Central Lancashire Online Knowledge (CLoK)

Title	Not all aberrations are equal: reading impairment depends on aberration type and magnitude
Type	Article
URL	https://clok.uclan.ac.uk/22405/
DOI	##doi##
Date	2011
Citation	Young, L.K., Liversedge, Simon Paul orcid iconORCID: 0000-0002-8579-8546, Love, G.D., Myers, R.M. and Smithson, H.E. (2011) Not all aberrations are equal: reading impairment depends on aberration type and magnitude. Journal of Vision, 11 (13). pp. 1-19.
Creators	Young, L.K., Liversedge, Simon Paul, Love, G.D., Myers, R.M. and Smithson, H.E.

It is advisable to refer to the publisher's version if you intend to cite from the work. ##doi##

For information about Research at UCLan please go to http://www.uclan.ac.uk/research/

Not all aberrations are equal: Reading impairment depends on aberration type and magnitude

Laura K. Young   Department of Physics, Durham University, UK

Simon P. Liversedge   School of Psychology, University of Southampton, UK

Gordon D. Love   Department of Physics, Durham University, UK

Richard M. Myers   Department of Physics, Durham University, UK

Hannah E. Smithson   Department of Experimental Psychology, University of Oxford, UK, & Department of Psychology, Durham University, UK

The eye's optical components are imperfect and cause distortions in the retinal image that cannot be corrected completely by conventional spectacles. It is important to understand how these uncorrected aberrations (those excluding defocus and primary astigmatism) affect visual performance. We assessed reading performance using text with a simulated monochromatic aberration (defocus, coma, or secondary astigmatism), all of which typically occur in the normal population. We found that the rate of decline in reading performance with increasing aberration amplitude was smaller for coma than for secondary astigmatism or defocus. Defocus and secondary astigmatism clearly had an impact on word identification, as revealed by an analysis of a lexical frequency effect. The spatial form changes caused by these aberrations are particularly disruptive to letter identification, which in turn impacts word recognition and has consequences for further linguistic processing. Coma did not have a significant effect on word identification. We attribute reading impairment caused by coma to effects on saccade targeting, possibly due to changes in the spacings between letters. Effects on performance were not accompanied by a loss of comprehension confirming that even if an aberration is not severe enough to make text illegible it may still have a significant impact on reading.

Keywords: reading, ocular aberrations, eye movements, spatial vision, phase distortions, physiological optics

Introduction

The eye's optical components are imperfect, causing distortions in the image produced at the retina. These distortions are not completely corrected by conventional spectacles and the remaining distortions can be corrected only by complicated optical surfaces or by wavefront correctors such as those used in adaptive optics (see Roorda, 2011, for a comprehensive review of the use of adaptive optics in vision science). In this paper, we consider how optical distortions of the retinal image affect eye movements made during reading. Reading differs from the common clinical tests of letter acuity in two important respects. First, in reading, letters are seen in the context of words so lateral masking between letters and between words becomes important. Second, in an acuity test, the subject is given an unlimited viewing time to recognize the letter, whereas in a reading task the processes of letter and word recognition are time critical.

Disruptions to word recognition do not necessarily completely prevent the subject from reading but do make the task more difficult, leading to longer reading times that can be distressing (see Legge, 2007, for example). Previous work has shown that optical defocus can impair reading performance (Legge, Pelli, Rubin, & Schleske, 1985) but defocus is not the only distortion present in real eyes. Here, we impose defocus, coma, and secondary astigmatism, which are specific types of distortions described by Zernike polynomials (see Appendix A) that are associated with common imperfections of the eye's optics, to test their effects on reading. While defocus relates to dioptric blur, coma and secondary astigmatism cause more complicated distortions, as described in Appendix A. Coma causes point sources to appear to have a tail (coma) and when applied to text this aberration smears the letters into one another. Secondary astigmatism is a higher order version of primary astigmatism, which causes orthogonal meridians of the eye to have different focal powers. This produces a more complicated distortion
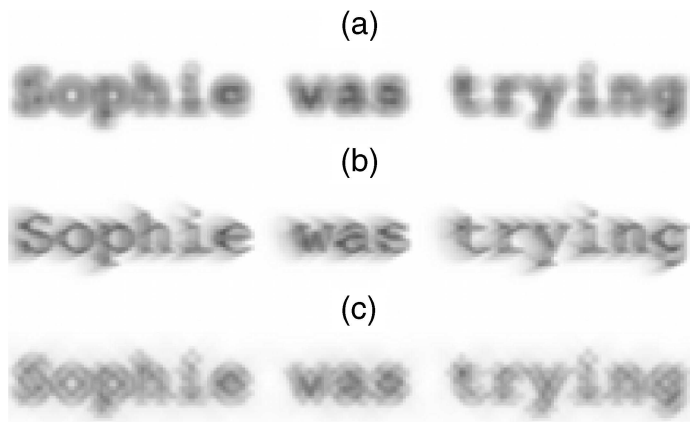
Figure 1. A sample sentence simulated with 0.3 μm of (a) defocus, (b) coma, and (c) secondary astigmatism. Only the beginning of the sentence is shown here. The full sentence read "Sophie was trying to draw her (sister/ferret) but was finding it very difficult" and was displayed on a single line. In this sentence, the target word is either "sister" that is a high-frequency word (38 occurrences per million words) or "ferret" that is a low-frequency word (1 occurrence per million words).

of the image akin to double vision. Examples of text samples used in this experiment, which demonstrate the effects of these aberrations, are given in Figure 1. During reading, our eyes make a series of characteristic movements that are driven by the visual processing demands of the task. By analyzing these movements, we can quantify the differences between types of distortion in terms of their effects on visual and linguistic processing. This study relies on detailed knowledge of optics and image formation and of eye movements and reading. We provide a brief introduction to these two areas in Appendix A.

Previous studies on higher order aberrations and letter acuity

Previous studies have investigated the effects of higher order aberrations on letter recognition although we know of no study that has investigated the effects of higher order aberrations on reading. Since word identification involves the parallel processing of letters, we might expect to see degradations in reading performance that are similar to those found in letter recognition. Oshika, Okamoto, Samejima, Tokunaga, and Miyata (2006) measured the ocular aberrations of their subjects before testing their contrast and letter sensitivity functions and found that coma-like aberrations had a significant influence. Applegate et al. found that the effect on visual acuity varied between different Zernike modes (Applegate, Sarver, & Khemsara, 2002) and combinations of Zernike modes (Applegate, Marsack, Ramos, & Sarver, 2003) such that those with low angular order (those near the center of the Zernike pyramid) affect acuity the most. Fang, Wang, and He (2009),

however, found that fourth radial order aberrations have the largest impact on the magnitude of the optical transfer function (OTF), the modulation transfer function (MTF), of myopic eyes. Results derived from the MTF are based purely on the optical properties of the eye, not on performance. The difference in these findings suggests that the nature of the visual task is important when determining which aberrations have the greatest impact. Using an adaptive optics system in conjunction with a Freiburg acuity test, Li et al. (2009) found a linear relationship between the root mean square (rms) amplitude of the aberration and visual acuity. Applegate, Ballentine, Gross, Sarver, and Sarver (2003) also found a decline in visual acuity with increasing amplitude of aberration but concluded that rms amplitude of aberration was not a good predictor of visual acuity. These findings suggest that degradations to visual performance relate to the amplitude of the aberration, but a better metric for predicting visual performance is required. It is therefore important to measure visual performance during different tasks to get a better understanding of the relationship between the wavefront measurement and the decline in performance.

Letters are typically identified using a visual channel that is sensitive to a particular band of spatial frequencies in the image (Solomon & Pelli, 1994). We know that contrast is important for letter identification and any aberration that reduces the contrast of spatial frequencies used for letter identification is likely to have a significant impact. However, aberrations can also change the phase of those spatial frequencies, altering the appearance of a letter. Recent work has shown that 180° phase shifts reduce visual acuity, whereas those with a phase shift of less than 180° (coma, for example) do not (Ravikumar, Bradley, & Thibos, 2010).

Previous studies on reading degraded text

Although there is a paucity of studies directly investigating the influence of different types of degraded text on reading, there have been some studies that have manipulated the visual characteristics of the text to make it more difficult to process. For example, Rayner, Fischer, and Pollatsek (1998) removed the spaces between the words of sentences and showed that readers were much slower to process the text. Removing spaces increases lateral masking between words and makes it more difficult for readers to identify word objects to which to target saccades. More recently, Juhasz, Liversedge, White, and Rayner (2006) employed an aLtErNaTiNg-CaSe manipulation to produce visual disruption to the text. They found that readers experienced more difficulty (e.g., increased fixation durations) when reading text with alternating case than normally presented text. Reingold and Rayner (2006; see also Drieghe, 2008) presented participants with sentences in which the contrast between the background and a target word was substantially reduced. This

manipulation reduced the visual quality of the stimulus, causing readers to fixate the word for longer. Disruption to reading performance has also been found in experiments using dioptric (Thorn & Thorn, 1996) and diffuse blur (Legge et al., 1985). Chung, Jarvis, and Cheung (2007) also studied the effect of dioptric blur, finding that acuity decreased from −0.16 to 0.58 logMAR and that reading acuity could be predicted from visual acuity.

The present study

In this paper, we present the eye movement data obtained during a reading task in which different rms amplitudes of defocus, coma, and secondary astigmatism have been added to the stimuli (see Figure 1). This experiment required subjects to read text covering a large field of view (about 15°). Subject's eyes, and therefore their pupils, made large movements and so the use of an adaptive optics system would have been impractical. We chose instead to study the effects of these aberrations by adding them in the rendering of the stimuli. The experiment reported here considered the average fixation duration and number of fixations over a whole sentence to test the difference in reading performance for these aberrations as a function of their rms amplitude. We also investigated the impact of these aberrations on lexical identification during reading by examining differences in fixation durations on high and low lexical frequency target words embedded in the same sentence. We assessed whether our subjects understood what they read, even in the presence of an aberration, by monitoring their responses to comprehension questions. Our data are related to predictions of reading performance based on the calculated distinguishability of letters in the presence of an aberration.

Methods

Selection of aberrations

It is not feasible to test several amplitudes of every Zernike mode so we selected three for investigation by considering their effects on letters and also their clinical relevance. We considered the effect of an aberration on the form of a letter and on the position of the center of a letter. Details of these calculations are given in Appendix B, and a summary of the results is given in the following sections.

