The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers

Article in Psychology and Aging · October 2006
Impact Factor: 2.73 · DOI: 10.1037/0882-7974.21.3.448 · Source: PubMed

CITATIONS
120

READS
569

5 authors, including:

Erik D Reichle
University of Southampton
71 PUBLICATIONS 3,904 CITATIONS
SEE PROFILE

Michael Stroud
Merrimack College
15 PUBLICATIONS 170 CITATIONS
SEE PROFILE

Carrick C Williams
California State University, San Marcos
28 PUBLICATIONS 1,296 CITATIONS
SEE PROFILE

Alexander Pollatsek
University of Massachusetts Amherst
232 PUBLICATIONS 12,229 CITATIONS
SEE PROFILE

All in-text references underlined in blue are linked to publications on ResearchGate, letting you access and read them immediately.
The Effect of Word Frequency, Word Predictability, and Font Difficulty on the Eye Movements of Young and Older Readers

Keith Rayner
University of Massachusetts at Amherst

Erik D. Reichle
University of Pittsburgh

Michael J. Stroud
University of Massachusetts at Amherst

Carrick C. Williams
Mississippi State University

Alexander Pollatsek
University of Massachusetts at Amherst

Young adult and older readers’ eye movements were recorded as they read sentences containing target words that varied in frequency or predictability. In addition, half of the sentences were printed in a font that was easy to read (Times New Roman) and the other half were printed in a font that was more difficult to read (Old English). Word frequency, word predictability, and font difficulty effects were apparent in the eye movement data of both groups of readers. In the fixation time data, the pattern of results was the same, but the older readers had larger frequency and predictability effects than the younger readers. The older readers skipped words more often than the younger readers (as indicated by their skipping rate on selected target words), but they made more regressions back to the target words and more regressions overall. The E-Z Reader model was used as a platform to evaluate the results, and simulations using the model suggest that lexical processing is slowed in older readers and that, possibly as a result of this, they adopt a more risky reading strategy.

keywords: reading, eye movements, aging and reading

A great deal has been learned over the past 25 years about the characteristics of eye movements of skilled readers (Rayner, 1978, 1998). However, remarkably little is known about the effect of aging on eye movements during reading. It is known that older adults read more slowly and make more fixations (and regressions) than younger readers (Kemper, Crow, & Kemtes, 2004; Kliegl, Grabner, Rolfs, & Engbert, 2004; Solan, Feldman, & Tujak, 1995). Furthermore, in simple oculomotor tasks, older participants’ saccade latencies are longer than those of younger participants (Munoz, Broughton, Goldring, & Armstrong, 1998; Pratt, Abrams, & Chasteen, 1997). Kliegl et al. found that older readers’ eye fixations were longer than those of younger readers, but they concluded that there were only minor effects of age and that the similarities between older and younger readers were more apparent than the differences.

In this article, we further examine the eye movement characteristics of older readers. Specifically, we asked older and younger readers to read sentences containing target words that varied either in frequency (low-frequency vs. high-frequency target words) or in predictability (low-predictable, medium-predictable, or high-predictable target words) to determine whether frequency and predictability interact with age when these target words are read. Furthermore, we varied the type font in which the sentences (and accordingly the target words) were presented. Half of the sentences appeared in a font that was easy to read (Times New Roman), whereas the other half of the sentences appeared in a font that was more difficult to read (Old English). We were interested in determining whether the differences between younger and older readers were largely difficulties in decoding letters, and thus, whether reading in a difficult font would make younger readers look like older readers when they read a more normal font. More generally, we were also interested in determining whether a more difficult font would change the general strategies that either group adopted in reading.

Word frequency and word predictability are known to have strong effects on how long readers look at a word. The basic finding with respect to word frequency is that words that occur
with high frequency in the language are fixed for less time than are low-frequency words even when word length is controlled (Inhoff & Rayner, 1986; Rayner, Ashby, Pollatsek, & Reicheck, 2004; Rayner & Duffy, 1986; Rayner, Sereno, & Raney, 1996). This word frequency effect is quite robust and has been demonstrated many times (see Rayner, 1998, for a review). Likewise, words that are more predictable from their prior context are fixed for less time than words that are not predictable (Balota, Pollatsek, & Rayner, 1985; Drieghe, Rayner, & Pollatsek, 2005; Ehrlich & Rayner, 1981; Rayner et al., 2004; Rayner & Well, 1996). Moreover, predictable words are skipped more frequently than unpredictable words (Balota et al., 1985; Drieghe et al., 2005; Ehrlich & Rayner, 1981; Rayner & Well, 1996).

Although frequency and predictability have clear effects on fixation times on words, less is known about how font difficulty influences how long readers look at a word. However, it stands to reason that a font that is more difficult to encode should cause readers to look at words longer than fonts that are easier to encode. Indeed, Reingold and Rayner (2006) recently demonstrated that when a word was more difficult to encode because it was faint (less contrast between the letters and background), readers fixated longer on target words. However, it is not at all clear how font difficulty would interact with either frequency or predictability.

Given that older adults tend to read more slowly than younger adults, it is important to investigate age differences in how individual words are processed to get a better understanding of the causes of the slowdown in their reading. Although there have been studies exploring age differences and the role of syntax on eye movements (Kemper et al., 2004), only one study (Kliegl et al., 2004) examined the impact of word properties on eye movements during reading. Cognitive aging theories attempt to explain age-related differences by relating performance on cognitive tasks, in this case reading, to general functioning or biological deficits or specific difficulties. For example, general functioning deficits, such as perceptual speed (Salthouse, 1992, 1996) or working memory (Kemtes & Kemper, 1997), are hypothesized to increase as one ages. These declines in performance of basic cognitive processes are hypothesized to be the explanatory mechanisms that account for predictable age-related differences. For example, Salthouse’s (1992, 1996) perceptual speed theory claims that older adults are slower in their processing speed than younger adults and that these differences in speed cause more complex cognitive functioning to show age differences in performance. Specifically, Salthouse (1996) claimed that differences in speed cause certain cognitive processes to fail to be completed in time for that information to be used by other processes. Other theories claim that typical age-related deficits arise from more general physiological problems. For example, Lindenberger and Baltes (1994; Baltes & Lindenberger, 1997) have claimed that age-related declines are the outcome of a common cause that they hypothesize to be the result of an aging brain. The aging brain is posited to affect all aspects of information processing from sensory input to intelligence.

In contrast to general deficits, age differences are claimed to be the result of deficits in specific cognitive processes (e.g., coordination of complex tasks; Verhaeghen, Klugel, & Mayr, 1997; or inhibition of irrelevant information; Hasher & Zacks, 1988). According to this type of theory, age-related differences should be evident when a task requires the use of the specific cognitive process, but when the specific process is not required for a task, no age-related difference should be present.

With respect to the focus of the current study, explanations for age-related differences in cognition have been applied to previous research in word identification, word naming, and reading. These studies have shown that word frequency and word predictability influence older and younger adults differently (for reviews, see Stine-Morrow, Miller, & Hertzig, 2006; Thornton & Lighet, 2006). For example, Spieler and Balota (2000) found that older adults were more influenced by the frequency of a word when asked to name it; they attributed this result to older adults’ more extensive experience with reading. Similarly, the predictability of a word in a spoken context has been shown to have a larger impact on older adults. Compared with younger adults, older adults were more influenced by the predictability of a word either when asked to identify the word in noise (Pichora-Fuller, Schneider, & Daneman, 1995) or when only a portion of a spoken target word was provided (Wingfield, Aberdeen, & Stine, 1991). This difference in the effect of predictability was not limited to spoken presentations, however. Speranza, Daneman, and Schneider (2000), in a reading experiment analogous to that of Pichora-Fuller et al., found that older adults were more influenced by the sentential context than younger adults in word identification in the presence of visual noise (both when presented with complete sentences and with word-by-word presentations).

Kliegl et al. (2004) examined the joint effects of frequency, predictability, and word length on the eye movements of older and younger readers using multiple regression analyses. These effects were examined both for all words of the reading material and for a subset of target words (one per sentence) representing a factorial design of frequency and word length (i.e., for nouns and verbs that were uncorrelated in length and frequency). They concluded that, aside from generally longer fixation durations for older adults, there were only a few age differences related to the effects of word frequency and word predictability. In the overall analyses, older readers showed larger frequency effects. In the target-word analyses, younger readers tended to skip a higher proportion of high-predictable words, whereas the older readers had a smaller probability of making multiple fixations on such words. Thus, the two groups appeared to be affected by frequency and predictability in slightly different ways.

Further examination of the effects of word frequency and word predictability as a function of age is important in the context of the development of computational models of eye movement control in reading such as E-Z Reader (Pollatsek, Reichle, & Rayner, 2006; Rayner et al., 2004; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 1999, 2003), SWIFT (Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005), Glenmore (Reilly & Radach, 2006), and SERIF (McDonald, Carpenter, & Shillock, 2005). For example, the E-Z Reader model relies on inputs influenced by the frequency and predictability of each word in the text to compute fixation times on words. Thus, it is essential to obtain information about the eye movement characteristics of older readers so as to ascertain that such models are not just describing the behavior of college-age readers and that the models are generalizable to other groups of readers. In the context of the SWIFT model, Laubrock, Kliegl, and Engbert (2006) simulated age differences in reading, with the model accounting for most of the differences in fixation time effects. Like Kliegl et al. (2004), they argued that differences
between older and younger readers were primarily quantitative (i.e., fixation times were longer and reading rates were slower for the older readers than the younger readers, but the pattern was the same). Although the fixation time data in the Kliegl et al. study appear to primarily reflect a quantitative difference, differences they found in refixation probability and skipping probability of words appear to reflect a qualitative difference. We discuss this issue further in the General Discussion section.

Our goal in the present research was to obtain further critical data to determine whether the eye movement characteristics of older readers are the same or differ from those of young readers. Given that older readers read slower than younger readers, there is some reason to suspect that their eye movements may be differentially affected by frequency and predictability. That is, Ashby, Rayner, and Clifton (2005) found that highly skilled and average readers differed with respect to how much reliance they placed on bottom-up processing (indexed by word frequency) and top-down processing (indexed by word predictability). Their results indicated that, although both the skilled and average readers fixated longer on low-frequency words when the target words were embedded in neutral contexts, the size of the frequency effect was larger for the average readers than the skilled readers. On the other hand, when the high- and low-frequency target words were embedded in highconstraint contexts, the two groups of readers showed different patterns of fixation time, as the average readers relied more heavily on predictability than frequency in processing the target words.

