Psychological Review
1998, Vol. 105, No. 1, 125–157

# Toward a Model of Eye Movement Control in Reading

Erik D. Reichle, Alexander Pollatsek, Donald L. Fisher, and Keith Rayner
University of Massachusetts at Amherst

The authors present several versions of a general model, titled the *E-Z Reader model,* of eye movement control in reading. The major goal of the modeling is to relate cognitive processing (specifically aspects of lexical access) to eye movements in reading. The earliest and simplest versions of the model (*E-Z Readers 1 and 2*) merely attempt to explain the total time spent on a word before moving forward (the *gaze duration*) and the probability of fixating a word; later versions (*E-Z Readers 3–5*) also attempt to explain the durations of individual fixations on individual words and the number of fixations on individual words. The final version (*E-Z Reader 5*) appears to be psychologically plausible and gives a good account of many phenomena in reading. It is also a good tool for analyzing eye movement data in reading. Limitations of the model and directions for future research are also discussed.

In his seminal book, *The Psychology and Pedagogy of Reading,* Huey (1908) wrote that

> to completely analyze what we do when we read would almost be the acme of a psychologist's achievements, for it would be to describe very many of the most intricate workings of the human mind, as well as to unravel the tangled story of the most remarkable specific performance that civilization has learned in all its history. (p. 6)

The goal of this article, therefore, is not to explain reading. Instead, our aims are more modest; we hope to give a reasonable account of how cognitive and lexical processing influences the eye movements of skilled readers. We believe, however, that this is an important and necessary enterprise for several reasons.

First, aspects of eye behavior, such as the durations of eye fixations on words or on regions of text, are commonly used to infer cognitive processes in reading (Just & Carpenter, 1987; Rayner & Pollatsek, 1989). Accordingly, it is important to understand the links between the observed data and the underlying cognitive processes as well as possible. Second, reading is arguably the most important and ubiquitous skill that people acquire for which they were not biologically programmed. If this skill could be understood, it might help to shed light on skill acquisition in general. Third, the general issue of eye movement control in visual perception is an important one. In many ecologically relevant visual tasks (e.g., looking at a visual scene), it is difficult to understand the relation of perceptual and cognitive variables to the control of eye movements because it is so unclear what the perceiver's goals and intentions are. In contrast, although people read prose silently for many purposes (pleasure, school, etc.), the task is reasonably well defined when they are reading for meaning (and not skimming), and, hence, an experimenter has a reasonable chance of being able to relate perceptual and cognitive processes to eye movements in an ecologically valid, complex continuous task.

The model that we present here assumes that eye movements in reading are largely driven by lexical access (broadly defined). We defend and qualify that position in detail later in the article, but for now we wish to make two main points. First, the assertion that there is any close link with cognitive processes and eye movements was at some point quite controversial (and still may be) because there is quite a narrow time window in which to program a saccadic eye movement in reading; thus it seemed implausible to many people 20–30 years ago that processes such as lexical access could affect eye movements (Bouma & deVoogd, 1974; Kolers, 1976). However, as we argue later, a great deal of evidence has accumulated, indicating that there is a close link between eye movements and cognitive processes. Second, we do not deny that global comprehension of the text influences how a reader progresses through the text; however, we take as a working hypothesis that it does not directly intervene on a large majority of fixations.

## Some Background:
## The Role of Eye Movements in Reading

### *Basic Components: Fixations and Saccadic Eye Movements*

Before plunging into the details of our theorizing, we need to present some background data on eye movements in reading (see Rayner, 1978; Rayner & Pollatsek, 1987, 1989, for more

---

Erik D. Reichle, Alexander Pollatsek, Donald L. Fisher, and Keith Rayner, Department of Psychology, University of Massachusetts at Amherst.

Portions of this work were presented at the Psychonomic Society Meeting in November 1995 and at the Mathematical Psychology Meeting in August 1996. This work was supported by National Institutes of Health Grant HD26765, National Institute of Aging Grant AG12461, and National Institute of Mental Health Grants MH16745 and MH01255. We thank George Avrunin for his help with the initial simulations and Ronan Reilly for helpful comments on a version of this article.

Correspondence concerning this article should be addressed to either Alexander Pollatsek or Keith Rayner, Department of Psychology, University of Massachusetts, Amherst, Massachusetts 01003. Electronic mail may be sent via Internet to either pollatsek@psych.umass.edu or rayner@psych.umass.edu.

detail). The first point that needs to be made is that (contrary to introspection) the eyes do not sweep continuously across the printed page but instead make sudden jumps from one location on the page to another location. That is, in any situation such as reading in which there is a static visual display, there are periods of 100 to 500 ms or so (called *fixations*) in which the eyes come to rest, interspersed with rapid, ballistic eye movements called *saccades*. In reading, most fixation durations are between about 200–350 ms, and most saccades go forward (i.e., in a left-to-right direction when reading English) about five to nine character positions. Saccades last about 15–40 ms (depending on the size of the eye movement) and merely serve the function of transporting the eye from fixation point to fixation point. Virtually no information is extracted from the printed page during the saccade (Wolverton & Zola, 1983), and furthermore, people are usually unaware either of making the movements or that there are periods in which visual information is not available.[1]

Not all saccades go from left to right in the text, of course. First of all, there are *return sweeps* that transport the reader from the end of one line to the beginning of the next. These are relatively complex and often involve a long saccade followed by a shorter corrective saccade. In addition, readers often (about 10% of the time) move backward in the text. These backward saccades are termed *regressions*. Many are small and go only a few character positions (sometimes moving to an earlier position on the same word), but a few are large and go back several words or even to earlier sentences. Surprisingly, readers are often unaware of making regressions (especially the small ones).

In summary, reading (or viewing any static display) is like a "slide show" in which a "slide" is on for about a quarter of a second (representing the world as seen on one fixation), followed by a brief interval in which the slide is off, and then followed by a new slide representing the world as seen from a new fixation point. As a result, the analysis of eye movements rests on two principal parameters: the durations of fixations and the locations of these fixations. (As the duration of a saccade is merely a function of its length, it is not a variable of much interest in modeling cognition.)

### *The Use of Eye Movements in Studying Reading*

Obviously, some sort of data reduction is needed in order to make sense out of the pattern of eye movements and to determine what it says about the process of reading. That is, a table of the sequential pattern of fixation durations and fixation locations (in $x$ and $y$ coordinates) for each reader and each passage of text would be virtually incomprehensible because of the staggering amount of detail. Accordingly, the eye movement record needs to be summarized in some way.

*Global averages.* One way to summarize the eye movement record is to present averages of various eye movement indices over a large segment of text such as a passage, a paragraph, or a set of sentences. These measures, such as the mean fixation duration, the mean length of forward saccades, and the probability that a saccade is a regression, have been shown to reflect, globally, the difficulty of the reading process. For example, if people are asked to read a more technical and less technical passage, readers will have longer mean fixation durations, shorter average forward saccade lengths, and more regressions on the more technical passage (Rayner & Pollatsek, 1989).

*Word-based measures.* Although such global measures have some value, more local measures are needed if one wants to use eye movements as an on-line measure to understand cognitive processes on a moment-to-moment basis (Blanchard, 1985). Accordingly, other measures have been developed that use a smaller region of text, such as a word or phrase, as the unit of analysis. Here, we focus on word-based measures, although similar measures have been developed for somewhat larger units such as phrases or clauses (see Rayner & Sereno, 1994; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989).

Put simply, these measures index whether a reader "looked at" (i.e., fixated) a word, and if so, how long he or she looked at it. A measure that we focus on, *gaze duration* (Just & Carpenter, 1980), is defined as the sum of the total fixation time on a word when it is encountered for the first time. To simplify discussion, we designate as *word n,* the word that we are interested in. More precisely, the gaze duration on word $n$ counts only instances when a word was fixated as a result of a forward saccade (i.e., from a word before it in the text—usually word $n - 1$ or $n - 2$). However, all fixations on word $n$ are counted until a subsequent word is fixated. For example, in the eye movement record displayed in Figure 1, the gaze duration on the word *creativity* in the fourth line would be 389 ms (the sum of 201 and 188 ms for the two fixations). In contrast, the gaze duration on the word *healthy* in the second line would only be 177 ms because Fixation No. 11 (the one following the first fixation on *healthy*) is three words to the right of *healthy*. The mean gaze duration that we use counts only those instances in which the word was fixated in the "normal way" in reading the text. That is, it does not count those instances in which the word is never fixated or those instances in which it is skipped over but then later fixated by a regression. In other words, gaze duration is the mean time spent fixating on a word on the reader's "first pass" through the region of text, conditional on the word being fixated.