Letter confusion

Spatial phase changes caused by aberrations can cause one letter to look like another, as shown in Movie 1 where increasing amounts of defocus are added to the letter m
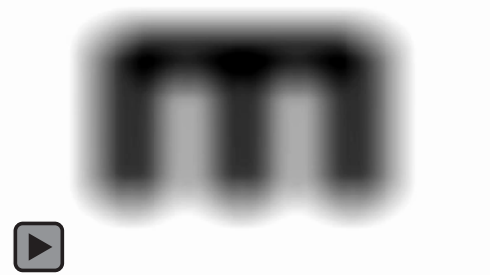
Movie 1. The letter m with increasing amplitude of defocus, showing a transition to it resembling a letter n.

until it resembles an n. Considering this, we investigated the potential for letters to be confused by a subject by comparing letters via a cross-correlation. Based on these results (given in Figure B3), we chose to investigate secondary astigmatism (Z_4^2) since it produced the greatest confusability factor when comparing aberrated and non-aberrated letters. From a clinical perspective, secondary astigmatism is thought to contribute to halo phenomena that are often experienced under night vision after successful laser-assisted in situ keratomileusis (LASIK) surgery (Villa, Gutiérrez, Jiménez, & González-Méijome, 2007).

Letter position

In addition to changes in the forms of letters, changes in the positions of the centers of letters were also considered. These are important during reading where lateral masking between letters within words can interfere with letter and word identification and changes in the word spacings can disrupt saccade targeting. Based on these results (given in Figure B4), we chose to investigate coma (Z_3^1) since it caused the letter center to vary the most. From a clinical point of view, coma is interesting because its amplitude is known to increase with age, which is believed to be due to the loss of compensatory effects between the corneal surfaces (Guirao, Redondo, & Artal, 2000; Lu et al., 2008). It is also prevalent in subjects with keratoconus, for which it is even being considered as a diagnostic quantity (Gobbe & Guillon, 2005).

In summary, the Zernike modes we have chosen to investigate are coma (Z_3^1) and secondary astigmatism (Z_4^2). Defocus was also chosen for a comparison with a low-order aberration and because it relates directly to dioptric blur.

Amplitude range

To choose a suitable range of amplitudes to test, we conducted a preliminary experiment in which we found that applying 0.3–0.5 μm rms defocus (over a 3.5 mm

pupil) allowed subjects to read text but with substantially increased difficulty. As Zernike polynomials are defined over a unit circle, the value of the coefficient is dependent on the radius of the pupil. Equivalent defocus, M_e, is an aberration metric that is independent of pupil size and is defined as the amount of defocus in diopters that produces the same wavefront variance as a given aberration. It is determined using

$$M_e = \frac{4\pi\sqrt{3}\mathrm{RMS}}{A} = \frac{4\sqrt{3}\mathrm{RMS}}{r^2}, \tag{1}$$

where RMS is the rms amplitude of the wavefront, A is the area of the pupil, and r is the pupil radius (Thibos, Hong, Bradley, & Cheng, 2002).

The main study used text samples aberrated with 0.3 μm, 0.35 μm, or 0.4 μm rms of one of the three types of aberration. These values are measured over a 3.5-mm pupil and so correspond to 0.68, 0.79, and 0.90 diopter (D) of equivalent defocus. It should be noted that the text size used in this experiment was much larger than the normal acuity limit and so the impact of these aberrations was lower than for a letter at the acuity limit. Considering the pixel scales in the image and in the PSF, if a letter at the acuity limit had been used, the aberration would have needed to be approximately three times lower to create the same amount of distortion of the image at the retina. This would correspond to 0.21 D, 0.26 D, and 0.3 D of equivalent defocus. A control condition was also tested where no aberration was applied and all results were compared to the control data. This avoided large errors caused by intersubject variability.

Subjects

Nineteen subjects participated in this study, twelve males and seven females, with a mean age of 28 years (SD = 7 years). All subjects were fluent in English, had at least 17 years of education, and had normal or corrected-to-normal vision. Subjects read with natural pupil dilation using any vision correction they would normally use for viewing a computer screen at a distance of 75 cm.

Apparatus

Aberrated text samples were presented on a CRT using a Cambridge Research Systems ViSaGe visual stimulus generator. Stimuli were presented as black text on a white background at a distance of 75 cm. The average luminance of the monitor was 104 cd/m^2 and the output of the display was gamma corrected. Eye movements were sampled at 250 Hz using a Cambridge Research Systems High-Speed Video Eyetracker (CRS HS-VET)

that tracks the pupil and two first Purkinje images. This system has a spatial resolution of 0.05° and an accuracy of 0.125°–0.25°. The CRT monitor and eye tracker were controlled using the CRS Matlab toolboxes.

Stimuli

A set of 52 sentence frames was constructed to occupy a single line and had, on average, 71 characters. The stimuli were generated in Courier font such that a single letter subtended 15 min of arc (equivalent to 20/60 acuity). Each sentence frame was generated twice, once each with either a high or a low lexical frequency target word. The target word consisted of 6 letters, was positioned approximately in the middle of the sentence frame, and was not used in more than one sentence frame. The Kučera–Francis written frequencies of the two types of target word were chosen to be significantly different ($t(51)$ = 11.23, p < 0.001) with an average high frequency of 160 occurrences per million words (SD = 189 occurrences per million words) and an average low frequency of 4 occurrences per million words (SD = 5 occurrences per million words). The 104 sentences (2 × 52 sentence frames) were prescreened for plausibility and the target words for predictability. Sentences in the high- and low-frequency conditions did not differ in respect of each. Individual subjects only viewed a particular sentence frame once, either with the high- or low-frequency target word. Sentences were presented at random using a Latin square approach such that each subject viewed four sentences per type and amplitude of aberration. Two of these sentences contained a high-frequency target word and two a low-frequency target word.

Simulating aberrations

The stimuli were distorted by convolving the text samples with an aberrated PSF (see Appendix A for an example of an aberrated PSF). Subjects viewed these simulated aberrations directly with natural pupil dilation. When a subject views these stimuli, the aberrations in their eye create additional distortions in the image formed at the retina. The effect of an optical system, in this case the eye, on an object can be described as

$$\mathcal{F}(I) = \mathcal{F}(O)\mathrm{OTF}_{\mathrm{eye}}, \tag{2}$$

where I is the image on the retina, O is the stimulus object, and $\mathcal{F}$ represents a Fourier transform. The OTF of the eye is given by

$$\mathrm{OTF}_{\mathrm{eye}} = \mathrm{OTF}_p\mathrm{OTF}_a, \tag{3}$$

where OTF$_p$ is the OTF of the eye's pupil, which causes diffraction, and OTF$_a$ is the OTF of the eye's aberrations.

In order to minimize the effect of the subject's eye on the retinal image, the magnitude of its OTF should be close to one for all spatial frequencies. First, we consider a diffraction-limited eye that is free of aberrations. The pupil of the eye has an approximately linearly decreasing MTF with a cutoff frequency that increases with increasing diameter. It is therefore desirable, in this case, to have as large a pupil as possible. However, in a real eye, the amplitude of aberration increases with increasing pupil diameter. We therefore want as small a pupil as possible to minimize the effect of the subject's aberrations on the retinal image. To ensure this effect was as small as possible, we measured the subject's aberrations immediately before the reading experiment using a Zywave aberrometer and used a bright stimulus to constrict the pupil. Pupil size was monitored during the experiment to check that the contribution of the subject's aberrations to the retinal image quality remained minimal. The average amount of higher order aberrations among subjects was 0.018 μm rms calculated over the pupil size during the experiment, which was 3.5 mm on average. Histograms of subjects' higher order aberrations are given in Figure 2. Previous work has shown that the visual system adapts neurally to the particular eye's monochromatic aberrations (Artal et al., 2004), at least for small amplitudes of aberration. So even though our subjects have different ocular aberrations, they are all using the same baseline (i.e., the aberrations they are used to).

It is known that accommodative behavior changes with visual acuity (see Heath, 1956, for example) and so we expect our subjects' accommodative state to vary with the different stimuli. If the accommodative response were incorrect for the stimulus distance, this would increase the amount of defocus and other higher order aberrations, particularly spherical aberration, introduced by the subject's eye. The stimuli used in this experiment are less than 1 D of equivalent defocus. From the work of Heath (1956), we have estimated the maximum error in accommodation due to 1 D of equivalent defocus, presented at 75 cm, to be 0.25 D or 0.1 μm rms over a 3.5 mm pupil. We expect the actual effect of accommodative error to be much smaller than we have stated here as there were other visual cues available to drive appropriate accommodation. Specifically, the edges of the monitor were visible and text samples were preceded and followed by high spatial frequency images such as the fixation cross and comprehension questions. Our subjects viewed the text binocularly and so had vergence cues available to them. Additionally, at a viewing distance of 75 cm, accommodation is likely to have only a small influence, and since pupil size was small, the depth of field is large, minimizing the effect of changes in accommodation. The effect that diffraction, higher order aberrations, and accommodative error have on the simulated image is demonstrated in Figure 3, which shows minimal changes.
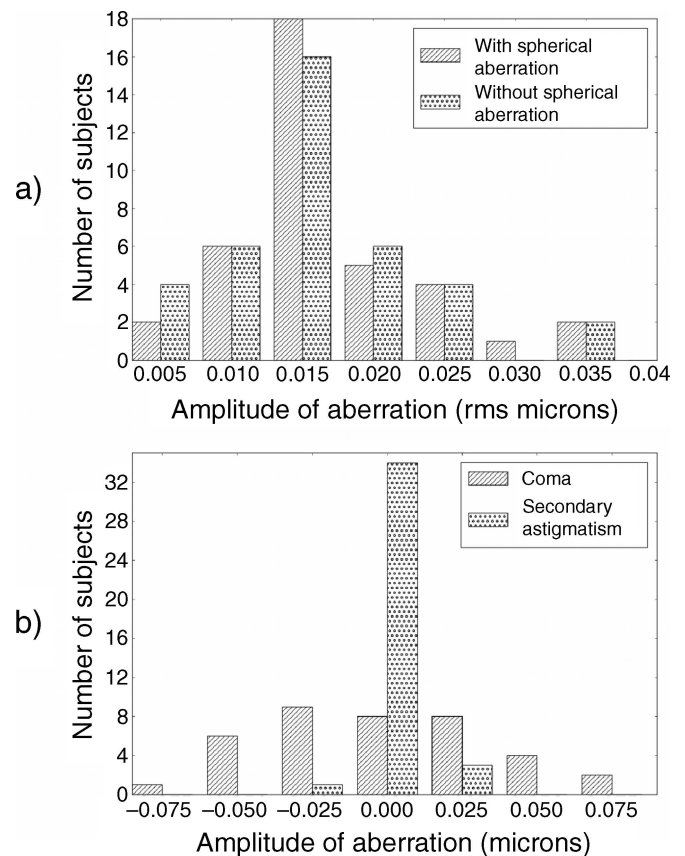

Figure 2. Histogram of the rms amplitude of subjects' measured aberrations over their average pupil diameter during the experiment. Pupil diameters during the experiment were always smaller than the diameter over which this measurement was taken. This was accounted for by recalculating the Zernike coefficients over the smaller diameter. Only higher order aberrations are included since subjects read using any prescribed vision correction, such as spectacles, which correct low-order aberrations. Results are shown for (a) rms amplitude of all higher order modes with and without the spherical component (Z_4^0) included (spherical aberration is affected by spectacle prescription) and (b) amplitudes of coma (Z_3^1) and secondary astigmatism (Z_4^2) separately (note that a negative value indicates an inverted wavefront). For comparison, the rms amplitudes of our simulated aberrations were 0.3 μm, 0.35 μm, and 0.4 μm.