Just as the average readers in the Ashby et al. (2005) study appeared to rely on predictability more than did the skilled readers, and hence use a different strategy than the skilled readers, do older readers adopt different strategies than younger readers to compensate for the fact that they generally read slower? We think that the two most plausible outcomes are the following. First, older readers may simply read slower than younger readers, but in all other aspects their reading may be similar to that of the younger readers. In this case, we would primarily have a quantitative difference between the two groups, but no qualitative difference. Second, the older readers may adopt some type of specialized strategy to compensate for their slower reading. For example, they may adopt a riskier reading strategy. Specifically, they might rely more heavily on partial visual information (perhaps from parafoveal vision), or they might rely on frequency and predictability information to effectively skip more target words in the sentences; given their extensive experience with reading text, they may be willing to guess what the next word in the text is more than younger readers. In this case, there would be a qualitative difference between the older and the younger readers. We hoped that by manipulating frequency, predictability, and font difficulty, we would be able to determine whether older readers’ eye movements during reading are qualitatively or quantitatively different from those of younger readers. After presenting data on the eye movements of older and younger readers, we use the E-Z Reader model (which is presented in detail later in this article) to attempt to elucidate the results.

Method

Participants

Sixteen young adults who were students at the University of Massachussets at Amherst and 16 older adults from the community participated in the experiment. The young adults (5 men and 11 women) averaged 23.9 years of age (range = 18 years to 34 years), and the older adults (6 men and 10 women) averaged 77.5 years of age (range = 70 years to 92 years). The groups did not differ in number of years of schooling (15.5 years for the young and 15.7 years for the older readers).

All of the older adults’ had corrected 20/20 vision for reading, and they wore their glasses in the experiment. All of them reported that they spent quite a bit of time each day reading newspapers and books (which they preferred to watching TV), and the amount of time that they spent reading each week was roughly comparable to that spent by the younger readers (with an average of 17.4 hr per week and a range of 10–28 hr per week).

Materials

Each participant read 120 sentences. The first 4 sentences were fillers used for warm-up purposes. Embedded in the list was a subset of 80 sentences taken from Juhasz, Liversedge, White, and Rayner (2006). This subset contained 40 high-frequency nouns (M = 143 words per million using CELEX written frequency) and 40 low-frequency nouns (M = 1.35 words per million) as target words (Baayen, Piepenbrock, & Gulikers, 1995). The high- and low-frequency words were paired and matched exactly on word length (M = 7.85; range = 7 to 10 characters). As shown in the following example, two neutral sentence frames were written for each word pair so that either word could plausibly fit into the sentence, and thus each participant would see all the target words in a counterbalanced design:

High frequency–1:

I took a tour of a famous building while I was on holiday.

Low frequency–1:

I took a tour of a famous catacomb while I was on holiday.

High frequency–2:

The police closed off the dangerous building yesterday.

Low frequency–2:

The police closed off the dangerous catacomb yesterday.

Sentences ranged from 49 to 60 characters. Participants were presented with each target word, each in a different sentence frame, for a total of 80 experimental sentences.

In addition to the 80 sentences in which frequency of a target word was varied, participants read 36 sentences taken from Rayner and Well (1996) in which the predictability of a target word (as determined by sentence completion norms) was varied. The high-predictable target words had a mean completion rate of 86% (range = 73%–100%), the medium-predictable target words had a completion rate of 41% (range = 13%–68%), and the low-predictable target words had a completion rate of 4% (range = 3%–8%). There were 12 sentences that contained either a high-predictable or a medium-predictable target word. 12 sentences that contained either a medium-predictable or a low-predictable target word, and 12 sentences that contained either a high-predictable or a low-predictable target word. As shown in the following example, the pairs of target words each fit naturally into the sentence frame (though they varied on predictability):

High predictability 1:

Harriet sang while my brother played the piano for my birthday.

Medium predictability–1:

Harriet sang while my brother played the flute for my birthday.

1 All of the older participants were part of a larger group of older volunteers in the Amherst area. They serve as a control group for comparison with older patients with Alzheimer’s disease in an ongoing study.
Medium predictability–2: They were startled by the sudden noise from the next room.
Low predictability–1: They were startled by the sudden voice from the next room.
High predictability–2: He scraped the cold food from his plate before washing it.
Low predictability–2: He scraped the cold food from his spoon before washing it.

Counterbalancing procedures ensured that each sentence (with each of the target words) was read evenly. Target words were matched for word length and frequency: The average word length was 6.0, 5.9, and 6.2 letters for the high-, medium-, and low-predictable target words, respectively; the average word frequency (Francis & Kučera, 1982) was 55, 60, and 61 per million, respectively. Counterbalancing procedures were also used to ensure that half of the sentences in each frequency and predictability condition were presented in Times New Roman and half were presented in Old English. Figure 1 shows examples of the two fonts.

**Apparatus**

The sentences were presented on a 17-in. (about 43-cm) ViewSonic 17PS monitor attached to a Pentium 166 MHZ computer interfaced with an SR Research Ltd. Eye-Link II eye tracking system. This system has high spatial resolution and a sampling rate of 500 Hz (2-ms temporal resolution). Although viewing was binocular, only the participant’s right eye was tracked. The sentences were displayed on the monitor on a single line with lowercase letters (except when capitals were appropriate) on the monitor. The letters were presented in black on a white background. Participants were seated 77 cm from the monitor, and three characters (presented in 16-point font) equaled 1° of visual angle. The Eye-Link system allows readers to move their heads (because the head is tracked as well as the eyes), but to ensure maximum accuracy, we used a chin rest.

**Procedure**

When participants first arrived at the experiment, some background information was obtained regarding the amount of time they spent reading each week. Then the eye movement system was calibrated for each participant. This typically took about 5 min. Calibration accuracy was determined by asking each participant to fixate, in turn, on three fixation points that appeared sequentially in three horizontal screen positions (where the sentence would subsequently appear). The experimenter checked the accuracy of the calibration prior to each sentence; if the calibration was not good, the participant was recalibrated.

Participants were asked to read sentences that appeared one at a time on the video monitor as their eye movements were recorded. They were asked to read each sentence normally for comprehension and to read as if they were reading a book or a newspaper. They were further told that they would be asked comprehension questions about the sentences that they read. These questions were randomly presented following one fourth of the sentences and were about the meaning of the sentence that was just read.

The answer was a yes–no response that was indicated by pushing one of two response keys that were in front of the participant.

**Results**

There was little difference between the young and older readers on the comprehension questions; the young readers were correct 85% of the time, and the older readers were correct 82% of the time ($t < 1$). We first present the comparison of the older and young readers on some overall global measures. The results for the frequency and predictability manipulations are then reported separately. For each of these manipulations, we also discuss the effects of age and font difficulty. For each manipulation, we examined four standard reading time measures (Rayner, 1998) for the relevant target word in a sentence: (a) first-fixation duration (the duration of the first fixation on the word), (b) single-fixation duration (the duration of the fixation when the reader made only one fixation on the target word), (c) gaze duration (the sum of all of the fixations on a word prior to moving to another word), and (d) total reading time (the sum of all fixations on a word, including regressions to the word). In addition, we computed the probability of fixating on the word during the first-pass reading of the sentence and the initial landing position of the eyes in the target word. There were very few cases of a track loss (less than 1% of the data). If a short fixation (under 80 ms) was on a character adjacent to one that was over 80 ms, the two were combined; otherwise, fixations less than 80 ms and longer than 800 ms were eliminated. Altogether about 2% of the data were lost or eliminated.

**Global Results**

Table 1 shows some global aspects of reading performance for the young and older readers. The means in Table 1 were analyzed via 2 (age) × 2 (font) analyses of variance (ANOVAs). As is clear in the table, the older readers took longer to read the sentences than the younger readers, $F(1, 30) = 4.184$, $p < .05$; their fixations were longer, $F(1, 30) = 5.85$, $p < .05$; and they made more regressions per sentence, $F(1, 30) = 4.46$, $p < .05$. Also, although the difference was not statistically significant, the older readers made more fixations per sentence, $F(1, 30) = 2.53$, $p = .12$. An ANOVA on the saccade length data (which included direction of saccade: forward vs. regressive saccades) indicated that, on average, the saccades of older readers were longer than the saccades of younger readers, both for forward (9.0 character spaces vs. 8.3 character spaces) and regressive saccades (13.7 characters vs. 11.4 characters), although the main effect of age was only marginally significant, $F(1, 30) = 3.03$, $p < .10$. On average, regressive saccades were longer than forward saccades (12.5 vs. 8.6 characters), $F(1, 30) = 17.6$, $p < .001$, but the interaction between group and forward versus backward saccades was not significant ($F < 1$). In addition, the variability in average saccade length was significantly larger for the older readers for both forward and backward saccades, $F(15, 15) = 3.40$, $p < .025$, and $F(15, 15) = 3.67$, $p < .01$, respectively.

In addition to the age effects, it is very clear that both groups of readers found the Old English font more difficult to read. Averaged over the two groups, it led to significantly longer reading time, $F(1, 30) = 24.91$, $p < .001$; longer fixations, $F(1, 30) = 196.64$, $p < .001$; and
shorter saccades, \( F(1, 30) = 41.4, p < .001 \), than the Times New Roman font. Moreover, the font effect was bigger for the older group, and the Group × Font interaction reached significance for sentence reading time, \( F(1, 30) = 5.72, p < .05 \), and fixation duration, \( F(1, 30) = 43.12, p < .001 \), and was marginally significant for number of fixations, \( F(1, 30) = 3.97, p = .055 \), and number of regressions, \( F(1, 30) = 3.51, p = .071 \).

In summary, the Old English font made encoding the words in the sentence more difficult. It is also interesting to note that, quantitatively, the young readers’ performance with the Old English font was quite similar to the older readers’ performance with the Times New Roman font except for their mean saccade length and number of regressions. In fact, when these two conditions (i.e., younger readers–Old English and older readers–Times New Roman) were directly compared, the only measures for which the age effect had an \( F \) value greater than 1 were number of regressions, \( F(1, 30) = 5.27, p < .05 \), and saccade length, \( F(1, 30) = 1.26, p > .20 \).

**Frequency Data Set**

To examine the effect of word frequency and the extent to which age and font difficulty influenced processing of the target word, we carried out, for each of the five dependent variables, a series of 2 (frequency) × 2 (font difficulty) × 2 (age) ANOVAs using participants and items as random effects variables. Frequency and font were within-subject manipulations, and age was a between-subjects manipulation. Unless otherwise noted, only significant effects are reported and any main effects or interactions that are not discussed did not approach significance.

**Fixation time.** Table 2 shows the fixation time results.\(^2\) Overall, readers fixated longer on low-frequency words than on high-frequency words. This result is the standard word frequency effect that has now been replicated many times in eye movement research (Rayner, 1998). The size of the effect was 27 ms in first-fixation duration, \( F(1, 30) = 28.28, p < .001, F(2, 1, 79) = 33.28, p < .001 \); 32 ms in single-fixation duration, \( F(1, 30) = 37.5, p < .001, F(2, 1, 79) = 33.62, p < .001 \); 67 ms in gaze duration, \( F(1, 30) = 51.78, p < .001, F(2, 1, 79) = 63.82, p < .001 \); and 96 ms in total fixation time, \( F(1, 30) = 45.50, p < .001, F(2, 1, 79) = 60.12, p < .001 \). Readers also fixated longer on the target word when it was in the more difficult Old English font than when it was in the Times New Roman font. The size of the font effect was 24 ms in first-fixation duration, \( F(1, 30) = 20.8, p < .001, F(2, 1, 79) = 20.92, 37 \) ms in single-fixation duration, \( F(1, 30) = 36.23, p < .001, F(2, 1, 79) = 24.95, p < .001; 69 \) ms in gaze duration, \( F(1, 30) = 45.48, F(2, 1, 79) = 51.03; and 104 ms in total time, \( F(1, 30) = 19.83, p < .001, F(2, 1, 79) = 57.89, p < .001 \).