A second, related, measure is the *total time* spent on a word, which would include not only the first-pass fixation time included in the gaze duration but also any additional time spent on the word when regressing back to it. Thus, the total time spent on *healthy* would be 373 ms (177 plus 196 ms). Two other measures focus more narrowly on the initial fixation duration on a word. One is the mean *first-fixation duration* (Inhoff, 1984), which is the first fixation on a word on the first pass (regardless of the number of fixations), again conditional on the word being

---

[1] Our picture of fixations and saccades is a bit of an oversimplification but is functionally correct for what we are discussing. First, it is not true that absolutely no information can be processed during saccades; however, in visual displays such as text, there are no data indicating that any useful information in reading is obtained during saccades. Second, the eyes do not stay absolutely still during fixations. There are (almost constantly) very rapid small movements called *nystagmus* and occasional slow drifts of the eyes (extending a character position or two) sometimes followed by brief microsaccades to bring the eyes back to the original position. However, it is reasonable to assume that visual information is almost continually extracted during a fixation and from a more or less constant viewing position.

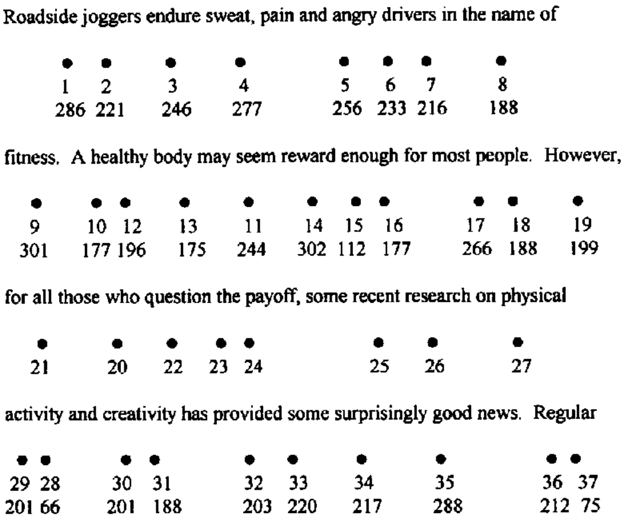


*Figure 1.* An excerpt from a passage of text with fixation sequence and fixation durations included. The dots below the words indicate the fixation locations, the first number below a dot indicates its ordinal number in the sequence, and the second number below a dot is the duration of the fixation (in milliseconds).

fixated. The second is mean *single-fixation duration* (Rayner, Sereno, & Raney, 1996), which is the mean duration of fixations on words that are fixated exactly once on the first pass through the text. Obviously, measures of mean second- and third-fixation durations can also be obtained; because words are usually fixated once, these measures are not as commonly used as first-fixation duration.

These measures indicate how long a reader fixates a word given that he or she fixated it. To make the record complete, one would also need to know (a) the probability that it was fixated on the first pass through the text, (b) the probability that it was skipped entirely, and (c) the probability that it was initially skipped and later regressed to. Although these measures give a pretty complete word-by-word description of the pattern of eye movements, they do not capture everything. They lose some sequential information, such as where regressions start and end, and some of the "fine grain" of the data, such as which letter of a word was fixated and the number of times a word was fixated once, twice, and so forth. However, if one could accurately account for how long each word was fixated and how many times it was skipped, one would have a pretty good account of the eye movement record. Moreover, if one could predict the individual fixation durations on words, one would have even a better understanding of the reading process.

The primary goal of this article is to understand the reading process at this level. Most basically, we wish to understand (a) how long readers look at each word in turn (gaze durations) before moving on and (b) the processes by which they skip words. We first attempt to model reading at that level. We then expand our model one step further and attempt to explain the causes of multiple fixations on a word and the durations of these individual fixations. We ignore the question of which letter on a word is being fixated, although we indicate how our modeling efforts could be expanded to account for that kind of finer grained analysis. First, however, we need to present some basic data on reading and eye movements in order to motivate our theorizing and to indicate how we evaluate our theory.

### *Relations Between Cognition and Word-Based Measures*

One indication that reading is largely a word-by-word process is that most words in text are in fact fixated. In normal text, about 80% of the content words (nouns, verbs, and adjectives) are fixated, and about 20% of the function words (articles, conjunctions, prepositions, and pronouns) are fixated (Carpenter & Just, 1983; Rayner & Duffy, 1988). More important, fixation times on words (gaze durations, first-fixation durations, and single-fixation durations) are sensitive to what is being fixated. Notably, the mean gaze duration, mean first-fixation duration, and mean single-fixation time on a word are all sensitive to the frequency of the word in the written language (henceforth *word frequency*). In normal text, the frequency and the length of the word (in number of letters) are highly correlated—high-frequency words tend to be short; however, a frequency effect is obtained even when word length is controlled (Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Rayner et al., 1996; Vitu, 1991). This means that lexical and/or semantic characteristics of a word—or something closely related to them—appear to be able to control the duration of the fixation on that word, and, thus, the relation between cognitive processes and eye movement control is fairly tight. Of course, this relation—higher frequency words are fixated on average for shorter periods of time—only forces the conclusion that lexical access of the word (or some related cognitive process) influences the duration of the fixation on at least some of the fixations. Conversely, when word frequency is controlled, word length influences gaze durations (Rayner & Fischer, 1996; Rayner et al., 1996).

Another variable that has been shown to have an effect on fixation time is the predictability of a word in the text. Predictability is usually assessed by showing a group of participants (other than the readers whose eye movements are being monitored) the text before the word of interest and having them guess what the next word would be. The predictability of a word is usually defined as the percentage of participants who guess that word given the prior words in the sentence. The typical experiment assessing the effects of predictability examines the fixation durations on two words that are placed in the same sentence frame that are matched on both *length* (number of letters) and frequency but vary in predictability. A robust finding is that mean gaze duration and mean first-fixation duration are affected by the predictability of the word (Balota, Pollatsek, & Rayner, 1985; Ehrlich & Rayner, 1981; Rayner & Well, 1996; Zola, 1984).

A second salient fact about eye movements in reading is that words tend to be skipped if they are frequent, short, or predictable. As we indicated earlier, function words that are frequent, short, and tend to be predictable are often skipped (O'Regan, 1979). The relative contribution of these three factors has not fully been analyzed in reading; however, predictability has been shown to have a strong influence, even when the other two factors have been controlled (Balota et al., 1985; Ehrlich & Rayner, 1981; Rayner & Well, 1996). This also speaks to the immediacy of eye movement control: A word can be processed before it is fixated so as to influence whether it is skipped or not.

In summary, a major fact about eye movements in reading is that they are influenced by cognitive variables on a moment-to-moment basis. For example, the fact that word frequency affects mean first-fixation duration indicates that (at least some of the time) the reader must have completed lexical access (or some closely related process) and that this process has played a part in the decision of when to move the eyes. Similarly, the fact that words are skipped on the basis of their predictability indicates that processing the word to the right of the fixated word (up to the point of knowing that it is the expected word) is influencing the decision of where to send the eye. We should emphasize that we are not claiming that these variables are the only text variables that influence the eyes on a moment-to-moment basis (discussed later); however, these are the variables upon which our model focuses.

### *Words Are Processed on More Than One Fixation*

Before outlining our model, we need to document one more key fact about reading that is crucial to our modeling: The processing of many words begins before they are fixated and is completed when they are fixated. That is, many words are processed on more than one fixation. The data on skipping make this assertion plausible as they indicate that many words are processed before they are fixated; however, it could be that the only words that are processed before they are fixated are the ones that are skipped.

There is a large body of data, however, indicating that *previewing* a word when it is in the parafovea reduces processing time on the word when it is later fixated (see Rayner, 1995; Rayner & Pollatsek, 1989). Space does not permit a full review of this literature, but the following gives the essential ideas. The basic technique used to establish this assertion is the *eye-contingent display-change technique* (McConkie & Rayner, 1975). In the version of this technique most relevant to the present argument, called the *boundary technique* (Rayner, 1975), a single word is changed contingent on where the eye is fixated. For example, consider the sentence in Figure 2. Before the reader fixates the word **bone,** a *preview word* is presented that may or may not be the word **bone.** However, when the reader crosses an invisible *boundary* (indicated by the | in the figure), the preview word changes into the *target word.* This display change is accomplished during the saccade to the target word and is not noticed by the reader.