As our stimuli are distorted in simulation, rather than optically, the transverse chromatic aberration will be different from that with an optical distortion due to the different light distribution at the eye's lens. All chromatic aberration could be eliminated by using monochromatic stimuli, but we decided against this in order to simulate natural viewing conditions. In addition, longitudinal chromatic aberration changes with accommodation and there is evidence that this may aid the accommodation response in some subjects (see Aggarwala, Nowbotsing, & Kruger, 1995; Kruger, Mathews, Aggarwala, & Sanchez, 1993; Kruger & Pola, 1986, for example).
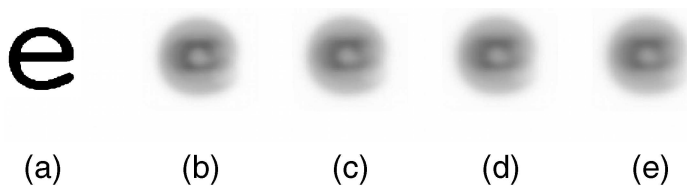
Figure 3. The letter e presented as (a) the original image, (b) the image simulated with 0.4 μm defocus (i.e., the stimulus), (c) the stimulus with diffraction effects from the subject's pupil, (d) the stimulus with diffraction effects and the subject's higher order aberrations (averaged over all subjects, the amplitude of higher order aberrations was 0.018 μm), and (e) the stimulus including diffraction, higher order aberrations, and accommodative error (equivalent to 0.1 μm defocus). The effects of an average subject's pupil, aberrations, and expected accommodative error on the simulated image are minimal. The retinal image (e) was therefore sufficiently similar to the image presented to the subject (b).

Procedure

Subjects were positioned in a headrest and asked to read the sentences silently so as to understand them. Each sentence was followed by a comprehension question that required a yes–no response via a button box. The subject was required to initiate the start of each trial so that they could control the pace of the experiment and take a break when required. Each sentence was displayed and data were collected only once it had been confirmed that the subject was fixating at the position of the first word. A 9-point calibration was performed and verified before every fourth trial or after the subject had moved from the headrest. Data were collected from subjects in single sessions of 45 min duration.

Eye movement data were analyzed using an event detection algorithm based on the work of Nyström and Holmqvist (2010). This algorithm first uses a Savitzky–Golay filter to smooth the data, thus avoiding noise amplification when calculating velocities (Savitzky & Golay, 1964). Next, saccades are determined by finding velocity peaks above the saccade detection threshold and their onsets are found by searching backward for the first sample to go below a lower saccade onset threshold. These thresholds are determined by a data-driven iterative approach based on the noise level over an individual trial and so they adapt to variations between and within subjects. The end of the saccade may or may not contain a glissade, which occurs when the eye overshoots its target, making the offset less obvious. The saccade offset is determined in a similar way to the saccade onset although a locally adaptive threshold is calculated so that glissades do not contaminate the data. Glissades are then separately detected as another class of eye movement. Everything that was not classified as noise, a saccade, or a glissade could be labeled as a fixation.

Trials were excluded from the analysis if the subject did not correctly answer the comprehension question or the eye movement data contained more than 20% noise, which could be due to a loss of tracking, blinks, or the subject looking away from the screen. Under these criteria, 9.3% of trials were excluded from the analysis. The accuracy of saccade detection was checked manually before any further analysis was carried out.

For each trial, the average fixation duration and the number of fixations were calculated as global measures of performance. For each participant, the average for the four trials in the control condition was subtracted from these measures to avoid large errors from intersubject variability. The weighted average according to the number of valid trials contributing to the estimate was then calculated over all subjects. Additionally, the sum of all fixation durations on the target word before leaving it (gaze duration) and the sum of all fixations on the target word (total reading time) were calculated for each trial. These local measures were then used for the lexical frequency analysis. Data were subjected to repeated measures analyses of variance (ANOVAs). When sphericity could not be assumed, a Greenhouse–Geisser correction was applied for epsilon values below 0.75, otherwise a Huynh–Feldt correction was used. In the reports below, we quote corrected degrees of freedom.

Results

Global measures
Average fixation duration

Figure 4 shows the increase in average fixation duration over the control condition vs. the amplitude of the aberration. The data were analyzed with a 3 (type of aberration) × 3 (amplitude of aberration) two-way repeated measures ANOVA. We found a significant main effect of the type of aberration ($F(2.0,143.3) = 59.9$, $MSE = 23177.4$, $p < 0.001$) with coma causing the smallest increase in average fixation duration (25.2 ms averaged over all amplitudes), followed by defocus (32.4 ms averaged over all amplitudes) and by secondary astigmatism that caused the greatest increase (46.3 ms averaged over all amplitudes). We also found a significant effect of the amplitude of the aberration ($F(2.0,143.3) = 15.7$, $MSE = 5840.6$, $p < 0.001$). As we expected, there was an increase in the average fixation duration as the amplitude of aberration was increased with, on average, an extra 15.9 ms exhibited at 0.3 μm of aberration, 31.7 ms at 0.35 μm, and 56.2 ms at 0.4 μm. Most importantly, there was a significant interaction between type and amplitude of aberration ($F(3.6,143.3) = 9.7$, $MSE = 3299.5$, $p < 0.001$) indicating that the effect of increasing the amplitude was different for different types of aberration. We explored the
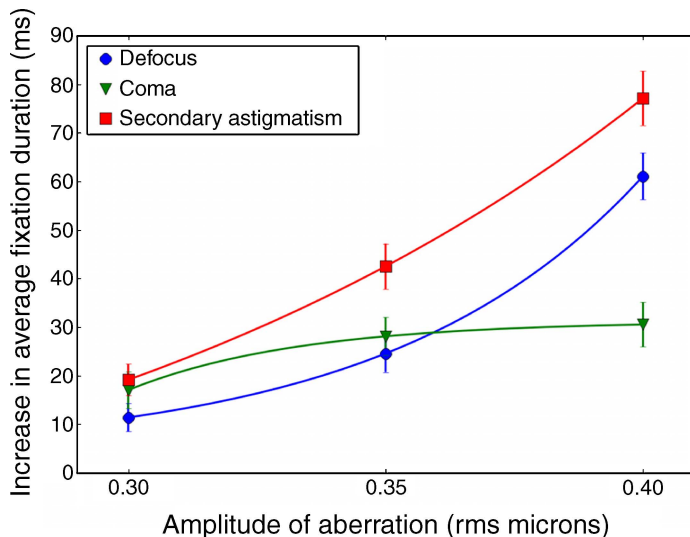
Figure 4. Increase in the average fixation duration from control data over a whole sentence with increasing amplitude of aberration. In the control condition, the average fixation duration was 189.5 ms (*SE* = 9.8 ms). Error bars represent the standard error of the mean of the measurements across 19 subjects. Three types of aberration are shown: defocus (Z_2^0), coma (Z_3^1), and secondary astigmatism (Z_4^2). Exponential growth curves have been fitted to the data via the least squares method. The growth constants were found to be 20.4 μm^{-1} for defocus (SSE < 0.001), −30.3 μm^{-1} for coma (SSE < 0.001), and 7.8 μm^{-1} for secondary astigmatism (SSE < 0.001).

significant interaction with an analysis of simple effects. Since there were only three groups, where one-way ANOVAs revealed a significant effect, post-hoc pairwise comparisons were made via Fisher's least significant difference (Howell, 1992, p. 356).

For defocus, there was a significant effect of the amplitude of aberration ($F(2.0, 36.0) = 40.3$, $MSE = 11934.0$, $p < 0.001$) causing an increase in average fixation duration from 200.9 ms at 0.3 μm, to 214.1 ms at 0.35 μm, up to 250.6 ms at 0.4 μm. Pairwise comparisons identified differences between 0.30 μm and 0.35 μm ($t(18) = 3.0$, $p \leq 0.007$), between 0.35 μm and 0.40 μm ($t(18) = 7.8$, $p < 0.001$), and between 0.30 μm and 0.40 μm ($t(18) = 6.3$, $p < 0.001$).

For secondary astigmatism, there was a similar significant effect of the amplitude of aberration ($F(2.0, 36.0) = 33.0$, $MSE = 16302.3$, $p < 0.001$). Once again, the average fixation duration increased with larger aberrations from 208.6 ms at 0.3 μm, to 232.0 ms at 0.35 μm, up to 266.6 ms at 0.4 μm. Again, pairwise comparisons identified differences between 0.30 μm and 0.35 μm ($t(18) = 3.2$, $p \leq 0.005$), between 0.35 μm and 0.40 μm ($t(18) = 7.7$, $p < 0.001$), and between 0.30 μm and 0.40 μm ($t(18) = 5.0$, $p < 0.001$).

For coma, there was a significant, but smaller effect, of the amplitude of aberration ($F(1.8, 33.0) = 3.9$, $MSE = $

898.5, $p \leq 0.034$). The average fixation duration for coma increased from 206.5 ms at 0.3 μm, to 217.6 ms at 0.35 μm, up to 220.1 ms at 0.4 μm. Clearly, these effects are much smaller in magnitude. In line with this, pairwise comparisons showed that the reliable effect was driven by differences between 0.30 μm and 0.35 μm ($t(18) = 2.6$, $p \leq 0.017$) and between 0.30 μm and 0.40 μm ($t(18) = 2.4$, $p \leq 0.028$) though not between 0.35 μm and 0.40 μm ($t(18) = 0.7$, $p \leq 0.525$).

These trends can be seen clearly in Figure 4, which shows that text becomes increasingly difficult to process as the amplitude is increased for defocus and secondary astigmatism. By contrast, the difficulty of processing the text plateaus at 0.35 μm for coma. This pattern of effects is consistent with the suggestion that when viewing text with coma the structure of the letters is still distinguishable, whereas with defocus and secondary astigmatism, their structure is altered. This can be seen in Figure 1, which shows that coma creates a smearing effect across the letters but leaves their form relatively intact.

Number of fixations

Figure 5 shows the increase in the number of fixations over the control condition vs. the amplitude of the aberration. Another two-way repeated measures ANOVA showed a significant main effect of the type of aberration ($F(1.7, 94.5) = 8.5$, $MSE = 60.7$, $p \leq 0.002$), with coma again causing the smallest detriment (no extra fixations on average over all amplitudes), followed by defocus (1 extra fixation on average over all amplitudes) and by secondary astigmatism that caused the largest increase in the number of fixations (2 extra on average over all amplitudes). Since there were only three groups, we explored these main effects with pairwise comparisons made via Fisher's least significant difference, which revealed that the main effects were driven by the following differences: There was a significant difference in the number of fixations between secondary astigmatism and defocus ($t(18) = 1.9$, $p \leq 0.008$) and between secondary astigmatism and coma ($t(18) = 3.5$, $p \leq 0.002$). There was only a marginal difference between defocus and coma ($t(18) = 3.0$, $p = 0.078$). The data indicate that defocus was least disruptive, causing readers to make fewest fixations, coma slightly more so, with secondary astigmatism most disruptive.