The main effects of frequency and font difficulty were expected. More interesting is that (a) there was also a main effect of age and (b) age interacted to some extent with both frequency and font. There was no hint of an interaction of font and frequency, nor of a three-way interaction (\( Fs < 1 \)). The former is somewhat surprising, as one might have expected the frequency effect to be bigger in the Old English font. Overall, the older readers’ first fixations on the target words were 316 ms compared with 268 ms for the younger readers, \( F(1, 30) = 6.5, p < .05, F(2, 1, 79) = 141.2, p < .001 \); their single-fixation durations averaged 328 ms compared with 279 ms for the younger readers, \( F(1, 30) = 5.33, p < .05, F(2, 1, 79) = 73.3, p < .001 \); their gaze durations averaged 390 ms compared with 326 ms for the younger readers, \( F(1, 30) = 4.34, p < .05, F(2, 1, 79) = 65.41, p < .001 \); and their total reading times for the target words averaged 499 ms compared with 371 ms for the younger readers, \( F(1, 30) = 5.09, p < .05, F(2, 1, 79) = 116.39, p < .001 \).

---

\( ^2 \) The means reported are based on the data without any trimming (other than track losses, which amounted to less than 1% of the data). When we trimmed the data by eliminating all data points more than 2.5 standard deviations from the mean, the pattern of data remained the same as presented in the tables (though the means and standard deviations were obviously smaller). Also, we report total reading time for completeness, but it is of less interest than the other three measures because they are all first-pass measures (and reflect immediate on-line processing), whereas total time includes later occurring regressions (and thus includes second-pass reading time).

---

### Table 1

| Sentence Reading Time (SRT), Average Fixation Duration (FD), Average Number of Fixations (NFix), Average Number of Regressions (NReg), and Average Saccade Length (SL) for the Young and Older Readers Reading Sentences in Times New Roman or Old English |
|----------------------------------|----|----|----|----|----|----|----|
| Sentence type                    | M  | SD | M  | SD | M  | SD | M  | SD |
| Young readers                    |    |    |    |    |    |    |    |    |
| Times New Roman                 | 2.394 | 451 | 246 | 27 | 9.8 | 1.4 | 0.64 | 0.15 |
| Old English                     | 2.830 | 448 | 263 | 26 | 10.7 | 1.2 | 0.65 | 0.18 |
| Older readers                    |    |    |    |    |    |    |    |    |
| Times New Roman                 | 2.976 | 1,373 | 260 | 49 | 11.1 | 3.9 | 0.91 | 0.40 |
| Old English                     | 4,213 | 2,446 | 303 | 65 | 13.6 | 6.3 | 1.01 | 0.49 |

**Note.** SRT and FD times are given in milliseconds. SRT was defined as the time from when the sentence appeared until the reader terminated the sentence via a button press.
The frequency effect was larger for the older readers than for the younger readers, although the difference was only marginally significant on some measures. For first-fixation duration, the frequency effect was 18 ms for the younger readers and 37 ms for the older readers, \( F(1, 30) = 3.65, p = .066, F(1, 79) = 4.02, p < .05 \); for single-fixation duration, it was 25 ms for the younger readers and 40 ms for the older readers, \( F(1, 30) = 4.36, p < .05, F(1, 79) = 4.18, p < .05 \); for gaze duration, it was 52 ms for the younger readers and 79 ms for the older readers, \( F(1, 30) = 4.79, p < .05, F(1, 79) = 3.23, p = .07 \); and for total time, it was 67 ms for the younger readers and 124 ms for the older readers, \( F(1, 30) = 3.98, p = .055, F(1, 79) = 3.98, p < .05 \). Similarly, there was a tendency for there to be a bigger font effect for the older readers. For first-fixation duration, the font effect was 16 ms for the younger readers and 32 ms for the older readers, \( F(1, 30) = 2.13, p = .155, F(1, 79) = 5.04, p < .05 \); for single-fixation, it was 26 ms for the younger readers and 48 ms for the older readers, \( F(1, 30) = 3.36, p = .077, F(1, 79) = 3.62, p = .061 \); for gaze duration, it was 49 ms for younger readers and 88 ms for the older readers, \( F(1, 30) = 3.78, p = .061, F(1, 79) = 6.89, p < .01 \); and for total time, it was 62 ms for the younger readers and 146 ms for the older readers, \( F(1, 30) = 3.23, p = .083, F(1, 79) = 12.74, p < .001 \).

**Fixation probability (skipping) and landing positions.** Table 2 also shows the probability of a first-pass fixation on the target words. We discuss the results in terms of skipping rate (which is the complement of the fixation probability). Overall, readers were more likely to skip high-frequency target words (skipping probability = .14) than low-frequency words (.08), \( F(1, 30) = 15.67, p < .001, F(1, 79) = 15.14, p < .001 \), and were more likely to skip a target word when the font was the less difficult Times New Roman font (.13) than when it was more difficult Old English font (.10). \( F(1, 30) = 4.26, p < .05, F(1, 79) = 6.39, p < .05 \). More interesting, however, older readers were somewhat more likely to skip a target word (.14) than younger readers (.08), \( F(1, 30) = 3.04, p = .092, F(1, 79) = 20.14, p < .001 \), and showed a different pattern for high- and low-frequency words. Although the older readers were much more likely to skip a high-frequency word (.19) than a low-frequency word (.09), the younger readers were equally likely to skip both (.08 for both high- and low-frequency target words). The interaction was significant, \( F(1, 30) = 15.66, p < .001; F(1, 79) = 21.98, p < .001 \). Font difficulty did not interact with age (Fs < 1).

Given that the older readers were more likely to skip the target word, it might be that they adopted a riskier reading strategy (O’Regan, 1990) involving moving their eyes further in the text with each saccade but then regressing more frequently. However, the global data analyses (see Table 1) indicated that although they did indeed regress more frequently than the younger readers, the difference wasn’t very large (about 63% of the younger readers’ fixations followed a regression compared with 8.2% for the older readers). Nevertheless, we thought it would be of value to examine the issue using the target-word data, as the situation was more controlled. We examined two measures. The first was the probability of refixating on the target word (i.e., the probability of making more than one fixation on the target word before moving out of that word), which might be higher for older readers if they adopted a general strategy of sending their eyes further and thus had to refixate some words because this led to less than optimal placement. However, there was no evidence for this, as the probability of refixating the target word was .22 for the younger readers and .21 for the older readers. The second measure was the probability of fixating on the target word given that it was initially skipped. When the target word was initially skipped, older readers regressed to it 28% of the time, whereas younger readers only did so 13% of the time, \( t(30) = 3.21, p < .01 \). Thus, although the older readers were more likely to skip a word (skipping the target word 14% of the time compared with 8% for the younger readers), they were also more likely to regress to skipped words. However, there was little frequency involvement in the different pattern of regressions for the younger and older readers. The regression rate back to the target words was about 1% greater for the low-frequency targets than for the high-frequency targets, which surprisingly was a significant difference, \( F(1, 30) = 4.18, p < .05 \), but there was no

<table>
<thead>
<tr>
<th>Target word type</th>
<th>FFD</th>
<th>SFD</th>
<th>GD</th>
<th>TT</th>
<th>PS</th>
<th>ILP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Young readers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNR-HF</td>
<td>252</td>
<td>29</td>
<td>254</td>
<td>32</td>
<td>278</td>
<td>35</td>
</tr>
<tr>
<td>TNR-LF</td>
<td>266</td>
<td>64</td>
<td>279</td>
<td>65</td>
<td>326</td>
<td>93</td>
</tr>
<tr>
<td>OE-HF</td>
<td>266</td>
<td>28</td>
<td>280</td>
<td>36</td>
<td>323</td>
<td>48</td>
</tr>
<tr>
<td>OE-LF</td>
<td>286</td>
<td>37</td>
<td>304</td>
<td>45</td>
<td>379</td>
<td>66</td>
</tr>
<tr>
<td>Older readers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNR-HF</td>
<td>280</td>
<td>52</td>
<td>280</td>
<td>52</td>
<td>305</td>
<td>63</td>
</tr>
<tr>
<td>TNR-LF</td>
<td>321</td>
<td>83</td>
<td>329</td>
<td>89</td>
<td>387</td>
<td>122</td>
</tr>
<tr>
<td>OE-HF</td>
<td>315</td>
<td>74</td>
<td>337</td>
<td>92</td>
<td>394</td>
<td>139</td>
</tr>
<tr>
<td>OE-LF</td>
<td>349</td>
<td>78</td>
<td>368</td>
<td>90</td>
<td>474</td>
<td>142</td>
</tr>
</tbody>
</table>

Note. FFD, SFD, GD, and TT times are given in milliseconds.
Frequency × Age interaction \((F < 1)\). These results indicate that the regression pattern wasn’t mirroring the skipping pattern in a simple way. Overall, this pattern indicates that older readers may have been using a more “risky” reading strategy (O’Regan, 1990), but one in which specific words are skipped rather than just involving longer saccades in general.

Table 2 also shows the eyes’ initial landing position in the target words. Whereas font did not influence where the eyes initially landed \((ps > .10)\), on average, the eyes landed a bit further into high-frequency words \((3.48\) characters\) than low-frequency words \((3.31\) characters\), \(F(1, 30) = 4.83, p < .05; F(1, 79) = 3.39, p = .069\). Also, younger readers tended to land further into the target word \((3.55\) characters\) than the older readers \((3.23\) characters\), \(F(1, 30) = 3.42, p = .074; F(1, 79) = 28.15, p < .001\).

