Journal of Experimental Psychology: Human Perception and Performance

The Distribution of Fixation Durations During Reading: Effects of Stimulus Quality

Sarah J. White and Adrian Staub

Online First Publication, September 12, 2011. doi: 10.1037/a0025338

CITATION

White, S. J., & Staub, A. (2011, September 12). The Distribution of Fixation Durations During Reading: Effects of Stimulus Quality. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication. doi: 10.1037/a0025338

Journal of Experimental Psychology:
Human Perception and Performance

© 2011 American Psychological Association
0096-1523/11/$12.00 DOI: 10.1037/a0025338

The Distribution of Fixation Durations During Reading: Effects of Stimulus Quality

Sarah J. White
University of Leicester

Adrian Staub
University of Massachusetts, Amherst

Participants' eye movements were recorded as they read single sentences presented normally, presented entirely in faint text, or presented normally except for a single faint word. Fixations were longer when the entire sentence was faint than when the sentence was presented normally. In addition, fixations were much longer on a single faint word embedded in normal text, compared to when the entire sentence was faint. The primary aim of the study was to examine the influence of stimulus quality on the distribution of fixation durations. Ex-Gaussian fitting revealed that stimulus quality affected the mean of the Normal component, but in contrast to results from single-word tasks (Plourde & Besner, 1997), stimulus quality did not affect the exponential component, regardless of whether one or all words were faint. The results also contrast with the finding (Staub, White, Drieghe, Hollway, & Rayner, 2010) that the word frequency effect on fixation durations is an effect on both of the critical distributional parameters. These findings are argued to have implications for the interpretation of the role of stimulus quality in word recognition, and for models of eye movement control in reading.

Keywords: reading, eye movements, ex-Gaussian, fixation durations, degradation, stimulus quality

In order to understand the processes underlying word recognition, it is important to understand how word recognition is influenced by manipulations of the physical properties of the visual stimulus. To this end, many studies have examined the effect of stimulus quality (e.g., visual contrast) on word recognition, both in single-word response tasks such as lexical decision and naming (e.g., Becker & Killion, 1977) and by recording eye movement behavior during natural reading (e.g., Reingold & Rayner, 2006). Unsurprisingly, both response times (RTs) in single-word tasks and fixation durations in reading are increased when a word is visually degraded.

In addition, a number of recent studies in the single-word recognition literature have examined the influence of stimulus quality on the distribution of RTs (e.g., Plourde & Besner, 1997; Yap & Balota, 2007; Yap, Balota, Tse, & Besner, 2008). These studies have used ex-Gaussian fitting (Balota & Yap, 2011; Ratcliff, 1979; this method is described in detail below) to determine how the increase in mean RT that is observed when a word stimulus is visually degraded is manifested in terms of the location and shape of individual subjects' RT distributions. Two effects have consistently emerged. First, stimulus degradation increases the duration of most responses, resulting in a shift to the right of the entire distribution of RTs. Second, stimulus degradation increases the frequency of especially slow responses, resulting in an RT distribution with a more pronounced right tail.

The present study investigates whether these distributional effects of stimulus quality generalize to fixation durations in normal reading. Does stimulus degradation increase the duration of most eye fixations? Does it increase the frequency of especially long fixations?

The study is motivated by two related considerations. First, of the two distributional effects of stimulus quality obtained in single-word recognition studies (i.e., a shifting of the distribution of RTs, and an increase in the weight of the right tail), it is arguably the case that the latter effect would be unexpected to arise in normal reading. It is plausible to assume that a manipulation of stimulus quality has an effect on word recognition primarily by lengthening an early perceptual stage of processing, perhaps devoted primarily to letter identification. Reingold and Rayner (2006) suggest that the effect of stimulus degradation in normal reading is to "disrupt early encoding of visual and orthographic features of the fixated word" (p. 742), and present eye movement data consistent with the claim that stimulus degradation has an effect only on a very early stage of processing. (This evidence is discussed in detail below.) On the assumption that this early perceptual encoding is mandatory, it is to be expected that stimulus degradation should add some increment of processing time to most or all eye fixations, resulting in a distributional shift. However, on the assumption that this early encoding is fast and automatic, it is not obvious why stimulus degradation should increase the duration of some fixations very dramatically, resulting in a distribution of fixa-

Sarah J. White, School of Psychology, University of Leicester; Adrian Staub, Department of Psychology, University of Massachusetts, Amherst.

We thank Louise Smith and Gurpreet Kaur for their assistance with data collection. Portions of this work were completed whilst the first author was on Study Leave granted by the University of Leicester. We also thank Denis Drieghe and Sachiko Kinoshita for their comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Sarah J. White, School of Psychology, University of Leicester, Lancaster Road, Leicester, LE1 9HN. E-mail: s.j.white@le.ac.uk or s.j.white@dunelm.org.uk

tions with an especially pronounced right tail.[1] The finding from the single-word literature that stimulus degradation does have this effect on RT distributions is, arguably, unexpected according to the view espoused by Reingold and Rayner. Thus, the present study tests the hypothesis that in normal reading, unlike in single-word tasks, the effect of stimulus degradation on fixation durations may be due entirely to an effect on the location of the distribution of fixation durations, with no effect on the weight of the right tail. If this pattern is indeed observed, the effect on the right tail that is observed in single-word studies will require an explanation in terms of the specific features of the tasks in question.

The second motivation for the present study is simply that, in the service of a more complete understanding of eye movement control in reading, it is of interest to know how the effect of a manipulation of the stimulus' physical properties on eye movements compares to effects of linguistic variables such as word frequency (Rayner & Duffy, 1986) and predictability (Ehrlich & Rayner, 1981). Recent studies have examined the effects of word frequency (Staub, White, Drieghe, Hollway, & Rayner, 2010) and word predictability (Staub, 2011a) on fixation duration distributions in reading; thus, it is now possible to ask whether the effect of stimulus quality is qualitatively similar to, or different from, the effects of these two variables. The answer to this question has the potential to inform models of eye movement control in reading such as E-Z Reader (Reichle, Rayner, & Pollatsek, 2003) and SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005).

The remainder of this introduction focuses on previous empirical work. It first provides a review of ex-Gaussian distributional approaches to understanding visual word recognition, and then discusses in more detail studies showing effects of stimulus quality on recognition of words in single-word tasks. Finally, it reviews the previous eye movement literature examining the influence of stimulus quality.

Using Ex-Gaussian Analyses to Understand Visual Word Recognition

Ratcliff (1979) established that empirical distributions of response times are well fit by the ex-Gaussian distribution, and a number of studies in the single-word recognition literature have subsequently adopted this technique (Andrews & Heathcote, 2001; Balota & Spieler, 1999; Balota & Yap, 2011; Balota, Yap, Cortese, & Watson, 2008; Kinoshita & Hunt, 2008; Plourde & Besner, 1997; Yap & Balota, 2007; Yap, Balota, Cortese, & Watson, 2006; Yap et al., 2008; Yap, Tse, & Balota, 2009). Staub et al. (2010) and Staub (2011a) established that fixation durations in reading are also well fit by the ex-Gaussian distribution. The ex-Gaussian distribution is the convolution of a Normal and an exponential distribution and can be characterized by three parameters: μ represents the location (mean) of the Normal component, σ represents the standard deviation of the Normal component, and τ reflects the mean and standard deviation of the exponential component. By fitting the ex-Gaussian distribution to the data from each participant in each condition, and then comparing the parameter values across conditions, we can establish whether differences in overall means are due to a shift in the Normal component (μ) and/or a specific difference in the slow tail of the distribution (τ).

A crucial point is that the ex-Gaussian analyses provide a much more complete description of the data than single mean values (Heathcote, Popiel, & Mewhort, 1991). The ex-Gaussian distribution is employed not only because it provides a good fit to distributions of response times (other distributions may do this too; see Van Zandt, 2000), but because it allows testing of whether manipulations have an effect on fixations of all durations (μ) or a particular effect on a subset of long fixations (τ). Furthermore, in the present study it was important to employ ex-Gaussian analyses in order to provide a direct comparison with similar methods employed in single-word response time studies, as outlined below. It must also be emphasized that the ex-Gaussian distribution is employed here to characterize the location and shape of the distribution of fixation durations, and that any attempt to link particular parameters to specific underlying processing mechanisms is speculative (see Matzke & Wagenmakers, 2009). Note also that other researchers have taken alternative approaches to examining the distribution of fixation durations during reading, such as comparing group distributions in the form of hazard functions, which indicate the probability of a fixation ending within a series of possible time periods (Feng, 2009; McConkie, Kerr, & Dyre, 1994; Yang & McConkie, 2001).

Staub et al. (2010) utilized datasets from White (2008) and Drieghe, Rayner, and Pollatsek (2008) to examine the effect of word frequency on the distribution of fixation durations during reading. Staub et al. showed that for fixation durations on low frequency words, the distribution was shifted to the right, and there was a particular increase in the number of long fixations, compared to fixation durations on high-frequency words (i.e., effects on both μ and τ). Staub et al.'s results may reflect aspects of the word-recognition process. For example, word frequency may influence the duration of a processing stage that is operative on all encounters with a word, while also influencing a processing stage that is operative primarily when a fixation is long (causing the effect on τ). It is important to note that Staub et al.'s results are in line with findings from single-word response tasks that show effects of word frequency on both μ and τ for RT distributions (Andrews & Heathcote, 2001; Balota & Spieler, 1999; Plourde & Besner, 1997; Yap & Balota, 2007). The similar effects of word frequency for single-word response tasks and fixation durations in natural reading provide support for the notion that single-word response tasks can provide a good correlate of the natural word recognition processes. Moreover, Staub (2011a) found that the effect of lexical predictability on fixation durations is entirely a μ effect, which may be seen as consistent with the finding from the single-word literature that semantic priming generally influences only μ (Balota et al., 2008; Yap et al., 2009). The present study provides a further test of whether distributional effects in the single-word recognition literature generalize to natural reading, by examining whether stimulus quality influences both the μ and τ parameters of fixation duration distributions in reading.

[1] Note that it is possible that differences in the weight of the right tail (τ) could be triggered by an early stage of processing. The focus here is to use the τ parameter to assess whether a subset of fixations are particularly influenced by stimulus quality, rather than trying to link ex-Gaussian parameters to particular processing stages/time periods.