We also found a significant effect of the amplitude of the aberration ($F(1.3, 94.5) = 12.3$, $MSE = 88.4$, $p \leq 0.001$). Consistent with what we expected, the increase in the number of fixations grew from 0 at 0.3 μm, to 1 at 0.35 μm, up to 3 at 0.4 μm. There was also a significant difference in the number of fixations between 0.30 μm and 0.35 μm ($t(18) = 2.6$, $p \leq 0.018$), between 0.35 μm and 0.40 μm ($t(18) = 3.4$, $p \leq 0.004$), and between 0.30 μm and 0.40 μm ($t(18) = 3.8$, $p \leq 0.001$). This shows that the number of fixations increased significantly across all amplitudes of aberration. Although there was no significant
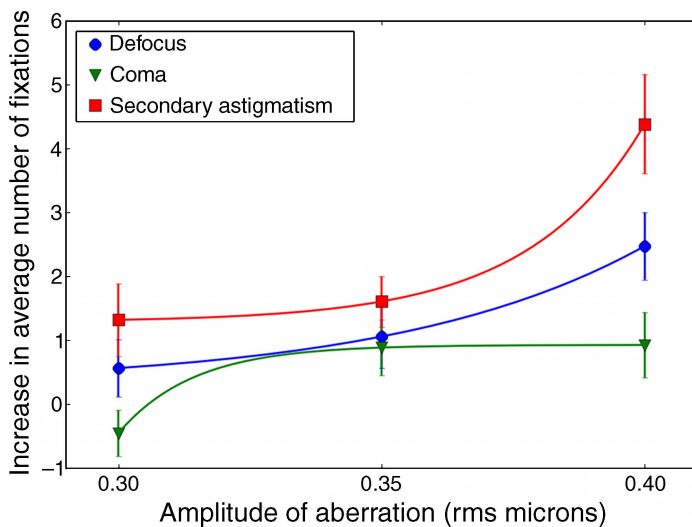
Figure 5. Same format as Figure 4. Increase in the average number of fixations from control data over a whole sentence with increasing amplitude of aberration. In the control condition, the average number of fixations was 11 (*SE* = 1). Exponential growth curves have been fitted to the data via the least squares method and the growth constants were found to be 21.1 μm^{-1} for defocus (SSE < 0.001), -70.1 μm^{-1} for coma (SSE < 0.001), and 45.3 μm^{-1} for secondary astigmatism (SSE < 0.001).

interaction ($F(2.1, 94.5) = 2.5$, $MSE = 17.1$, $p \leq 0.091$), the data show similar numerical trends to those shown in Figure 4.

To briefly summarize, these global measures indicate that the amplitude of aberration reliably influenced both the number and duration of fixations during reading. We also saw a significant influence of the type of aberration on both of these measures, with poorest performance for secondary astigmatism and best performance for coma. For fixation duration, we additionally note that increasing the amplitude was differentially effective in impairing performance with different types of aberration.

Comparison with confusion analysis

We compared the confusion metric for the three Zernike modes at the amplitudes used in the experiment with the measured increase in average fixation duration. Figure 6 shows that there was a correlation between the predicted and measured performances that followed an exponential growth. The growth constant of the curve depended on the Zernike mode with coma having the lowest constant (-296.4). Defocus and secondary astigmatism had similar growth constants and we found that both sets of data could be described by the same curve, which had a growth constant of 198.8. This suggests that the mechanism by which secondary astigmatism and defocus affect performance might be similar. Coma, however, showed a larger

increase in average fixation time for a smaller confusion metric, suggesting performance is affected differently. Again, we suggest that coma aberration leaves the forms of letters relatively intact and so increases their confusability to a much lesser extent than defocus or secondary astigmatism.

Local measures

The gaze durations and total reading times for high- and low-frequency words are given in Figures 7 and 8. The difference between these values gives the size of the lexical frequency effect.

Average gaze durations

The average gaze durations were subjected to a 3 (type of aberration) × 3 (amplitude of aberration) × 2 (lexical frequency) three-way repeated measures ANOVA. As expected, there was a significant main effect of the amplitude of aberration ($F(1.6, 378.4) = 33.5$, $MSE = 153954.2$, $p < 0.001$) with durations increasing on average
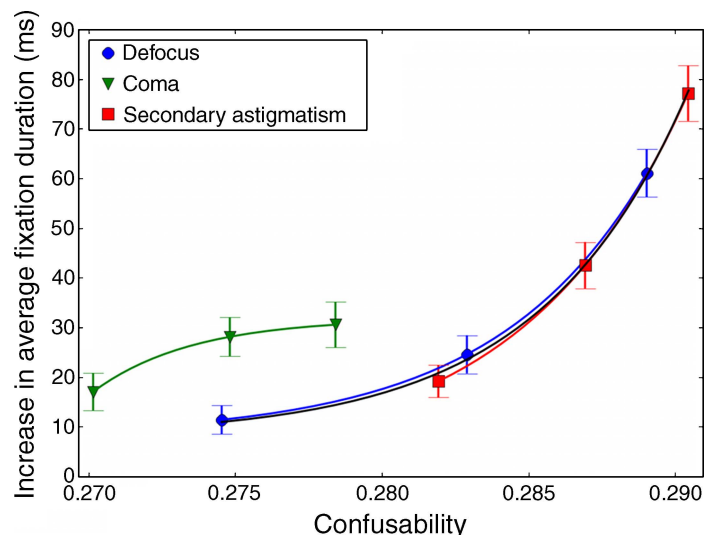

Figure 6. Relationship between the predicted confusability of letters (aberrated letters compared with aberrated letters) and the increase in average fixation duration for defocus (Z_2^0), coma (Z_3^1), and secondary astigmatism (Z_4^2). In the control condition, the average fixation duration was 189.5 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 the same curve with a growth constant of 198.8 (SSE = 4.535), as shown in black.
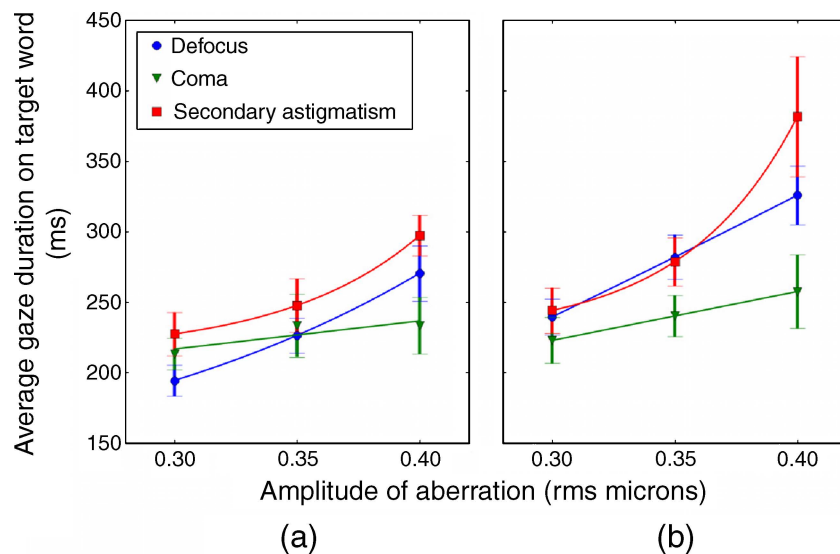
Figure 7. Gaze durations for (a) high and (b) low lexical frequency target words. Error bars represent the standard error of the mean of the measurements across 19 subjects.

from 223.6 ms at 0.3 μm, to 251.4 ms at 0.35 μm, up to 294.3 ms at 0.4 μm. This pattern mirrors that reported earlier. We also found a significant effect of the lexical frequency of the word ($F(1.0,378.4) = 20.9$, $MSE = 112180.1$, $p < 0.001$), as was expected, with fixation durations of 238.2 ms on average for high-frequency words and 274.7 ms for low-frequency words. There was a marginal effect of the type of aberration ($F(1.9,378.4) = 3.2$, $MSE = 46022.3$, $p \leq 0.057$), with subjects fixating the longest in the presence of secondary astigmatism (279.5 ms on average), followed by defocus (256.4 ms on average), and the shortest fixations were those in the presence of

coma (233.4 ms on average). A marginal interaction between the type and amplitude of the aberration was also found ($F(2.3,378.4) = 2.9$, $MSE = 29404.2$, $p \leq 0.063$), indicating that, as in the global results, the increase in fixation duration was different for the different types of aberration. None of the other two-way interactions nor the three-way interactions were significant.

Total reading times

The total reading times were also subjected to a 3 (type of aberration) × 3 (rms amplitude of aberration) × 2 (lexical
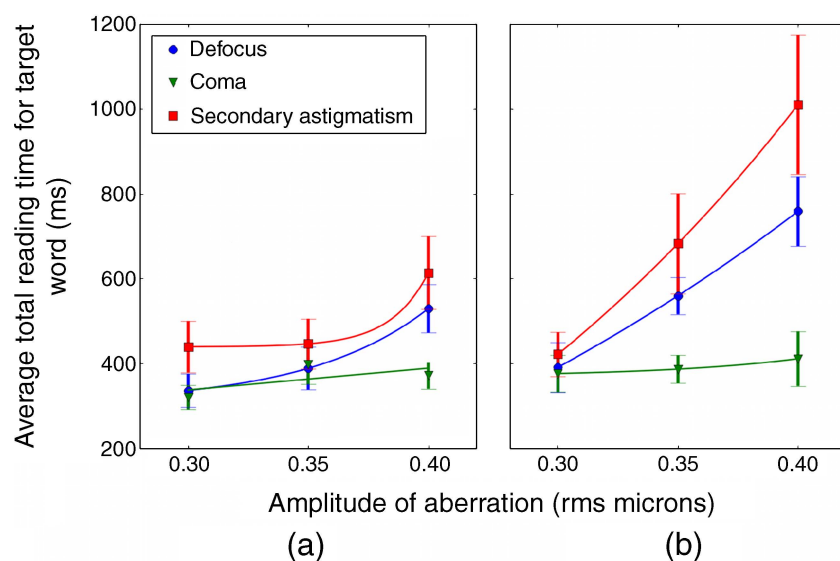

Figure 8. Total reading times for (a) high and (b) low lexical frequency target words. Error bars represent the standard error of the mean of the measurements across 19 subjects.

frequency) three-way repeated measures ANOVA. In line with what we expected, there was a significant main effect of the amplitude of aberration ($F(1.6,233.8) = 10.9$, $MSE = 1531859.0$, $p \leq 0.005$), with durations increasing on average from 380.7 ms at 0.3 μm, to 477.0 ms at 0.35 μm, up to 615.8 ms at 0.4 μm. Consistent with our other results, there was also a significant effect of the type of aberration ($F(1.6,233.8) = 6.5$, $MSE = 1055920.3$, $p \leq 0.002$), with subjects fixating the longest in the presence of secondary astigmatism (602.4 ms on average), followed by defocus (493.9 ms on average), and the shortest fixations were those in the presence of coma (377.2 ms on average). Unsurprisingly, we also found a significant effect of the lexical frequency of the word ($F(1.0,233.8) = 1.7$, $MSE = 134175.6$, $p \leq 0.001$) with subjects fixating on high-frequency words for 427.0 ms on average and on low-frequency words for 555.4 ms. Importantly, there was a significant interaction between the type of the aberration and the lexical frequency of the word ($F(1.4,233.8) = 15.2$, $MSE = 2024411.0$, $p \leq 0.040$). There was also a significant interaction between the amplitude of the aberration and the lexical frequency of the word ($F(1.9,233.8) = 1.2$, $MSE = 127891.5$, $p \leq 0.009$). In agreement with our previous findings, a marginal interaction between the type and amplitude of the aberration was also found ($F(1.7,233.8) = 3.9$, $MSE = 637574.6$, $p \leq 0.054$). The three-way interaction was not significant.