Given that landing position in a word is heavily influenced by the launch site of the saccade (McConkie, Kerr, Reddix, & Zola, 1988; Rayner et al., 1996), we also examined the launch sites of the saccades into the target word. Older readers’ launch sites were further \((6.2\) letter spaces\) from the beginning of the target word than those of young readers \((5.6\) letter spaces\), \(F(1, 30) = 6.04, p < .05; F(1, 79) = 10.55, p < .01\). Thus, the fact that the older readers landed earlier in the target word is plausibly due to the fact that they were, on average, fixated further away on the launch site fixation than the younger readers. Of interest, there was no significant difference in launch sites between high- \((6.0\) letter spaces\) and low-frequency \((5.8\) letter spaces\) words \((ps > .19)\). However, there was a significant difference between fonts \((\text{Times New Roman} = 6.1\) letter spaces, \text{Old English} = 5.7 letter spaces\), \(F(1, 30) = 7.94, p < .01; F(1, 79) = 6.8, p < .05\). Not surprisingly, when the target word was skipped, the launch site was much closer to the beginning of the target word \((2.5\) letter spaces from the beginning of the word) than when it was not skipped \((5.9\) letter spaces\). When the word was skipped, the only variable that was significant was font, as the eyes were farther from the target word for \text{Times New Roman} \((3.4\) letter spaces\) than \text{Old English} \((1.7\) letter spaces\), \(F(1, 30) = 27.21, p < .001; F(1, 79) = 79.74, p < .001\).

We thought it would be instructive to compare the older readers in the \text{Times New Roman} condition and the younger readers in the \text{Old English} condition—conditions where the global analyses suggested overall reading difficulty was essentially equated. As with the global duration measures, there were no differences between these two conditions that were close to being significant on any of the fixation duration measures in the participant analyses: first-fixation duration, \(F(1, 30) = 1.89, p = .18; F(1, 79) < 1\); single-fixation duration, \(Fs < 1\); gaze duration, \(Fs < 1\); total time, \(Fs < 1\). However, for the probability of skipping the target word, there was both a significant difference between these conditions, \(F(1, 30) = 8.19, p < .01; F(1, 79) = 26.2, p < .001\), and a significant interaction with frequency, \(F(1, 30) = 4.41, p < .05; F(1, 79) = 5.49, p < .025\). This suggests that the differences in skipping are not merely attributable to differences in being able to visually process the text being read.

**Summary of frequency effects.** Consistent with a great deal of other research (Rayner, 1998), there were clear effects of frequency. And, not surprisingly, there were clear effects of font difficulty. The most interesting age differences were that the older readers’ fixation times were longer than those of younger readers and the size of the frequency effect was larger for the older readers. In addition, the older readers were more likely to skip high-frequency words than low-frequency words, whereas the younger readers showed no difference. Moreover, the older readers re-gressed to skipped words more frequently than young readers, which is consistent with the view that the older readers adopted a riskier reading strategy. Given that there were no differences between the high- and low-frequency target words in predictability, it might also be tempting to assume that the skipping results suggest that older readers use parafoveal information more effect-ively than do younger readers. According to this argument, they were able to identify the high-frequency words in parafoveal vision and to then program a saccade to skip the word and move to the next word in the text. We suspect that this was not the case, and we return to this issue in the Discussion section.

**Predictability Data Set**

To examine the effect of word predictability when word frequency is equated, and the extent to which age and font difficulty influenced the processing of these target words, we carried out, for each of the dependent variables, a series of 3 \((\text{predictability}) \times 2 \,(\text{font difficulty}) \times 2 \,(\text{age})\) ANOVAs using participants and items as random effects variables. Predictability and font were within-subject manipulations, and age was a between-subjects manipulation. Unless otherwise noted, only significant effects are reported and any main effects or interactions that are not discussed did not approach significance.

**Fixation time.** Table 3 shows the fixation time results. Consistent with Rayner and Well (1996), there was an effect of predictability on fixation times. Readers’ first-fixation durations were shorter on high-predictable words \((269\) ms\) than on medium-predictable \((289\) ms\) or low-predictable words \((288\) ms\), \(F(1, 60) = 5.42, p < .01; F(2, 70) = 5.38, p < .01\). Similarly, single-fixation times on the high-, medium-, and low-predictable words were \(273\) ms, \(287\) ms, and \(298\) ms, respectively, \(F(2, 60) = 4.03, p < .05; F(2, 70) = 8.86, p < .001\), and gaze duration times were \(289\) ms, \(316\) ms, and \(325\) ms, respectively, \(F(2, 60) = 3.06, p = .054; F(2, 70) = 3.29, p < .05\). Post hoc t-tests confirmed that for all three measures, the fixation times on the high-predictable target differed \((ps < .01)\) from the medium- and low-predictable target word (which did not differ from each other). Total time showed a slightly different pattern, as the high-, medium-, and low-predictable words averaged \(336\) ms, \(374\) ms, and \(395\) ms, respectively, \(F(2, 60) = 5.17, p < .01; F(2, 70) = 7.42, p < .001\). However, post hoc t tests again indicated that the total times on the high-predictable target word differed \((ps < .01)\) from the other two conditions (which again did not differ from each other).

As with the frequency manipulation, readers, on average, looked longer at the target words in the \text{Old English} font than in the \text{Times New Roman} font: \(300\) ms versus \(264\) ms for first-fixation duration, \(F(1, 30) = 28.85, p < .001; F(1, 35) = 73.02, p < .001\); \(310\) ms versus \(263\) ms for single-fixation duration, \(F(1, 30) = 9.09, p < .01; F(1, 35) = 30.27, p < .001\); \(332\) ms versus \(287\) ms for gaze duration, \(F(1, 30) = 28.43, p < .001; F(1, 35) = 124.55, p < .001\); and \(412\) ms versus \(325\) ms for total time, \(F(1, 30) = 20.3, p < .001; F(1, 35) = 102.1, p < .001\). There was a tendency for font and predictability to interact (especially in the items analysis): first-fixation duration, \(F(2, 60) = 2.38, p = .10; F(2, 70) = 7.3, p < .001\); single-fixation duration, \(F(2, 60) = 1.62, p = .21; F(2,
Table 3

First-Fixation Duration (FFD), Single-Fixation Duration (SFD), Gaze Duration (GD), Total Reading Time (TT), Probability of Skipping (PS), and Initial Landing Position (ILP) for High-Predictable (HP), Medium-Predictable (MP), and Low-Predictable (LP) Target Words in Times New Roman (TNR) and Old English (OE)

<table>
<thead>
<tr>
<th>Target word type</th>
<th>FFD M</th>
<th>FFD SD</th>
<th>SFD M</th>
<th>SFD SD</th>
<th>GD M</th>
<th>GD SD</th>
<th>TT M</th>
<th>TT SD</th>
<th>PS M</th>
<th>PS SD</th>
<th>ILP M</th>
<th>ILP SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNR-HP</td>
<td>230</td>
<td>47</td>
<td>234</td>
<td>52</td>
<td>250</td>
<td>73</td>
<td>297</td>
<td>111</td>
<td>.26</td>
<td>.21</td>
<td>2.7</td>
<td>1.1</td>
</tr>
<tr>
<td>TNR-MP</td>
<td>267</td>
<td>66</td>
<td>272</td>
<td>68</td>
<td>274</td>
<td>69</td>
<td>306</td>
<td>87</td>
<td>.22</td>
<td>.17</td>
<td>2.9</td>
<td>1.0</td>
</tr>
<tr>
<td>TNR-LP</td>
<td>252</td>
<td>56</td>
<td>257</td>
<td>62</td>
<td>273</td>
<td>60</td>
<td>318</td>
<td>92</td>
<td>.20</td>
<td>.19</td>
<td>2.9</td>
<td>0.6</td>
</tr>
<tr>
<td>OE-HP</td>
<td>265</td>
<td>41</td>
<td>258</td>
<td>44</td>
<td>278</td>
<td>65</td>
<td>313</td>
<td>77</td>
<td>.24</td>
<td>.17</td>
<td>3.1</td>
<td>1.0</td>
</tr>
<tr>
<td>OE-MP</td>
<td>258</td>
<td>50</td>
<td>262</td>
<td>56</td>
<td>274</td>
<td>55</td>
<td>317</td>
<td>121</td>
<td>.18</td>
<td>.16</td>
<td>3.0</td>
<td>1.0</td>
</tr>
<tr>
<td>OE-LP</td>
<td>282</td>
<td>48</td>
<td>292</td>
<td>53</td>
<td>316</td>
<td>76</td>
<td>399</td>
<td>102</td>
<td>.15</td>
<td>.13</td>
<td>2.9</td>
<td>1.1</td>
</tr>
<tr>
<td>TNR-LP</td>
<td>277</td>
<td>70</td>
<td>274</td>
<td>71</td>
<td>291</td>
<td>69</td>
<td>318</td>
<td>80</td>
<td>.24</td>
<td>.22</td>
<td>2.7</td>
<td>1.3</td>
</tr>
<tr>
<td>TNR-MP</td>
<td>294</td>
<td>98</td>
<td>291</td>
<td>105</td>
<td>326</td>
<td>100</td>
<td>381</td>
<td>178</td>
<td>.22</td>
<td>.16</td>
<td>2.6</td>
<td>0.9</td>
</tr>
<tr>
<td>TNR-LP</td>
<td>266</td>
<td>38</td>
<td>306</td>
<td>55</td>
<td>307</td>
<td>89</td>
<td>329</td>
<td>113</td>
<td>.30</td>
<td>.24</td>
<td>2.7</td>
<td>1.2</td>
</tr>
<tr>
<td>OE-HP</td>
<td>305</td>
<td>96</td>
<td>327</td>
<td>86</td>
<td>336</td>
<td>125</td>
<td>413</td>
<td>189</td>
<td>.35</td>
<td>.27</td>
<td>2.8</td>
<td>1.2</td>
</tr>
<tr>
<td>OE-MP</td>
<td>338</td>
<td>74</td>
<td>321</td>
<td>115</td>
<td>388</td>
<td>109</td>
<td>493</td>
<td>225</td>
<td>.25</td>
<td>.26</td>
<td>2.3</td>
<td>0.8</td>
</tr>
<tr>
<td>OE-LP</td>
<td>353</td>
<td>116</td>
<td>339</td>
<td>90</td>
<td>404</td>
<td>143</td>
<td>536</td>
<td>332</td>
<td>.19</td>
<td>.25</td>
<td>2.6</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Young readers

Older readers

Note: FFD, SFD, GD, and TT times are given in milliseconds.

70) = 2.53, \( p = .087 \); gaze duration, \( F(1, 60) = 2.53, p = .088 \); \( F(2, 70) = 6.47, p < .01 \); and total time, \( F(1, 60) = 4.14, p < .05 \); \( F(2, 70) = 4.51, p < .05 \).