One finding (Balota et al., 1985; Rayner, 1975) is that the gaze duration on the target word is shorter when the preview is identical to the target than when it is very different (e.g., the gaze duration on **bone** when previewed by **bone** is shorter than when **bone** is previewed by **food** or **name**). More interestingly, the gaze duration on **bone** is also facilitated (relative to the same control) by a preview of **borc** (although the benefit is somewhat less than when **bone** is the preview). This indicates that it must have been something other than the complete identification of the preview that shortened the gaze duration on **bone** when it was fixated. We should emphasize that readers are unaware of the display change and are unaware of the semantics of the preview word. This is shown by the fact that the effects of the completely anomalous word **name** and the nonword **foak** in the parafovea are about the same as when the reasonable word **food** is in the parafovea.

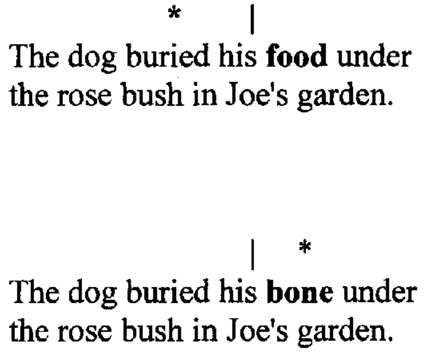


*Figure 2.* A representation of the display change during a boundary experiment. The top display represents the text display before the boundary is crossed by the eyes, and the bottom display represents the text display after the boundary is crossed. The invisible boundary is represented by a vertical line in the figure, the location of the eye is represented by an asterisk, and the preview and target words are indicated by boldface. Other possible previews besides **food** are **bone, name, borc,** and **foak.**

As indicated earlier, space does not permit a full summary of what is known about the nature of this *preview benefit* (i.e., shorter processing time on a word when there is a preview of it than when there is a control letter string in the parafovea). However, a few basic facts are worth citing. First, a virtually identical pattern of results is obtained when the paradigm is changed to one in which the preview and target words are seen in isolation and the participant's task is to name the target word (see Pollatsek, Lesch, Morris, & Rayner, 1992). Second, the effect is not at the level of visual features because the same preview benefit obtains regardless of whether the case of the letters is changed from fixation to fixation (McConkie & Zola, 1979; Rayner, McConkie, & Zola, 1980). Third, it does not appear that there is any benefit from a semantically similar preview (Rayner, Balota, & Pollatsek, 1986). Fourth, we are not sure of all the factors involved in preview benefit, but it is clear that both orthographic and phonological similarity of preview and target are relevant, indicating that whatever access process is reflected in fixation duration measures involves computation of both orthographic and phonological codes (Henderson, Dixon, Petersen, Twilley, & Ferreira, 1995; Pollatsek et al., 1992).

## Preliminary Outline of the Model

### *Some General Design Principles*

Before discussing the specific features of our modeling endeavors, it might be of some use to outline some general design principles that guided us. The first obvious design principle is that the parameters of a model should be plausible in terms of what is known about eye movements and reading. That is, the estimates of lexical access time or eye movement planning time that a model produces should relate to what is known about lexical access time and eye movement planning time. As noted earlier, the eye usually remains on a word for about 200–250 ms. This means that the eye movement is typically programmed roughly 25–100 ms after the eye first lands on a word, as the time between the initiation of an eye movement program and the execution of the eye movement is on the order of 150–175 ms (Rayner, Slowiaczek, Clifton, & Bertera, 1983). Thus, it appears that cognitive events are often influencing the programming of an eye movement within the first 100 ms after the eye has landed on a word. Given that the primary cognitive event, lexical access, takes on the order of 100–300 ms (Rayner & Pollatsek, 1989), the manner in which a cognitive event can influence the initiation of an eye movement program is not immediately obvious. As argued later, parafoveal preview is an important ingredient in being able to explain this.

The second design principle that we adopted is that the mechanisms for guiding the eyes should be relatively "dumb." That is, we wanted to avoid assumptions in which some sort of central executive was habitually involved in the decisions of *where* and *when* to move the eyes. The reason for this is that the central cognitive task in reading is to understand the text. It seems unreasonable to think that the system for programming the eyes should be set up to compete with text understanding for limited processing resources (Rayner & Morris, 1992). Moreover, readers are unaware of consciously directing their eye movements in reading except on rare occasions (discussed later). This makes the modeling task interesting as one has to explain apparently "intelligent" behavior, such as skipping predictable words, from "dumb" mechanisms.

A third design principle that we adopted is to assume that the signal to move the eyes is the successful completion of a psychological process (such as lexical access). In particular, we wanted to avoid having to posit that decisions to move or not move the eyes are based on noticing that one is having difficulty in extracting information. The reason for this is that such processes, if accurate, are likely to be slow. For example, one might want to posit that a second fixation is made on a word because lexical access has not been achieved before a certain deadline (Henderson & Ferreira, 1990). The problem with such an assumption is the following. If one sets the deadline for "not achieving lexical access yet" to be conservative (late), it would mean that the duration of the fixation before this refixation would have to be long as it would involve this slow decision followed by 150 ms or so of additional time to execute the eye movement program. (In fact, first fixations that are followed by refixations tend to be relatively short.) On the other hand, if one refixated on the basis of an early deadline of failure to achieve lexical access in order to make the process reasonably quick, one would be in danger of programming unneeded refixations on a majority of words.

### *The Core Model*

Our model is a quantification and elaboration of a model originally proposed by Morrison (1984) that we (Pollatsek & Rayner, 1990; Rayner & Pollatsek, 1989) have modified subsequently. The Morrison model proposes that language processing (which we interpret as lexical access) is the basic engine that drives eye movements. There are two key assumptions in Morrison's model. The first is as follows:

> If word *n* is fixated and attended to, then when lexical access of that word is complete, attention moves to word *n* + 1 and an eye movement program is simultaneously initiated to fixate word *n* + 1.

The eye movement program is completed, and an eye movement is executed with some latency determined by characteristics of the eye movement system. There is both behavioral and physiological evidence for such movement of covert spatial attention before moving the eyes (Posner, 1980; Wurtz, Goldberg, & Robinson, 1982). In Morrison's model, the shift of covert attention is tightly linked to eye movements; in essence, the movement of covert attention is the command to move the eyes.

If this were the only assumption of the model, it would successfully predict that fixation times on words would be a function of variables that influence lexical access time, but the reader would be doomed to fixate each word in turn. Thus, something more is needed to explain why words are skipped. Morrison's (1984) second assumption provides a plausible "dumb" mechanism that can explain skipping. He posits that eye movements can be programmed in parallel.

> When the reader shifts attention to word *n* + 1, lexical processing of word *n* + 1 begins. If this is completed before the eye movement to word *n* + 1 is executed, then covert attention moves to word *n* + 2 and an eye movement is programmed to word *n* + 2 (in parallel with the program to word *n* + 1).

At this point, one of three things can happen: (a) If the time between the initiation of the program to fixate word *n* + 1 and the initiation of the program to fixate word *n* + 2 is "short," the program to fixate word *n* + 1 is cancelled, and only the later program is executed (this can be thought of as *backward masking* in the motor system). (b) If the time between the initiation of the two programs is "long," then the executions of the saccades to words *n* + 1 and *n* + 2 are both executed in a normal manner (but the fixation on word *n* + 1 would most likely be short because the eye movements were programmed in parallel). (c) If the time between initiation of the two programs is "intermediate," then a single eye movement is executed, but the saccade will go to a target intermediate between the targets of the two programs.

There is evidence for these mechanisms in a series of experiments by Becker and Jürgens (1979; see also Abrams, 1992) who used a much simpler situation than reading. In their experiments, participants fixated a target, and then the location of the target shifted. Participants were supposed to fixate the new location of the target. However, on some trials, the location of the target shifted a second time. What Becker and Jürgens observed, consistent with the assumptions indicated in the previous paragraph, was that (a) when the two shifts of the target were close to one another in time, participants would make only one eye movement (to the final location of the target); (b) when the two movements were relatively far apart in time, participants would make two saccades (one to the first shifted location and the second to the second shifted location), with the first of these fixations being relatively short (about 50 ms); whereas (c) if the time between the two shifts of the target was intermediate, participants would often make a single saccade to a location intermediate between the two shifted locations.