On a priori grounds, we decided to investigate the interaction between the type and the amplitude of aberrations for high and low lexical frequency words separately. A 3 × 3 two-way repeated measures ANOVA showed no significant effect of this interaction on either the average gaze durations or the total reading times of high-frequency words. There was, however, a significant effect of this interaction on the total reading times of low-frequency words ($F(2.1,79.8) = 3.4$, $MSE = 231524$, $p \leq 0.038$), indicating that the increase in total reading time of low-frequency words was different for the different types of aberration. This interaction did not reach significance for average gaze durations ($F(3.4,108.9) = 2.0$, $MSE = 16039$, $p = 0.118$).

We investigated the significant interaction with an analysis of simple effects. Where one-way ANOVAs revealed a significant effect, pairwise comparisons were again made via Fisher's least significant difference since there were only three groups.

For defocus, there was a significant effect of the amplitude of aberration on the total reading times of low-frequency words ($F(1.9,34.2) = 8.8$, $MSE = 565732$, $p \leq 0.001$). In the presence of defocus, the total reading times of low-frequency words increased from 390.8 ms at 0.3 μm, to 559.7 at 0.35 μm, up to 758.9 at 0.4 μm. Pairwise comparisons showed differences between 0.30 μm and 0.35 μm ($t(18) = 2.4$, $p \leq 0.028$), between 0.35 μm and 0.40 μm ($t(18) = 2.3$, $p < 0.037$), and between 0.30 μm and 0.40 μm ($t(18) = 3.8$, $p < 0.001$).

For secondary astigmatism, there was a significant effect of the amplitude of aberration ($F(1.4,25.2) = 6.3$, $MSE = 2841004$, $p \leq 0.012$). In the presence of secondary astigmatism, the total reading times of low-frequency words increased from 421.6 ms at 0.3 μm, to 682.8 at 0.35 μm, up to 1010.5 at 0.4 μm. Pairwise comparisons revealed differences between 0.30 μm and 0.35 μm ($t(18) = 2.3$, $p \leq 0.036$), between 0.35 μm and 0.40 μm ($t(18) = 2.3$, $p < 0.037$), and between 0.30 μm and 0.40 μm ($t(18) = 3.8$, $p < 0.001$).

For coma, the one-way ANOVA was not significant and so no pairwise comparisons were made.

To summarize, the low-frequency words took longer to identify than high-frequency words, and when increasing amplitudes of defocus or secondary astigmatism were applied to the text, lexical identification became increasingly difficult. When text was viewed with coma aberration, lexical identification was not similarly increasingly inhibited with increased amplitude of aberration. Once again, we consider that this indicates that coma leaves the form of letters and words relatively intact.

These local results with respect to the type and amplitude of the aberration show the same pattern as our global results for the average fixation duration but with an exaggerated effect on low- compared with high-frequency words. The main effects of lexical frequency indicate clearly that lexical identification occurred even when readers were viewing aberrated text. This reinforces our conclusion based on the comprehension data that readers successfully processed the text. Furthermore, significant two-way interactions involving lexical frequency for total reading time suggest that different types of aberration differentially affect lexical identification and high amplitudes of aberrations have a particularly strong effect. Differences between average gaze durations and total reading times are interesting because they reveal differences in behavior. An increase in average gaze durations would suggest that subjects made additional refixations on the target word before moving on to the next word, whereas an increase in the total reading times would indicate that subjects revisited the target word. The absence of two-way interactions involving lexical frequency for gaze duration suggests that the size of the lexical frequency effect is affected by the type and amplitude of the aberration in terms of later refixations on the word, after the first pass through the sentence.

Discussion

Although Zernike coefficients are the standard quantity for measuring aberrations in the eye, they are not necessarily suitable for describing visual performance. Understanding how visual performance is affected by

higher order aberrations is important when considering a correction for them. We wanted to know which aberrations contribute to a degradation in performance in a high-level visual task that includes both perceptual and cognitive components. Two factors that can affect fixation durations on a word during reading are the ease of word identification and uncertainty about where to target the next saccade. As pointed out in Appendix A, word identification is a primary determinant of when a reader moves their eyes during reading (Liversedge & Findlay, 2000; Rayner, 1998). Degradation of the orthographic input slows word identification and other subsequent language processing stages. Saccadic targeting can also be affected by the visual clarity of the target, in this case the next word in the sentence to be fixated, and slower initiation of a saccade results in an increased fixation duration. The global analysis of average fixation duration subsumes both these factors. The lexical frequency manipulation allows us to isolate effects associated with word identification. Optical distortions, like those we have simulated, degrade visual information in a number of ways. Most simply, they cause a reduction in contrast; however, they additionally cause changes in form by, for example, introducing new contours. They also increase the spatial extent of a stimulus, and in the case of text, they cause the letter and word boundaries to become less clear. The effect of coma could be informally described as smearing the text, leaving the form of letters relatively undisturbed but filling in the spaces between words and between letters within words, effectively increasing lateral masking. This characteristic of coma leads to a relatively low value of our letter-based confusion metric for a given rms amplitude. Defocus and secondary astigmatism have different characteristics. Like coma, they increase the spatial extent of stimuli, but in addition to this, they have a profound effect on spatial form, as quantified by a higher confusion metric for these aberrations. In interpreting our results, we take these differences into consideration.

From the global results presented here, we can see that the effect of higher order aberrations on reading performance depends on both the type and the amplitude of the aberration. The most striking result is that coma does not affect reading performance as much as defocus or secondary astigmatism. This is in contrast to the findings of Applegate, Marsack et al. (2003) and Applegate et al. (2002) that Zernike modes with low angular order affect visual acuity more than those with high angular order and to the findings of Oshika et al. (2006) who concluded that coma had a significant impact on letter contrast sensitivity. We also found that secondary astigmatism had a greater effect on the average fixation duration than defocus for the same rms amplitude, which agrees with the work of Fang et al. (2009) who found that fourth-order aberrations affected optical quality the most. This finding suggests that it is as important (if not more so for improving reading performance) to consider correcting secondary astigmatism as it is defocus if the rms amplitude is high.

Since our results differ from those of Applegate et al. and Oshika et al., we suggest that the impact of higher order aberrations on visual performance is likely to be dependent on the visual task and measure of performance.

We compared our results to our prediction of performance based on how similar letters are made to look in the presence of an aberration. We found a correlation between our measure of confusability and the increase in average fixation duration that could be described by an exponential curve. It is interesting to note that the relationships for defocus and secondary astigmatism are similar, whereas the relationship for coma is clearly different. We believe that this indicates a different source of performance loss for coma than for defocus and secondary astigmatism. For a given amplitude of aberration, the letters are more distinguishable for coma and the average fixation durations are shorter. However, in comparison to defocus and secondary astigmatism, the average fixation durations are longer for lower confusability values. Given the qualitative difference between the distortions produced by different aberrations, we suggest that performance loss in the case of coma might be more heavily influenced by saccade planning than by word identification (as we discuss below with respect to lexical frequency effects). Our data do not allow us to test this interpretation directly. With the current number of position measurements per target word, we were unable to reliably compare the changes in fixation location relative to the changes in the centers of gravity of the stimuli. In a recent experiment that we will report fully in a separate publication, we have shown that the relationship between our confusion metric and the contrast threshold for letter identification is the same for all three types of aberration. We suggest that this reinforces our interpretation because, in the absence of saccade planning and lateral masking effects that do not occur with single letter presentations, the performance loss for coma showed a dependence on amplitude that followed the same exponential form as it did for defocus and secondary astigmatism. We suggest that there are differences between the effects these aberrations have on performance, even for related visual tasks such as reading and letter recognition.

Lexical frequency effects can be taken as a direct index of the ease with which a word is identified. By considering how lexical frequency effects change between different types and amplitudes of aberration, we can assess the impact of different aberrations on word identification per se. An increased lexical frequency effect under conditions of aberration implies that low-frequency words (those that are less common and therefore most difficult to identify) became even more difficult to identify when their constituent orthography was degraded. At the level of individual letters, this relates directly to the confusion analysis we have performed. Orthographic familiarity affects the ease of word identification that, as previously stated, contributes significantly to the fixation duration on the current word.

The responses to the comprehension questions confirm that the increases in average fixation duration are not an artifact of the text being made completely illegible. Clearly, participants fully understood the sentences. The eye movement data additionally show that a lexical frequency effect was observed even in the presence of an aberration, indicating that subjects did successfully identify the words in the sentences (even if they had to revisit the words in order to do this). The interactions between either the type of aberration or the amplitude of the aberration and the lexical frequency of the word were significant for the total reading time but not for the gaze duration. In other words, the size of the lexical frequency effect was modulated by the type and amplitude of aberration only in relation to second-pass fixations on the word, not the first-pass fixations. For coma, the size of the lexical frequency effect was constant across all amplitudes of aberration, for both gaze durations and total reading times. This suggests that subjects successfully lexically identified target words in the presence of coma during the first-pass reading and did not need to spend a substantial amount of time refixating on them in order to identify them. For defocus and secondary astigmatism, the lexical frequency effect was unaffected by the amplitude of aberration during the first pass through the sentence. However, there was an increase in the size of the lexical frequency effect with increasing amplitude of these aberrations that occurred in the refixations on the target word. We interpret this as a failure to initially correctly identify the words of the sentence during the first visit. Presumably, subjects either preliminarily guessed the identity of the target word or instead left the target without settling on its identity, in order to fixate words downstream in the sentence to (potentially) recruit further linguistic information that could facilitate its identification. Either way, the data suggest that subjects made refixations on the target word to try and extract more orthographic information to either confirm or to unambiguously identify them. Clearly, the visual information obtained on the first attempt was insufficient for word identification. The total reading times show that the lexical frequency effect was differentially modulated by the nature and level of aberration. An increase in the total reading times of low-frequency words shows that when subjects were presented with words that are difficult to identify, the addition of an aberration increases this difficulty. This increased with the amplitude of aberration for defocus and secondary astigmatism indicating that these aberrations had a particularly strong effect on this psychological subprocess.