As with the target words used in the frequency manipulation, older readers’ fixations, on average, were longer on the target words than those of the younger readers: 305 ms versus 259 ms for first-fixation duration, \( F(1, 30) = 5.34, p < .05 \); \( F(1, 35) = 132.19, p < .001 \); 310 ms versus 263 ms for single-fixation duration, \( F(1, 30) = 5.5, p < .05 \); \( F(1, 35) = 99.23, p < .001 \); 342 ms versus 278 ms for gaze duration, \( F(1, 30) = 8.85, p < .05 \); \( F(1, 35) = 83.92, p < .001 \); and 412 ms versus 325 ms for total reading time, \( F(1, 30) = 3.86, p < .05 \); \( F(1, 35) = 87.74, p < .001 \). Also, as with the frequency target words, there was an interaction between age and font, with a larger font effect for the older readers: 53 ms versus 18 ms for first-fixation duration, \( F(1, 30) = 6.96, p < .05 \); \( F(1, 35) = 28.6, p < .001 \); 39 ms versus 16 ms for single-fixation duration, \( F(1, 30) = 1.51, p = .23 \); \( F(1, 35) = 4.06, p = .052 \); 68 ms versus 33 ms for gaze duration, \( F(1, 30) = 4.95, p < .05 \); \( F(1, 35) = 36.63, p < .001 \); and 138 ms versus 70 ms for total time, \( F(1, 30) = 6.99, p < .05 \); \( F(1, 35) = 42.95, p < .001 \). However, unlike the frequency manipulation, there was no hint of an interaction between age and predictability (Fs < 1).

Fixation probability (skipping) and landing position. Table 3 also shows the probability of a first-pass fixation on the target words. Predictability influenced skipping, although the effect was only significant in the items analysis, \( F(1, 30) = 2.25, p = .115 \); \( F(2, 70) = 4.26, p < .05 \). The basic pattern in the data (see Table 3) mimicked the fixation time data with readers more likely to skip high-predictable (.27) than either medium- (.22) or low-predictable (.21) target words. Age also was a factor in skipping, with the older readers skipping target words more frequently (.26) than the younger readers (.21), but once again the effect was only significant in the items analysis, \( F(1, 30) = 1.01, p > .10 \); \( F(1, 35) = 9.02, p < .01 \). Age did not interact with predictability (Fs < 1), but there was a hint of an interaction between predictability and font, \( F(2, 60) = 2.64, p = .08 \); \( F(2, 70) = 4.03, p < .05 \), with the difference in font exerting little influence for the high- and medium-predictable target words but the easier font leading to more skips than the difficult font for the low-predictable target words.

Given that the older readers were more likely to skip the target word (as with the frequency data), we again examined the probability of refixating on the target word and the probability of refixating on the target word given that it was initially skipped. The probability of refixating the target word was .10 for the younger readers and .12 for the older readers. When the target word was initially skipped, older readers regressed to it 19% of the time, whereas younger readers regressed to it 18% of the time. Neither of these differences was significant. Thus, in contrast to the frequency data, there was no indication that the older readers were more likely to regress to skipped words.

Table 3 also shows the eyes’ initial landing position in the target words. There was an age effect (but only by items); the older readers tended to land earlier in the target word (2.6 characters) than did the younger readers (2.9 characters), \( F(1, 30) = 2.65, p = .116 \); \( F(1, 35) = 12.93, p < .001 \). As with the frequency data, analyses of the launch site data revealed that this effect was due to the older readers’ being fixated, on average, further away (6.9 letter spaces) from the beginning of the target word than were the younger readers (6.2 letter spaces), \( F(1, 30) = 5.86, p < .05 \); \( F(1, 35) = 5.11, p < .01 \).

\(^3\) The skipping rates were much higher for the predictability data than the frequency data because the target words were shorter in the former case, and word length has a big effect on skipping (Rayner, 1998).
than the Old English (2.8 letters), the difference on the word with the Times New Roman (3.3 letter spaces) was again that readers were further away from the beginning of the word when they subsequently skipped the target word (3 letter spaces) than when they did not (6.6 letter spaces). The only significant effect when the target word was skipped was that readers were further away from the beginning of the word with the Times New Roman condition (3.3 letter spaces) than the Old English condition (2.8 letters), $F(1, 30) = 4.43, p < .05; F(1, 35) = 4.89, p < .05$.

Again, we compared the performance for the older readers in the Times New Roman condition with the younger readers in the Old English condition. As in the analysis of the frequency data, the differences on all the fixation duration measures were far from significant (all $F$s $< 1$). However, possibly because of greater noise in these data, the differences in skipping rates were significant only in the item analyses, $F(1, 30) = 2.24, p < .15; F(1, 35) = 7.84, p < .01$. Thus, although the evidence is admittedly weaker here, the pattern of these analyses suggests that the overall differences between the two groups of readers in skipping rates for the predictability target words was not due to differences in the difficulty of visually processing the text in front of them.

**Summary of predictability effects.** Consistent with a great deal of other research (Rayner, 1998), there were clear effects of predictability. And, consistent with the frequency manipulation, there were clear effects of font difficulty and the older readers’ fixation times were longer than younger readers’ times. However, unlike the frequency manipulation, there was no interaction between age and predictability, and thus the groups appeared to be using predictability information similarly. As with the target words used in the frequency analysis, there was a difference in skipping rates, as the older readers were more likely to skip target words (independent of their predictability) than were younger readers. This result, again, could be used to argue either that the older readers were using parafoveal information more effectively than the younger readers or that they were engaged in a more risky reading strategy. Unlike with the frequency data, there was no evidence that the older readers were regressing more frequently to skipped words. Consistent with prior research (Rayner, Binder, Ashby, & Pollatsek, 2001), predictability did not influence the initial landing position in a word, though it did influence the overall probability of skipping a word.

**Group Differences in Oculomotor Control?**

We suspect that overall, the older readers’ data imply that they adopted a riskier reading strategy than did younger adults (see later text). However, one could perhaps account for the differences in word skipping because of poorer oculomotor control in the older readers. To test this possibility, we examined the landing position distributions for the two groups of readers for all seven letter target words (collapsed over the frequency and predictability items) in our sample. As can be seen in Figure 2, the landing position distributions (all fixations that landed on either the target word or space preceding it) were quite similar between the two groups. Not surprisingly, given the mean landing positions reported in Tables 2 and 3, the peak of the distribution is a bit earlier in the word for the older readers than the younger readers. However, both distributions are fairly representative of and consistent with previously reported data (McConkie et al., 1988; Rayner, 1978; Rayner et al., 1996). We take these data as being consistent with the notion of older readers’ adopting a more risky strategy and inconsistent with the notion of poorer oculomotor control among those readers. Moreover, given that older readers tended to land earlier on the word, it is unlikely that their higher skipping rates were due to poor oculomotor control. That is, if the increased skipping rates for older readers were largely due to intending to fixate on the target word but overshooting it, one would have expected to see an increase in fixation probabilities for the older readers near the end of the word as well.

**Discussion**

The results of the present study confirm that word frequency and word predictability independently exert a strong influence on how long readers fixate a word and do so independent of the length of the word. In addition, they provide clear evidence that how difficult a word is to decipher (as determined by the font difficulty) also influences how long readers fixate on a word. More interestingly, the present data demonstrate that older readers are also very much influenced by these variables. Indeed, for the most part, the results suggest that any differences between younger and older readers in terms of fixation times on a word are primarily quantitative. The older readers’ fixations were consistently longer in reading the sentences in general, and, like young readers, their fixation time on a word was influenced by its frequency, its predictability, and how difficult the font was to decipher. Our data are quite consistent with data reported by Kliegl et al. (2004) and Laubrock et al. (2006).

There were, however, two interesting qualitative differences between the older and younger readers: (a) The older readers were more likely to skip a target word than were the younger readers, and (b) the older readers also made more regressions than did the younger readers (findings also reported by Kliegl et al., 2004; Laubrock et al., 2006).

---

4 It is possible that tendencies to undershoot the target word could have influenced skipping for the older readers, but we think it is somewhat unlikely that it had a big effect. If older readers have more oculomotor error, the landing distribution should be much flatter than it is in Figure 2 because undershoots would result in more fixations at the end of any given word and overshoots would result in more fixations at the beginning.
Laubrock et al., 2006). These two effects contributed to the tendency for older readers’ saccades to be longer than those of the younger readers. Moreover, our analyses comparing older readers with Times New Roman and younger readers with Old English suggested that these effects still occurred when visual processing of the text was about equal for the two groups. This is similar to findings that although age-related sensory changes are correlated with cognitive processing difficulties, they do not account for age-related difficulties in cognitive functioning (Lindenberger, Scherer, & Baltes, 2001). In the frequency data set, the higher skipping rate for older readers was primarily for the high-frequency words; in the predictability data set, they skipped target words (independent of their predictability) more than the younger readers did. As we noted earlier, this result could mean that older readers are better able, on the current fixation, to identify the word to the right of the currently fixated word and to program a saccade to skip that word. However, given that it is known that older participants do not process nonfoveal information as effectively as do younger participants (Ball, Beard, Roenker, Miller, & Griggs, 1988; Sekuler, Bennett, & Mamelak, 2000), this explanation seems rather implausible.

An alternative, more plausible, explanation of the result is that the older readers adopt a more risky reading strategy than do the younger readers. That is, in an attempt to speed up their reading, older adults may rely more heavily on partial parafoveal information. This may allow them to move through text more quickly but may have the consequence that they may have to regress to earlier portions of the text more frequently to clarify text that was not correctly processed earlier. We are not the first to suggest that older and younger adults use different reading strategies to compensate for age-related differences. Stine-Morrow and colleagues (Miller & Stine-Morrow, 1998; Stine, 1990; Stine-Morrow, Loveless, & Soederberg, 1995) have found that older adults tend to allocate their processing resources differently than younger adults during reading (see Stine-Morrow et al., 2006, for a review of strategic differences in reading). Older adults appear to allocate more processing resources to the integration of new concepts as they are introduced in text, whereas younger adults tend to wait until the end of the sentence to integrate the new concepts (Miller & Stine-Morrow, 1998). Miller and Stine-Morrow attributed this pattern to an attempt to compensate for age-related differences in working memory capacity. In other words, older adults may break up discourse into smaller conceptual units than do younger adults to accommodate the limitation that smaller working memory spans place on comprehension (Thornton & Light, 2006).

One technique that can examine the relative contributions of an age-related strategy shift in reading performance is to simulate the different reading patterns of the two age groups using a computational model. In order to simulate their results in the context of the SWIFT model, Laubrock et al. (2006) reported that the parameter-estimating algorithm that was used to find optimal model parameter values resulted in parameter values that suggest that older readers have a smaller and more asymmetrical perceptual span than younger readers. They noted that Rayner (1986) found that beginning readers had a smaller perceptual span than more skilled readers and suggested that older readers might have a smaller span because of deficits—decrements in visual acuity. In order to assess the theoretical implications of the findings we obtained, we used a strategy we have used previously (Pollatsek et al., 2006; Rayner et al., 2004; Rechile et al., 1999) of simulating the general pattern of data within the context of the E-Z Reader model. As becomes evident later in this article, the assumptions that we made to simulate the data differ from those made by Laubrock et al. (2006).

In the next section, we briefly describe the E-Z Reader model. One reason for using the model is that it gives a good quantitative account of eye movement data. Furthermore, because the model is relatively transparent, it is a good platform for diagnosing the results, as it is relatively easy to assess failures of the model and why they occur.