The Effect of Stimulus Quality on Recognition of Single Words

Many studies using tasks based on responses to single words have demonstrated that stimulus degradation impairs word processing regardless of the nature of the stimulus quality manipulation (e.g., contrast reduction, Becker & Killion, 1977; random dots, Meyer, Schvanedveldt, & Ruddy, 1975; alternating target and symbol mask, Balota et al., 2008; Yap et al., 2008). However the nature of the stimulus quality manipulation can affect the results; for example Stanovich and West (1983) demonstrated that the type of stimulus quality manipulation (stimulus intensity vs. inserted asterisks) influenced the nature of the interaction between stimulus quality and context. In the present study, contrast reduction was employed, as it is perhaps one of the more naturalistic methods of manipulating stimulus quality, and has been previously employed in the eye movement literature.

In order to investigate the impact of stimulus quality on particular aspects of the word recognition process, a range of studies using single words have tested whether the effect of stimulus quality on word recognition is modulated by other factors such as word frequency. However, recent work has highlighted that the pattern of effects is also influenced by the nature of the task. For example, studies have shown additive effects of stimulus quality and word frequency in lexical decision (e.g., Becker & Killion, 1977; Stanners, Jastrzembski, & Westbrook, 1975), whereas more recent studies have shown interactive effects for reading aloud (O'Malley, Reynolds, & Besner, 2007; Yap & Balota, 2007) and semantic classification (Yap & Balota, 2007). Generally, the differences in results across a variety of single-word tasks emphasize the importance of examining effects of stimulus quality across a range of tasks. This also provides strong justification for testing effects of stimulus quality on word processing during natural reading.

The focus of the present study is the influence of stimulus quality on the distribution of responses. Plourde and Besner (1997) showed that lexical decision RTs to stimuli with reduced contrast were both shifted to the right and had a particularly high frequency of long RTs (that is, effects on both μ and τ) compared to higher contrast stimuli. Yap and Balota (2007) and Yap et al. (2008) manipulated stimulus quality by presenting the target word and a random letter string in rapid alternate succession. Similar to the results of Plourde and Besner, both Yap and Balota and Yap et al. showed effects of stimulus quality on μ and τ across a series of single-word and nonword tasks.

However, similar to the studies discussed above showing differences in effects of stimulus quality on mean response times across tasks, work on the effects of stimulus quality on the distribution of RTs has also revealed differences as a function of the nature of the task. Although there were additive effects of stimulus quality and word frequency on means and ex-Gaussian parameters in lexical decision with legal nonwords (Plourde & Besner, 1997; Yap & Balota, 2007) there were interactive effects when the nonwords were pseudohomophones such as "brane" (Yap et al., 2008). Yap and Balota (2007) also found interactive effects of these variables on ex-Gaussian parameters in naming and semantic categorization tasks. Clearly, the variability in effects of stimulus quality across different tasks, for both mean response times and the distribution of response times, indicates that the nature of the task can have an important role in determining how a stimulus quality effect is manifested.

The Effect of Stimulus Quality on Eye Movements in Reading

Eye movement behavior during reading has been shown to be influenced by the legibility of the text, both in terms of font type (Paterson & Tinker, 1947; Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006; Slattery & Rayner, 2010; Tinker & Paterson, 1955) and contrast (Hackman & Tinker, 1957; Paterson & Tinker, 1944, 1947; Tinker & Paterson, 1955). Presenting text at low contrast reduces reading rate (Legge, Rubin, & Luebker, 1987), such that text is read with more and longer fixations (Legge, Ahn, Klitz, & Luebker, 1997).

More recent studies (Drieghe, 2008; Reingold & Rayner, 2006; Wang & Inhoff, 2010) have investigated the effect of stimulus quality in more detail by manipulating the contrast of a single critical word within a sentence. When the stimulus quality of only one word is manipulated, the time course of the influence of stimulus quality on reading behavior (influences on previous and subsequent fixations) can be much more finely assessed. Reingold and Rayner (2006) showed that reading times were longer on faint compared to normally presented words, and that this effect was largely localized to the faint word, with little effect on reading time on the following word.

Reingold and Rayner (2006) argued that this result supports the assumption that stimulus quality influences only the initial stage of word processing known as L_1 within the E-Z Reader model of eye movement control during reading (Reichle et al., 2003). In the E-Z Reader model, word recognition occurs in two stages, L_1 and L_2. There is some debate about the nature of word processing that may occur in these stages, but it is assumed that while the duration of L_1 might be modulated by perceptual factors (such as acuity), the input to the subsequent word processing stage, L_2, is not (Reichle, Pollatsek, & Rayner, 2007). The completion of L_1 triggers the beginning of saccade programming to the next word. Word processing continues while the saccade is being programmed until the L_2 stage has completed. Attention then moves to the following word, and processing of L_1 for the following word commences (in the form of parafoveal preview until the saccade is executed). The partitioning of word processing into these two stages results in a prediction of a dissociation: Factors that influence L_1 for a given word should influence reading time on that word, but not on the next word; but factors that influence L_2 should influence reading time on the next word, by modulating the amount of parafoveal preprocessing of that word. Hence, if perceptual factors influence only the first stage of word processing, L_1, then these factors should influence reading time on the word in question, but not on the next word. These predictions are consistent with the results of Reingold and Rayner (2006).

The results obtained by Drieghe (2008) and Wang and Inhoff (2010) are generally similar. Drieghe (2008) also found the effect of stimulus quality to be localized, even with shorter words. Wang and Inhoff (2010) again showed a clear effect of contrast on critical words, and although contrast did affect first fixation duration on the following word, there was no effect of stimulus quality on gaze duration (the sum of all fixation durations on the first encounter with the word) on the following word.

Drieghe (2008) also showed effects of the stimulus quality of a word on the probability of skipping the following word. Words were more likely to be skipped when they followed normal, compared to faint, words. However, given that effects of stimulus quality did not spill over to the following word (as shown by reading times), Drieghe argued that foveal difficulty modulated skipping probability directly, rather than influencing processing of the following word. In line with this suggestion, Wang and Inhoff (2010) showed no modulation by stimulus quality of the parafoveal preview of the following word. It is worth noting that other variables thought to affect later stages of word recognition do clearly influence processing of the following word; for example, both word frequency (Henderson & Ferreira, 1990; White, Rayner, & Liversedge, 2005) and case alternation (Wang & Inhoff, 2010) do modulate parafoveal processing of the following word.

Experiment

The present experiment tests whether the influence of stimulus quality on the distribution of responses in single word tasks (effects on both μ and τ) also holds for fixation durations during natural reading. If stimulus quality affects fixation durations in the same manner as RTs, this would support the conclusion that both the μ and τ effects reflect intrinsic aspects of perceptual and word recognition processes. However, if stimulus quality affects only the location of fixation duration distributions, and not the exponential component, this might indicate that the tendency for degraded stimuli in single-word tasks to produce some especially long RTs reflects task-specific processes.

When only a critical word has a different visual format, it is possible that this may act to highlight the word, which may influence how it is processed. Therefore, in the present study, the effect of stimulus quality was compared when either the entire sentence or only a critical word was visually degraded (faint), compared to a normal presentation condition. The studies that have investigated effects of stimulus quality for a single critical word within a sentence have focused on testing the time course of effects in terms of influences on the critical word and preceding and following words. To aid comparison with this work, the Results include the section "Local Analyses: Reading Behavior Before and After the Critical Word."

Note that the ex-Gaussian techniques employed here involve fitting distributions separately for each participant for each condition, which requires a large number of data points (Heathcote, Brown, & Mewhort, 2002; Speckman & Rouder, 2004). Consequently it was necessary to design the experiment such that the maximum number of observations per participant per condition (60) was much higher than that in most studies of eye movement behavior during reading. In order to minimize the possibility that the stimulus quality of the critical word might provide any linguistic cues, the characteristics of the critical word varied widely (see details below). A wide range of critical word types were also adopted in order that the results might be generalized to processing of many types of words.

Method

Participants. Thirty undergraduates at the University of Leicester participated for course credit. Participants were native English speakers with normal or corrected-to-normal vision and were naïve in relation to the purpose of the experiment.

Apparatus. Eye movements were monitored using an EyeLink 1000 eye tracker (SR Research Ltd.). Pupil location was sampled at a rate of 1000Hz with a spatial accuracy of <0.3 degrees. Viewing was binocular though only movements of the right eye were recorded. The sentences were presented on a ViewSonic P227fb monitor with a refresh rate of 7ms (150Hz). The viewing distance was 80 cm and 2.3 characters subtended one degree of visual angle.

Materials and design. There were 180 experimental items, each including a critical word. There were three conditions: In the "normal" condition the entire sentence was presented at high contrast; in the "faint" condition the entire sentence was presented at low contrast; in the "faint-critical" condition, only the critical word was presented at low contrast while the other words were presented normally (at high contrast).

Each experimental item was composed of a single sentence on a single line of text (maximum 60 characters); the words in each item were identical across conditions. The critical word was always positioned at least three words into the sentence, and never at the sentence-final position. The critical words were selected to have a range of word lengths ($M = 6.8$, $SD = 1.5$, $Min = 5$, $Max = 11$), syntactic categories, and word frequencies ($M = 168$, $SD = 370$, $Min = 0$, $Max = 2948$). (Word frequencies were calculated in counts per million using the CELEX English word form corpus (Baayen, Piepenbrock, & Gulikers, 1995). It is important, therefore, that the presence of a single visually degraded word provided no cue as to the nature of the word. The words prior to (n − 1) and following (n + 1) the critical word also had a wide range of lengths (n −1: $M = 4.0$, $SD = 1.9$, $Min = 1$, $Max = 11$; n + 1: $M = 3.8$, $SD = 2.1$, $Min = 1$, $Max = 10$) and frequencies. Words n − 1 and n + 1 were often function words, hence on average these were much more frequent than the critical words (n − 1: $M = 23{,}694$, $SD = 26{,}845$, $Min = 0$, $Max = 60{,}958$; n + 1: $M = 14{,}724$, $SD = 18{,}246$, $Min = 0$, $Max = 60{,}958$).