As already suggested in the discussion of global results, we predicted that defocus and secondary astigmatism should affect letter distinguishability more than coma. Global results show that the relationship between reading impairment and letter confusability is the same for defocus and secondary astigmatism but different for coma. We suggest that our local results reinforce this interpretation,

since they show that defocus and secondary astigmatism have an increasing impact on word identification, whereas coma does not. There is clearly an effect on global fixation durations due to coma although this effect does not present itself in local fixation durations and therefore word identification times. We tentatively conclude that coma has a stronger effect on saccade planning, whereas defocus and secondary astigmatism have a greater impact on linguistic processing. We speculate that spatial phase changes are the underlying cause of changes in reading strategy for text that had been distorted with secondary astigmatism and to a lesser extent defocus. These result in spurious resolution and the creation of sharp but false features. These false features provide incorrect orthographic cues that cause impaired lexical identification. Recent work by Ravikumar et al. (2010) has shown that phase errors caused by higher order aberrations have an effect on visual acuity and that this effect is small for coma.

Conclusion

We conclude that during a task such as reading, there is a significant impact from the addition of higher order aberrations that depends on both the type and magnitude of the aberration. We found that the greatest impairment to reading was caused by secondary astigmatism and the smallest impairment by coma. From our results, we also infer that secondary astigmatism and defocus significantly impacted word identification, whereas coma did not. The addition of defocus or secondary astigmatism results in less efficient extraction of orthographic information from the image, which in turn hinders efficient lexical identification. This encumbrance was greater for less familiar than more familiar words. Effects on reading performance for coma might be attributed to disruptions in saccade planning and to lateral masking rather than to word identification. In a recent experiment, we have shown that with single letter identification the performance loss caused by coma can be predicted by the same relationship as that for defocus and secondary astigmatism. This further suggests that the difference in the effects on reading performance caused by coma, rather than by defocus or secondary astigmatism, are caused by disruptions to processes that deal with letter strings, such as those that control eye movements and those that integrate signals from neighboring letters.

These aberrations have a clear impact on the retinal image and it is important to understand how this affects subsequent visual and linguistic processing. We have examined only three Zernike modes in this paper but intend to further this work by investigating other modes and combinations of modes. In particular, lateral masking effects can be tested by studying the different orientations of coma (Z_3^1 and Z_3^{-1}). We have shown that the effect of

aberrations can depend on the visual task in question. The effects of aberrations on other higher level visual tasks such as object recognition are also likely to be important.

Appendix A

Brief primer on optics and image formation

If the eye's optics were perfect, then light rays entering from an object that is localized at a single point would be focused on to the retina, forming the sharpest possible point image. The resolution of this image is limited only by diffraction, which is the process by which light bends as it passes through an aperture, in this case the eye's pupil. One way to describe light rays is to define a locus of points in which the rays are in phase with each other, which is known as a wavefront. This wavefront is a two-dimensional surface that is perpendicular to the direction of the light rays. The simplest example of this is for distant objects where the divergence of the light rays is so small that the rays appear parallel and the wavefront is a plane. The focusing components of an ideal eye would change the plane wavefront to a spherical wavefront so the light rays converge at the retina.

If the eye is myopic, it produces too much focal power and the light rays are brought to a focus in front of the retina leading to dioptric blur at the retina. Dioptric blur is a real optical distortion and is different to Gaussian and diffuse blur. The differences are characterized by the OTF, which describes the way in which contrast at each spatial frequency is transmitted by the optical system. A perfect OTF transmits contrast at all spatial frequencies, with no reduction in contrast. Gaussian and diffuse blur have OTFs that reduce contrast at high spatial frequencies but which leave the polarity of the contrast (white on black or black on white) unchanged. The OTF of dioptric blur oscillates between positive and negative values, changing the polarity of the contrast in the image as a function of spatial frequency. This can have drastic implications for information about the form of objects, and Movie 1 illustrates difficulties in letter recognition that can arise from dioptric blur.

Imperfections in the eye's optical components distort the wavefront, causing aberrations that cannot be corrected by adjusting focus. The light rays do not converge into a perfect focal point either at the retina or anywhere else. The image blur resulting from this does not look like dioptric blur and can cause complex changes in the retinal image. Examples of wavefronts and how they change the propagation of light rays are given in Figure A1. In this figure, the point spread function (PSF) is shown, which depicts what a point of light looks like after it has passed through the lens.

We quantify the distortion of the wavefront as the wavefront error, which is the rms deviation from a perfect
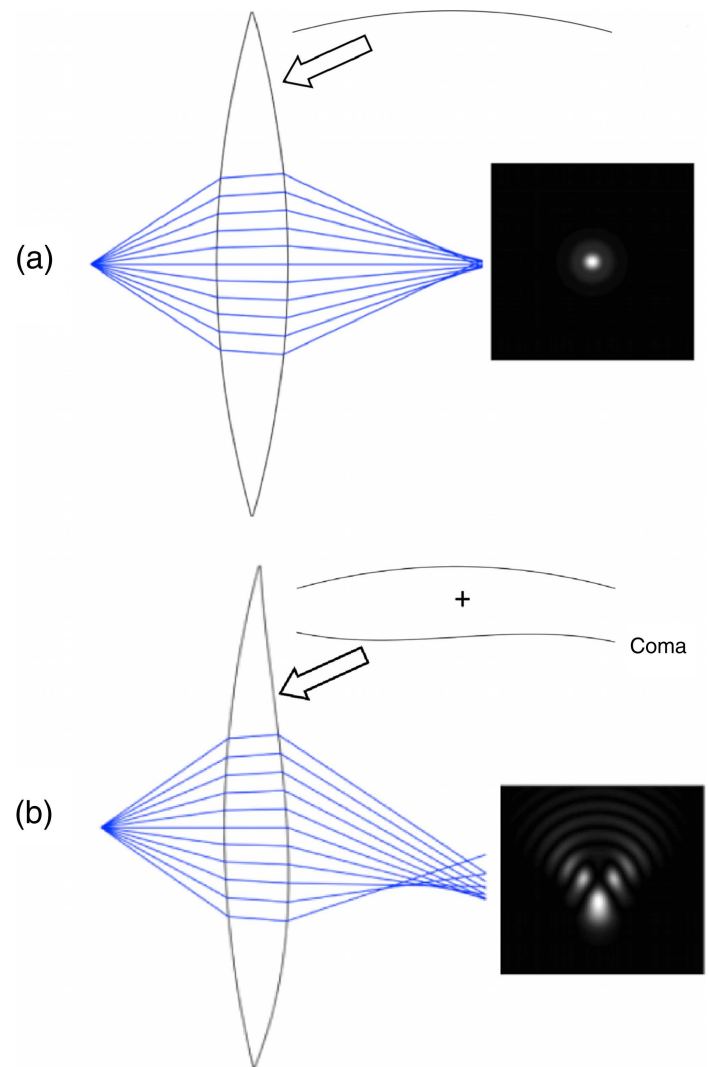

Figure A1. Ray trace showing imaging of a point through (a) a biconvex lens (Zernike defocus only) and (b) a lens with coma aberration. When light rays pass into a material with a higher refractive index, they are slowed and so their paths bend. In the case of an aberration-free lens, the light rays bend such that they are focused into a point with a blur disk due to diffraction. In this real ray trace, the rays do not converge perfectly due to the slightly spherical nature of the lens surfaces. In the lens with coma aberration, light is not focused into a point but is spread over an area causing the image to appear blurred.

plane wave. The distorted wavefront can be broken down into components by fitting it with a series of polynomials, a mathematical process that is analogous to breaking down an image into spatial frequency components. In ophthalmology, the wavefront is typically described by Zernike polynomials (Thibos, Applegate, Schwiegerling, & Webb, 2000), which are represented in Figure A2. Zernike polynomials are orthogonal and so we can understand the wavefront as containing a certain amplitude of a particular polynomial independent of the other polynomials. However, these polynomials are not orthog-
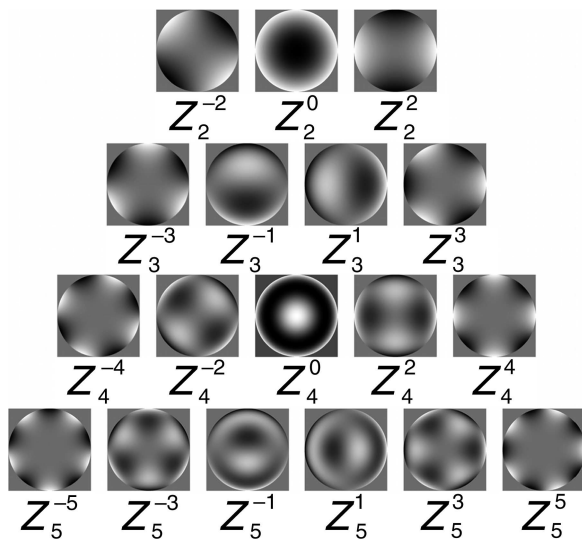
Figure A2. The Zernike pyramid showing wavefront modes from second radial order to fifth radial order. The orders are given as $Z_{\text{radial order}}^{\text{angular order}}$, where Z_2^0 represents defocus, Z_3^1 represents coma, and Z_4^2 represents secondary astigmatism. Spectacles correct for defocus and primary astigmatism, $Z_2^{\pm2}$, whereas higher order aberrations (those with radial order greater than 2) can only be corrected with complicated optical surfaces or with an adaptive optics system.

onal in terms of their effect on the image. We know that equal amplitudes of different polynomials do not produce the same qualitative amount of blur in the image nor do they have the same qualitative effect on visual form information (see Figure 1 for examples). To interpret the ocular wavefront measurement, it is important to know how each of these polynomials, and combinations of polynomials, affects the retinal image. Crucially, it is important to understand how these distortions affect visual performance, which is not necessarily predictable from existing metrics of retinal image quality.

