cheng, Bradley, Ravikumar, & Thibos, 2010) and optically (Chen, Singer, Guirao, Porter, & Williams, 2005; Rocha, Benard, & Legras, 2007; Rouger, Benard, & Legras, 2010; Zhao et al., 2009). Although it is well understood that the strength of the effect an

aberration has on visual acuity varies with the type of aberration, it is important to determine why some aberrations are more detrimental than others. Figure 1 (right panel) shows each Zernike mode from second to fifth radial order applied to the letter “e”, and it is clear that some modes have a strong effect on the form of the letter, and therefore its legibility, that is different than the rather simple effect of low-pass filtering that is familiar from Gaussian blur transformations. Kwon and Legge (2013) recently showed that as spatial resolution is reduced by low-pass filtering images of letters, subjects require a higher contrast for letter recognition. They suggested that human performance in such conditions was primarily determined by the information content of the stimuli, which degrades as the blur is increased. How then does this relate to the degradation produced by ocular aberrations? One measure of the performance of an optical system is its optical transfer function (OTF). Contrast changes are represented by the amplitude of the OTF (the modulation transfer function, MTF), and the phase changes are represented by the argument of the OTF (the phase transfer function, PTF). Aberrations that cause spurious resolution, which results from spatial phase changes in the image, have a particularly large impact. Ravikumar, Bradley, and Thibos (2010) tested the effects of phase errors on the recognition of single letters, letter clusters, and facial expressions under four conditions. These conditions were phase rectification (i.e., PTF set to zero), low-pass filtering (i.e., MTF set

to zero for frequencies above the first minimum), low-pass filter plus negative lobes (i.e., including the phase-reversed frequencies), and low-pass filter plus positive lobes (i.e., including the frequencies with the correct phase). They showed that, with sufficient contrast, 180° phase changes can significantly impair visual acuity (size threshold for identification) whereas phase changes of less than 180° had a smaller impact. This implies that even-ordered aberrations, which only cause 0° and 180° phase shifts, are more likely to be detrimental to visual performance than odd-ordered aberrations that cause phase changes between these values. These phase changes are likely to have a particularly detrimental effect when they occur at spatial frequencies that are important for letter identification (see the next paragraph). They additionally showed that the improvement in visual acuity with phase rectification when viewing letter clusters was larger than with single letters. The letter clusters they used were spaced such that crowding effects were likely to occur, and so this improvement is indicative of an effect of crowding associated with phase changes in the stimulus.

There are several possible ways to measure visual performance in a letter-identification task, and we consider here our rationale for choosing to measure contrast thresholds rather than size thresholds. Changing the size of the stimulus alters the spatial frequency composition of the image. We know that an ocular aberration causes spatial frequency-dependent changes in the retinal image, quantified by its OTF. Changing the size of the letter also changes the effect of the aberration on an image of that letter. Because the MTF typically decreases as a function of increasing spatial frequency, a given rms amplitude of aberration will cause a larger degradation to the image quality of small letters than large letters. Both the MTF and PTF can be nonmonotonic, raising the possibility of complex interactions between the effect of an aberration and the size of the image (see Appendix 1). So while size-thresholds relate to natural viewing conditions in which object viewing distances change, we have chosen to maintain a fixed letter size and to modulate the contrast to avoid introducing size-dependent effects. This was particularly important in this work as we compare the results to empirical measures of reading performance in which letter size also remained constant.

The spatial frequencies present in the stimulus are an important consideration for object recognition. With letters in particular, Solomon and Pelli (1994) showed that identification is mediated by a visual channel sensitive to a specific spatial-frequency band. The center of this band is typically around three cycles per letter and has a bandwidth of one to two octaves. Importantly, the bandwidth of this channel is not

predicted by the spatial-frequency content of the letter stimuli alone because it is narrower than that derived from an ideal-observer model and instead must be constrained by the properties of the human visual system. Later, Majaj, Pelli, Kurshan, and Palomares (2002) showed that the center frequency of this channel is selected bottom-up by the properties of the signal, not top-down by the observer. Importantly, they showed that when a band-pass filtered letter is kept at a constant size, the center frequency of the channel scales less than proportionally with the center frequency of the pass-band. This may imply that aberrations, which act to spatially filter the image, could potentially cause a shift in the channel frequency to one that may be suboptimal for the identification task.

In a recent paper, we reported the results of an experiment that tested the effects of defocus (Z_2^0), coma (Z_3^1), and secondary astigmatism (Z_4^2) on the eye movements made during reading (Young et al., 2011). We found a categorically different pattern of effects for coma compared with defocus and secondary astigmatism, suggesting that they had their effects by disrupting different mechanisms in reading. Increasing the rms amplitude of the three types of aberration caused an increase in the average fixation time (the average time a reader spends looking at a single word) and in the number of fixations. We investigated the way in which this increase depended on image degradation by considering a template-matching model of letter confusability (as described in the Methods sections and also more fully in our previous publication). The model calculates confusability defined by the average strength of cross-correlation between letters rendered with a given aberration. Cross-correlation models of letter recognition have previously shown good correlation with empirical measures of performance with degraded letters (Kwon & Legge, 2013; Watson & Ahumada, 2012). The relationship between our confusability metric and the increase in average fixation duration was the same for defocus and secondary astigmatism but different for coma, which gave a higher increase in the average fixation duration at a lower confusability value. We proposed that this difference was due to the asymmetry in the point spread function (PSF) of coma that was aligned with the horizontal lines of text and which therefore caused smearing of adjacent letters. This implies that the different effect of coma compared to defocus and secondary astigmatism may occur only with letter strings in a reading task and not with identification of single letters. It is this hypothesis that we test in this paper.

We have tested the effects of defocus, coma, and secondary astigmatism on single-letter recognition. As far as we are aware, no existing studies have measured the contrast threshold for letter recognition in the presence of these aberrations. We additionally com-

pared our results to the pattern of impairment that we found previously for reading performance and considered how well differences between aberrations were accounted for by considering image degradation as assessed via a cross-correlation metric. We found a consistent relationship between the ability to identify letters and our confusability model for all three types of aberration, which is different from the case of reading.

Introducing ocular aberrations

In this study, we introduce ocular aberrations by presenting rendered images to normal observers. So it is important both to consider the limitations of rendering in achieving a desired retinal image, which we cover in the Methods section for our specific stimuli, and the effects of visual experience with particular aberrations, which we consider here.

Adaptation to blur can profoundly affect perceived image quality, and performance can be influenced by an observer's experience viewing the world with a particular set of aberrations. In particular, adaptation to higher-order aberrations, as well as to defocus (Mon-Williams, Tresilian, Strang, Kochhar, & Wann, 1998) and to astigmatism (Sawides, Gamba, Pascual, & Marcos, 2010), is possible. Artal et al. (2004) showed that subjects experienced the greatest perceived image quality when viewing a stimulus via their own aberrations rather than a rotated version of them even though the amplitude of the wavefront error was the same. More recently, Sawides, de Gracia, Dorronsoro, Webster, and Marcos (2011) showed that a subject's perceived best focus changed proportionally with the amplitude of induced higher-order aberrations, indicating adaptation to the blur they create. However, they suggest that this demonstrates adaptation to the global amount of blur and does not test whether adaptation occurs to local changes caused by specific aberration patterns. Moreover, there remains the question of whether observers can learn the specific pattern of changes to the stimulus that are caused by aberrations. Measuring visual performance in normal subjects in the presence of added aberrations is therefore not the same as measuring performance in subjects for whom the aberration occurs naturally and has been present for a long period of time. However, aberrations cause a loss of information, and it is the association between the type of information loss and the type of performance loss that we wish to characterize. Additionally, if it is possible that participants have learned to, in some sense, compensate for their own aberrations, then the performance we measure is the best performance we can expect from our subjects.