The sentences were presented in Courier font in gray on a lighter background. For all conditions the luminance of the background was 57.8cd/m^2. The normal text, presented at a high contrast, had a luminance of 4.9 cd/m^2. The faint text, presented at low contrast, had a luminance of 47.5cd/m^2. Note that the size of the text and the luminance of the normal and faint text were matched as closely as possible to those used by Reingold and Rayner (2006).

The three conditions were manipulated within participants and items following a Latin square design. Three lists of 188 sentences were constructed and 10 participants were randomly assigned to each list. Eight of these sentences were filler items of which six were always presented at the start of the experiment for practice. All other items were presented in a pseudorandom order for each participant, with a maximum of two items in the same condition appearing consecutively. Fifty sentences were followed by questions to test for comprehension.

Procedure. Participants first undertook an acuity test. They were asked to read letters from a Bailey Lovie chart, monocularly for each eye. All of the participants included here had fewer than

three errors for either eye in reporting stimuli with a logMAR value of 0.1. Three participants were replaced because they failed this acuity criterion.

For the main experiment, participants were instructed to read the sentences for comprehension and to respond "yes" or "no" to question statements using a button response. A chinrest and forehead rest minimized head movements. The eye tracker was calibrated using a three-point horizontal calibration. The calibration was checked every fourth trial at each of the three calibration positions, and recalibrations were undertaken if necessary. Participants had to accurately fixate on a fixation cross at the position of the start of the line of text before each item was presented. Participants pressed a button to indicate when they had finished reading.

Analyses. Two sets of analyses are presented. Global analyses compare overall reading behavior for the entire sentence for the two conditions that manipulated the characteristics of all of the words in the sentence (normal and faint conditions). Local analyses provide a more detailed examination of reading behavior for the critical word for all three of the conditions. Note that the faint-critical condition is not included in the global analyses as only a subset of the fixations within a sentence are influenced by the manipulation in this condition (as shown in the section "Local Analyses: Reading Times Before and After the Critical Word"). For both global and local analyses, the effects of stimulus quality on overall mean reading times and fixation probabilities are first presented, followed by more detailed examination of the ex-Gaussian distribution of fixation durations across conditions. For the local analyses, distributions of first fixation durations, single fixation durations and gaze durations are presented. Note that the global analyses, including all fixations, provide a particularly strong test of the effect of stimulus quality on the distribution of fixation durations as they generate many observations per participant per condition (Normal: $M = 565$, $SD = 142$; Faint: $M = 586$, $SD = 147$). For both sets of analyses, fixations less than 80 and longer than 1200 ms were excluded (fixations were not merged in order to preserve the integrity of the distributions).

The QMPE software (Cousineau, Brown, & Heathcote, 2004; Heathcote et al., 2002) was used to fit the data for each participant and each condition to the ex-Gaussian distribution separately for each measure. The software uses maximum likelihood estimation to determine a set of parameters. As outlined above, the μ parameter represents the mean of the normal component, σ represents the standard deviation of that normal component, while τ reflects the mean and standard deviation of the exponential component. Note that the fitting procedure involves dividing the observed values into a series of quantile bins. In line with Heathcote et al. (2002) and Rouder, Lu, Speckman, Sun, and Jiang (2005) the maximum number of quantiles were used in the fitting procedures employed here such that each data point was effectively in a separate quantile bin. The resulting parameters were analyzed as dependent variables across the conditions.

Results and Discussion

The average error rate on the comprehension questions was 2.2%, indicating that participants read and understood the sentences.

Global analyses. Paired samples t tests were undertaken to compare the normal and faint (entire sentence) conditions, using participants (t_1) and items (t_2) as random factors. Total sentence reading times were significantly longer in the faint ($M = 3{,}678$, $SD = 1{,}151$) compared to the normal ($M = 3{,}419$, $SD = 1{,}036$) condition [$t_1(29) = 5.88$, $p < .001$; $t_2(179) = 8.65$, $p < .001$]. For all other global measures, 9.2% of fixations were excluded due to blinks immediately prior to or following a fixation and 2.6% were excluded due to being less than 80 or longer than 1,200 ms. A further 10% were excluded because they were not preceded by a valid ($\geq$80ms, $\leq$1200ms) fixation (note that this procedure removes the first fixation on each sentence).

Similar to total sentence reading times, average fixation durations were longer in the faint ($M = 250$, $SD = 114$) compared to the normal ($M = 234$, $SD = 114$) condition [$t_1(29) = 8.43$, $p < .001$; $t_2(179) = 14.00$, $p < .001$]. There were also significantly more fixations in the faint ($M = 9.8$, $SD = 3.8$) compared to the normal ($M = 9.4$, $SD = 3.6$) condition [$t_1(29) = 2.89$, $p < .01$; $t_2(179) = 3.52$, $p < .01$] though there was no significant difference in saccade lengths between the normal ($M = 7.94$, $SD = 6.64$) and faint ($M = 7.85$, $SD = 6.60$) conditions [$t_1 < 1$; $t_2(179) = 1.12$, $p = .264$]. There was also no difference in the proportion of regressive saccades per sentence between the normal ($M = .23$, $SD = .14$) and faint ($M = .23$, $SD = .14$) conditions (ts < 1). The longer reading times and greater number of fixations for the faint compared to normal conditions clearly show that the normal text was easier to process than the faint text.

Global analyses: Ex-Gaussian distributions. Ex-Gaussian analyses are presented for all fixation durations for the normal and faint conditions. Note that these analyses were undertaken across all fixations for each participant within each condition, regardless of item. As explained in the introduction, μ reflects the mean of the normal component of the distribution, σ reflects the standard deviation of the normal component and τ reflects the mean and standard deviation of the exponential component. There was a significantly larger μ parameter for the faint (158) compared to normal (142) text [$t_1(29) = 9.43$, $p < .001$], with a corresponding larger σ for faint (35) compared to normal (28) text [$t_1(29) = 6.49$, $p < .001$], but no difference in τ between the normal (92) and faint (91) conditions ($t < 1$). Note also that separate ex-Gaussian analyses of fixations following progressive and regressive saccades both showed effects of stimulus quality on μ and σ (ps $<$ .01) but not τ ($t < 1$). Importantly, findings suggest that the longer average fixation durations for faint compared to normal text reported above can be accounted for by a shift in the location of the normal component of the distribution with no difference in the frequency of very long fixations; that is, no difference in the prominence of the distribution's right tail.

In order to further examine the effect of stimulus quality on the distribution of fixation durations, we also present the data in the form of vincentile plots (Ratcliff, 1979; Vincent, 1912). The procedure for computing vincentiles is as follows. For each participant in each condition the observations are ordered and then divided into 10 bins, such that the first bin contains the shortest 10% of data points (vincentile 1), the second bin the next shortest 10% (vincentile 2), and so forth. The mean for each vincentile is calculated, and these means are then averaged across participants to generate overall means (and standard errors) at each vincentile, in each condition, which are then plotted. A shift of the entire distribution is characterized by parallel curves, with an equal separation across all vincentiles. In contrast, a difference in the exponential

component is represented by differences in the steepness of the slope of the curves on the right side of the graph; that is, a greater separation between the conditions at the higher vincentiles.

In addition, we focus in detail on how an experimental manipulation affects different parts of the fixation duration distribution by directly plotting the differences between vincentiles. Figure 1 shows three possible patterns. A difference between conditions that is restricted entirely to the μ parameter will result in a difference of constant size across the full range of vincentiles. A difference that is restricted entirely to the τ parameter will result in a difference that gradually increases, moving from the fastest vincentiles to the slowest, with this rate of increase (i.e., the slope of the curve) increasing at the right edge of the plot. Finally, we highlight a third possible situation, which arose in the present data (for the local analyses, below). When both μ and σ are substantially increased in one condition, with no difference in τ, differences will again increase on the right side of the plot, but the curve will be convex, rather than concave; the difference will rapidly increase as one moves from the fastest to the middle vincentiles, with little increase from the middle vincentiles to the slowest. Note that all three of the effects illustrated in Figure 1 correspond to an identical 20 ms effect on the mean.

The vincentile plot for the global analyses is shown in Figure 2a. There is a separation across all the vincentiles, such that fixations across the full range of durations are slightly longer when the text is faint compared to normal. Figure 2a also displays (as open triangles) the vincentiles predicted by the mean of the best fitting parameters for each of the conditions. The predicted vincentiles were generated by first simulating ex-Gaussian distributions based on the mean of the best fitting parameters. For each condition, 20,000 random samples were taken, each created by summing a sample of the normal distribution (based on the μ and σ parameters) and a sample from the exponential distribution (rate parameter $1/\tau$). For each condition, the 20,000 samples were then used to generate the predicted vincentiles shown in Figure 2a. It appears that the predicted 10th vincentile slightly undershoots the observed 10th vincentile in both conditions (a common pattern for highly skewed distributions; see Staub et al., 2010), but otherwise, the predicted vincentiles are extremely close to the observed values, and indeed, the triangles tend to be obscured by the corresponding circles. Thus, the best fitting parameters do capture the shapes of the observed distributions.

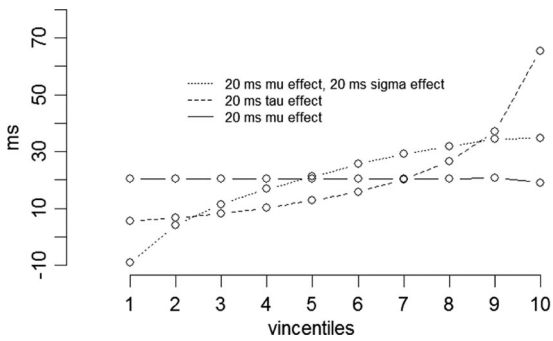

Figure 1. Differences between vincentiles corresponding to a 20-ms effect on the mean, with the effect restricted to the μ parameter (solid line); τ parameter (dashed line); or restricted to the μ parameter, but with an additional 20-ms increase in σ (dotted line).