Brief primer on eye movements and reading

When reading a language such as English, the eyes make a series of ballistic rotations, called saccades, each of which moves the point of gaze forward by approximately 6–9 character spaces (Rayner, 1998). Between saccades, the eyes remain quite still, fixating for approximately 250 ms. During fixations, visual information is extracted to allow for identification of the word under fixation (Liversedge & Findlay, 2000; Rayner, 1998; Starr & Rayner, 2001). Additionally, since visual input is suppressed during the eye movement, the landing point of the eye movement must be preprogrammed during the preceding fixation. This is achieved via the use of

parafoveal information about words to the right of fixation and, in particular, their length (McConkie & Rayner, 1976; Rayner & McConkie, 1976). Arguably, the duration of a fixation on a particular word is determined both by the extent to which it was parafoveally preprocessed prior to fixation, as well as the ease with which it is identified and interpreted within the context of the sentence or paragraph up to that point. Thus, the fixation duration on a word is affected by the characteristics of that word such as its lexical frequency (how common that word is in language). It is well known that words that are less common in language take longer to identify (Inhoff & Rayner, 1986; Rayner, 1998; Rayner & Duffy, 1986; Rayner, Liversedge, & White, 2006). Characteristics such as word length and the contextual predictability of the word also increase fixation durations, and longer or less predictable words may also require an additional fixation. Another important point to understand is that readers do not process text symmetrically about the point of fixation. It is the case that visual acuity reduces symmetrically with increased horizontal distance from the fovea (the small area of the retina that delivers detailed information to the human visual processing system). Intuitively, therefore, one might imagine that the same amount of text would be processed to the left and to the right of the point of fixation. However, this is not the case. The perceptual span (Rayner, 1975) is the number of characters a reader can process during a fixation and is 14–16 to the right of fixation (Den Buurman, Roersema, & Gerrissen, 1981; McConkie & Rayner, 1975; Rayner & Bertera, 1979; Rayner & Duffy, 1986; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981) and 3–4 to the left (McConkie & Rayner, 1976; Rayner, Well, & Pollatsek, 1980; Underwood & McConkie, 1985). The asymmetry of the perceptual span reflects the importance of attention in reading and how this is centrally associated with what we are processing moment to moment during any particular fixation.

It is also known that word spacing affects eye movements and fixation durations, and when the spaces are altered in English text, it is much harder to read than when they are not (Fisher, 1976; Malt & Seamon, 1978; Morris, Rayner, & Pollatsek, 1990; Pollatsek & Rayner, 1982; Rayner, 1998; Spragins, Lefton, & Fisher, 1976). This could be due to lateral masking effects, in which a letter becomes difficult to identify due to its close proximity to other letters or due to the removal of word boundary information that hinders saccadic targeting (saccades are targeted toward the middle of words; e.g., see White & Liversedge, 2004). Physical blurring of letters due to optical imperfections can also affect word spacing such that the spatial extents of words overlap. Disruption to saccade planning, in turn, has consequences for fixation durations, particularly if saccade landing positions are not in the optimal place within a word for its identification. Often when the eyes land in a non-optimal position within
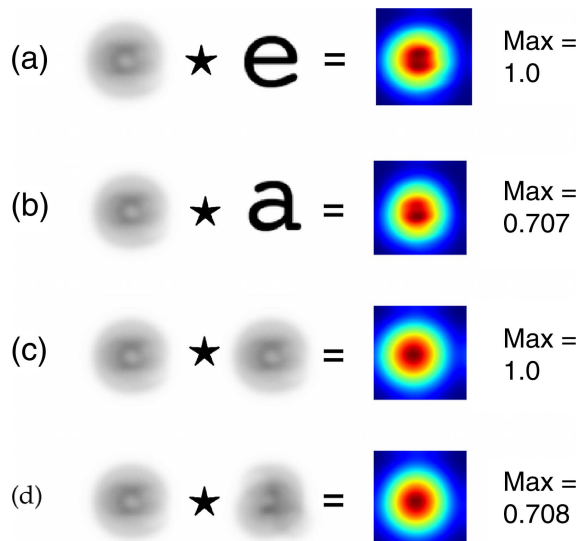
Figure B1. Examples of the cross-correlations (*) performed when calculating the confusion matrix. The color maps used to produce these images span the minimum to maximum range within each image. The value on the right indicates the maximum value of the cross-correlation, normalized such that the confusion matrix has a value of one along the diagonal. On the left is a letter e with 0.4-μm defocus simulated, which could easily be confused with a letter a. This is cross-correlated with (a) an unchanged letter e and (b) an unchanged letter a, where the peak values of the cross-correlation produce results for the confusion matrix in Figure B4a). It is also cross-correlated with (c) itself and (d) a letter a with 0.4-μm defocus simulated, where the peak values produce results for the confusion matrix in Figure B4b.

a word, a corrective eye movement is necessary, resulting in increased processing time on a word.

Considering the average fixation duration (i.e., total sentence reading time in relation to the number of fixations made over a whole sentence) provides a measure of overall reading performance. However, it is also possible to consider effects associated with the processing of specific words, and whether such effects occur due to the linguistic characteristics of that word (e.g., its frequency within the language), or because of its visual appearance (e.g., whether it is or is not visually degraded). For example, if two sentences are presented such that the only difference between them is a single high or low lexical frequency target word (matched in length and predictability), the difference in fixation duration on that word may be attributed to the difference in processing time required to identify it. Furthermore, if the text is blurred, and the blur changes the appearance of letters in the word such that they look like other letters (i.e., they become more confusable), then visual processing of the word will be more difficult and subsequent word identification will be disrupted. As a consequence, fixation durations will be increasingly affected. This may be particularly problematic if the word is similar to many

other words in terms of its constituent letters and has many lexical neighbors (words that differ by a single letter; Coltheart, Davelaar, Jonasson, & Besner, 1977).

Appendix B

Letter confusion analysis

We investigated the potential for letters to be confused by a subject. We have chosen not to use feature matching techniques for this analysis since our stimuli are distorted significantly and are therefore much less likely to contain identifiable features. Instead, we compared letters in the font used in our experiment by performing cross-correlations, giving a measure of similarity based on linear transforms of the stimuli. An example of this is given in Figure B1. For each Zernike mode, an amount of aberration was applied in simulation to each letter of the alphabet. The resulting letters were compared in two ways. First, we cross-correlated aberrated letters with other aberrated letters to represent a subject trying to identify letters by looking for differences between them. Second, we cross-correlated aberrated letters with non-aberrated letters to represent a subject trying to identify letters by using known letter shapes. The maximum value of the cross-correlation was taken to be a measure of "confusability" as this procedure allows for positional effects. These values were entered in a 26-by-26 matrix, which was subsequently
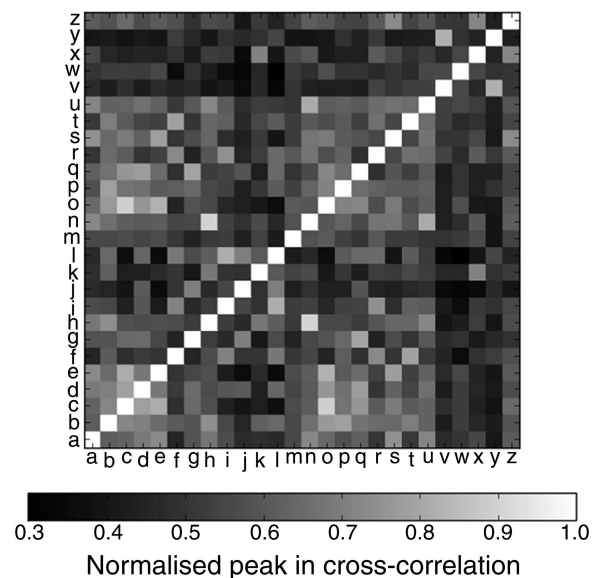

Figure B2. Confusion matrix comparing each letter of the alphabet with every other letter when no aberration is applied to letters in Courier font. The values in the matrix are the peak value of the cross-correlation of one letter with another and the matrix is normalized such that the values along the diagonal are 1. The scale is such that a value of 1 indicates that the letters are identical.
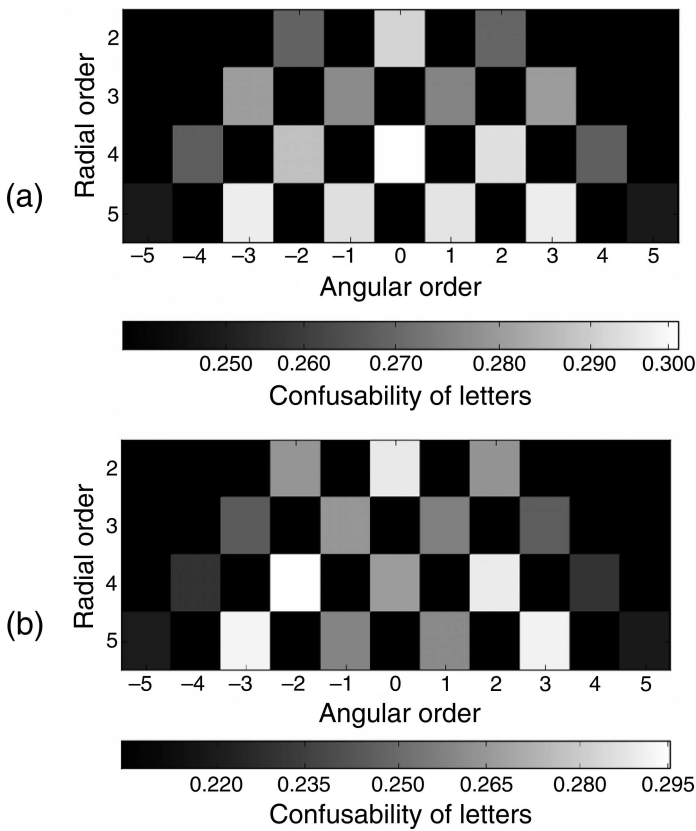
Figure B3. Representation of the Zernike pyramid indicating the confusability of letters caused by each mode where a value of 1 indicates that letters are completely indistinguishable. Higher order aberrations are those with a radial order greater than 2. In (a), each aberrated letter is compared with the set of aberrated letters representing a subject attempting to recognize letters by looking for differences between the letters they are presented with. In (b), each aberrated letter is compared with the set of non-aberrated letters representing a subject attempting to recognize letters by making comparisons with known letter shapes. The Zernike modes that cause the most confusion between aberrated and non-aberrated letters are Z_2^0 (defocus), $Z_4^{\pm2}$ (secondary astigmatism), and $Z_5^{\pm3}$ (secondary trefoil).

normalized such that the values along the diagonal equaled one by multiplying matrix elements by

$$n(x, y) = \frac{1}{\sqrt{c(x,x) \times c(y,y)}}, \qquad (B1)$$

where $n(x, y)$ is the normalization constant for matrix element (x, y) and $c(x, x)$ and $c(y, y)$ refer to the diagonal matrix elements prior to normalization for the letters corresponding to x and y. An example of this normalized confusion matrix, calculated for letters with no aberration, is shown in Figure B2.

In order to derive a single confusability value, we took a weighted mean of the matrix. These weights accounted for the probabilities of letters occurring in language so that,

for example, distinguishability of a letter e has a greater effect on the result than that of a letter z. For this, we used the letter counts of Jones and Mewhort (2004), which used approximately 183 million English words and counted upper and lowercase letters separately. This confusability value was used as a metric for the extent to which a particular mode of aberration made letters less distinguishable. These values are represented in the conventional Zernike pyramid in Figure B4. Based on these results, we chose to investigate secondary astigmatism (Z_4^2) since it produced the greatest confusability factor when comparing aberrated and non-aberrated letters.