Methods

Selection of aberration conditions

We chose to study defocus (Z_2^0), coma (Z_3^1), and secondary astigmatism (Z_4^2) from a prediction of which Zernike modes we expected to have an impact on reading as described in our previous paper (Young et al., 2011). This prediction was based on the maximum value of a cross-correlation between the images of the letters rendered with an aberration (which we used as a confusability metric) and consideration of how asymmetry in the PSF could cause phase and contrast changes to occur preferentially in the direction of text.

In this experiment, we used large letters (3 cm, subtending 1° viewed at a distance of 1.7 m, corresponding to a Snellen acuity of 20/240, or 7 mm/20 pt font at a typical reading distance of 40 cm) so that letter identification would be limited by the aberration and not by our subjects' acuity limits. As the letters were large, the amplitude of the aberration also needed to be large to have an impact equivalent to that which would be typical for letters presented at or nearer the acuity limit, such as those in a book, for example. In preliminary experiments, we found that rms amplitudes in the interval 0.5–0.9 μm (2.2–4.0 D of equivalent defocus) gave a significant variation in performance. Five amplitudes were tested in this range: 0.5, 0.6, 0.7, 0.8, and 0.9 μm .

The experimental data we report here are compared to our previously published data on reading. The two experiments were carried out using different sizes of letters, 1° for letter recognition and 0.25° for reading performance. To ensure that we tested the same range of stimulus degradation for letter recognition and for reading, we used different amplitudes of aberration in the two experiments. Because there is no simple relationship between angular image size and the effect of different amplitudes of aberration, we used normalized cross-correlation to confirm the similarity of stimuli used in the two experiments. Letters of a 1° size were rendered with 0.5, 0.6, 0.7, 0.8, and 0.9 μm of each type of aberration and compared against 0.25° letters that were rendered with a given amplitude of aberration and rescaled to 1° . The amplitude at which letters had the maximum normalized cross-correlation was taken to be the equivalent amplitude of aberration for letters of 0.25° as used in the reading performance experiment. Comparisons between letter images gave a normalized cross-correlation of at least 0.98, and Figure 2 shows that the equivalent amplitudes overlap with the actual amplitudes used. If the stimuli used in the letter recognition experiment were presented at the same size as those in the reading performance experiment, the amplitudes that should have been

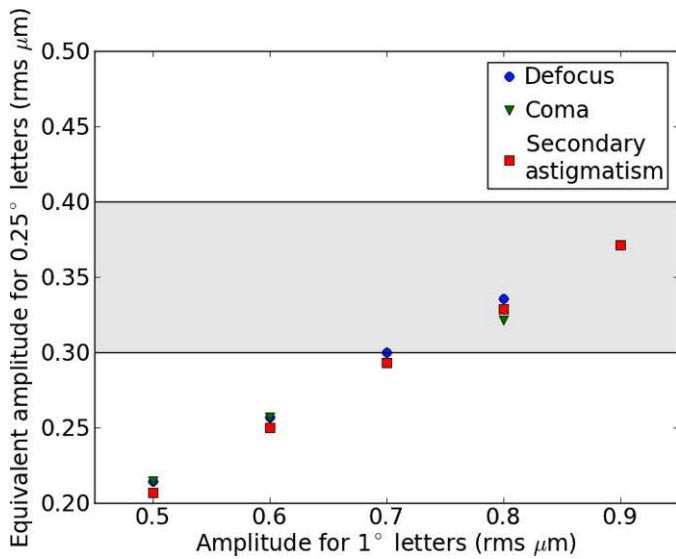


Figure 2. The equivalent amplitudes of aberration that produce the same apparent (correlation > 0.98) letter image at 0.25° as letters at 1° . Horizontal lines indicate the range of amplitudes used in the reading performance experiment (0.30 , 0.35 , and $0.40 \mu\text{m}$).

applied overlap with the amplitudes of those actually tested.

Stimuli

The stimuli used in this experiment were single lowercase letters in courier font, examples of which are given in Figure 3. In each trial, the contrast of the stimulus was determined by the observer's previous performance, according to an adaptive staircase (see below).

Subjects

Five subjects participated in this study, three male and two female, with a mean age of 32 years ($SD = 7$ years). Three subjects did not require any vision correction for viewing letters on a monitor at a distance of 1.7 m, and two used the vision correction that they normally use for viewing at this distance. All subjects were familiarized with the font prior to the experiment and were told that letters would appear with the frequency with which they normally occur in the English language.

Apparatus

Images of aberrated letters were generated using custom-written Python code that performed a convolution with the appropriate PSF. Images were presented

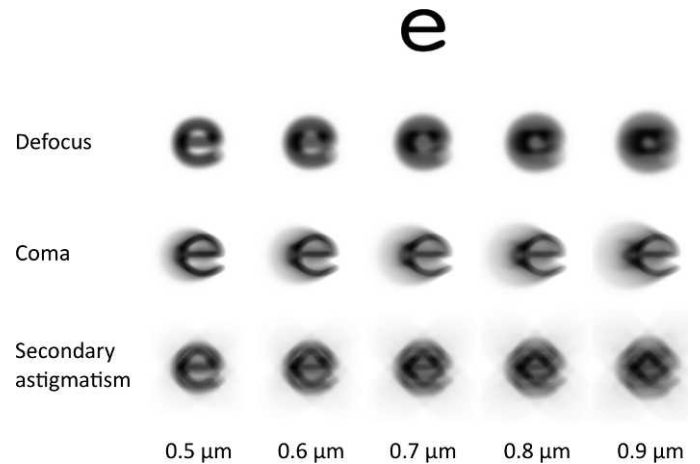


Figure 3. Examples of the stimuli used in the letter experiment. The rows show different aberration types, and the columns indicate different rms amplitudes of aberration. Images are scaled such that they represent the size used in the experiment (1°) at 40 cm.

on a CRT using a Cambridge Research Systems ViSaGe visual stimulus generator using the CRS Matlab toolboxes, and the look-up table was carefully specified to maintain a resolution of 8 bits per gun across the full range of contrast. Stimuli were presented as black text on a white background at a distance of 1.7 m and viewed through a 2.5 mm pupil. The luminance output of the CRT display was linearized (gamma corrected), and the average luminance was 27 cd/m^2 .

Although images are presented at a fixed distance, we must consider the possible effect that inducing an aberration has on a subject's accommodative state. Defocus is known to be a cue for accommodation (see Heath, 1956, for example), but higher-order aberrations can also provide such cues (Stark et al., 2009). The maximum amplitude of aberration in this experiment is $0.9 \mu\text{m}$ rms wavefront error, which is 4 D of equivalent defocus. From the work of Heath (1956), we have estimated the maximum accommodative error due to 4 D of equivalent defocus at a presentation distance of 1.7 m to be 0.1 D or $0.02 \mu\text{m}$ for a 2.5 mm pupil. We expect that at this pupil size the depth of focus is large enough for changes in accommodation to have little effect on the retinal image. Additionally, other cues to drive a correct accommodative response were available as a full-contrast sharp-edged box surrounded the stimulus.