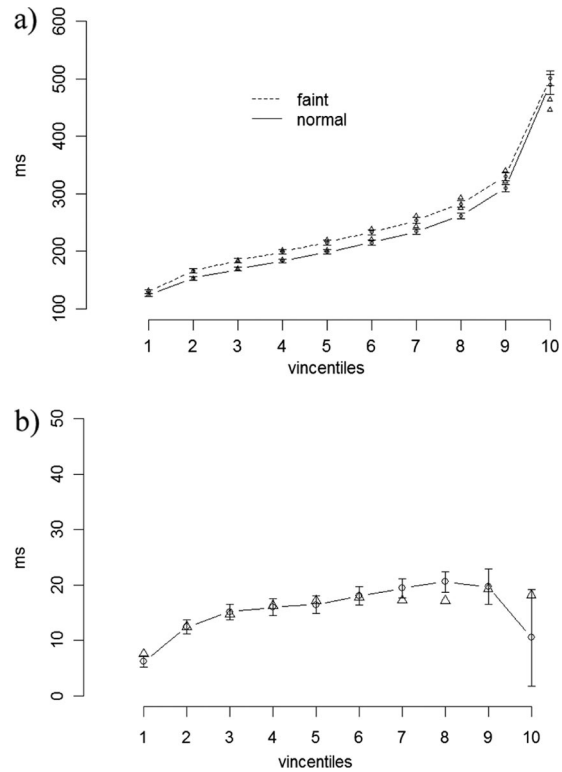

Figure 2. Panel a: Vincentile plots for all fixation durations (global analyses) for text presented entirely in faint or normal font. Panel b: Mean differences between faint and normal vincentiles. Error bars show standard error of the mean. Predicted vincentiles and differences based on mean ex-Gaussian parameters are shown by triangles.

Figure 2b shows the mean difference between the observed vincentiles for the two conditions. Note that, while there is some variability in the size of the difference across the vincentiles, the differences are not larger on the right side of the graph. The somewhat convex shape of this plot may be due, at least in part, to the differences in σ between conditions, as illustrated in Figure 1. Note also that the predicted difference based on the best fitting ex-Gaussian parameters is within a few ms of the observed difference across the full range of vincentiles. In sum, the mean ex-Gaussian parameters and the vincentile plots both suggest that a global manipulation of stimulus quality shifts the distribution of fixation durations to the right, but has no particular influence on the slow tail of the distribution.

It may be noted that the values for the τ parameter in this analysis are larger than those reported below for first fixations and single fixation durations, as well as those reported by Staub et al. (2010) and Staub (2011a). This corresponds to the relatively steep slope of the curve on the right side of Figure 2a. The larger τ parameter for fixations across the entire sentence reflects a larger proportion of particularly long fixations that are likely to be associated with comprehension difficulty. Also, fixations at the ends of sentences tend to be longer (Just & Carpenter, 1980; Kuperman, Dambacher, Nuthmann, & Kliegl, 2010), and these would be included in the global, but not local, analyses.

In order to investigate the effects of stimulus quality on word processing in more detail, local analyses of reading behavior on the

critical word are presented below. Mean reading time measures and fixation probabilities are first presented, followed by ex-Gaussian analyses of first-pass reading times for the critical word.

Local analyses. Blinks on first-pass reading of the critical word, or on the prior fixation, led to exclusion of 4% of trials. The following measures include only fixations made on first pass (before moving to the right of the word): first-fixation duration (the duration of the first fixation on the word), single-fixation duration (the duration of the first fixation on the word for cases in which only one fixation is made on the word on first pass), gaze duration (the sum of fixation durations on the word before leaving it), and refixation probability (the proportion of cases in which the word was refixated before leaving it). In addition, we report the probability of making a regression out of the word on first pass, the probability of making a subsequent regression into the critical word, and total time (sum of fixation durations on the word).

One-way repeated measures Analyses of Variance (ANOVAs) were undertaken across the three conditions using participants (F_1) and items (F_2) as random factors. Analyses were corrected for sphericity where necessary, but uncorrected degrees of freedom are presented to avoid confusion. Table 1 shows the mean reading times and fixation probabilities for the critical word for each of the conditions.

There were significant main effects of condition for all of the reading time measures [first-fixation duration: $F_1(2,58) = 91.31$, $MSE = 282$, $p < .001$; $F_2(2,358) = 204.59$, $MSE = 636$, $p < .001$; single-fixation duration: $F_1(2,58) = 96.14$, $MSE = 454$, $p < .001$; $F_2(2,358) = 200.51$, $MSE = 1173$, $p < .001$; gaze duration: $F_1(2,58) = 106.54$, $MSE = 624$, $p < .001$; $F_2(2,358) = 171.42$, $MSE = 1929$, $p < .001$; total time: $F_1(2,58) = 40.82$, $MSE = 2797$, $p < .001$; $F_2(2,358) = 87.66$, $MSE = 6808$, $p < .001$]. For all of the measures, reading times were longest in the faint-critical condition and shortest in the normal condition. Specifically, reading times were significantly longer in the faint-critical than in the faint condition [first-fixation duration: $t_1(29) = 8.12$, $p < .001$; $t_2(179) = 11.62$, $p < .001$; single-fixation duration: $t_1(29) = 9.01$, $p < .001$; $t_2(179) = 11.80$, $p < .001$; gaze duration: $t_1(29) = 10.28$, $p < .001$; $t_2(179) = 12.11$, $p < .001$; total time: $t_1(29) = 5.02$, $p < .001$; $t_2(179) = 8.14$, $p < .001$] and significantly longer in the faint condition than in the normal condition [first-fixation duration: $t_1(29) = 7.20$, $p < .001$; $t_2(179) = 8.49$, $p < .001$; single-fixation duration: $t_1(29) = 6.79$, $p < .001$; $t_2(179) = 8.28$, $p < .001$; gaze duration: $t_1(29) = 5.66$, $p < .001$; $t_2(179) = 5.99$, $p < .001$; total time: $t_1(29) = 4.51$, $p < .001$; $t_2(179) = 5.12$, $p < .001$]. Clearly therefore, reading times were also significantly longer in the faint-critical compared to the normal condition (ts > 8.1, ps $< .001$).

There was also a significant main effect of condition on the probability of refixating the critical word [$F_1(2,58) = 14.16$, $MSE = 0.004$, $p < .001$; $F_2(2,358) = 21.05$, $MSE = 0.017$, $p < .001$]. The critical word was more likely to be refixated in the faint-critical condition than either the normal [$t_1(29) = 4.49$, $p < .001$; $t_2(179) = 5.92$, $p < .001$] or faint condition [$t_1(29) = 3.92$, $p < .001$; $t_2(179) = 5.02$, $p < .001$] but there was no difference in refixation probability between the normal and faint conditions (ts < 1). There was also no reliable main effect on word skipping; although the items analyses showed a significant main effect [$F_2(2,358) = 3.57$, $MSE = 0.006$, $p < .05$] there was no significant difference across participants [$F_1(2,58) = 2.73$, $MSE = 0.001$, $p = .074$]. As some of the critical words were quite long, further analyses were undertaken for the probability of word skipping for a reduced set of 87 items for which the critical word was either five or six letters long. For this dataset, although the critical word was numerically less likely to be skipped in the faint-critical condition (0.06) compared to either the normal (0.09) or faint (0.08) conditions, there was no significant main effect of condition [$F_1(2,58) = 1.38$, $MSE = 0.005$, $p = .260$; $F_2(2,172) = 2.50$, $MSE = 0.007$, $p = .085$]. The absence of a significant effect of stimulus quality on word skipping is consistent with Wang and Inhoff (2010). However, in both the present study and that of Wang and Inhoff, the skipping rates were low overall and so the null effects could be due to a floor effect.

There was also no effect of condition on the probability of making a regression out of the critical word on first pass (Fs < 1) and, in contrast to the results of Wang and Inhoff (2010), there was no effect of stimulus quality on the probability of making a regressive saccade into the critical word subsequent to fixating it on first pass [$F_1(2,58) = 1.29$, $MSE = 0.006$, $p = .283$; $F_2(2,358) = 2.21$, $MSE = 0.021$, $p = .111$].

Additional analyses were undertaken to examine whether stimulus quality influenced saccade targeting to the critical word on first pass (see Table 1). There was no effect of condition on the position of the initial fixation within the word (Fs < 1); however, there were main effects of condition on both saccade length

Table 1
Local Measures for the Critical Word for Each of the Conditions. Standard Deviations in Parentheses

Measure	Normal	Faint	Faint-critical
Skip probability	0.07	0.06	0.05
First-fixation duration (ms)	220 (72)	240 (76)	272 (91)
Single-fixation duration (ms)	220 (71)	242 (75)	284 (90)
Gaze duration (ms)	263 (121)	288 (134)	344 (153)
Total time (ms)	369 (226)	409 (270)	477 (280)
Refixation probability	0.19	0.20	0.27
First pass regression out probability	0.15	0.14	0.14
Regression in probability (after FP fix)	0.26	0.26	0.28
Landing position (characters)	3.01 (1.81)	3.01 (1.72)	2.98 (1.72)
Saccade length (characters)	7.76 (2.83)	7.53 (2.69)	7.03 (2.64)
Launch site (characters)	4.73 (3.04)	4.53 (2.95)	4.04 (2.90)

[$F_1(2,58) = 20.30$, $MSE = 0.28$, $p < .001$; $F_2(2,358) = 27.01$, $MSE = 1.03$, $p < .001$] and launch site [$F_1(2,58) = 16.22$, $MSE = 0.24$, $p < .001$; $F_2(2,358) = 20.89$, $MSE = 1.20$, $p < .001$]. Saccades into the critical word were significantly shorter in the faint-critical condition compared to both the faint [$t_1(29) = 4.15$, $p < .001$; $t_2(179) = 4.64$, $p < .001$] and normal conditions [$t_1(29) = 5.24$, $p < .001$; $t_2(179) = 6.61$, $p < .001$]. Saccades were also significantly shorter in the faint compared to normal condition [$t_1(29) = 2.62$, $p < .05$; $t_2(179) = 2.73$, $p < .01$]. The absence of an effect on landing positions, together with differences in saccade lengths, are explained by differences in launch site. Launch sites were significantly nearer in the faint-critical compared to the normal [$t_1(29) = 5.00$, $p < .001$; $t_2(179) = 5.87$, $p < .001$] and faint [$t_1(29) = 4.37$, $p < .001$; $t_2(179) = 4.06$, $p < .001$] conditions, explaining the shorter saccade lengths in this condition. Launch sites also tended to be nearer for the faint compared to normal condition; this difference was significant across items [$t_2(179) = 2.34$, $p < .05$] but not participants [$t_1(29) = 1.67$, $p = .106$].