Notice that Morrison's (1984) model explains the data on skipping that we discussed earlier. If lexical access of word $n + 1$ is rapid (as with high-frequency words that are highly predictable), then the time between the attention shift to word $n + 1$ and the attention shift to word $n + 2$ will be short, and, hence, there is a reasonable likelihood that word $n + 1$ will be skipped. In short, Morrison's model can simply and elegantly account (at least qualitatively) for two of the major facts about eye movements in reading: (a) fixation time on a word is a function of word frequency and length, and (b) high-frequency words and predictable words are likely to be skipped. Thus, we think this is a promising beginning for a model of eye movements in reading. However, there are some problems and limitations with Morrison's model that we need to discuss before outlining our quantitative modeling efforts. First we discuss two major limitations that we do not attempt to deal with in this article.

### *Limitations of the Core Model That We Do Not Address in Our Modeling*

*Ignoring higher order processing.* The first limitation is that Morrison's (1984) model posits that lexical access (or a related process) always drives the eyes. We know this cannot be true. For example, when participants read a sentence such as "Since Jay always jogs a mile seems like a short distance," readers show disruption unrelated to accessing the form or meaning of individual words (Frazier & Rayner, 1982). Instead, the cause of the disruption is that the reader has been led down a syntactic "garden path" and has constructed a parse of the sentence in which "a mile" is the object of "jogs." When the reader then processes the word "seems," this word cannot be incorporated into the syntactic structure, and thus the reader needs to reparse the sentence to come up with the correct syntactic structure (in which "a mile" is the subject of "seems"). While the exact interpretation of readers' behaviors in such sentences is a matter of some dispute (Frazier & Clifton, 1996; MacDonald, Pearlmutter, & Seidenberg, 1994), the point that is essential for our present purposes is that readers indeed show quite massive disruption when encountering "seems" and that this disruption is unrelated to lexically accessing that word. In fact, the consequences of "garden pathing" are variable, with readers sometimes fixating "seems" a long time, sometimes going back to the beginning of the sentence and rereading it, and at other times apparently adopting other, more complex, "repair" strategies (Frazier & Rayner, 1982).

A second example of a process that Morrison's (1984) model does not account for is resolving anaphoric relations, such as finding the appropriate antecedent for a pronoun. Several studies have shown effects of pronominal assignment on eye movements by using passages in which a gender stereotyped profession (such as nurse) is mentioned first and then, later, either the pronoun *he* or *she* appears, for which there is no other possible antecedent (Ehrlich, 1983; Ehrlich & Rayner, 1983). The finding is that readers are disrupted when the gender of the pronoun mismatches the gender stereotype of the prior noun. Again, the disruption is not plausibly related to lexical access of the pronoun but to processes engaged after the denotative meaning of the pronoun has been extracted. The effect of this disruption is usually not simple, however; in general, all that can be said is that a fixation duration somewhere in the following phrase may be lengthened.

The above-indicated two examples are representative of situations in which processes "higher" than lexical access have been shown to affect the movement of the eyes through the text. Three comments are in order. The first is that such effects are likely to be quite difficult to model. In some cases (e.g., pronominal reference), the effect is hard to localize, and in many cases, once the disruption occurs, readers can adopt a variety of strategies in an attempt to repair the misunderstanding. In addition, attempting to predict such effects precisely would require a much better model of on-line parsing and text comprehension than anyone has proposed to date.

The second comment is that the higher order influences on eye movements in reading that have clearly been demonstrated appear to be disruptive. We take as a working hypothesis that our modification of Morrison's (1984) model represents the "default" control of the eyes in reading. This default control is overridden at certain times when higher order text comprehension processes decide that something "doesn't compute." At that point, a signal goes to the default process to stop, and then a more complex process takes over until the difficulty is repaired. While we are not sure that this is the only way that higher order processes directly intervene in eye movement control, we think it is a reasonable working hypothesis given the available data on reading.

The third comment is that our modeling efforts do deal with indirect influences of higher order processes on eye movements. That is, in our view, the lexical access process includes top-down influences from text processing as well as the bottom-up process of identifying the word from the marks on the printed page. Explaining this interactive processing is currently a major subject of debate as to its sources (e.g., localized priming effects, syntactic constraints, and discourse constraints). As a result, we deal with such effects under the general rubric of predictability with the caveat that such predictability effects could very well be quite complex and be at different levels.

These considerations lead us to narrow the scope of our modeling. As a result, our models explain only part of the reading process. Nonetheless, we think that something like lexical access is controlling the eyes on a large majority of the fixations in reading and that successfully accounting for these fixations will be a major accomplishment. In addition, if our working hypothesis is correct, an adequate model of "normal" eye movement control in reading will allow for a better formulation and analysis of the effects of higher order text comprehension processes on reading. These considerations also lead us to narrow the domain of the data we analyze. We assumed that regressions to prior words mostly come from such higher order processes; thus, we excluded from our corpus of data any sentence in which a reader regressed to a word before the one that was currently fixated.

*We are not modeling the precise location of fixations.* Morrison's (1984) model is vague about exactly where a saccade lands: He assumes that a command is sent to fixate a word and a saccade is programmed somewhere on the target word. For the most part, we leave our predictions at the same level of generality and will merely predict which word is fixated on each fixation. Unlike modeling higher order processes, adding this next level of complexity to our modeling efforts would not be a big conceptual step. McConkie, Kerr, Reddix, and Zola (1988) have shown that there are general laws that explain a large percentage of the variability in these data. Specifically, their large corpus of data is explained by a relatively simple model: (a) readers program saccades to land in the middle of a word (O'Regan, 1981; Rayner, 1979); but (b) there is bias, and, on average, saccades tend to undershoot the target; (c) there is also random variability in the location of the end point of the saccade; and (d) both the bias and random variability increase the farther the "launch site" of the saccade is from the target word. While adding a module that incorporates these assumptions to the type of model we propose is not conceptually difficult, it would make evaluation of the rest of the model close to impossible. However, we feel fairly confident that if our modeling efforts are successful at the level we intend, then adding a module that incorporates these characteristics of McConkie et al.'s model is likely to fit the reading data well at this increased level of specificity.

We need to make one more point about McConkie et al.'s (1988) data and modeling, which indicates a small limitation in our modeling even at the word level. Our modeling assumes that all saccades intended for a word actually land on the word. Their data, however, suggest that this is not the case; that is, the distributions of "landing positions" on a word that McConkie et al. reported as looking like truncated normal histograms. Thus, there is the clear suggestion that the missing tails of the distributions (saccades that undershoot the word or overshoot it) are part of the normal process. These tails, however, are relatively small, and so our assumption that saccades always hit the intended word target may not be grossly in error.

### *Limitations of the Morrison (1984) Model We Hope to Address*

One feature of Morrison's (1984) model is that it predicts that the amount of information extracted from the parafovea is independent of the difficulty of processing the word in the fovea. This prediction follows from the tight locking of the shift of covert attention with the programming of the eye movement. That is (according to Morrison's model), when the reader accesses the fixated word, covert attention moves immediately to the next word ($n + 1$) and the saccade follows with a latency determined by properties of the eye movement system. Thus, the time spent attending to the parafoveal word will merely be equal to the mean latency of executing the eye movement program and will be unaffected by the time it took to access word $n$.

There are data, however, that indicate that this assumption is false and that preview benefit decreases as the difficulty of foveal processing increases. Henderson and Ferreira (1990; see also Kennison & Clifton, 1995) manipulated foveal difficulty by manipulating either the frequency or syntactic complexity of a target word. In both cases, the amount of preview benefit was close to zero in the more difficult condition. In addition, Rayner (1986) found that fourth-grade readers extracted information from a narrower region of text when the text was difficult than when it was easy, and Inhoff, Pollatsek, Posner, and Rayner (1989) showed that adult readers extracted information from a smaller region when the text was distorted (e.g., the letters in the text went from right to left). The way that Morrison's model needs to be modified to account for these data seems clear: shifts of covert attention need to be decoupled from eye movement programming. Our outline of the model indicates how we did that.

We should note that the failure of Morrison's (1984) model to predict that preview benefit is modulated by foveal difficulty is symptomatic of a more general problem with the model: Morrison's model predicts that all the effects of "difficulty" of processing a word are seen in the fixation time on that word and no effects are passed "downstream" to subsequent processing. Modulation of preview benefit is one indication that there can be delayed effects of processing. A second, commonly noted, indication is that there are "spillover effects" in which the frequency of a word influences fixation durations on subsequent words (Rayner & Duffy, 1986; Rayner et al., 1989). Morrison's model cannot account for these effects either. We defer further discussion of spillover effects until later in the article, however.