In this experiment, as in our previous one, we test the effects of aberrations by adding them to the stimuli computationally by convolution of the aberrated PSF with the image of a letter. As we do not correct for our subject's own ocular aberrations, they will also influence the retinal image. It is known that aberrations increase with increasing pupil size, so we have minimized these effects by using an artificial pupil

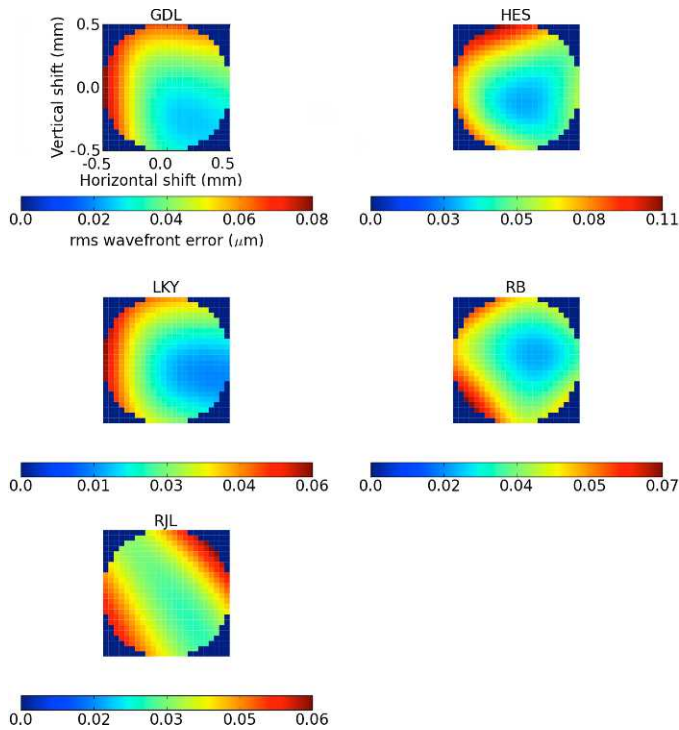


Figure 4. The rms wavefront error over a 2.5 mm pupil when decentered with respect to subjects' 3.5 mm natural pupil.

placed in front of the subject's eye (within 10 mm), thus restricting their effective pupil size to 2.5 mm. Wavefront measurements were taken of the subjects' right eyes (the same eye with which the task was performed) with a Zywave aberrometer so that we could quantify the additional effect of their ocular aberrations on the retinal image. For our subjects, the average rms amplitude of higher-order aberrations over a 2.5 mm pupil was $0.06 \mu\text{m}$ ($SD = 0.03 \mu\text{m}$), which is small compared to the amplitudes of aberration we are testing (7%–12%). We additionally calculated the confusability value associated with subjects' higher-order aberrations over a 2.5 mm pupil. The average value was 0.193 ($SD = 0.006$) comparing aberrated against aberrated letters (the corresponding confusability in the absence of aberrations is 0.189) and was 0.188 ($SD = 0.004$) comparing aberrated against nonaberrated letters (the corresponding confusability in the absence of aberrations is 0.185). The subjects' own aberrations would therefore have only a minimal effect on image quality. The position of the artificial pupil in our setup precluded real-time monitoring of the alignment of the artificial pupil with the subject's pupil. To quantify the effects of misalignments, we took samples of our subjects' measured wavefronts via a laterally shifted 2.5 mm aperture. The rms amplitude of these pupil samples was calculated as a function of the amplitude of the shift. In a previous experiment under similar lighting conditions, our subjects had a pupil diameter of 3.5 mm on average (Young et al.,

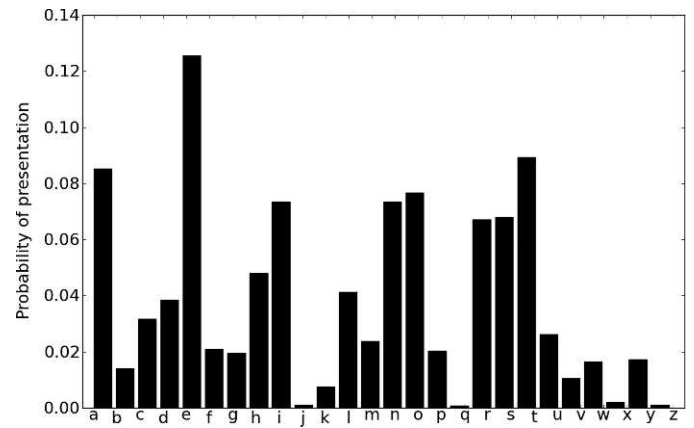


Figure 5. Probability of presentation of letters in the experiment measuring contrast threshold for letter identification, calculated using the frequency counts of Jones and Mewhort (2004).

2011), and we used this diameter for the wavefront samples. Subjects were asked to align themselves with the artificial pupil such that the image appeared sharpest, and they were held in a chin and forehead rest to maintain this position. They were also asked to position themselves such that there was no vignetting of the image, and consequently, rms values have been calculated for all shifts for which the artificial pupil did not clip the edge of the subject's pupil. The results, given in Figure 4, show that the upper limit on rms wavefront error associated with a possible misalignment between the artificial pupil and the eye's pupil ranged from 0.06 to $0.11 \mu\text{m}$. These values were only reached at the extremes of misalignment, and the rms wavefront error was smaller than this over the central range.

Procedure

A single letter was presented for 200 ms to the right eye, and the left eye was patched. This was preceded by a fixation cross that directed the subject's gaze to the correct location and then a uniform white field for 200 ms to prevent afterimages interfering with the presented stimuli. Subjects were asked to report which letter they had seen using a keyboard, and audio feedback was given to indicate whether the response had been correct. A session lasted for 30 min, during which 16 experimental conditions were tested: five amplitudes (0.5, 0.6, 0.7, 0.8, and $0.9 \mu\text{m}$ rms) of each of the three types of aberration (defocus, coma, and secondary astigmatism) and a control condition in which no aberration was applied. For each trial, a letter was chosen at random, and the probability of selection was weighted by the frequency counts of Jones and Mewhort (2004) given in Figure 5. A maximum likelihood staircase algorithm (ML-PEST; Harvey,

1997) was implemented using the Matlab Palamedes Toolbox (Prins & Kingdom, 2009), and this was used to find the Weber contrast threshold for letter recognition. The ML-PEST algorithm operates by selecting the most likely threshold value to present in each trial based on responses to previous trials. As the algorithm converges, the probability density function (PDF) for the threshold value narrows, and its center shifts to the most likely value. As the PDF is normalized by its sum, the maximum value of the narrowing PDF tends to one. The staircase ran for a minimum of 20 trials and until the maximum value of the PDF reached 0.8 (chosen by simulation to give minimal errors in threshold estimation). The algorithm converged on a threshold that corresponded to a percentage correct of 64%. The staircases for all experimental conditions were interleaved so that in any trial the type and amplitude of aberration was chosen at random. Each subject completed eight sessions, providing eight repeated measurements of performance for every combination of amplitude and type of aberration. The contrast threshold for letter recognition of a subject within a single session was determined for each experimental condition, and the threshold for the control condition was subtracted. This reduced inter-subject and intersession variability. An average was calculated over all sessions for each subject before data were averaged over all subjects.

Confusability metric

The results of this experiment are compared to our previously defined confusability metric (see selection of aberrations in Young et al., 2011). This metric was calculated based on a cross-correlation between the set of letter stimuli. Letters were compared either as aberrated letters against nonaberrated letters (to represent a subject identifying aberrated letters based on matching letters to an unaberrated template) or aberrated letters against aberrated letters (to represent a subject identifying aberrated letters by looking for differences between them). All 26 letters of the alphabet were compared against each other, producing a 26-by-26 element matrix, in which each element contains the maximum value of a cross-correlation between the relevant pair of letters. The columns of this matrix were weighted by the frequency with which letters occurred in the experiment (which was the frequency with which they occur in the English language; Jones & Mewhort, 2004). The average value of this matrix was taken as a single measure of confusability for the particular type and amplitude of aberration and at the particular letter sizes used in the two experiments. Both types of comparison are given although there are only subtle differences between them.