Note that further analyses of the data indicate that the nearer launch sites for the faint-critical condition are also characterized by lower skipping probability for the word prior to the critical word (word n−1) in the faint-critical condition (0.31) compared to the normal (0.39) or faint (0.37) conditions (ts > 3.5, ps $< .001$). It is possible that word n-1 was more likely to be fixated in the faint-critical condition simply due to the influence of low-level visual factors on saccade targeting, with word n−1 being more salient (higher contrast) than the critical word, and hence perhaps attracting fixations that might otherwise have landed on the critical word. The greater likelihood of fixating word n−1 in the faint-critical condition would necessarily shorten the launch sites and saccade lengths for subsequent saccades to the critical word in this condition. In contrast, the trend for shorter saccades in the faint compared to normal conditions is consistent with there being generally more fixations in the faint compared to normal conditions as a whole (as detailed in the global analyses).

Local analyses: Reading behavior before and after the critical word. Additional analyses examined if the effects of stimulus quality in the faint-critical condition were localized to the critical word when compared with the normal condition. There was no difference between the faint-critical and normal conditions in the duration of the fixation preceding the first pass fixation on the critical word (normal: $M = 214$, $SD = 94$; faint-critical: $M = 210$, $SD = 78$) [$t_1(29) = 1.49$, $p = .148$; $t_2(179) = 1.43$, $p = .155$]. There were also no differences in first-fixation (FF) or gaze duration (GD) on word n−1 between the faint-critical (FF: $M = 212$, $SD = 70$; GD: $M = 241$, $SD = 114$) and normal (FF: $M = 214$, $SD = 77$; GD: $M = 245$, $SD = 126$) conditions (ts < 1).

There was no difference in the duration of the fixation following the critical word in the faint-critical ($M = 245$, $SD = 112$) compared to normal ($M = 241$, $SD = 102$) conditions [$t_1 < 1$; $t_2(179) = 1.79$, $p = .171$]. There were also no differences in first-fixation or gaze duration on word n + 1 between the faint-critical (FF: $M = 246$, $SD = 115$; GD: $M = 278$, $SD = 153$) and normal (FF: $M = 245$, $SD = 111$; GD: $M = 279$, $SD = 151$) conditions (ts < 1). In line with the findings of Reingold and Rayner (2006) and Drieghe (2008), the results show that the stimulus quality of an individual word is localized to that word; difficulty associated with stimulus quality did not influence prior fixation durations or spill over to subsequent fixations.

Drieghe (2008) showed that the word following a faint word was less likely to be skipped than if it followed a normally presented word, whereas in the present study, as in Wang and Inhoff (2010), there was no difference in the probability of skipping word n + 1 between the normal (0.47) and faint-critical (0.48) conditions [$t_1 < 1.1$; $t_2(179) = 1.16$, $p = .246$]. Although the length of word n + 1 varied considerably across the items, even when the data set was restricted to the 100 items that were three to six characters long, there was still no effect on skipping probabilities (normal: 0.42; faint-critical: 0.43, ts < 1). However, note that the manipulation of stimulus quality in Drieghe's study had a much more substantial effect on reading times on the critical word compared to the present study (186 ms influence on single fixation duration compared to 64 ms here), which might be attributed to the faint text being nearer to the threshold of perception in Drieghe's study compared to the faint text used here. As noted in Drieghe's discussion, the effect of stimulus quality on skipping of the following word may emerge due to its impact on foveal processing difficulty; the findings here indicate that perhaps such effects are confined only to particularly high levels of difficulty.

Local analyses: Ex-Gaussian distributions. Ex-Gaussian analyses are presented for first fixation, single fixation and gaze durations, with parameter values shown in Table 2. Compared to first fixation and single fixation, the distributions for gaze duration are characterized by a higher frequency of very long durations (larger τ parameter), which is likely to be due to the subset of gaze durations for which there was a refixation. Consequently, note that any differences in τ for the gaze duration measure might reflect differences in refixation probability (as shown for the faint-critical condition as detailed above) rather than differences in the distribution of individual fixation durations.

All of the measures showed significant main effects of condition on μ [first-fixation duration: $F_1(2,58) = 38.38$, $MSE = 593$, $p < .001$; single-fixation duration: $F_1(2,58) = 94.88$, $MSE = 589$, $p < .001$; gaze-duration: $F_1(2,58) = 74.94$, $MSE = 724$, $p < .001$] and σ [first-fixation duration: $F_1(2,58) = 38.13$, $MSE = 128$, $p < .001$; single-fixation duration: $F_1(2,58) = 51.84$, $MSE = 228$, $p < .001$; gaze-duration: $F_1(2,58) = 44.51$, $MSE = 299$, $p < .001$]. All three of the measures produced significantly larger μ parameters for the

Table 2
Mean Ex-Gaussian Parameters for Local Reading Time Measures for the Critical Word

Measure and condition	μ	σ	τ
First-fixation duration (ms)			
Normal	165	30	56
Faint	184	34	57
Faint-critical	212	54	59
Single-fixation duration (ms)			
Normal	165	31	56
Faint	187	37	55
Faint-critical	235	62	49
Gaze duration (ms)			
Normal	154	26	109
Faint	175	32	114
Faint-critical	221	59	123

faint-critical compared to the faint condition [first-fixation duration: $t_1(29) = 4.62$, $p < .001$; single-fixation duration: $t_1(29) = 7.43$, $p < .001$; gaze-duration: $t_1(29) = 6.82$, $p < .001$], and for the faint compared to normal condition [first-fixation duration: $t_1(29) = 5.22$, $p < .001$; single-fixation duration: $t_1(29) = 7.31$, $p < .001$; gaze-duration: $t_1(29) = 6.75$, $p < .001$]. For all of the measures the σ parameter was also larger for the faint-critical compared to the faint condition [first-fixation duration: $t_1(29) = 6.51$, $p < .001$; single-fixation duration: $t_1(29) = 6.16$, $p < .001$; gaze-duration: $t_1(29) = 5.81$, $p < .001$]. The σ parameter was also significantly larger in the faint compared to normal conditions for single-fixation duration [$t_1(29) = 2.87$, $p < .01$] and gaze-duration [$t_1(29) = 2.40$, $p < .05$] with no significant difference for first-fixation duration [$t_1(29) = 1.62$, $p = .116$]. For all of the measures, the μ and σ parameters were also significantly larger for the faint-critical compared to normal conditions (ts > 7.7, ps $< .001$).

There were no significant main effects of condition on the τ parameter for first-fixation duration ($F_1 < 1$) or single-fixation duration [$F_1(2,58) = 1.16$, $MSE = 323$, $p = .320$]. For gaze duration there was no significant effect of condition on the τ parameter [$F_1(2,58) = 2.60$, $MSE = 571$, $p = .083$], though τ was numerically larger in the faint-critical compared to the normal and faint conditions. As noted above, this numerical difference is likely due to the proportion of refixation cases being significantly higher for the faint-critical compared to the other conditions, producing a larger subset of longer gaze durations in the faint-critical condition.

The data are again presented in the form of vincentile plots. Figure 3a plots the first fixation vincentiles in each of the three conditions, together with predicted values based on the best fitting ex-Gaussian parameters. The fit is even more precise than for the global analysis, so that the triangles representing the predicted values overlap the circles representing the observed values at all vincentiles in all conditions. Figure 3b plots the differences between the vincentiles for the faint and normal conditions, and the differences between the faint-critical and faint conditions, together with the corresponding predicted differences. While the observed (and predicted) difference between the faint and normal conditions is essentially constant across the vincentiles, it is clear that the difference between the faint-critical and faint condition does become larger for the slower vincentiles. Note, however, that this is as predicted based on the best fitting ex-Gaussian parameters in these conditions; the particular curvature for the faint-critical compared to faint conditions arises from a combination of an increase in μ and a dramatic increase in σ, rather than from an increase in τ (see Figure 1).

The corresponding plots for single-fixation duration are shown in Figures 4a and 4b. The predicted vincentiles are very close to the observed vincentiles. Again, the difference between the normal and faint conditions is nearly constant across all vincentiles; the exception is the slowest (10th) vincentile, at which the difference becomes larger, and is less well predicted by the mean ex-Gaussian parameters. As with first-fixation duration, the difference between the faint-critical and faint conditions does become larger moving from the fastest vincentiles to the slowest. Again, the convex form of this increase is captured well by the best fitting ex-Gaussian parameters, for which both μ and σ are much larger in the faint-critical condition; in this case, τ is actually numerically smaller in the faint-critical condition (49 ms) than in the faint condition (55 ms).

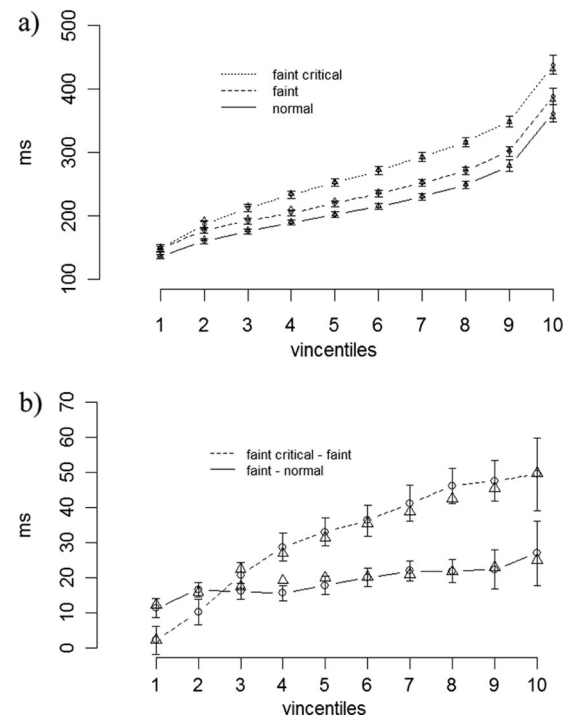

Figure 3. Panel a: Vincentile plots for first fixation duration (local analyses) for text presented in normal, faint, or faint-critical conditions. Panel b: Mean differences between faint and normal vincentiles, and between faint-critical and faint vincentiles. Error bars show standard error of the mean. Predicted vincentiles and differences based on mean ex-Gaussian parameters are shown by triangles.