The other limitation of Morrison's (1984) model that we hope to overcome is that (as it stands) it predicts that a word is never *refixated* (refixation means that the reader makes a second or third fixation on a word before moving to another word). As indicated earlier, however, our strategy in presenting our modeling efforts is to deal with refixations only after we have successfully modeled the total time a word is fixated on the first pass (gaze duration), ignoring whether this time comes from a single fixation or is the sum of multiple fixations. As shown later, when we attempt to explain refixations, we merely add one additional process, and, thus, the basic structure that we present in the first part is preserved. Moreover, the initial predictions that we make about gaze durations and skipping are little affected by adding this additional process.

### *Relation of Our Modeling Efforts to Other Models of Reading*

It might help if we also briefly review prior efforts to model eye movements in reading in order to clarify our modeling enterprise. Prior models of eye movement control can be broken down into two types of models: qualitative and quantitative models. The qualitative models, such as those proposed by Morrison (1984) and O'Regan (1990, 1992), are largely verbal descriptions of eye movement control. Both models have generated a considerable amount of research. We have already described Morrison's model in some detail because it serves as the basis of our modeling efforts. It is beyond the scope of this article to discuss O'Regan's model in detail. However, briefly, his model focuses primarily on where readers fixate in words, and it places less premium on cognitive processes influencing the reading process than does Morrison's model. Another qualitative model of reading that has widely been referenced (Just & Carpenter, 1980) focuses on the duration of fixations and finesses most of the detail about the location of fixations. Like Morrison, Just and Carpenter ignore refixations within words and focus solely on gaze duration on a word. They also ignore the question of whether a word is fixated or skipped because they define gaze duration differently than we have. Their definition of gaze duration counts a trial when a word is skipped as a zero fixation time on that word, and the mean gaze duration includes those trials. That is, their measure of gaze duration is a composite measure, indexing both the duration of actual fixations on the word and how often it is skipped.

In addition to the qualitative models, there have been some attempts to produce quantitative models. Specifically, there have been several recent attempts to model the characteristics of saccadic eye movements during reading (Legge, Klitz, & Tjan, 1997; Suppes, 1990). These models have focused on low-level aspects of reading and have not concerned themselves with the duration of fixations. The two models, however, are quite different. Legge et al. attempt to explain the details of where readers fixate cognitively, assuming an intelligent guiding mechanism that is controlled by lexical access and the details of which letters can be processed because of acuity limitations and other considerations. In contrast, Suppes assumes rather "dumb" mechanisms and focuses on the stochastic properties of the variability of saccadic eye movements. One other quantitative model that deserves mention is that of Thibadeau, Just, and Carpenter (1982), which is a formal production system that provides a more quantitative account of the Just and Carpenter (1980) model. Like its predecessor it focuses on a composite gaze duration measure and ignores important details like the probability of refixations and of skipping.

These models are all reasonable attempts to explain part of the eye movement record. However, our modeling goes a significant step beyond them by trying to account simultaneously for the details of individual fixation durations and the location of individual fixations (at the level of which word is fixated). As (to the best of our knowledge) there are no extant models that successfully account for eye movements in reading at the level of our modeling efforts, our main focus is not in testing one model against another. Instead, we view our modeling as a quest for a relatively simple model that is competent to account for eye movement control in reading. In a sense, our models can be thought of as attempts at "existence theorems": proof that one can account for the details of eye movements in reading with a plausible model that assumes cognitive control.

## Successive Approximations to a Model of Eye Movement Control in Reading

Our general approach is "minimalist" (McKoon & Ratcliff, 1992). In each of the models discussed later, we attempted to posit as few processes as possible and have used as few free parameters as possible to determine whether the kind of model we are proposing is basically on the right track. We believe that it is better to have a model that may be a bit oversimplified (but basically correct) than to have a model that is so complex that one cannot penetrate it to understand why it is or is not working.

We discuss a number of successive approximations of a computational model of eye movement control and present fits of our simulations to observed data. For convenience, we refer to the general model as the *E-Z Reader* model.[2] The general E-Z Reader model is an elaboration of Morrison's (1984) model with one important change: The E-Z Reader model decouples the shift of attention and the program to make an eye movement. This change is primarily motivated by the finding that foveal difficulty influences the degree to which information from the parafovea is extracted (Henderson & Ferreira, 1990).

We believe, however, that there are other data that indicate that attentional shifts and eye movement programming may not be as tightly coupled as Morrison (1984) envisaged. First, in the standard spatial cuing paradigm (Posner, 1980), shifts of attention occur even though the participant maintains fixation. One could possibly explain such a phenomenon by positing that an eye movement was automatically programmed during the attentional shift but then was canceled by a conscious mechanism driven by the demands of the task. However, it is more difficult to explain the finding that covert attention can be drawn to a different location than the target of a saccadic movement (Posner, 1980; Stelmach, Campsall, & Herdman, 1997).

Second, although the results of several studies suggest that there is an obligatory attentional shift to the target of a saccade before its execution (Hoffman & Subramaniam, 1995; however see Stelmach et al., 1997), an analysis of the costs associated with dissociating attention from the saccadic target indicated only a minimal decrease in saccadic accuracy and latency (Kowler, Anderson, Dosher, & Blaser, 1995). Moreover, Klein and Pontrefact (1994) have shown that, in some circumstances, a saccadic eye movement program can be prepared to a location with no apparent shift in covert attention (even though there is no obvious cost of shifting attention in their task). Of course, one has to be cautious in generalizing from these results to reading because they are all derived from laboratory tasks that are, to varying degrees, unnatural: Participants are required to maintain fixation much longer than in most real-word circumstances and are maintaining or shifting attention consciously according to instructions, as a result of sudden stimulus changes, or both, rather than the largely unconscious endogenous shifts of attention that occur in reading.

Finally, most of the evidence showing a close link between attention and eye movements has implicated *input selection* (Treisman, 1969), or those processes related to spatial selection, rather than *analyzer selection,* or those processes related to target identification. This distinction is important because the decoupling of eye movements from attention in our model naturally lends itself to such an interpretation: Although saccadic programming includes covert shifts of spatial attention to the upcoming target location (i.e., input selection), "shifting of attention" refers to the process of disengaging the mechanisms responsible for word identification so that they can be used elsewhere (i.e., analyzer selection). Of course, our basic assumption could be correct—shifts of attention are decoupled from eye movement programming in reading—but the particu-

---

[2] The name "E-Z Reader" originated from a television program, *The Electric Company,* that the first author watched in kindergarten. During one of the program's skits, a very hip character named Easy Reader (played by Morgan Freeman) walked around reading different kinds of signs, impressing upon viewers just how cool it is to read. After the model was dubbed by the first author, the second author pointed out that, unbeknownst to the first author, the character's name was a spoof on the movie *Easy Rider.*

lar mechanisms we propose later for how this takes place are wrong. Our goal, however, was to posit the simplest plausible rationale and mechanism for how such a decoupling might occur.

## E-Z Reader 1

Our plan was to begin our modeling with a very small set of assumptions to determine which processes are absolutely necessary to account for eye movement control in reading. In *E-Z Reader 1,* we make no reference to either controlled or conceptually driven processes and assume that the eyes are driven forward only by the lexical properties of individual words. The model consists of five processes: (a) a familiarity check of a word ($f$), (b) completion of lexical access of the word ($lc$), (c) a labile stage of saccade programming that can be canceled by a subsequent saccade ($m$), (d) a subsequent nonlabile stage of saccade programming ($M$), and (e) the actual saccadic eye movement ($s$).

These processes can be segregated into two functionally distinct modules (Fodor, 1983). The familiarity and lexical identity of a word are computed in a module responsible for word recognition. The process by which a word's familiarity, $f$, is computed (before lexical access) is likely to be the product of many factors, including frequency of occurrence in printed text, length, age of acquisition, frequency of usage, recency of usage, and the number and frequency of word *neighbors* (i.e., words that differ from a word by a single letter). Because these factors are related to the proficiency and probability of successful lexical resolution, the level of familiarity seems like a good candidate to be the signal to the eye movement system that lexical access is imminent and that a saccade should be planned to the subsequent word.