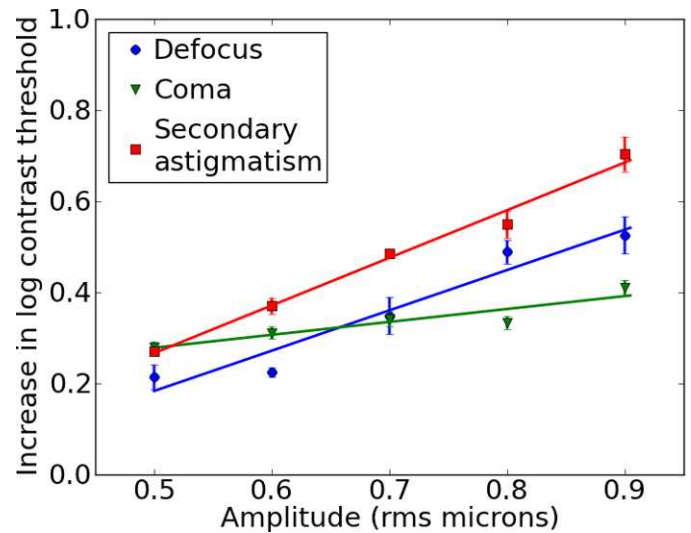


Figure 6. Increase in the contrast threshold for letter identification from the control condition as a function of aberration amplitude. The mean log contrast threshold in the control condition was -1.48 ± 0.03 . Error bars represent the standard error on the mean, calculated over five subjects.

Results

Contrast threshold for letter recognition

Figure 6 shows the increase in log contrast measured from the control condition as the amplitude of aberration was increased. The data were analyzed with a three (types of aberration) by five (amplitudes of aberration), two-way repeated measures ANOVA, and the significant effects are summarized in Table 1. All main and simple effects were significant at the 5% level, demonstrating that the increase in contrast threshold, i.e., the amount of performance loss, was affected by the type of aberration and the amplitude of aberration. Furthermore, the rate of increase in contrast threshold associated with increasing the amplitude of an aberration differed between the three types of aberration. As the number of observers in this study is small, we additionally examined individual performances, which all showed very similar trends to those in Figure 6. To

	<i>F</i>	<i>MSE</i>	<i>p</i>
Type of aberration	$F(4, 41) = 41.3$	0.211	<0.001
Amplitude of aberration	$F(4, 41) = 61.1$	0.531	<0.001
Type and amplitude interaction	$F(8, 41) = 5.1$	0.051	0.016

Table 1. Data were subjected to a 3×5 , two-way repeated measures ANOVA. The significant interaction indicates that the rate of increase in contrast threshold with increasing amplitude of aberration differed between the three types of aberration.

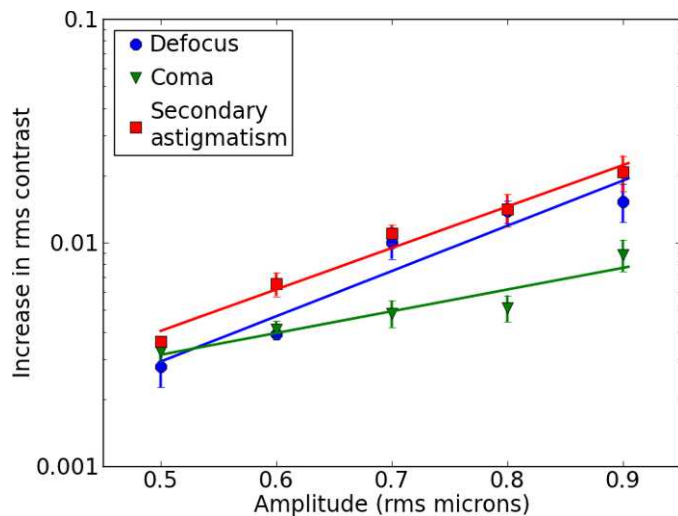


Figure 7. Increase in the rms contrast threshold for letter identification from the control condition (calculated using the letter “x”) as a function of aberration amplitude. The mean rms contrast threshold in the control condition was 0.0055 ± 0.0003 . Error bars represent the standard error on the mean, calculated over five subjects.

aid comparison with other published measures of contrast threshold for letter identification, Figure 7 additionally shows the same data expressed in terms of the increase in rms contrast from the control condition (calculated using the letter “x”, Kwon & Legge, 2013). We note that these trends are very similar to those in Figure 6.

Confusion analysis

We compared our experimental data with the confusion analysis in which letters were compared via a cross-correlation. The results are displayed in Figures 8 and 9, which show that for each type of aberration there is a relationship between our predicted measure of the confusability of letters (Young et al., 2011) and the data obtained in this experiment. Furthermore, these relationships are similar among the types of aberration. So knowing the confusability metric and the growth constant, we can predict an increase in contrast threshold for letter recognition, at least in the range of rms amplitudes we tested for these three aberrations.

Comparison to previous measures of reading performance

We recently showed that this is not the case when considering reading performance (Young et al., 2011). Figures 10 and 11 (which are replotted from Young et al., 2011) show the relationship between our predicted

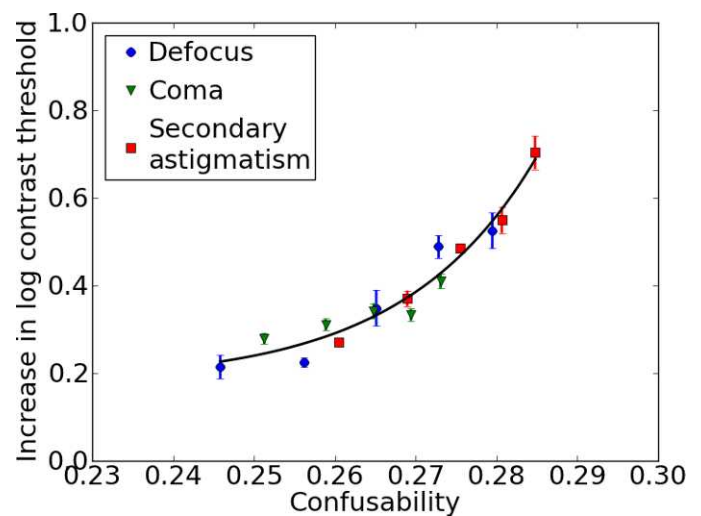


Figure 8. Relationship between the predicted confusability of letters (aberrated letters compared with unaberrated letters) and the increase in contrast threshold for letter identification (shown in Figure 6). In the control condition, the average log contrast threshold was -1.48 ± 0.03 . Error bars represent the standard error of the mean of the measurements across five subjects. An exponential curve has been fitted to the data using the least squares method, and the growth constant is 61.1 ($SSE = 0.013$). The Spearman’s Rank correlation coefficient is 0.95 ($p < 0.001$).