Finally, the plots for gaze duration are shown in Figures 5a and 5b. Once again the fits between the predicted and observed vincentiles are excellent. Figure 5b shows that the difference in gaze duration between the faint and normal conditions becomes only very slightly larger as one moves from the fastest vincentiles to the slowest. The increase in gaze duration in the faint-critical condition, compared to the faint condition, follows a less easily characterized pattern, reflecting the fact that there are increases in all three of the ex-Gaussian parameters in the faint-critical condition, due to a combination of increased fixation durations and increased probability of multiple fixations.

The distributional analyses of the local eye movement data may be summarized as follows. As would be expected based on the global analysis comparing fixation durations in the normal and faint conditions, the reading time difference on the critical word, between these two conditions, is due to a shifting of the fixation duration distribution to the right in the faint condition. Thus, all analyses converge in suggesting that globally faint text causes most fixations to be lengthened.

The effect of encountering a single faint word is more complex. While the distribution of fixation durations on the critical word is even further shifted to the right, it is also the case that there is substantially greater variance in the fixation duration distribution. Together, these two effects result in longer fixations being length-

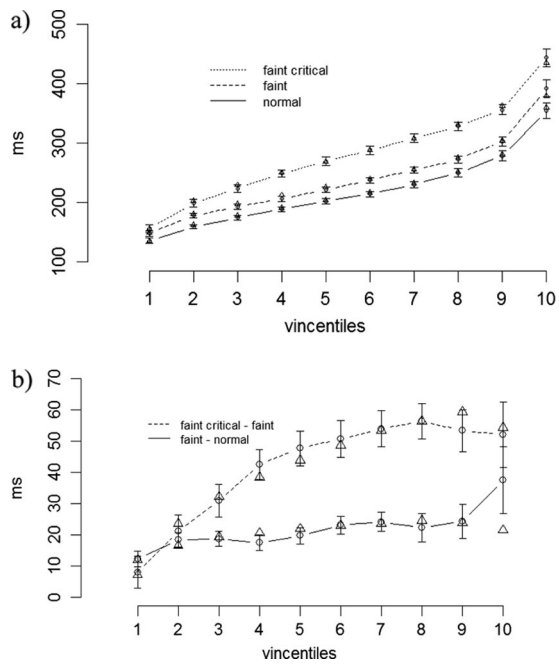

Figure 4. Panel a: Vincentile plots for single fixation duration (local analyses) for text presented in normal, faint, or faint-critical conditions. Panel b: Mean differences between faint and normal vincentiles, and between faint-critical and faint vincentiles. Error bars show standard error of the mean. Predicted vincentiles and differences based on mean ex-Gaussian parameters are shown by triangles.

ened more than shorter ones. Notably, however, encountering a single faint word in normal text does not result in a particular influence on the slow tail of the distribution.

Finally, the best fitting ex-Gaussian parameters consistently predicted the observed vincentiles very precisely. This result confirms that fixation duration distributions are indeed well fit by the ex-Gaussian, as also reported by Staub et al. (2010) and Staub (2011a), which provides a useful descriptive parameterization of the distribution.

General Discussion

Similar to previous work, the present study showed a clear effect of stimulus quality on reading times for a critical word (Drieghe, 2008; Reingold & Rayner, 2006; Wang & Inhoff, 2010) and for entire sentences (e.g., Legge et al., 1987, 1997). In line with Reingold and Rayner's findings, the effects of stimulus quality for the critical word were largely localized to that word. The large number of observations per participant in the present study enabled the fixation durations to be fit to the ex-Gaussian distribution, and hence allowed an examination of the extent to which mean differences between conditions are characterized by a shift in the location of the Normal component of the distribution (μ) or a difference in the exponential component of the distribution (τ). The results of the present study were clear. When the entire text was faint, the effect of stimulus quality on fixation durations was characterized by a shift, and not by any particular difference in the frequency of very long fixations. When only a single word was faint, there was a sizable shift effect, coupled with a sizable increase in the standard deviation of the Normal component (σ), while again there were no convincing differences in τ. Crucially, the present results contrast with those from single-word response tasks, which show effects of stimulus quality on the exponential component of the distribution as well as effects on the location of the distribution (Plourde & Besner, 1997; Yap & Balota, 2007; Yap et al., 2008). The General Discussion begins by discussing the important differences in findings between these tasks. The implications of the results for the role of stimulus quality during normal reading are then considered, followed by a discussion of the implications of the distributional analyses for models of eye movement control during reading.

Effects of Stimulus Quality on the Distribution of Response Times and Fixation Durations

Previous studies have shown that effects of stimulus quality vary as a function of methodology, with word frequency and stimulus quality having additive or interactive effects (on both means and ex-Gaussian parameters) as a function of task characteristics (Yap & Balota, 2007). The present study demonstrates that the general effect of stimulus quality on the distribution of responses also differs between single-word response tasks and fixation durations during reading. Previous studies that examined response times in single-word tasks have shown an effect of stimulus quality on both the μ and τ parameters of response time distributions (Plourde & Besner, 1997; Yap & Balota, 2007; Yap

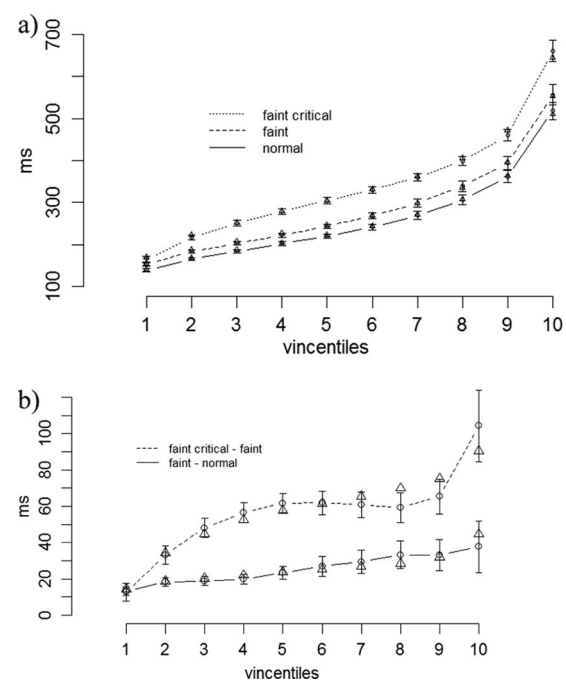

Figure 5. Panel a: Vincentile plots for gaze duration (local analyses) for text presented in normal, faint, or faint-critical conditions. Panel b: Mean differences between faint and normal vincentiles, and between faint-critical and faint vincentiles. Error bars show standard error of the mean. Predicted vincentiles and differences based on mean ex-Gaussian parameters are shown by triangles.

et al., 2008), whereas the present study only showed a clear effect of stimulus quality on the location of the distribution of fixation durations (μ). Together, the findings from the previous literature and the present findings indicate that a) for visually degraded single words (i.e., in lexical decision and naming tasks) there is a general increase in RT, and also a particular increase in the frequency of very long RTs; b) for globally low-contrast text, there is a general increase in fixation duration, with essentially no sign of any particular influence on the slow tail of the distribution; c) for a low-contrast word presented in otherwise normal text, there is a general increase in fixation duration, as well as an increase in variance; there is also an increase in refixation probability, resulting in a statistically marginal suggestion of an influence on the exponential component in the gaze-duration measure.

Thus, the results from eyetracking and single-word tasks could be partially reconciled if the faint-critical condition is regarded as the closest analogue to the single-word tasks, and if gaze duration is regarded as the measure that is most appropriately compared to single-word RT. However, the pronounced differences between the results from the faint-critical condition and the results from the (globally) faint condition raise doubts about whether the effects in the former condition reflect stimulus quality itself, as we discuss below. It appears that while the results in the faint-critical condition do reflect the effect of stimulus quality, they even more clearly reflect the difference in stimulus quality from the surrounding text. Indeed, in light of the present results, it seems possible that the stimulus quality effect, which is dramatically reduced in absolute magnitude when the entire sentence is faint, would be further attenuated in a blocked design, in which sentences in faint text were not intermixed with normal text.

The influence of stimulus quality on responses in single-word tasks has been shown to vary as a function of the type of task and stimuli, and the present results suggest an important dissociation between the effect of stimulus quality in single-word tasks and in normal reading. This raises the question of which findings reflect the word recognition process itself, and which might result from the nature of the task. Responses in single-word studies tend to be much longer (~500–700 ms) than the duration of individual fixations in natural reading (~200–300 ms), and indeed, considerably longer than gaze durations (~250–400 ms). In addition, there is generally a much more pronounced right tail (a larger τ component) in the distributions of single word RTs than in the distributions of fixation durations or gaze durations (Staub et al., 2010). Staub et al. also found, when comparing eye movement and lexical decision experiments with the same target words, that the size of the frequency effect on the τ parameter was much larger in lexical decision than in eye movements, though the latter was also significant. Note that Schilling, Rayner, and Chumbley (1998) also showed larger effects of word frequency on lexical decision responses compared to fixation durations (whereas the results for naming and eye tracking were more similar). Considering these previous results together with the present results, a reasonable hypothesis is that the cognitive processes underlying the exponential component of the RT distribution in single-word tasks play a less prominent role in normal reading, so that the τ parameter of the fixation duration distribution is small in general, and is less susceptible to experimental modulation, resulting in either relatively small (word frequency) or nonexistent (stimulus quality) τ effects. One possibility is that the processes involved in making a controlled, potentially attention-demanding behavioral response are reflected in the τ parameter, and because these processes play little role in online eye movement control, the τ parameter is both less prominent in fixation duration distributions and less susceptible to experimental manipulation.

Another possibility, however, is that differences in text presentation in single-word tasks and in sentence reading may influence how words are processed. For example, the availability of a parafoveal preview is a crucial difference between most single-word tasks and natural reading (White, Johnson, Liversedge, & Rayner, 2008). In natural reading, words are preprocessed parafoveally prior to fixation, which facilitates processing of those words when they are subsequently fixated (Rayner, 1975). Perhaps the impact of stimulus quality on word processing is different in natural reading (such that there is no particular influence on a subset of long responses) when the word recognition system first receives a parafoveal preview.