The E-Z Reader Model

Model Overview

The E-Z Reader model (Pollatsek et al., 2006; Rayner et al., 2004; Rechile et al., 1998, 1999, 2003) is an attempt to model how the language processing system, the oculomotor system, and the attention system are coordinated during reading. The two most important assumptions of the model are (a) that lexical processing of words occurs in a strictly serial manner and (b) that lexical processing of the fixed word is the primary engine for the decision to move the eyes to the next word. Thus, lexical processing of word \( n \) completes before lexical processing of word \( n + 1 \) begins. However, it should be noted that preliminary, low-level preattentive processing occurs in parallel across the visual field before lexical processing, which requires the allocation of attention to the features of a given word, begins. In contrast, the SWIFT model (Engbert et al., 2002, 2005) assumes (a) that attention is allocated as a gradient to several words at a time so that lexical processing of those words can occur in parallel and (b) that a signal from an autonomous “timer” that “fires” at random intervals and that has nothing to do with lexical processing is largely responsible for moving the eyes to new viewing locations. Because of the serial nature of lexical processing in E-Z Reader and because the decisions in the model that determine when the eyes move are based on lexical processing, we maintain that the model is more tractable than SWIFT and that it is therefore more useful for understanding how assumptions about cognitive processes during reading relate to the observed data (Rayner, Pollatsek, & Rechile, 2003). Indeed, in the simulations that are reported later in this article, the conceptual transparency of our model allowed us to adjust its parameter values by hand (i.e., without the use of parameter-estimating algorithms, as required in the simulation using SWIFT; Laubrock et al., 2006) so that the model’s eye movement behavior qualitatively resembled that of older readers. The simple fact that we were able to do this indicates how transparent the model is and how useful it is as a heuristic device for generating hypotheses.

Figure 3 is a schematic diagram of the model. As the figure shows, word identification in the E-Z Reader model is assumed to occur in three stages. The first is a preattentive stage of visual processing, which is labeled \( V \) in Figure 3. During this stage of processing, the features on the printed page are extracted in parallel from across the visual field so that this information can be used for subsequent lexical processing and so that low-spatial frequency information about word boundaries can be used in programming saccades. In the model, the time that is needed to complete \( V \) is set equal to 50 ms, which corresponds to several
Figure 3. Schematic diagram of the E-Z Reader model of eye movement control during reading.

The time that is needed to complete L₁ on a given word, \( t(L₁) \), is a function of the natural logarithm of its frequency of occurrence in printed text (as tabulated by corpus norms; Francis & Kučera, 1982) and its predictability within a given sentence context (as determined through separate cloze-task norms). The values of \( t(L₁) \) are specified by Equation 1, where \( \text{freq}_n \) is the frequency of word \( n \) and \( \text{pred}_n \) is its predictability:

\[
t(L₁) = \begin{cases} 
0 & \text{with probability } p = \text{pred}_n \\
\alpha_1 - \alpha_2 \ln(\text{freq}_n) - \alpha_3 \text{pred}_n & \text{with probability } 1 - p.
\end{cases}
\]

Equation 1 can be interpreted as follows. Upon shifting attention to word \( n \), the word has a certain probability, \( p \), of being guessed from the context of the sentence in which it is contained. For the sake of simplicity, we assume that the probability of guessing word \( n \) from its preceding sentence context is equal to its cloze predictability (i.e., \( p = \text{pred}_n \)). Thus, if a word is successfully guessed from its context (corresponding to the upper branch of Equation 1 during a given Monte Carlo simulation run of the model), then the time that is needed to complete L₁ is set equal to 0 ms. To foreshadow our simulation results just a bit, one might expect age-related differences in the propensity to engage in this type of “word guessing” behavior if one assumes that, by virtue of having more reading experience, older readers are more adept at making use of linguistic context to help them constrain the identities of parafoveal words.

In the majority of instances, however, word \( n \) is not guessed but must instead be identified using the information that is available on the printed page. If a word is not successfully guessed, then the time that is needed to complete L₁ is equal to the value determined by the lower branch of Equation 1. In the lower branch of Equation 1, \( \alpha_1 \) is a free parameter that determines the maximum amount of time (in milliseconds) that is needed to complete L₁ for a word, and \( \alpha_2 \) and \( \alpha_3 \) are free parameters that modulate how much a word’s frequency and predictability (respectively) attenuate this time. (In our previous simulations, the best fitting values of \( \alpha_1, \alpha_2, \) and \( \alpha_3 \) were 122, 4, and 10, respectively; these values minimized a composite root-mean-square deviation between four sets of observed and simulated dependent measures; see Reichle et al., 1998, Appendix.)

The mean time that is needed to complete L₂, \( t(L₂) \), is a fixed proportion of the mean time that is needed to complete L₁ as specified by the lower branch of Equation 1. The values specified by the lower—and not the upper—branch of Equation 1 are used because the lexical information that becomes available with the completion of L₂ must be generated so that it can be used for additional, higher level linguistic processing; in instances in which a word is “guessed” from context—that is, \( t(L₁) = 0 \) ms—the information corresponding to L₂ has to be generated through

\(^5\) When we use the term guessing the next word throughout this article, we do not mean to imply any type of conscious strategy on the part of the reader. Rather, the processing system is unconsciously engaging a strategy of skipping words on the basis of partial visual information about the skipped word.

Recent empirical estimates of the time that is needed for visual information to propagate from the retina to the visual cortex (Foxe & Simpson, 2002; Van Rullen & Thorpe, 2001). Lexical processing is completed in two stages, labeled L₁ and L₂ in Figure 3. The first stage, L₁, is a rapidly available index of the word’s familiarity and serves as the trigger for programming an eye movement. The completion of the L₁ stage culminates in a judgment that identification of the fixated word is imminent and that it is “safe” to program a saccade to the next word. That is, through years of experience, the systems involved in programming eye movements learn that a saccadic program initiated at this time is unlikely to produce a saccade that would move the eyes off of the word too soon (i.e., before it had been identified) or after waiting too long (i.e., long after the word had already been identified). (For a discussion of how the relationship between early lexical processing and saccadic programming might be learned, see Reichle & Laurent, 2006.)

The completion of lexical processing is represented in the model by the completion of L₂, which corresponds to the actual process of specifying the word’s orthographic, phonological, and/or semantic codes to the degree necessary for subsequent linguistic processing. We do not necessarily think of L₁ and L₂ as discrete stages in lexical processing, but the modeling is more convenient if they are formalized that way. As the word is fully identified with the completion of the L₂ stage, this produces an attention shift to the next word, so that lexical processing of that word can begin. One important implication of this last assumption is that saccadic programming is decoupled from attention shifts: The completion of L₁ is the signal to begin programming a saccade to the next word, whereas the completion of L₂ is the signal to shift attention to the next word. (These two stages of lexical processing and the bases for their distinction are discussed in detail elsewhere; Pollatsek et al., 2006; Rayner, Pollatsek, & Reichle, 2003.)
top-down processing, so that \( t(L_2) \) still requires some amount of time to be completed. The mean time to complete \( L_1 \) is given by Equation 2, where \( \Delta \) is a parameter that has been set to .5.

\[
t(L_1) = \Delta L_1.
\] (2)

Equations 1 and 2 specify the mean times needed to complete \( L_1 \) and \( L_2 \); the actual times for a given Monte Carlo run of the model are determined by sampling values from gamma distributions having means specified by Equations 1 and 2 and having standard deviations equal to .22 of the means. Finally, the actual time that is needed to complete \( L_1 \) (after a value has been sampled from a gamma distribution) is also modulated by visual acuity, as given by Equation 3:

\[
t(L_1)' = t(L_1) e^{n2 \text{letter-fixation}/in}.
\] (3)

In Equation 3, \( t(L_1)' \) is the time that is needed to complete \( t(L_1) \) as a function of the mean eccentricity (i.e., distance in character spaces) between the fixation location (i.e., the center of vision, or fovea) and each of the letters in the word that is being processed. \( N \) indicates the number of letters in the word being processed, and \( e \) (= 1.15) is a free parameter that modulates the degree to which eccentricity influences the rate of lexical processing. Using Equation 3, longer words and words that are far from the center of vision will (on average) take longer to identify than shorter words and words that are closer to the center of vision.

All of the remaining assumptions of the E-Z Reader model are related to the programming and execution of saccades. These assumptions are not central to the issues being addressed in this section, so we will not discuss them further (for a complete discussion of these issues, see Pollatsek et al., 2006; Reichle et al., 2003). However, it is important to note that skipping of words is explained within the model when the \( L_1 \) stage for word \( n + 1 \) completes before the saccade programmed to word \( n + 1 \) has reached a critical nonlabile stage (i.e., a point of no return). In such situations, the program to move the eyes to word \( n + 1 \) is canceled when the oculomotor system begins programming a saccade to move the eyes to word \( n + 2 \). Given the aforementioned assumptions, the model can explain a large amount of data, including the effects of word frequency, predictability, and length, as well as a number of reading-related phenomena (see Pollatsek et al., 2006).

**Simulations**

Our primary goal with the simulations was to account for the data we obtained wherein (a) older readers’ fixations were longer than those of the younger readers, (b) the older readers skipped words more frequently and made more regressions than did the younger readers, (c) older readers made about the same number of fixations as the younger readers. Logically, (b) and (c) imply that, on average, the mean saccade length for older readers was larger than for younger readers; however, as indicated earlier, this latter effect was not significant, largely because of the variability in the older age group. To keep the simulations as simple as possible, we changed the values of at most four parameters between the age groups. These changes are consistent with the general assumptions that, compared with younger readers, older readers (a) require more time to process or identify words because the rate of lexical processing is slowed (Faust, Balota, Spieler, & Ferraro, 1999; Spieler & Balota, 2000), (b) have limited visual acuity in the parafovea (Ball et al., 1988; Sekuler et al., 2000), and (c) are more likely to engage in a risky reading strategy of guessing words that are in the parafovea (O’Regan, 1990). The probability values for the two age groups are shown in Table 4. The values for the younger readers are the same values that we have used in all of our recent standard simulations (Pollatsek et al., 2006). The values for the older readers were selected through trial and error so as to produce the correct qualitative pattern of results.

We conceptualized each of the parameter-value changes that we implemented with the older readers as follows:

1. \( \alpha_1 \): The larger value of the intercept parameter for the word-identification times reflects a general slowing in older readers’ ability to process/identify words.

2. \( \alpha_2 \): The larger value of the word frequency slope parameter causes the effect of word frequency to be more pronounced with the older readers. This reflects the many additional years of reading practice that the older readers have had relative to the younger readers. (An alternative method of capturing this age-related difference would be to use different estimates of word frequency for the two groups, with the logic being that—on average—words will have been experienced more often by the older readers.)