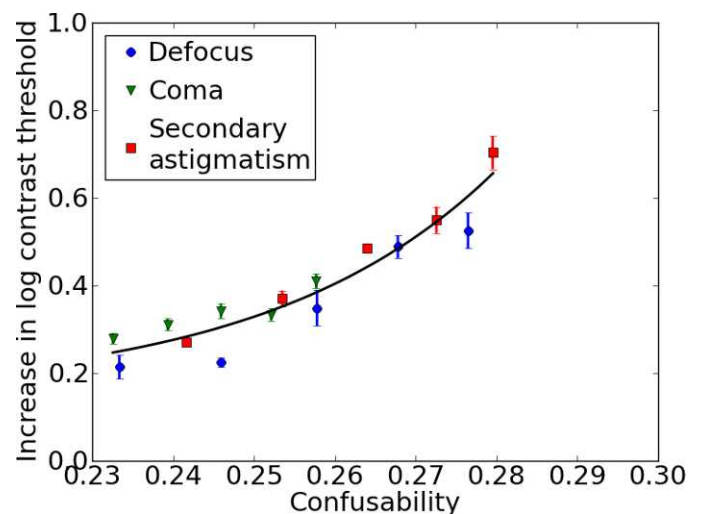


Figure 9. Relationship between the predicted confusability of letters (aberrated letters compared with unaberrated letters) and the increase in contrast threshold for letter identification (shown in Figure 6). In the control condition, the average log contrast threshold was -1.48 ± 0.03 . Error bars represent the standard error of the mean of the measurements across five subjects. An exponential curve has been fitted to the data using the least squares method, and the growth constant is 35.4 ($SSE = 0.024$). The Spearman’s Rank correlation coefficient is 0.92 ($p < 0.001$).

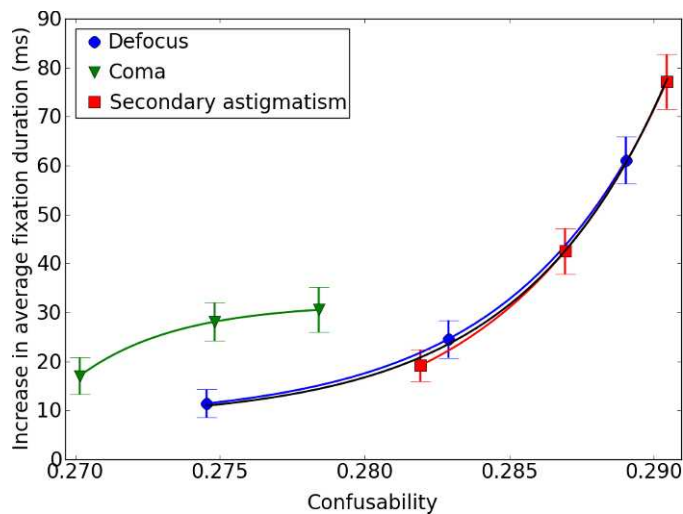


Figure 10. Relationship between the predicted confusability of letters (aberrated letters compared with aberrated letters) and the increase in average fixation duration. In the control condition, the average fixation duration was 189.5 ms ($SE = 9.8$ ms). Error bars represent the standard error of the mean of the measurements across 19 subjects. Exponential curves have been fitted to the data using the least squares method. The growth constants are 190.1 for defocus ($SSE < 0.001$), -296.4 for coma ($SSE < 0.001$), and 179.1 for secondary astigmatism ($SSE < 0.001$). The data for secondary astigmatism and defocus can also be fit with a single curve with a growth constant of 198.8 ($SSE = 4.535$) as shown in black (replotted from Young et al., 2011).

measure of confusability and the increase in average fixation duration during reading. These results show that, compared to defocus and secondary astigmatism, coma gave a different relationship between the amplitude of aberration and the increase in average fixation duration on a word. The cross-correlation metric of confusability was based either on comparisons within the set of aberrated letters (Figures 8 and 10) or on comparisons between aberrated letters and an unaberrated letter template (Figures 9 and 11). For both letter identification and reading, performance is best accounted for (smaller SSE in the curve fits) by considering comparisons within the set of aberrated letters. Therefore, the data provide some suggestion that, even with a limited amount of practice, comparisons are based on the set of stimuli to be discriminated rather than with reference to some stored representation of unaberrated letters.

Discussion

Studying the effects optical aberrations have on an image and how those effects translate into performance

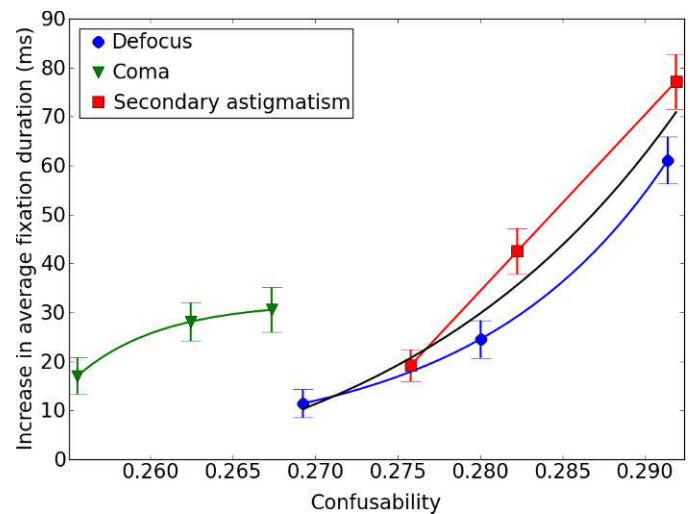


Figure 11. Relationship between the predicted confusability of letters (aberrated letters compared with unaberrated letters) and the increase in average fixation duration. In the control condition, the average fixation duration was 189.5 ms ($SE = 9.8$ ms). Error bars represent the standard error of the mean of the measurements across 19 subjects. Exponential curves have been fitted to the data using the least squares method. The growth constants are 87.2 for defocus ($SSE < 0.001$), -193.6 for coma ($SSE < 0.001$), and 0.02 for secondary astigmatism ($SSE < 0.001$). The data for secondary astigmatism and defocus can also be fit with a single curve with a growth constant of 56.7 ($SSE = 174.553$) as shown in black (replotted from Young et al., 2011).

impairments has implications for fundamental aspects of object recognition. It is important to understand the relationship between spatial phase and contrast information in a stimulus and the perceptual response that leads to the identification (or failure of identification) of that stimulus. In this paper, we have examined this relationship by studying the impairments to performance caused by aberrations, which have a quantifiable effect on phase and contrast. There are differences in the effects such phase and contrast changes have on the forms of letters, and the impact of these effects depends on the context within which those letters are presented. For letters in isolation, form changes alone can describe the change in performance as demonstrated by our cross-correlation analysis. However, when letters are presented as words, functional differences arise that can be explained by considering how the phase and contrast effects propagate through the hierarchy of processes that underlie word recognition. Such interactions have important consequences for understanding visual impairment. We speculate that one example might be the onset of reading disorders that appear as a child develops. Children are typically taught to recognize large letters, with which aberrations may have only a subtle effect, but are then expected to use these

representations to identify smaller letters, when reading a book, for example. Additionally, as the eye develops, the effects of aberrations may change, perhaps leading to difficulties in maintaining an internal representation of a letter image.