Effects of Stimulus Quality During Natural Reading

Three issues are discussed here: (1) the differential impact of stimulus quality when only a single word is faint compared to an entire sentence; (2) localized effects of stimulus quality when only a single word is faint; and (3) general implications of the effect of stimulus quality on the distribution of fixation durations for accounts of saccade programming during reading.

The effect of stimulus quality on gaze duration when a single word was presented at low contrast was more than triple the gaze duration effect on the same word when the entire text was presented at low contrast (81 ms vs. 25 ms). Thus, the present results demonstrate that studies which manipulate the stimulus quality of just a single word (Drieghe, 2008; Reingold & Rayner, 2006; Wang & Inhoff, 2010) may elicit results which reflect processes other than those associated only with the demands of processing a low-contrast stimulus. Longer reading times for single low-contrast words could be due to the different presentation format inducing surprise, or perhaps drawing attention to the critical word. However, if a single faint word does draw attention, note that this is likely to be different to other highlighting formats, such as presenting words in bold, as bold text also influences subsequent fixations (spillover effects), whereas the effects of stimulus quality are largely localized to the critical word (Reingold & Rayner, 2006). Another possibility is that reading times may be longer for single words than for text presented entirely in low contrast simply due to the reader having to adjust to sampling the text at a different intensity.

Another interesting difference between the two faint conditions is that although both showed significant effects of stimulus quality on the standard deviation of the Normal component (σ) of the distribution of fixation durations, these effects were much larger when only one word was faint compared to the entire sentence. It is possible that the degree of "surprise" elicited by a single faint word, or the adjustment required for reading a single faint word, is highly variable. For example, the extent of the surprise or adjustment might have been modulated by the time since the last single faint word, the number of single faint words encountered in the experiment prior to that trial, or even the extent of parafoveal preview (modulated by launch site). The discussion of the theoretical implications below focuses largely on μ and τ effects, but

it is important to remember that ideally models of eye movement control would also explain differences in σ.

The second striking finding about the effects of stimulus quality on word processing in natural reading is that, in line with previous studies (Drieghe, 2008; Reingold & Rayner, 2006; Wang & Inhoff, 2010), the effects of stimulus quality when only a single word is manipulated are largely localized to that word, with no significant influences on previous or subsequent fixations. As Reingold and Rayner emphasized, these findings are consistent with the assumption in the E-Z Reader model (Reichle et al., 2003) that stimulus quality affects only the L_1 stage of processing.

The third important finding, the influence of stimulus quality on the location of the distribution of fixation durations, has implications for how stimulus quality influences saccade programming during normal reading. The clear influence of stimulus quality on the μ parameter illustrates that stimulus quality can influence fixations of all durations, such that even very short fixations are modulated by stimulus quality. These findings demonstrate that stimulus quality does not simply modulate fixation durations as a result of saccade cancellation and reprogramming (which would influence only longer fixations). Note that Staub et al. (2010) made a similar argument with respect to word frequency: The fact that word frequency affects μ demonstrates that frequency exerts an influence on fixations of all durations, and that frequency effects cannot simply be mediated by occasional saccade reprogramming as some models previously suggested (Yang & McConkie, 2001, 2004). In the next section we discuss in more detail how models of eye movement control during reading might be adapted in order to account for differential influences of word characteristics on the distribution of fixation durations.

Differential Influences on the Distribution of Fixation Durations During Reading

In contrast to the effects of stimulus quality shown here, Staub et al. (2010) showed that word frequency influenced the exponential component of the distribution of fixation durations, not just the location of the Normal component. In addition, Staub (2011a) showed that, similar to stimulus quality, word predictability influences the location of fixation duration distributions with no particular influence on the frequency of very long responses. Thus, the distributional effects of three of the principle variables that are known to influence mean fixation duration have now been examined (stimulus quality, word frequency, word predictability), and only one of them, frequency, appears to influence the exponential component. These results indicate that stimulus quality and word predictability have qualitatively different influences on the distribution of fixation durations during reading compared to word frequency. Although models of eye movement control during reading such as E-Z Reader (Reichle et al., 2003) and SWIFT (Engbert et al., 2005) predict that factors such as stimulus quality and word frequency influence fixation durations, these models do not predict any difference in how such factors influence the distribution of fixation durations. We will first explain how these two models account for the distribution of fixation durations generally, before considering how they might ultimately account for how different word characteristics produce differences in the distribution of fixation durations.

In the E-Z Reader model, the time to complete the first stage of word processing (L_1) determines the duration of fixations, hence the distribution of L_1 finishing times also predicts the distribution of fixation durations during reading. Staub et al. (2010) demonstrated that because L_1 finishing times are modeled using gamma distributions, differences in finishing times are necessarily characterized by differences in both the location and the slow tail of the distributions. Critically, in its present form, the E-Z Reader model would predict that any factor affecting L_1 would have the same impact on its distribution of finishing times. Given that the variables of word frequency, predictability, and stimulus quality are all thought to influence fixation durations as a function of L_1 finishing times, the model therefore currently predicts that they should all influence both the location and slow tail of the distribution of fixation durations.

In the SWIFT model (Engbert et al., 2005), the duration of fixations is determined by a random timer that can be delayed (through a mechanism called foveal inhibition) as a function of lexical activation. Stimulus quality is likely to impact on lexical activation, either by modulating lexical processing rate in a similar way to visual acuity, or by impacting more directly on word processing difficulty. Similar to the E-Z Reader model, the SWIFT model currently makes no specific predictions about the effect of different variables on the distribution of fixation durations.

Nevertheless, it is important to emphasize that these models have so far been primarily designed to predict the effect of variables on mean fixation durations, not the distributions. We now consider how these models might ultimately account for different influences on the distribution of fixation durations. We focus on two possibilities: First, that differences in L_1 finishing times in E-Z Reader or levels of lexical activation in SWIFT might be accounted for by a more complete account of the word recognition process. Second, that different word characteristics might have different influences on the distribution of fixation durations as a result of pre- or postlexical mechanisms.

Different accounts have been proposed for precisely how stimulus quality influences word recognition mechanisms. Some predict that stimulus quality impacts directly on word processing at the orthographic input lexicon, directly influencing the rate of activation for specific lexical representations (Borowsky & Besner, 1993). In contrast, others suggest that the output of the letter level might be thresholded such that stimulus quality would impact directly only at the feature and letter levels (Besner & Roberts, 2003). Regardless of the mechanisms used to account for effects of stimulus quality, many models of word recognition currently do not make clear predictions for how stimulus quality and other factors might differentially influence the distribution of responses to words (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981; Paap, Newsome, McDonald, Schvaneveldt, 1982; Plaut, McClelland, Seidenberg, & Patterson, 1996). Ratcliff, Gomez and McKoon (2004) highlighted the importance of accounting for the distribution of response times and Norris (2009) recently simulated the effect of word frequency on response-time distributions. Given the current interest in this topic within the single-word literature, it seems likely that models of word recognition will be developed to account for the influence of a range of variables on response distributions. Ultimately, models of eye movement control during

reading might link to models of word recognition to provide a more detailed account of the reading process (Grainger, 2003; Liversedge & Blythe, 2007). The input from a word recognition system might determine the L_1 and L_2 finishing times in E-Z Reader (or lexical activation in the SWIFT model), such that these processing stages might then be characterized by different distributions as a function of different variables. Models of eye movement control during reading would then predict the appropriate distributions of fixation durations as a function of how different variables influence the duration of the word recognition process itself.

A second possibility is that word variables may have unique influences on the reading process independently of word recognition mechanisms. Two possibilities are suggested here: First, word frequency may also be associated with processing beyond the level of word recognition (that is, beyond stages such as L_1 and L_2). Currently, the postlexical integration (I) stage proposed in the latest version of E-Z Reader, known as E-Z Reader 10 (Reichle, Warren, & McConnell, 2009), is not influenced by lexical frequency. However, it is plausible to suppose that, in addition to influencing the difficulty of word recognition itself, lexical frequency may influence the difficulty of integrating a word into its sentence context. (Note, however, that Staub [2011b] has suggested that this is not the case, as indicated by the pattern of effects in experiments that manipulate both frequency and integration difficulty.) A second possibility is that, though Reingold and Rayner (2006) proposed that stimulus quality influences L_1, it is plausible to suppose that it may instead, or in addition, influence the rate of early preattentive visual processing during the prelexical visual processing stage (V) that is proposed by the model. This hypothesis is especially attractive given that the rate of early visual processing could perhaps be modulated by stimulus quality in the same way that, in present versions of the model, it is modulated by the dropoff in acuity as a word is centered farther from the fovea.

Conclusions

The present study provides the first examination of the effect of stimulus quality on the distribution of fixation durations during natural reading. Stimulus quality (especially when consistent across a sentence) influences the location of the Normal component of the distribution of fixation durations during reading, rather than having a particular influence on the frequency of long fixations. The findings have important implications when compared with previous studies. First, the results indicate that the effect of stimulus quality on the exponential component of single word RT distributions may be linked to the nature of the task rather than the word recognition process itself. Second, given that for fixation durations word frequency but not stimulus quality has a particular influence on the slow tail of the distribution of fixation durations, models of eye movement control ultimately need to account for how different variables can have qualitatively different influences on the distribution of fixation durations during reading. Incorporating natural reading factors within models of word recognition (such as parafoveal preview) may enable word recognition models to more accurately account for word processing in natural reading. Similarly, incorporating models of word recognition into models of eye movement control may provide a more complete account of how word characteristics can affect eye movement behavior. Within such combined accounts, eye movement control modules would enable word inputs to vary as a function of fixation position (e.g., providing parafoveal previews), while word recognition modules could play a more central role in influencing the duration of stages that reflect word processing such as L_1 and L_2 in E-Z Reader and lexical activation in SWIFT.