As in Morrison's (1984) model, we assume that completion of lexical access, $lc$, is the signal to shift covert attention to the next word. Thus, E-Z Reader decouples covert attention shifts from eye movement programming by having (a) completion of the familiarity stage be the signal for the initiation of an eye movement program and (b) completion of lexical access be the signal for a shift of covert attention. This decoupling is central to the general E-Z Reader model. The division of lexical access into two discrete serial stages, $f$ and $lc$, is largely a modeling convenience. All we are assuming is that some computation from the word identification process is made before full identification of the word that can trigger an eye movement program. Lexical access refers to the process of identifying a word's orthographic and/or phonological pattern so that semantic information can be retrieved. Although familiarity and lexical access are likely to be affected by many of the same factors, the latter process is inherently more difficult because of the greater specificity needed to locate a unique representation in memory.

For example, the familiarity of an orthographic input pattern might reflect the pattern's overall similarity to the collective contents of the lexicon. Such a measure of familiarity, in conjunction with an appropriate criterion, would provide a means to ascertain quickly and accurately whether a particular orthographic pattern corresponds to a known word and hence whether lexical access is imminent. In contrast, lexical access may require that the orthographic input match, or make contact with, a single representation in the lexicon. For instance, the orthographic pattern might be used to retrieve phonological and/or semantic information or simply information that in some way indicates that the pattern corresponds to a known word form.

The distinction between (a) matching on the basis of global similarity and (b) retrieval through reintegration, or pattern completion, is common to many models of memory (Eich, 1985; Hintzman, 1988; Humphreys, Bain, & Pike, 1989; Murdock, 1993; Raaijmakers & Shiffrin, 1981). For example, in MINERVA 2 (Hintzman, 1988) an input pattern can be compared with the collective contents of memory to produce a scalar value, *echo intensity,* that reflects how similar, or familiar, the pattern is to all of the other items in memory. An input pattern can also be used to retrieve additional information, *echo content,* from memory through redintegration. The process of redintegration is more computationally demanding than determining global similarity because (a) the former process requires an additional step, and (b) the pattern that is produced by redintegration is subject to interference from other, similar patterns so that it is often necessary to "clean up" the pattern by an iterative (time consuming and error prone) "deblurring" process. However, despite the fact that one process is more computationally demanding than the other, both processes are completed by the same system. This fact underscores our view that the familiarity check and lexical access could be the product of a single-word recognition module.

The remaining processes are components of a module responsible for programming and executing saccades. The division between labile and nonlabile motor programming components ($m$ and $M$, respectively) is consistent with Becker and Jürgens's (1979) results indicating that the computations necessary to make a saccade to a target location can be canceled or modified if a new target location is presented during the early stages of computation. If the new location is indicated late in programming, the computations cannot be interrupted and the original saccade will be executed. The logic of this partitioning and a more detailed discussion of the individual model processes are presented after a brief description of our basic modeling approach.

Figure 3 contains two flow diagrams that summarize our verbal description of E-Z Reader 1 and indicate how the various components of the model are interrelated. However, for the purposes of quantitative modeling, another level of description, an *order-of-processing* (OP) diagram, is needed (see Fisher & Goldstein, 1983, for an introduction to this topic). Each box in the OP diagram in Figure 4 is not a component of the model. Instead, it represents a discrete state that the model can occupy at any given point in time in terms of which processes have been completed and which are ongoing. The individual processes (which were represented by boxes in Figure 3) are represented by lowercase letters, and each is indexed by a subscript to indicate the relative ordinal position of the word to which it is being applied. Ordinal position is indexed relative to the word currently being retrieved from lexical memory—word $n$. (Note that in Figure 4 and in the following discussion of the figure, $n$ is the index of the word that is currently being processed and not the index of the word being fixated.) The box representing each state contains ongoing processes. Arrows are labeled by letters that represent processes that have been completed in the

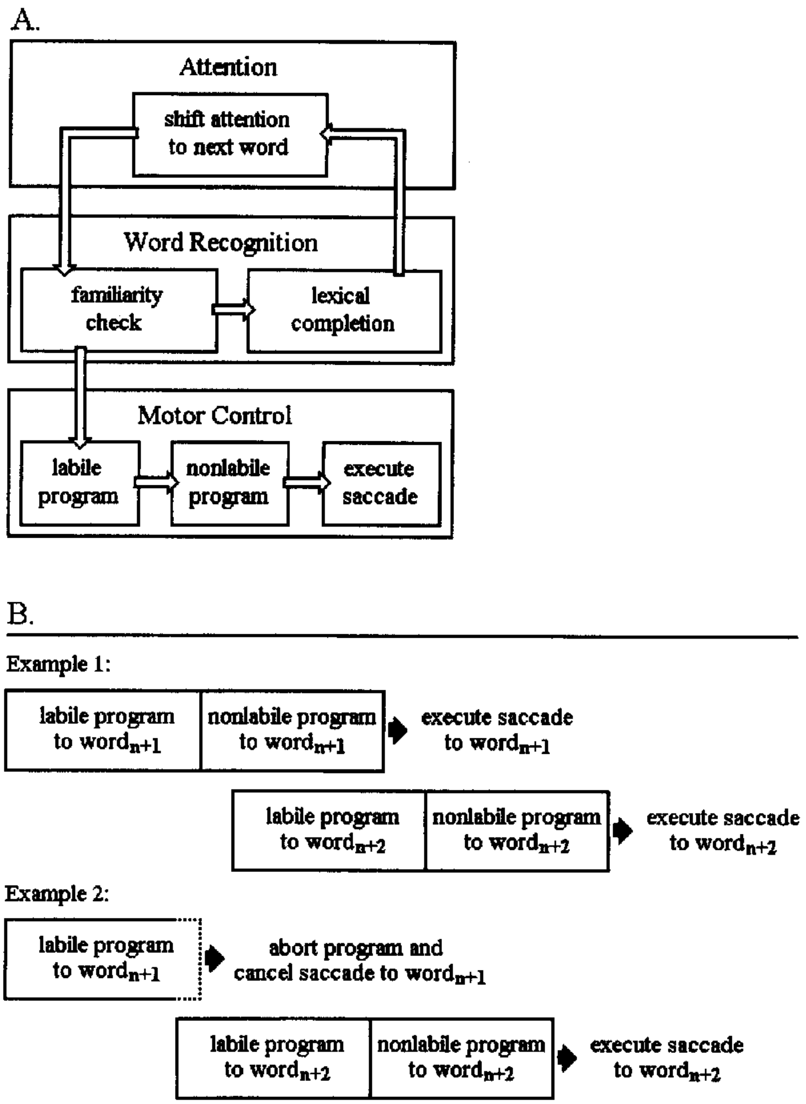


*Figure 3.* (A) This shows a schematic diagram of the component processes of E-Z Reader. The diagram indicates how hypothesized cognitive operations influence eye movements and covert attention and how attentional processes in turn influence cognitive operations. In the model, an attentional shift to an upcoming word initiates a familiarity check on that word. After the familiarity check has finished, lexical access continues, and a labile program to make a saccade is initiated. The completion of lexical access results in an attentional shift to the next word, and the completion of the labile stage of saccadic programming results in a nonlabile stage of programming, followed by a saccade. (B) This shows a schematic diagram illustrating the relationship between the labile and nonlabile stages of saccadic programming. In the first example, the labile stage of programming a saccade to word $n + 2$ is initiated during the nonlabile programming of a saccade to word $n + 1$; consequently, both saccades are executed (i.e., a saccade is made to word $n + 1$, then another saccade is made to word $n + 2$). In the second example, the labile stage of programming a saccade to word $n + 2$ is initiated during the labile programming of a saccade to word $n + 1$; the latter program is therefore aborted, and a single saccade is made to word $n + 2$ (i.e., word $n + 1$ is skipped).

transition between the states. Also, note that some arrows are dashed and are labeled by $n = n + 1$. In these transitions, attention has moved from word $n$ to word $n + 1$, and thus the index $n$ is updated. Finally, because the onset of a saccade always occurs immediately after the nonlabile stage of programming, $M$, has been completed, saccades are only represented implicitly in the OP diagram: That is, an arrow indicating that $M$ is complete implicitly signifies that $s$ is complete as well, with a total elapsed time, which is the sum of the $M$ and $s$ durations. In other words, $M$ and $s$ can formally be considered a single process in which the duration is the sum of the durations of the two subcomponents.