Advances in optical instrumentation for vision correction, such as the use of adaptive optics, for example, have shown that there are higher-order aberrations in the eye that can be corrected (Liang & Williams, 1997; Liang et al., 1997), albeit by more complicated means than conventional spectacles. Devices such as phase plates (Yoon et al., 2004) and contact lenses (López-Gil, Chateau, Castejón-Monchón, Artal, & Benito, 2003) have also shown potential for providing a static portable correction. The use of wavefront-guided laser refractive surgery for this purpose has also been studied (Netto, Dupps, & Wilson, 2006), but the feasibility of this is still unclear because the effect of the healing process is difficult to predict. In general, these techniques are difficult because, for example, they require accurate and stable alignment of the correcting system to the eye. It is important to study the effects of higher-order aberrations on visual performance to quantify the benefit we can hope to achieve from such a correction and to understand the types of aberration that cause the greatest impact so that a correction system can concentrate on them. We have tested defocus, coma, and secondary astigmatism in isolation, but real measurements of ocular aberrations show combinations of Zernike modes. However, we draw attention to specific features in the PSFs of the aberrations that we have chosen, e.g., the asymmetry in the PSF of coma, and our intention is to understand the effects of these features on a visual task. The relationship between such features and performance can be generalized to wavefronts measured in the normal population using a metric, such as our confusability metric.

We have tested the effects of defocus, coma, and secondary astigmatism on letter recognition, measured by the contrast threshold for correct identification. We have found that secondary astigmatism causes the largest increase in log contrast required for identification, increasing by a factor of 2.61 from the smallest to the largest amplitude of aberration. Defocus was less detrimental to letter identification over the same amplitude range, causing only a factor of 2.45 increase in contrast threshold. The effect of coma on letter recognition was even less severe (it increased by a factor of 1.47) over the same amplitude range although the increase in contrast threshold rose more sharply between 0.8 and 0.9 μm . Notably, the relationship between the amplitude of defocus and the increase in log contrast threshold appears sigmoidal rather than linear as had previously been found for size-threshold measurements (Applegate et al., 2003; Cheng et al.,

2010). On examination of Figure 3, there is a significant difference between the appearance of the letter “e” at 0.6 μm and at 0.7 μm for defocus, corresponding to phase shifts occurring in the critical band of spatial frequencies for letter recognition (approximately three cycles per letter). This is the same amplitude range over which the increase in log contrast threshold for letter identification in the presence of defocus rises sharply to give its sigmoidal shape. This difference between our results and previous results is likely due to the nature of the visual task because during a contrast threshold measurement there is no way to avoid phase shifts at critical spatial frequencies in the image (as could be accomplished by increasing the letter size for example). A small change in letter size could have a large impact on legibility if it shifted the spatial frequencies at which the phase shifts occur to a spatial frequency band that is not critical for recognition. However, at the same amplitude of aberration, any increase in contrast threshold cannot remove effects on letters from such phase shifts, making the task of recognition significantly more difficult. This effect could be worse for certain letters as, for example, a reversal of phase at critical frequencies may cause an “e” to look like a “c” or an “m” to look like an “n,” increasing the number of incorrect responses due to confusion. This would explain the greater variance in the data above 0.7 μm of defocus because subjects were tested using all letters of the alphabet chosen at random.

We suggest that these results reflect the nature of the effects these aberrations have on letters. Defocus and secondary astigmatism cause spatial phase changes that are either 0° or 180° , resulting in the contrast of certain spatial frequencies being inverted (i.e., black to white and vice versa). Coma causes phase changes that are not confined to 0° or 180° , which results in certain spatial frequencies being shifted with respect to one another. This bunching up or spreading out of spatial frequencies is less disruptive to the forms of letters than phase reversals are. Phase changes that lead to spurious resolution can create features that are inconsistent with the original form of the letter. This leads to subjects being unable to determine the letter and so guessing its identity, or worse, it results in them incorrectly identifying one letter as another. This may be particularly problematic when the contrast is low if these incorrect features have a higher contrast than the correct ones. Incorrect features with a higher contrast than correct features may contribute significantly in feature integration processes, and this can have knock-on effects for letter identification. Additionally, recall that Majaj et al. (2002) found that the spatial frequency channel that mediates letter identification is driven by the spatial frequencies in the stimulus. As these aberrations act to spatially filter the image, changing the relative distribution of power across frequencies, it

is not an unreasonable hypothesis that spurious resolution could drive the channel to an inappropriate frequency, reinforcing its influence on letter identification.

Cross-correlation-based metrics of visual performance in letter-based tasks have previously been employed for degraded letter stimuli and have shown good correlation with empirical data (Watson & Ahumada, 2012; Kwon & Legge, 2013). Here, we show that our model of confusability that is based on cross-correlations between letter images works well for predicting measures of contrast threshold for letter recognition in the presence of simulated optical aberrations. As Kwon and Legge (2013) suggested, increasing the blur in letter images increases the similarity between them, increasing their confusability. In this paper, we have tested three distinct types of blur that have very different effects on the forms of the letters and the features contained within them. Despite this, there is a consistent relationship between our metric and the empirical data, and this relationship is the same for all three types of aberration.

In a recent paper, we reported the effects of defocus, coma, and secondary astigmatism on reading performance. We found that the increase in average fixation duration on a word was related to our predicted measure of confusability and that this relationship was the same for defocus and secondary astigmatism but not for coma. In the case of coma, a lower confusability value resulted in a relatively larger increase in the average fixation duration. By studying the difference between fixation durations on common or uncommon words, we measured the effects of these aberrations on word recognition. We showed that increasing the amplitude of either defocus or secondary astigmatism caused uncommon words to be increasingly difficult to identify, suggesting that these aberrations have an impact on word recognition processes. In the case of coma, such lexical processes appear unaffected: Increasing the amplitude of the aberration increasingly impairs feature and letter identification, but this does not propagate into word identification processes. From these findings, we speculated that coma was more likely to be having its effect through an impact on the planning of eye movements or on lateral masking effects (either through physical overlap of contrast or through crowding) or both. Neither lateral masking nor eye-movement related effects should be implicated in impairments to single-letter identification. Consistent with this interpretation, when comparing our experimental results to our predicted measure of the confusability of letters we find that the effects of coma on letter identification are closely matched to those of defocus and secondary astigmatism. This reveals an interaction that arises with letter strings in a reading task but not with individual letters.

We cannot consider one single measure of visual performance as a standard measure of the effect of a monochromatic aberration on vision as effects are unlikely to be transferrable between tasks. This was pointed out by Ravikumar et al. (2010), who showed that phase rectification gives different performance increases depending on the task. As another example, Sawides et al. (2010) studied the effects of adaptive optics correction on real-life tasks and found an increase in subjective sharpness and an improvement in face, but not facial expression, recognition. It is beyond the scope of this paper to discuss the appropriateness of a clinical measure of visual acuity in relation to real-life tasks. However we can conclude that even when two visual tasks use the same component stimulus set, as with letter recognition and reading, there may be significant functional differences between them in the effects caused by different types of aberration. This is best exemplified by comparing the present results to our previous finding that defocus and secondary astigmatism specifically affect word recognition while coma does not and by our current analysis that suggests that coma aberration has a stronger effect when multiple stimuli are combined into a more complex stimulus, such as letters to words.

Conclusion

Letter identification was impaired the most by secondary astigmatism followed by defocus and was least affected by coma. The contrast threshold measures of letter identification reported here show a different pattern of results to those found with size-threshold measures (see Applegate et al., 2003, for example).

In the case of single-letter recognition, there was a common relationship between a prediction of performance, based on a cross-correlation between the distorted letter stimuli and our experimental measure of performance (increase in log contrast threshold for letter identification). This was not the case when we compared our prediction to an experimental measure of reading performance (increase in average fixation duration; Young et al., 2011). We conclude that, unsurprisingly, the effects of different types of monochromatic aberration are dependent on the visual task. This is the case even when comparing two tasks using the same stimulus set, such as letter recognition and reading, which involves a letter recognition subprocess. It is important to consider not only the nature of the distortions caused by a type of aberration and the nature of the task, but also to consider the interaction between the two. The effects of some distortions are only revealed for certain tasks (in this case, reading).