References

Andrews, S., & Heathcote, A. (2001). Distinguishing common and task-specific processes in word identification: A matter of some moment? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 514–544. doi:10.1037/0278-7393.27.2.514

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database.* [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General, 128,* 32–55. doi:10.1037/0096-3445.128.1.32

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science, 20,* 160–166. doi:10.1177/0963721411408885

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language, 59,* 495–523. doi:10.1016/j.jml.2007.10.004

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 389–401. doi:10.1037/0096-1523.3.3.389

Besner, D., & Roberts, M. A. (2003). Reading nonwords aloud: Results requiring change in the dual route cascaded model. *Psychonomic Bulletin & Review, 10,* 398–404. doi:10.3758/BF03196498

Borowsky, R., & Besner, D. (1993). Visual word recognition: A multistage activation model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 813–840. doi:10.1037/0278-7393.19.4.813

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256. doi:10.1037/0033-295X.108.1.204

Cousineau, D., Brown, S., & Heathcote, A. (2004). Fitting distributions using maximum likelihood: Methods and packages. *Behavior Research Methods, Instruments, & Computers, 36,* 742–756. doi:10.3758/BF03206555

Drieghe, D. (2008). Foveal processing and word skipping during reading. *Psychonomic Bulletin & Review, 15,* 856–860. doi:10.3758/PBR.15.4.856

Drieghe, D., Rayner, K., & Pollatsek, A. (2008). Mislocated fixations can account for parafoveal-on-foveal effects in eye movements during reading. *Quarterly Journal of Experimental Psychology, 61,* 1239–1249. doi:10.1080/17470210701467953

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning & Verbal Behavior, 20,* 641–655. doi:10.1016/S0022-5371(81)90220-6

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112,* 777–813. doi:10.1037/0033-295X.112.4.777

Feng, G. (2009). Time course and hazard function: A distributional analysis of fixation duration in reading. *Journal of Eye Movement Research, 3*(2):3, 1–23.

Grainger, J. (2003). Moving eyes and reading words: How can a computational model combine the two? In J. Hyönä, R. Radach, & H. Deubel

(Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 457–470). Amsterdam: Elsevier.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read out model. *Psychological Review, 103,* 518–565. doi:10.1037/0033-295X.103.3.518

Hackman, R. B., & Tinker, M. A. (1957). Effect of variations in color of print and background upon eye movements in reading. *American Journal of Optometry and Archives of American Academy of Optometry, 34,* 354–359.

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review, 9,* 394–401. doi:10.3758/BF03196299

Heathcote, A., Popiel, S. J., & Mewhort, D. J. K. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin, 109,* 340–347. doi:10.1037/0033-2909.109.2.340

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 417–429. doi:10.1037/0278-7393.16.3.417

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329–354. doi:10.1037/0033-295X.87.4.329

Kinoshita, S., & Hunt, L. (2008). RT distribution analysis of category congruence effects with masked primes. *Memory & Cognition, 36,* 1324–1334. doi:10.3758/MC.36.7.1324

Kuperman, V., Dambacher, M., Nuthmann, A., & Kliegl, R. (2010). The effect of word position on eye movements in sentence and paragraph reading. *Quarterly Journal of Experimental Psychology, 63,* 1838–1857. doi:10.1080/17470211003602412

Legge, G. E., Ahn, S. J., Klitz, T. S., & Luebker, A. (1997). Psychophysics of reading—XVI. The visual span in normal and low vision. *Vision Research, 37,* 1999–2010. doi:10.1016/S0042-6989(97)00017-5

Legge, G. E., Rubin, G. S., & Luebker, A. (1987). Psychophysics of reading—V. The role of contrast in normal vision. *Vision Research, 27,* 1165–1177. doi:10.1016/0042-6989(87)90028-9

Liversedge, S. P., & Blythe, H. I. (2007). Lexical and sublexical influences on eye movements during reading. *Language and Linguistics Compass, 1,* 17–31. doi:10.1111/j.1749-818X.2007.00003.x

Matzke, D., & Wagenmakers, E.-J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review, 16,* 798–817. doi:10.3758/PBR.16.5.798

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Pt. 1. An account of basic findings. *Psychological Review, 88,* 375–407. doi:10.1037/0033-295X.88.5.375

McConkie, G. W., Kerr, P. W., & Dyre, B. P. (1994). What are "normal" eye movements during reading: Toward a mathematical description. In J. Ygge, & G. Lennerstrand (Eds.), *Eye movements in reading* (pp. 315–328). Oxford, UK: Pergamon Press.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1975). Loci of contextual effects on visual word recognition. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 98–118). New York, NY: Academic Press.

Norris, D. (2009). Putting it all together: A unified account of word recognition and reaction-time distributions. *Psychological Review, 116,* 207–219. doi:10.1037/a0014259

O'Malley, S., Reynolds, M. G., & Besner, D. (2007). Qualitative differences between the joint effects of stimulus quality and word frequency in reading aloud and lexical decision: Extensions to Yap and Balota (2007). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 451–458. doi:10.1037/0278-7393.33.2.451

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation verification model for letter and word recognition: The word superiority effect. *Psychological Review, 89,* 573–594. doi:10.1037/0033-295X.89.5.573

Paterson, D. G., & Tinker, M. A. (1944). Eye movements in reading optimal and non-optimal typography. *Journal of Experimental Psychology, 34,* 80–83. doi:10.1037/h0056763

Paterson, D. G., & Tinker, M. A. (1947). The effect of typography upon the perceptual span in reading. *American Journal of Psychology, 60,* 388–396. doi:10.2307/1416919

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115. doi:10.1037/0033-295X.103.1.56

Plourde, C. E., & Besner, D. (1997). On the locus of the word frequency effect in visual word recognition. *Canadian Journal of Experimental Psychology, 51,* 181–194. doi:10.1037/1196-1961.51.3.181

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86,* 446–461. doi:10.1037/0033-2909.86.3.446

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review, 111,* 159–182. doi:10.1037/0033-295X.111.1.159

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81. doi:10.1016/0010-0285(75)90005-5

Rayner, K., & Duffy, S. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14,* 191–201. doi:10.3758/BF03197692

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements on young and older readers. *Psychology and Aging, 21,* 448–465. doi:10.1037/0882-7974.21.3.448

Reichle, E. D., Pollatsek, A., & Rayner, K. (2007). Modeling the effects of lexical ambiguity on eye movements during reading. In R. P. G. Van Gompel, M. F. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 271–292). Oxford, England: Elsevier.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26,* 445–526. doi:10.1017/S0140525X03000104

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review, 16,* 1–21. doi:10.3758/PBR.16.1.1

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the E-Z Reader model. *Psychological Science, 17,* 742–746. doi:10.1111/j.1467-9280.2006.01775.x

Rouder, J. N., Lu, J., Speckman, P., Sun, D., & Jiang, Y. (2005). A hierarchical model for estimating response time distributions. *Psychonomic Bulletin & Review, 12,* 195–223.

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition, 26,* 1270–1281. doi:10.3758/BF03201199

Slattery, T. J., & Rayner, K. (2010). The influence of text legibility on eye movements during reading. *Applied Cognitive Psychology, 24,* 1129–1148. doi:10.1002/acp.1623

Speckman, P. L., & Rouder, J. N. (2004). A comment on Heathcote, Brown, and Mewhort's QMLE method for response time distributions. *Psychonomic Bulletin & Review, 11,* 574–576. doi:10.3758/BF03196613

Stanners, R. F., Jastrzembski, J. E., & Westbrook, A. (1975). Frequency and visual quality in a word-nonword classification task. *Journal of*

Verbal Learning and Verbal Behavior, 14, 259–264. doi:10.1016/S0022-5371(75)80069-7

Stanovich, K. E., & West, R. F. (1983). On priming by a sentence context. *Journal of Experimental Psychology: General, 112,* 1–36. doi:10.1037/0096-3445.112.1.1

Staub, A. (2011a). The effect of lexical predictability on distributions of eye fixation durations. *Psychonomic Bulletin & Review, 18,* 371–376. doi:10.3758/s13423-010-0046-9

Staub, A. (2011b). Word recognition and syntactic attachment in reading: Evidence for a staged architecture. *Journal of Experimental Psychology: General.* Advance online publication. doi:10.1037/a0023517

Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1280–1293. doi:10.1037/a0016896

Tinker, M. A., & Paterson, D. G. (1955). The effect of typographical variations upon eye movements in reading. *Journal of Education Research, 49,* 171–184.

Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin and Review, 7,* 424–465. doi:10.3758/BF03214357

Vincent, S. B. (1912). The function of the viborissae in the behavior of the white rat. *Animal Behavior Monographs, 1,* (5).

Wang, C.-A., & Inhoff, A. W. (2010). The influence of visual contrast and case changes on parafoveal preview benefits during reading. *Quarterly Journal of Experimental Psychology, 63,* 805–817. doi:10.1080/17470210903147494

White, S. J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 205–223. doi:10.1037/0096-1523.34.1.205

White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of word beginning letters. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1261–1276. doi:10.1037/0096-1523.34.5.1261

White, S. J., Rayner, K., & Liversedge, S. P. (2005). Eye movements and the modulation of parafoveal processing by foveal processing difficulty: A re-examination. *Psychonomic Bulletin & Review, 12,* 891–896. doi:10.3758/BF03196782

Yang, S.-N., & McConkie, G. W. (2001). Eye movements during reading: A theory of saccade initiation time. *Vision Research, 41,* 3567–3585. doi:10.1016/S0042-6989(01)00025-6

Yang, S.-N., & McConkie, G. W. (2004). Saccade generation during reading: Are words necessary? *European Journal of Cognitive Psychology, 16,* 226–261. doi:10.1080/09541440340000231

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 274–295. doi:10.1037/0278-7393.33.2.274

Yap, M. J., Balota, D. A., Cortese, M. J., & Watson, J. M. (2006). Single versus dual process models of lexical decision performance: Insights from RT distributional analysis. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1324–1344. doi:10.1037/0096-1523.32.6.1324

Yap, M. J., Balota, D. A., Tse, C.-S., & Besner, D. (2008). On the additive effects of stimulus quality and word frequency in lexical decision: Evidence for opposing interactive influences revealed by RT distributional analyses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 495–513. doi:10.1037/0278-7393.34.3.495

Yap, M. J., Tse, C.-S., & Balota, D. A. (2009). Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical ambiguity. *Journal of Memory and Language, 61,* 303–325. doi:10.1016/j.jml.2009.07.001

Received February 9, 2011
Revision received July 15, 2011
Accepted July 21, 2011 ■