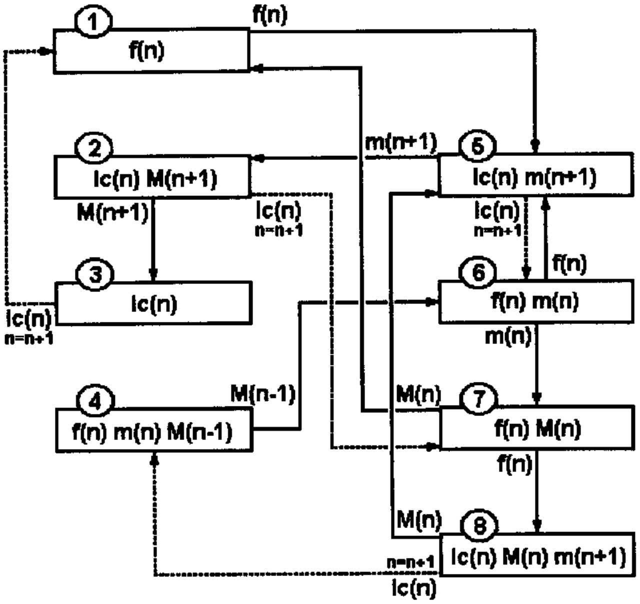
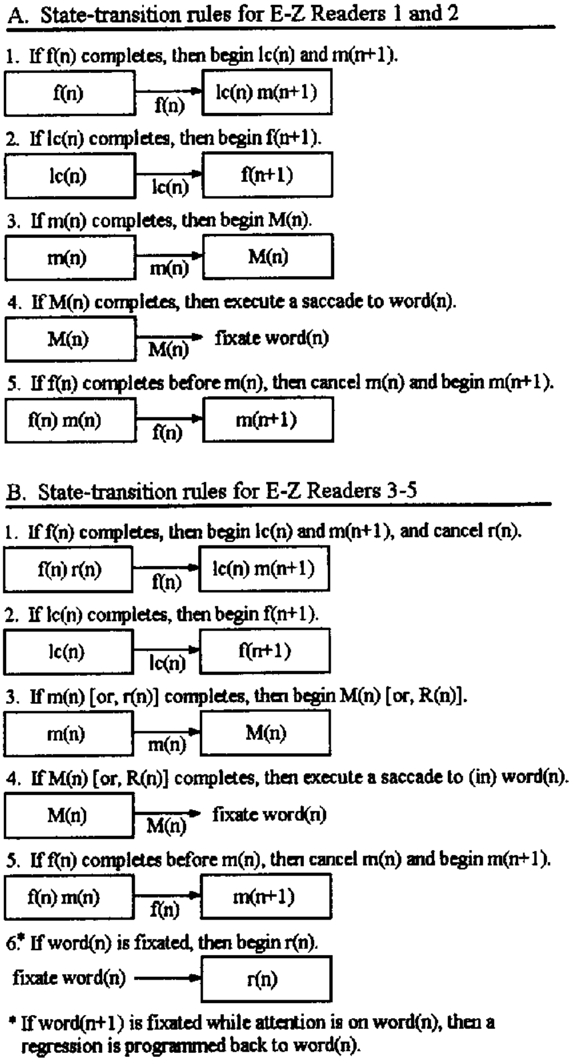


*Figure 4.* An order-of-processing diagram representation of E-Z Readers 1 and 2. The boxes are possible states that the model could be in, with the ongoing processes represented in the box. Each arrow is labeled by the process that has completed, and dotted arrows indicate that attention has shifted forward (indicated by $n = n + 1$ on the diagram). Note that $n$ indexes the attended word, not the fixated word. (The numbers given to the boxes are essentially arbitrary.) $f$ = familiarity check of the word; $lc$ = completion of lexical access of the word; $m$ = a labile stage of saccade programming that can be canceled by a subsequent saccade; $M$ = a subsequent nonlabile stage of saccade programming.

The OP diagram representation embodies some of the major assumptions of the model. First, processes can be executed in parallel. In State 2, for example, completion of lexical access of word $n$ is occurring contiguously with the nonlabile programming of a saccade to word $n + 1$. Second, note that the state transitions in Figure 4 are governed by a small number of rules (see Figure 5). The model is thus similar to a finite-state grammar because these transition rules define a limited state space of processes that can be operating in parallel at any given point in time. Finally, the transition rules are iterative. This is represented throughout the OP diagram by the arrows between states. For example, completion of lexical access of word $n$ in State 3 transfers the model to State 1, where the word recognition module starts computing the familiarity of word $n + 1$.

With this general overview, the individual processes can now be discussed in greater detail. The word recognition module begins computing the familiarity checking stage of lexical access for word $n$ immediately after lexical access of word $n - 1$; thus, lexical access is the "engine" driving the eyes forward. However, to simplify the formal aspects of the modeling, we have divided the process of lexical access into two serial stages: (a) the familiarity stage, $f_n$, and (b) $lc_n$, which represents the process of completing lexical access after the familiarity stage is complete. This formalism implies that the computation of $f_n$ invariably precedes lexical access. This seems reasonable because the computation of global similarity is easier than item retrieval (Eich, 1985; Hintzman, 1988; Humphreys et al., 1989; Murdock, 1993; Raaijmakers & Shiffrin, 1981).

One metaphor for these processes is suggested by random-walk and diffusion models of information accrual (Ratcliff, 1978, 1988): On shifting attention to a word, information about the word's identity begins to accumulate at a rate proportional to the word's frequency, neighborhood density, and so forth. The amount of information required to ascertain whether the word will eventually be recognized (i.e., whether the word is in the reader's vocabulary) is relatively modest—enough to


A. State-transition rules for E-Z Readers 1 and 2

1. If f(n) completes, then begin lc(n) and m(n+1).

2. If lc(n) completes, then begin f(n+1).

3. If m(n) completes, then begin M(n).

4. If M(n) completes, then execute a saccade to word(n).

5. If f(n) completes before m(n), then cancel m(n) and begin m(n+1).

B. State-transition rules for E-Z Readers 3-5

1. If f(n) completes, then begin lc(n) and m(n+1), and cancel r(n).

2. If lc(n) completes, then begin f(n+1).

3. If m(n) [or, r(n)] completes, then begin M(n) [or, R(n)].

4. If M(n) [or, R(n)] completes, then execute a saccade to (in) word(n).

5. If f(n) completes before m(n), then cancel m(n) and begin m(n+1).

6.* If word(n) is fixated, then begin r(n).

* If word(n+1) is fixated while attention is on word(n), then a regression is programmed back to word(n).

*Figure 5.* State-transition rules for E-Z Readers 1 and 2 (A) and E-Z Readers 3-5 (B). Below each rule is an example showing how the rule is represented in the order-of-processing diagram notation. Note that $n$ indexes the attended word, not the fixated word. $f$ = familiarity check of the word; $lc$ = completion of lexical access of the word; $m$ = a labile stage of saccade programming that can be canceled by a subsequent saccade; $M$ = a subsequent nonlabile stage of saccade programming; $r$ = the labile stage of intraword saccadic programming that can be cancelled by a subsequent saccade; $R$ = the nonlabile stage of intraword saccadic programming.

exceed some relatively low-level threshold (i.e., the duration of $f_n$). However, additional information is required to determine a word's identity—enough to exceed a more rigorous criterion (i.e., lexical access). A second metaphor for $f_n$ and $lc_n$ is in terms of an activation–verification model of lexical access (e.g., Paap, Newsome, McDonald, & Schvaneveldt, 1982): $f_n$ would represent the initial activation stage of lexical access, and $lc_n$ would represent the later verification stage. However, we do not wish to tie $f_n$ and $lc_n$ to a particular model of lexical access. The important point is that there are two discrete time points at which different signals are given to the attentional and eye movement systems (see Figure 3).

In E-Z Reader 1, the mean durations of both the $f_n$ and $lc_n$ stages for a particular word are assumed to be functions of that word's natural frequency of occurrence (as tabulated in the norms of Francis & Kučera, 1982). Token—rather than type—frequency counts were used because our initial endeavor was to explain eye movements by using only data-driven processes. This is likely to be an oversimplification as many factors other than word frequency plausibly influence the durations of $f_n$ and $lc_n$ as well. However, our goal is to start as simply as possible and with as few free parameters and add complications only when they are needed. The assumption that the duration of $lc_n$, the second stage of lexical access, is also a function of word frequency is a critical one for reasons that are made clearer in the following paragraphs. It seems reasonable, however, because the process of full lexical access should be more dependent on the frequency of the word than an initial stage, which would be more strongly influenced by other factors, such as its orthographic properties.