Keywords: ocular aberrations, letter recognition, spatial vision, phase distortions, physiological optics, reading

Acknowledgments

Laura K. Young acknowledges support from a Science and Technology Facilities Council Studentship. Hannah E. Smithson and Laura K. Young additionally acknowledge support from the John Fell Fund. We appreciate the discussions with Richard Myers and Rob Lee and would like to thank them for their contributions. We also wish to thank Simon Liversedge for his expertise and support with the reading study.

Commercial relationships: none.

Corresponding author: Laura K. Young.

Email: laura.young@psy.ox.ac.uk.

Address: Department of Experimental Psychology, University of Oxford, Oxford, UK.

References

- Applegate, R. A., Ballentine, C., Gross, H., Sarver, E. J., & Sarver, C. A. (2003). Visual acuity as a function of Zernike mode and level of root mean square error. *Optometry and Vision Science*, *80*, 97–105.
- Applegate, R. A., Sarver, E. J., & Khemsara, V. (2002). Are all aberrations equal? *Journal of Refractive Surgery*, *18*, S556–S562.
- Artal, P., Chen, L., Fernández, E. J., Singer, B., Manzanera, S., & Williams, D. R. (2004). Neural compensation for the eye's optical aberrations. *Journal of Vision*, *4*(4):4, 281–287, <http://www.journalofvision.org/content/4/4/4>, doi:10.1167/4.4.4. [PubMed] [Article]
- Chen, L., Singer, B., Guirao, A., Porter, J., & Williams, D. R. (2005). Image metrics for predicting subjective image quality. *Optometry and Vision Science*, *82*(5), 358–369.
- Cheng, X., Bradley, A., Ravikumar, S., & Thibos, L. N. (2010). Visual impact of Zernike and Seidel forms of monochromatic aberrations. *Optometry and Vision Science*, *87*(5), 300–312.
- Harvey, L. O. (1997). Efficient estimation of sensory thresholds with ML-PEST. *Spatial Vision*, *11*(1), 121–128.
- Heath, G. G. (1956). The influence of visual acuity on accommodative responses of the eye. *American Journal of Ophthalmology and Archives of American Academy of Optometry*, *33*(10), 513–524.
- Jones, M. N., & Mewhort, D. J. K. (2004). Case-sensitive letter and bigram frequency counts from large-scale English corpora. *Behaviour Research Methods, Instruments and Computers*, *36*, 388–396.
- Kwon, M., & Legge, G. E. (2013). Higher-contrast requirements for recognizing low-pass-filtered letters. *Journal of Vision*, *13*(1):13, 1–15, <http://www.journalofvision.org/content/13/1/13>, doi:10.1167/13.1.13. [PubMed] [Article]
- Liang, J., Grimm, B., Goetz, S., & Bille, J. F. (1994). Objective measurements of wave aberrations of the human eye with the use of a Hartmann-Shack wavefront sensor. *Journal of the Optical Society of America A*, *11*, 1949–1957.
- Liang, J., & Williams, D. R. (1997). Aberrations and retinal image quality of the normal human eye. *Journal of the Optical Society of America A*, *14*, 2873–2883.
- Liang, J., Williams, D. R., & Miller, D. T. (1997). Supernormal vision and high-resolution retinal imaging through adaptive optics. *Journal of the Optical Society of America A*, *14*, 2884–2892.
- López-Gil, N., Chateau, N., Castejón-Monchón, J. F., Artal, P., & Benito, A. (2003). Correcting ocular aberrations by soft contact lenses. *The South African Optometrist*, *62*, 173–177.
- Majaj, N. J., Pelli, D. G., Kurshan, P., & Palomares, M. (2002). The role of spatial frequency channels in letter identification. *Vision Research*, *42*, 1165–1184.
- Mon-Williams, M., Tresilian, J. R., Strang, N. C., Kochhar, P., & Wann, J. P. (1998). Improving vision: Neural compensation for optical defocus. *Proceedings of the Royal Society London B*, *265*, 71–77.
- Netto, M. V., Dupps, W., & Wilson, S. E. (2006). Wavefront guided ablation: Evidence for efficacy compared to traditional ablation. *American Journal of Ophthalmology*, *141*, 360–368.
- Prins, N., & Kingdon, F. A. A. (2009). *Palamedes: Matlab routines for analyzing psychophysical data*. Available from <http://www.palamedestoolbox.org>.
- Ravikumar, S., Bradley, A., & Thibos, L. N. (2010). Phase changes induced by optical aberrations degrade letter and face acuity. *Journal of Vision*, *10*(14):18, 1–12, <http://www.journalofvision.org/content/10/14/18>, doi:10.1167/10.14.18. [PubMed] [Article]
- Rocha, K. M., Benard, Y., & Legras, R. (2007). Effect of monochromatic induced aberrations on visual

performance with real-life tasks under adaptive optics correction. *Journal of Refractive Surgery*, 26(8), 953–959.

Rouger, H., Benard, Y., & Legras, R. (2010). Effect of monochromatic induced aberrations on visual performance measured by adaptive optics technology. *Journal of Refractive Surgery*, 26(8), 578–587.

Sawides, L., de Gracia, P., Dorronsoro, C., Webster, M., & Marcos, S. (2011). Adapting to blur produced by ocular high-order aberrations. *Journal of Vision*, 11(7):21, 1–11, <http://www.journalofvision.org/content/11/7/21>, doi:10.1167/11.7.21. [PubMed] [Article]

Sawides, L., Gamba, E., Pascual, D., & Marcos, S. (2010). Visual performance with real-life tasks under adaptive-optics ocular aberration correction. *Journal of Vision*, 10(5):19, 1–12, <http://www.journalofvision.org/content/10/5/19>, doi:10.1167/10.5.19. [PubMed] [Article]

Solomon, J. A., & Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature*, 369, 395–397.

Stark, L. R., Kruger, P. B., Rucker, F. J., Swanson, W. H., Schmidt, N., Hardy, C., et al. (2009). Potential signal to accommodation from the Stiles-Crawford effect and ocular monochromatic aberrations. *Journal of Modern Optics*, 56(20), 2203–2216.

Watson, A. B., & Ahumada, A. J. (2012). Modelling acuity for optotypes varying in complexity. *Journal of Vision*, 12(10):19, 1–19, <http://www.journalofvision.org/content/12/10/19>, doi:10.1167/12.10.19. [PubMed] [Article]

Yamane, N., Miyata, K., Samejima, T., Hiraoka, T., Kiuchi, T., Okamoto, F., et al. (2004). Ocular higher-order aberrations and contrast sensitivity after conventional laser in situ keratomileusis. *Investigative Ophthalmology and Visual Science*, 45(11), 3986–3990, <http://www.iovs.org/content/45/11/3986>. [PubMed] [Article]

Yoon, G., Jeong, T. M., Cox, I. G., & Williams, D. R. (2004). Vision improvement by correcting higher-order aberrations with phase plates in normal eyes. *Journal of Refractive Surgery*, 20, S523–S527.

Young, L. K., Liversedge, S. P., Myers, R. M., Love, G. D., & Smithson, H. E. (2011). Not all aberrations are equal: Reading impairment depends on aberration type and magnitude. *Journal of Vision*, 11(13):20, 1–19, <http://www.journalofvision.org/content/11/13/20>, doi:10.1167/11.13.20. [PubMed] [Article]

Zhao, H.-X., Xu, B., Li, J., Dai, Y., Zhang, Y.-D., & Jiang, W.-H. (2009). Effects of different Zernike

terms on optical quality and vision of human eyes. *Chinese Physics Letters*, 26(5), 054205, 1–3.