3. \( p \): Technically, this is not a free parameter in the model (see Equation 1); that is, in the most recent versions of E-Z Reader (Pollatsek et al., 2006), it is simply assumed that readers are able to guess an upcoming word with a probability equal to that word’s mean cloze-task predictability. In the simulations of the older readers, we made the additional assumption that older readers are (on average) more likely to engage in this type of risky (i.e., error-prone) guessing behavior. Our results in addition indicate that this behavior is also influenced by word frequency, with older readers being more likely to guess high-frequency words more often than low-frequency words. To implement this behavior, we modified the model so that the probability of guessing word \( n \) (i.e., the probability that \( L_1 \) is set equal to 0 ms in the upper “branch” of Equation 1) is defined as follows:

\[
p = \max\{\text{pred}_w + \kappa \ln(\text{freq}_w)/\max[\ln(\text{freq}_w),1]\}.
\] (4)

Thus, the probability of guessing a given word \( (p) \) is affected by that word’s predictability \( (\text{pred}_w) \), a parameter that modulates a reader’s overall tendency to guess
words ($\kappa$), and the word’s frequency (scaled to the interval 0 to 1). The value of $\kappa$ was set equal to a larger (i.e., nonzero) value for older readers under the assumption that they are more likely to guess high-frequency words from parafoveal information than are younger readers. The maximum, $\max[x]$, of 1 or the quantity defined by the terms to the left of the comma is then selected so that $p$ is limited to the range $[0, 1]$.

Let make two general comments about the aforementioned assumptions. First, we did not make any attempt to find optimal parameter values because the simulations were done using the Schilling, Rayner, and Chumbley (1998) sentences and not the actual sentences used in the experiment. Second, our assumptions provide an interesting contrast to the assumptions that Laubrock et al. (2006) had to adopt by using their parameter-fitting algorithm to fit the SWIFT model to the data from their older readers (e.g., a more asymmetrical attention gradient with older readers).

The simulation results shown in Table 5 demonstrate the effects of manipulating three different variables: (a) the overall rate of lexical processing and the degree to which it is affected by word frequency (i.e., the $\alpha_1$ and $\alpha_2$ parameters), (b) the slowdown in lexical processing that results from limited visual acuity (i.e., the $\kappa$ parameter), and (c) the tendency to guess words in the parafovea (i.e., the parameter $\kappa$ in Equation 4, which determines the value of $p$ in the first upper branch of Equation 1). Table 5 shows the global measures of overall reading performance that E-Z Reader predicts for the younger readers (i.e., using the model’s standard best fitting parameter values), as well as seven different combinations of the three aforementioned variables. Thus, the simulations in Table 5 allow one to get a feeling for the effects of making each of the four parameter changes listed earlier. (Showing all possible simulations making one to four of the parameter changes would be hopelessly confusing.) Each of the simulations was based on the mean performance of the model across 1,000 Monte Carlo runs.

What can be learned from the simulations? First, the general slowing down in reading that is observed with older readers can be produced two ways: by increasing the values of the slope and intercept parameters for lexical processing (i.e., $\alpha_1$ and $\alpha_2$) or by

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Age group</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_1$</td>
<td>Young</td>
<td>122</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>Older</td>
<td>180</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Young</td>
<td>0</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Older</td>
<td>0.8</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Young</td>
<td>1.15</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Older</td>
<td>1.20</td>
</tr>
</tbody>
</table>

Note. $\alpha_1$ = intercept parameter for determining maximum mean L1 duration (see the bottom branch of Equation 1); $\alpha_2$ = slope parameter for determining how mean L1 duration is modulated by word frequency; $\kappa$ = parameter that modulates the overall tendency to guess words from their preceding sentence context (see Equation 4); $\epsilon$ = parameter that modulates the effect of retinal eccentricity.

### Table 5 Simulation Results

<table>
<thead>
<tr>
<th>Age group</th>
<th>Simulation</th>
<th>Fixation duration</th>
<th>No. of fixations</th>
<th>Saccade length</th>
<th>Probability of regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young</td>
<td>1</td>
<td>235</td>
<td>8.25</td>
<td>7.18</td>
<td>.023</td>
</tr>
<tr>
<td>Older</td>
<td>2</td>
<td>278</td>
<td>8.54</td>
<td>6.81</td>
<td>.076</td>
</tr>
<tr>
<td>Older</td>
<td>3</td>
<td>248</td>
<td>7.48</td>
<td>7.80</td>
<td>.051</td>
</tr>
<tr>
<td>Older</td>
<td>4</td>
<td>256</td>
<td>8.22</td>
<td>7.13</td>
<td>.037</td>
</tr>
<tr>
<td>Older</td>
<td>5</td>
<td>267</td>
<td>6.61</td>
<td>8.66</td>
<td>.018</td>
</tr>
<tr>
<td>Older</td>
<td>6</td>
<td>305</td>
<td>8.37</td>
<td>6.77</td>
<td>.102</td>
</tr>
<tr>
<td>Older</td>
<td>7</td>
<td>275</td>
<td>7.37</td>
<td>7.80</td>
<td>.069</td>
</tr>
<tr>
<td>Older</td>
<td>8</td>
<td>238</td>
<td>6.60</td>
<td>8.69</td>
<td>.009</td>
</tr>
<tr>
<td>Older</td>
<td>7'</td>
<td>269</td>
<td>8.39</td>
<td>7.64</td>
<td>.107</td>
</tr>
</tbody>
</table>

Note. Fixation duration is given in milliseconds; saccade length is given in character spaces. Simulation 1: younger readers ($\alpha_1 = 122$ ms, $\alpha_2 = 4$ ms, $\kappa = 0$, $\epsilon = 1.15$); Simulation 2: older readers with slowed lexical processing ($\alpha_1 = 180$ ms, $\alpha_2 = 7$ ms, $\kappa = 0$, $\epsilon = 1.15$); Simulation 3: older readers with slowed lexical processing and increased “guessing” of parafoveal words ($\alpha_1 = 180$ ms, $\alpha_2 = 7$ ms, $\kappa = .8$, $\epsilon = 1.15$); Simulation 4: older readers with slowed parafoveal processing ($\alpha_1 = 122$ ms, $\alpha_2 = 4$ ms, $\kappa = 0$, $\epsilon = 1.2$); Simulation 5: older readers with slowed parafoveal processing and increased guessing of parafoveal words ($\alpha_1 = 122$ ms, $\alpha_2 = 4$ ms, $\kappa = .8$, $\epsilon = 1.20$); Simulation 6: older readers with slowed lexical and parafoveal processing ($\alpha_1 = 180$ ms, $\alpha_2 = 7$ ms, $\kappa = 0$, $\epsilon = 1.20$); Simulation 7: older readers with slowed lexical and parafoveal processing and increased guessing of parafoveal words ($\alpha_1 = 180$ ms, $\alpha_2 = 7$ ms, $\kappa = .8$, $\epsilon = 1.20$); Simulation 8: older readers with increased guessing of parafoveal words ($\alpha_1 = 122$ ms, $\alpha_2 = 4$ ms, $\kappa = .8$, $\epsilon = 1.15$); Simulation 7': older readers with slowed lexical and parafoveal processing, increased guessing of parafoveal words, and some proportion ($p = .1$) of saccades being directed back toward words that are incorrectly guessed ($\alpha_1 = 180$ ms, $\alpha_2 = 7$ ms, $\kappa = .8$, $\epsilon = 1.2$). $\alpha_1$ = intercept parameter for determining maximum mean L1 duration (see the bottom branch of Equation 1); $\alpha_2$ = slope parameter for determining how mean L1 duration is modulated by word frequency; $\kappa$ = parameter that modulates the overall tendency to guess words from their preceding sentence context; $\epsilon$ = parameter that modulates the effect of retinal eccentricity.
increasing the value of the eccentricity parameter (i.e., \( \varepsilon \)). Relative to what happens with the standard, best fitting parameter values for younger readers, either of these changes will cause the overall reading rate to decrease (and fixation times to increase). The effects of slowed lexical processing can be seen most directly by comparing Simulation 2 (in which \( \alpha_1 = 180 \) and \( \alpha_2 = 7 \)) to Simulation 1 (in which \( \alpha_1 = 122 \) and \( \alpha_2 = 4 \)); similarly, the effects of decreasing visual acuity can be seen most directly by comparing Simulation 4 (\( \varepsilon = 1.2 \)) to Simulation 1 (\( \varepsilon = 1.15 \)). Second, the increased skipping rates and longer saccade lengths that are observed with older readers are dependent on the model’s assumption that older readers are more likely to guess words in the parafovea than are the younger readers. This can be seen by comparing Simulation 8, in which \( \alpha_1 = \alpha_2 = 0 \), to Simulation 1, in which \( \alpha_1 = 0 \).

Overall, the assumptions that best simulated the older readers’ performance were those used in Simulation 7. That is, Simulation 7 yielded longer fixation durations, a higher probability of regression, and longer saccade lengths for the older readers than the younger readers, which was the same pattern that was produced by the older readers, relative to the younger readers, when reading the Times New Roman font. However, careful examination of Table 5 in comparison to the actual data in Table 1 reveals that the simulated older readers are making too few fixations relative to what was observed with the older readers (see Table 1) and relative to the simulated younger readers. One way to remedy this shortcoming is to assume that some proportion of the words that are guessed and consequently skipped by the older readers are guessed incorrectly; this assumption would presumably result in an increased number of regressions back to the skipped words and thereby increase both the number of fixations and the number of regressions that are predicted for the older readers. Both of these changes should bring the predictions for the older readers more in line with what was actually observed.

To test the feasibility of our explanation for why older readers might make more fixations and regressions than younger readers, we ran one final simulation (labeled Simulation 7’ in Table 5) that included the same assumptions as Simulation 7, but which also included the additional assumption that guessed words might occasionally be misidentified, thereby making it necessary for the reader to process those words a second time. This was done by adding the assumption that each guessed word had a fixed probability of being incorrectly identified (\( p = 0.1 \)), which then made it necessary to reprocess those words (i.e., both \( L_1 \) and \( L_2 \) had to be repeated on the word, with the \( L_1 \) time being determined by the lower branch of Equation 1). The proportion of incorrectly guessed words—like the values of the other parameters that were considered in Simulations 1 through 7—was adjusted via trial and error to give reasonable quantitative fits. The intended target of any saccade that was initiated prior to the completion of \( L_1 \) (which would move the eyes to the right) was also directed back toward the misidentified words. The intuition behind these two assumptions is that readers, upon realizing that they have misidentified a word, will have to process the word a second time, and that they will sometimes do this by making a regression back to the word. In this final simulation, we therefore did not exclude simulation trials that resulted in regressions. The fourth and sixth columns of Table 5 (respectively) show that the additional assumptions of Simulation 7’ were sufficient to increase both the overall number of fixations and the overall probability of making regressions. We therefore contend that the assumptions that were necessary for the E-Z Reader model to account for the eye movement behavior of older readers are both plausible and informative; the former assertion is supported by the fact that the model was able to reproduce the pattern of results that was observed with older readers using only very simple assumptions, and the latter assertion is supported by the fact that the model provides an interesting alternative hypothesis to the one Laubrock et al. (2006) needed in order to get SWIFT to simulate older readers—that older readers have a less symmetrical perceptual span than younger readers. Future research will be necessary to explicitly test the validity of these two hypotheses.