Letter position analysis

Zernike aberrations that have asymmetric PSFs, such as coma, will cause a shift in the center of an image. To test the implications of this, each Zernike mode was applied to the letters of the alphabet in simulation and the centers of gravity (the center of the letter determined by intensity) of resulting images were calculated for the direction parallel to the line of text. The differences between these centers of gravity and those of corresponding non-aberrated letters express the changes in letter spacing that occur with an aberration. For each mode, the standard deviation of this difference was calculated over all 26 letters. These results are represented in Figure B3. Coma (Z_3^1) caused the letter
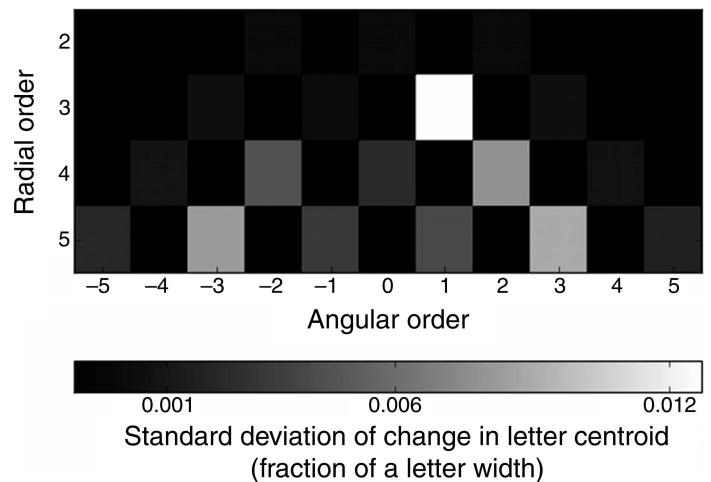

Figure B4. Representation of the Zernike pyramid indicating the standard deviation of the shift in the centers of gravity of letters (from the no aberration condition) caused by 0.4 μm rms of each mode. Shifts are expressed as a fraction of the average width of a letter. The Zernike mode that causes the most variation in letter position in a direction parallel to the text (and therefore the greatest variation in the spacing between letters) is coma (Z_3^1) with a standard deviation of 1.3% of the width of a letter. Secondary trefoil (4.2% for Z_5^{-3} and 3.1% for Z_5^3) and secondary astigmatism (4.5% for Z_4^{-2} and 7.8% for Z_4^2) cause smaller variability in the centers of letters.

center to vary the most, and although the standard deviation was relatively small (1.3% of the average width of a letter), changes in the centers of letters of up to 12.3% the width of a letter could occur in the text. Secondary trefoil and secondary astigmatism also caused some variability in the center of gravity but only about half as much as coma ([Figure B4](#)).

Acknowledgments

Commercial relationships: none.
Corresponding author: Laura K. Young.
Email: l.k.young@durham.ac.uk.
Address: Department of Physics, Durham University, Science Laboratories, South Road, Durham, DH1 3LE, UK.

References

Aggarwala, K. R., Nowbotsing, S., & Kruger, P. B. (1995). Accommodation to monochromatic and white-light targets. *Investigative Ophthalmology and Visual Science, 36,* 2695–2705. [[PubMed](#)] [[Article](#)]

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. [[PubMed](#)]

Applegate, R. A., Marsack, J. D., Ramos, R., & Sarver, E. J. (2003). Interaction between aberrations to improve or reduce visual performance. *Journal of Cataract and Refractive Surgery, 29,* 1487–1495. [[PubMed](#)]

Applegate, R. A., Sarver, E. J., & Khemsara, V. (2002). Are all aberrations equal? *Journal of Refractive Surgery, 18,* S556–S562. [[PubMed](#)]

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](#)]

Chung, S. T., Jarvis, S. H., & Cheung, S. H. (2007). The effect of dioptric blur on reading performance. *Vision Research, 47,* 1584–1594. [[PubMed](#)]

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

Den Buurman, R., Roersema, T., & Gerrissen, J. F. (1981). Eye movements and the perceptual span in reading. *Reading Research Quarterly, 16,* 227–235.

Drieghe, D. (2008). Foveal processing and word skipping during reading. *Psychonomic Bulletin and Review, 15,* 856–860. [[PubMed](#)]

Fang, L., Wang, Y., & He, X. (2009). Evaluation of optical quality in white light from wavefront aberrations for a myopic population of human eyes. *Clinical and Experimental Optometry, 92,* 313–319. [[PubMed](#)]

Fisher, D. F. (1976). Spatial factors in reading and search: The case for space. In R. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes* (pp. 417–427). Hillsdale, N.J.: L. Erlbaum.

Gobbe, M., & Guillon, M. (2005). Corneal wavefront aberration measurements to detect keratoconus patients. *Contact Lens & Anterior Eye, 28,* 57–66. [[PubMed](#)] [[Article](#)]

Guirao, A., Redondo, M., & Artal, P. (2000). Optical aberrations of the human cornea as a function of age. *Journal of the Optical Society of America A, 17,* 1697–1702. [[PubMed](#)]

Heath, G. G. (1956). The influence of visual acuity on accommodative responses of the eye. *American Journal of Optometry, 33,* 513–524. [[PubMed](#)]

Howell, D. C. (1992). *Statistical methods for psychology* (3rd ed.). California: Duxbury Press.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics, 40,* 431–439. [[PubMed](#)]

Jones, M. N., & Mewhort, D. J. (2004). Case-sensitive letter and bigram frequency counts from large-scale English corpora. *Behavior Research Methods, Instruments and Computers, 36,* 388–396. [[PubMed](#)]

Juhasz, B. J., Liversedge, S. P., White, S. J., & Rayner, K. (2006). Binocular coordination of the eyes during reading: Word frequency and case alternation affect fixation duration but not fixation disparity. *Quarterly Journal of Experimental Psychology, 59,* 1614–1625. [[PubMed](#)]

Kruger, P. B., Mathews, S., Aggarwala, K. R., & Sanchez, N. (1993). Chromatic aberration and ocular focus: Fincham revisited. *Vision Research, 33,* 1397–1411. [[PubMed](#)] [[Article](#)]

Kruger, P. B., & Pola, J. (1986). Stimuli for accommodation: Blur, chromatic aberration and size. *Vision Research, 26,* 957–971. [[PubMed](#)]

Legge, G. E. (2007). *Psychophysics of reading in normal and low vision*. Hillsdale, N.J.: L. Erlbaum.

Legge, G. E., Pelli, D. G., Rubin, G. S., & Schleske, M. M. (1985). Psychophysics of reading: I. Normal vision. *Vision Research, 25,* 239–252. [PubMed]

Li, S., Xiong, Y., Li, J., Wang, N., Dai, Y., Xue, L., et al. (2009). Effects of monochromatic aberration on visual acuity using adaptive optics. *Optometry and Vision Science, 86,* 868–874. [PubMed]

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Science, 4,* 6–14. [PubMed]

Lu, F., Wu, J., Shen, Y., Qu, J., Wang, Q., Xu, C., et al. (2008). On the compensation of horizontal coma aberrations in young human eyes. *Ophthalmic and Physiological Optics, 28,* 277–282. [PubMed]

Malt, B. C., & Seamon, J. G. (1978). Peripheral and cognitive components of eye guidance in filled-space reading. *Perception & Psychophysics, 23,* 399–402. [PubMed]

McConkie, G., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society, 8,* 365–368.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17,* 578–586.

Morris, R. K., Rayner, K., & Pollatsek, A. (1990). Eye movement guidance in reading: The role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 268–281. [PubMed]

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eye tracking data. *Behavior Research Methods, 42,* 188–204. [PubMed]

Oshika, T., Okamoto, C., Samejima, T., Tokunaga, T., & Miyata, K. (2006). Contrast sensitivity function and ocular higher-order aberrations in normal human eyes. *Ophthalmology, 113,* 1807–1812. [PubMed]

Pollatsek, A., & Rayner, K. (1982). Eye movement control in reading: The role of word boundaries. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 817–833.

Ravikumar, S., Bradley, A., & Thibos, L. (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]

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422. [PubMed]

Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science, 206,* 468–469. [PubMed]

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity and lexical ambiguity. *Memory & Cognition, 14,* 191–201. [PubMed]

Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38,* 1129–1144. [PubMed]

Rayner, K., Inhoff, A. W., Morrison, R., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 167–179. [PubMed]

Rayner, K., Liversedge, S. P., & White, S. J. (2006). Eye movements when reading disappearing text: The importance of the word to the right of fixation. *Vision Research, 46,* 310–323. [PubMed]

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research, 16,* 829–837. [PubMed]

Rayner, K., Well, A. D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception & Psychophysics, 27,* 537–544. [PubMed]

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the E-Z Reader model. *Psychological Science, 17,* 742–746. [PubMed]

Roorda, A. (2011). Adaptive optics for studying visual function: A comprehensive review. *Journal of Vision, 11*(5):6, 1–21, http://www.journalofvision.org/content/11/5/6, doi:10.1167/11.5.6. [PubMed] [Article]

Savitzky, A., & Golay, M. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36,* 1627–1639.

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

Spragins, A. B., Lefton, L. A., & Fisher, D. F. (1976). Eye movements while reading and searching spatially transformed text: A developmental examination. *Memory & Cognition, 4,* 36–42. [PubMed]

Starr, M. S., & Rayner, K. (2001). Eye movements during reading: Some current controversies. *Trends in Cognitive Science, 5,* 156–163. [PubMed]

Thibos, L. N., Applegate, R. A., Schwiegerling, J., Webb, R. H., & VSIA Standards Taskforce Members (2000). Standards for reporting the optical aberrations of eyes. In V. Lakshminarayanan (Ed.), *OSA Trends in optics and photonics*, Vision science and its applications (vol. 35, pp. 232–244). Optical Society of America, Washington D.C.

Thibos, L. N., Hong, X., Bradley, A., & Cheng, X. (2002). Statistical variation of aberration structure and image quality in a normal population of healthy eyes. *Journal of the Optical Society of America A, 19,* 2329–2348. [PubMed]

Thorn, F., & Thorn, S. (1996). Television captions for hearing-impaired people: A study of key factors that affect reading performance. *Human Factors, 38,* 452–463. [PubMed]

Underwood, N. R., & McConkie, G. W. (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly, 20,* 153–162.

Villa, C., Gutiérrez, R., Jiménez, J. R., & González-Méijome, J. M. (2007). Night vision disturbances after successful Lasik surgery. *British Journal of Ophthalmology, 91,* 1031–1037. [PubMed] [Article]

White, S. J., & Liversedge, S. P. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology, 16,* 52–78.