Appendix 1

Aberrations cause spatial frequency–dependent phase and contrast changes in an image. Resizing an image, as is necessary in size-threshold acuity measurements, changes the spatial frequency spectrum of that image. However, the convolution of the aberrated PSF with a resized image is not the same as resizing the convolution of that same PSF with the original image as demonstrated by Figure A2. The spectrum of the resulting convolution is described in terms of its amplitude (contrast) and its phase:

$$S = Ae^{iP}, \quad (1)$$

where S is the complex spectrum of the resulting image, A is its amplitude, and P is its phase. As the convolution is the multiplication of the spectra of the PSF and of the image, it is trivial to show that the amplitude spectrum of the result is the multiplication of their amplitude spectra and that the phase spectrum of the result is the sum of their phase spectra:

$$S = S_{PSF}S_{Image} \quad (2)$$

$$= A_{PSF}e^{iP_{PSF}}A_{Image}e^{iP_{Image}} \quad (3)$$

$$= A_{PSF}A_{Image}e^{i(P_{PSF}+P_{Image})} \quad (4)$$

where S_{PSF} is the complex spectrum of the PSF (the OTF of the system producing that PSF), A_{PSF} is its amplitude spectrum (the MTF of the system), P_{PSF} is its phase spectrum (the PTF of the system), S_{Image} is the complex spectrum of the image, A_{Image} is its amplitude spectrum, and P_{Image} is its phase spectrum. Therefore,

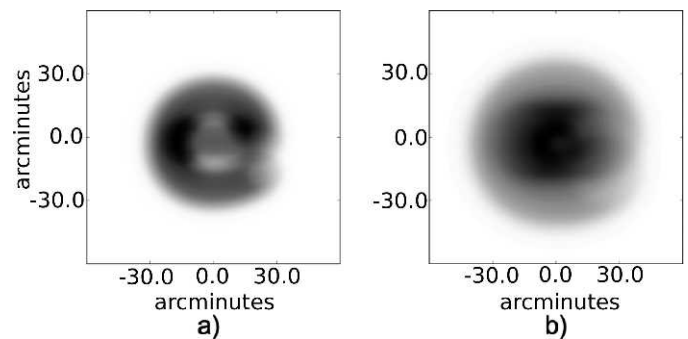


Figure A1. Demonstration of the difference between applying 0.6 μm of defocus in the rendering of (a) a 1.0° letter “e” and (b) a 0.5° letter “e”, which is then resized to 1.0°.

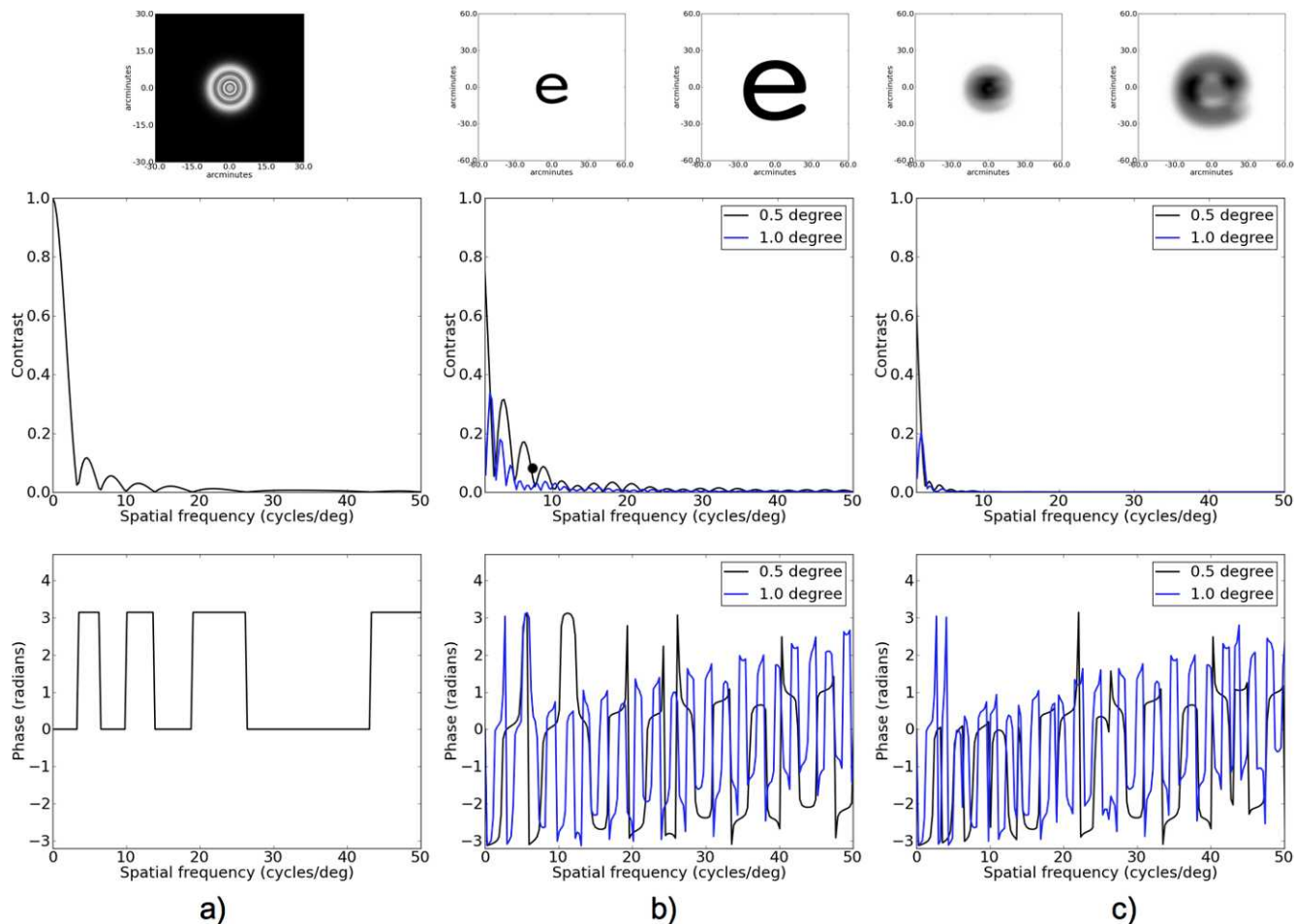


Figure A2. Demonstration of the spatial-frequency-dependent effects of resizing the letter “e.” The top row shows the images of (a) the PSF corresponding to 0.6 μm of defocus for a 2.5 mm pupil, (b) a 0.5° and a 1.0° letter “e” and (c) the convolutions between those letters and the PSF. The middle row shows the corresponding amplitude spectra, and the bottom row shows the corresponding phase spectra.

$$A = A_{PSF}A_{Image} \tag{5}$$

and

$$P = P_{PSF} + P_{Image}. \tag{6}$$

Figure A1 shows the amplitude and phase spectra of an example aberration, of different sizes of letter “e” and of the resulting convolution. In this example, a PSF was calculated from a wavefront of 0.6 μm defocus and convolved with a 0.5° and a 1.0° letter “e,”

demonstrating the origin of the differences between the two images in Figure A2.

In measuring contrast threshold instead of size threshold, the amplitude spectrum of the stimulus is merely scaled, but most importantly, phase spectrum of the stimulus is unchanged. This means that effects due to phase shifts cannot be nullified by increasing the contrast of the stimulus whereas they can be nullified by reducing the viewing distance, effectively increasing its size.