Finally, we have not systematically explored the effects of frequency and predictability for older readers within the model using the sentences used in the present study. However, we did run simulations of the older and young readers to examine the frequency effect using the Schilling et al. (1998) corpus. The results of these simulations are shown in Table 6. Not surprisingly (given prior successes with the model), the simulation of the young readers nicely captures the size of the frequency effect in the three fixation time measures and the skipping probability. The seven simulations of the older data vary in terms of how well they capture the size of the frequency effect, but Simulation 7 once again does the overall best job. Finally, for the sake of completeness, Table 6 also shows the size of the frequency effects that were predicted by Simulation 7’, which included the additional assump-

---

8 We do not think it would be difficult to further simulate the Old English font effects, but we also think the most important pattern of results to account for is the more normal Times New Roman.

9 Because interword regressions have—until now—always been beyond the theoretical scope of the E-Z Reader model, all of our previously published simulations have excluded simulation trials in which the model made interword regressions. (These predicted interword regressions were always due to saccadic error: cases in which the model “intends” to re fixation the beginning of a word but then overshoots the intended saccade target.)

10 To completely simulate the frequency and predictability effects for specific items is a labor-intensive procedure. The first step in doing this is to calculate the frequency, length, and predictability of each word in the sentence corpus. Although frequency and length are relatively easy to obtain, the predictability values are more difficult because they must be obtained through a separate cloze-task experiment. The next step is to then calculate the dependent measures of interest (e.g., gaze durations) for each word in the corpus so that these measures can be used to define a goodness-of-fit metric that can be used in selecting parameter values. The final step is to then use some type of algorithm to search through the parameter space and evaluate various combinations of parameter values to select those values that, according to the goodness-of-fit metric, provide the best fit between the model and the data. Because it would have been prohibitively expensive to complete all of these steps, the simulation results that are reported in Table 6 were completed using an alternative procedure that allowed us to examine the model’s performance on specific words using a sentence corpus (Schilling et al., 1998) and best-fitting model parameters from previously published simulations (Pollatsek et al., 2006).

To do this, we first identified two groups of words for which the model predicted frequency effects that were of the correct size for the younger readers—words having frequencies of 1 to 10 per million versus words having frequencies of 11 to 100 per million. (These two groups of words correspond to Frequency Classes 1 and 2 in our previous simulations; Pollatsek et al., 2006). The two classes of words were then used as the low and high-frequency target words (respectively) to examine word frequency effects with older readers.
tion that 10% of the guessed words are misidentified and must therefore be reprocessed, resulting in some interword regressions back to the misidentified words. A comparison of Simulations 7 and 7' indicated that the additional assumption of the latter did not significantly change the predicted pattern of results, with the size of the predicted frequency effects being both larger than those observed with younger readers, and of comparable size to those that were observed with the older readers.\footnote{The one possible caveat here is that Simulation 7' underpredicts the size of the frequency effect on the older readers' gaze durations. Although one could increase the size of this predicted effect by changing the model's parameters (i.e., the values of $\alpha_1$, $\alpha_2$, and/or $\nu$), we did not think this minor discrepancy warranted such efforts.}

### General Discussion

Using the E-Z Reader as a platform to understand the general pattern of results that we obtained, it appears that there are two principal ways that older readers differ from younger readers. First, they fixate longer on individual words in a sentence, and this translates into overall longer sentence reading times (i.e., they read slower than young readers). Second, perhaps to compensate for their slower reading, they adopt a more risky reading strategy than younger readers. Specifically, their processing system appears to guess word $n + 1$ and, consequently, skip over it much more so than young readers. They do, however, tend to go back and fixate skipped words more frequently than do younger readers,\footnote{This statement is certainly the case for the frequency data, less so for the predictability data.} suggesting that their guesses are not always correct. It is also instructive that these two differences, the first (fixation time) being a quantitative difference and the second (risky strategy) being a qualitative difference, do not result in differences in comprehension between the two groups: The older readers answered the comprehension questions as effectively as the young readers.

With respect to the quantitative differences in fixation time, although the older readers' fixation times were longer overall, like the younger readers they still showed very strong frequency, predictability, and font effects. Indeed, there was an interaction of age with both frequency and font: Older readers' frequency and font effects were larger than those of the younger readers. Similar interactions of age and frequency, with larger effects for older participants, have been reported in naming and lexical decision tasks (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Spieler & Balota, 2000). This interaction is easily predicted by natural changes in parameters in lexical processing time, such as in our simulations with E-Z Reader.

It is not clear why we failed to find age-related differences in the impact of predictability on reading times or skipping rates. Previous findings using other techniques indicate that predictability influenced word identification in degraded speech (Pichora-Fuller et al., 1995; Wingfield et al., 1991) and degraded reading (Speranza et al., 2000). However, we failed to find any indication that predictability had a larger impact for older adults' reading. Our failure to find age-related differences in predictability are more in line with those of Kliegl et al. (2004), who found few age-related differences in eye movement measures of reading resulting from changes in predictability. Possibly, in normal reading, both age groups are able to maximally extract information from the presentation of text, and thus any advantage that predictability provides in degraded situations is not present.

The fact that frequency and predictability influenced fixation times for both the older and the younger readers is quite consistent with a great deal of prior research (see Rayner, 1998). One interesting difference between the present results and those reported previously by Rayner and Well (1996) emerged in the effect of predictability. That is, Rayner and Well reported a dissociation between their fixation time measures and their skipping measure: Whereas the high- and medium-predictability words did not differ from each other with respect to the fixation time measures (though they both yielded shorter fixation times than low-predictable words), the high-predictable words differed from the medium- and low-predictable words on skipping rates. In the present study, no such dissociation was found: The pattern for the fixation time measures and the skipping measure mimicked each other. We are not sure why this difference occurred between the present study and the earlier Rayner and Well study as the same materials were used and the same younger participant pool was used. A different eye-tracking system was used, but given that the overall data set is not that different across the two eye trackers, we doubt that this can explain the difference. One thing that is certain, however, and this is true across both studies, is that high-predictable target words lead to shorter fixation times and higher skipping rates than low-predictable target words. Medium predictability target words appear to yield more variable data patterns.

\footnote{Note. FFD, SFD, and GD times are given in milliseconds.}

### Table 6

<table>
<thead>
<tr>
<th>Variable</th>
<th>Young</th>
<th>Older</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ob</td>
<td>S1</td>
</tr>
<tr>
<td>FFD</td>
<td>16</td>
<td>21</td>
</tr>
<tr>
<td>SFD</td>
<td>25</td>
<td>28</td>
</tr>
<tr>
<td>GD</td>
<td>48</td>
<td>49</td>
</tr>
<tr>
<td>Pr Skip</td>
<td>.01</td>
<td>.08</td>
</tr>
</tbody>
</table>

The Size of the Frequency Effect for the Observed (Ob) and Simulated (S) Data for Young and Older Readers for First-Fixation Duration (FFD), Single-Fixation Duration (SFD), Gaze Duration (GD), and Skipping Probability (Pr Skip)
Although we were fairly successful in simulating the quantitative and qualitative differences between the older and younger readers, we believe that further experiments are needed to further explicate differences between the two age groups. In particular, it would seem particularly relevant to determine whether the size of the perceptual span of older readers differs from that of younger readers. For example, Laubrock, Kliewer, and Englert (2005) recently argued that older readers have a slightly smaller and more asymmetrical perceptual span than younger readers. Another interesting test would be to determine if older readers’ initial encoding processes differ from those of younger readers. That is, a number of studies (Ishida & Ikeda, 1989; Liversedge et al., 2004; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981; Rayner, Liversedge, & White, 2006; Rayner, Liversedge, White, & Vergilino-Perez, 2003) have demonstrated that if younger readers have 50 to 60 ms to process a fixated word before it is masked or disappears, reading proceeds quite smoothly. Would older readers show a similar result, or would they need the words to be available longer before they disappear during a fixation?

Finally, the suggestion that older adults use a different strategy during reading is similar to results from another skilled activity—typing. Older adults tend to have slower finger tapping and choice reaction times, but skilled older typists do not differ in their overall typing speed compared with younger adults (Bosman, 1993; Saltz, 1984). Saltz (1984) found that older adults’ typing speed (specifically interkeystroke intervals) were more affected by the number of characters of the upcoming text that were visible, and he interpreted these findings as indicating that older adults compensated for their slowed response times by planning—looking farther ahead in the text than younger adults. This is similar (although not perfectly analogous) to the strategic differences found in the current study. Older adults, perhaps in an attempt to speed up reading, attempt to look farther ahead in the text. This strategy may result in greater reliance on partial parafoveal information, which allows older adults to process words in the parafovea more effectively (resulting in their skipping more words because these words have already been identified). Although this strategy may work most of the time, there is an increased need to regress to some words in order to obtain missed information. This strategy would speed up reading for older adults; however, as our data and previous data indicate, older adult’s reading speed does not reach the level of younger adults (in contrast to typing, where there is no age difference in speed). This conclusion is supported by the results of our simulations; a direct comparison of the overall reading rates predicted in Simulation 7’ (in which the older readers used the strategy of guessing parafoveal words) with the reading rates predicted in Simulation 6 (in which the older readers did not use this strategy) indicated that the strategy sped up reading by 20.66%, even though the strategy also made it necessary to go back and reprocess 10% of the words (i.e., words that had been guessed incorrectly). Although this analysis ignores any additional costs that might result from misidentifying 10% of the words in each sentence (e.g., such errors might result in additional problems with syntactic processing), it does suggest why older readers may compensate for slower lexical and parafoveal processing by adopting such a risky reading strategy—the strategy may allow the older readers to maintain a faster reading rate.

Our findings that older and younger readers demonstrated similar patterns of fixation durations on target words indicate that, although the older adults tended to be slower overall, the reading process was similar in both groups. Thus, although there may be a generalized slowing of older adults’ reading as evidenced by longer reading and fixation times, it does not lead to a dramatic decrement in reading performance. With respect to the theories of cognitive aging, the overall age difference in reading speed could be the result of generalized physical decline as in common cause theory (Baltes & Lindenberger, 1997) or a decline in perceptual speed (Saltz, 1996). However, we also found that older adults and younger adults showed the same patterns of fixation times for frequency and predictability, indicating that any general decline in cognitive functioning does not appear to interact with reading performance at the word level (see also Kliewer et al., 2004). In other words, reading performance, although slower overall for older adults, is relatively unimpaired with respect to younger adults. Further research will be required to determine whether any other characteristics of words show age-related differences in reading. In general, our research fits well with other research indicating that language processing is only minimally affected by aging (e.g., Davidson, Zacks, & Ferreira, 2003; Waters & Caplan, 2005).

References


Received July 27, 2005
Revision received April 19, 2006
Accepted April 21, 2006

465