To get a feeling for the model, consider how a simulation would work. At the beginning of the sentence, the model is in State 1 in Figure 4. The $f_n$ in the box indicates that the familiarity stage on Word 1 is not complete ($n = 1$ in this case). When just starting the sentence, the first change of state has to be completion of the familiarity stage on Word 1, leading to State 5 ($f_n$ is not represented in State 5 because it has been completed). Now two processes are initiated, $lc_n$ and $m_{n+1}$, which means that either could finish first, and thus there are two possible alternatives for leaving State 5: going to State 2 or going to State 6.

Consider the transition to State 6: Lexical access of Word 1 ($n = 1$) is completed, which removes $lc_n$ and increments $n$ (because the reader is now attending to the next word); as a result, the ongoing processes, $f_n$ and $m_n$ in Box 6, are with reference to Word 2. That is, the transition represents an attention shift to Word 2 ($n + 1 = 2$), with the reader beginning to process Word 2 even though Word 1 is still fixated. The transition to State 7 indicates completion of the labile stage of the motor program to Word 2, and a transition back to State 1 indicates the completion of the nonlabile stage of the motor program and a saccade to Word 2 (remember that $n$ has been updated to 2). Note, however, that the $f$ process on Word 1 began when State 1 was initially entered, but it has already begun for Word 2 prior to when State 1 is reentered. This is because time has elapsed between entering State 6 and returning to State 1. In this time, the reader has been attending to (and processing) Word 2 so that the familiarity check stage has progressed toward its completion although it has not completed. In short, the model has obtained preview benefit. (The path from State 5 to State 2 to State 7 to State 1 is similar except that lexical access of Word 1 occurs after completion of the labile motor program, and thus there would be less preview benefit.)

Obviously, we do not have the space to describe all of the paths in the figure. However, one more path might be illuminating. We return to State 6, where Word 1 is fixated and has lexically been accessed and thus processing of Word 2 has begun, but the labile stage of the motor program to Word 2 is not completed. (Remember that $n$ was updated to 2 when we entered State 6.) The alternate path out of State 6 goes back to State 5. The arrow representing this transition indicates that the familiarity check on Word 2 is completed, which initiates $m_{n+1}$, a labile motor program to Word 3 that cancels $m_n$, the labile motor program to Word 2. Thus, when we return to State 5 (with $n = 2$ even though Word 1 is still fixated), lexical access of Word 2 and the labile stage of the motor program to Word 3 are the ongoing processes moving to completion. If we then go back through States 6, 7, and 1 (or alternate routes such as States 2, 3, and 1), when State 1 is reached, a saccade to Word 3 is executed and Word 2 has been skipped. The path from State 6 to State 5 thus enables words to be skipped.

This run through of the model should help to give a qualitative feel for the model. A quantitative application of the model, however, requires that one know the durations of the various processes. Our strategy throughout was to use the data to obtain the best fitting values for these durations. In order for that not to be an empty exercise, we used as few free parameters as seemed psychologically plausible (e.g., E-Z Reader 1 has five parameters). Our specific assumptions are as follows.

In E-Z Reader 1, the mean time for the completion of process $f_n$, $t(f_n)$ was assumed to be a linear function of the natural log of the frequency of word $n$,

$$t(f_n) = f_b - [f_m \cdot \ln(\text{freq}_n)], \qquad (1)$$

where $f_b$ and $f_m$ are the intercept and slope parameters, respectively. The natural logarithm was used because several word recognition studies (Balota & Chumbley, 1984; Chumbley & Balota, 1984) have demonstrated linear relationships between the natural logarithm of word frequency and the mean response latencies to process the words. Variability was introduced into the model by assuming that the time to complete $f_n$ on a given trial, $T(f_n)$, was a gamma distribution with the mean equal to $t(f_n)$ and the standard deviation equal to $0.33 \cdot t(f_n)$. The mean time for the completion of lexical access, $lc_n$, was also assumed to be a linear function of log frequency. To minimize the number of parameters in the model, we assumed that $t(lc_n)$ was a constant multiple of $t(f_n)$. That is,

$$t(lc_n) = \Delta \cdot t(f_n), \qquad (2)$$

where $\Delta$ is a fixed parameter greater than zero. We also assumed that $T(lc_n)$ was a gamma distribution with the standard deviation equal to about one third of the mean. (We assume that $T(f_n)$ and $T(lc_n)$ are gamma distributed with standard deviations equal to one third of the mean in all subsequent models as well.)[3] Equation 2 produces an increasing disparity between the time to complete the familiarity check and the time to complete lexical access as word frequency decreases because the slope of $t(lc_n)$ is greater than zero. This is desirable because the preview benefit gained through parafoveal processing decreases as $t(lc_n)$ increases. (Note that total lexical access time for word $n$ is equal to $t(f_n) + t(lc_n)$.)

Figure 6 illustrates why this is the case. The bottom line in Figure 6 represents the function relating the mean familiarity check time, $t(f_n)$, to the log of word frequency. The middle line represents the mean lexical access time, and the vertical distance between the middle and bottom lines is the function relating $t(lc_n)$ to the log of word frequency. As indicated earlier, this function makes the slope of the (total) lexical access time greater than that of the familiarity check time. The top line, which is parallel to $t(f_n)$, represents the mean time that the eye movement program initiated by $f_n$ is actually executed. The vertical distance between the top and bottom lines, which is independent of word frequency, is the sum of the mean eye movement programming times, $t(m_n) + t(M_n)$. The difference between the middle and top lines (shaded in the figure) represents the amount of time allowed for processing information in the parafovea because lexical access (represented by the middle line) is the trigger for attention to move off the fixated word and begin processing the parafoveal information. As seen in Figure 6, this time (represented by the vertical distance in the shaded region) decreases as word frequency decreases. (While this example is simplified, as we have assumed processing of the fixated word to begin when the word was fixated, the general principle still applies even when the fixated word itself has partially been processed in the parafovea.)

In E-Z Reader 1, the durations of $m_n$ and $M_n$ were also assumed to be gamma distributions with a standard deviation equal to one third of the mean. For simplicity, the mean durations of $m_n$, $M_n$, and the saccade execution stage, $s$, were fixed at 150 ms, 50 ms, and 25 ms, respectively, in E-Z Reader 1 and all subsequent models. In addition, in this and all subsequent versions of the model, all processes except for $s$ were assumed to have a standard deviation equal to one third of the mean, but the $s$ process was assumed to be fixed from trial to trial to reduce the computational complexity of our efforts.

E-Z Reader 1 was applied to data collected by Schilling, Rayner, and Chumbley (in press) in an eye-tracking experiment. Participants in the study read 48 sentences. Each sentence was 8–14 words in length. Three kinds of data were tabulated for each word in the corpus: (a) the natural frequency of occurrence (from Francis & Kučera, 1982), (b) the mean gaze duration, and (c) the mean proportion of times that the word was skipped. The mean gaze durations and proportion of skips did not include data from sentences that included regressions to a prior word. As indicated, E-Z Reader posits that difficulties in completing higher order processing are the major cause of such interword regressions, and, thus, explaining regressions is beyond the scope of the model. Because the difficulty reflected in the regression may have started before the regressive fixation, we decided to be conservative and exclude any trial (i.e., sentence) from our analysis in which an interword regression was made, rather than merely exclude the regression and subsequent eye movements on the trial. (As a result, 36% of the trials were used in the analyses.) The data obtained from this corpus were then collapsed across five frequency classes of words (defined in Table 1) to produce 10 means—5 for gaze durations and 5 for the proportion of skips. The first and last words of each sentence were not included in the analysis because (a) the first word is initially fixated by a reading-irrelevant movement from a fixation cross and (b) the fixation on the last word is concluded by a button push indicating comprehension of the sentence rather than by an eye movement.

E-Z Reader 1 was fitted to the corpus by using multiple grid searches to determine the best fitting values for the model parameters. A more complete description of the fitting process is contained in the Appendix. A few salient features are worth commenting on. First, Monte Carlo type simulations were needed to estimate best fitting parameters as there is no closed-form solution for any of the predicted values for observables from the model parameters (e.g., because processing of words begins before they are fixated, there is no simple relation between the time values assumed for the durations of $f$, $lc$, $m$, and $M$ and predicted gaze durations). Second, because we were trying to simultaneously fit fixation durations and probabilities of fixation, we needed to develop a composite measure of goodness of fit that encompassed both sets of data and one that was sensible. The one we chose was the root-mean-square normalized difference (error) between the observed and predicted means of the five frequency classes. The normalization process involved taking each difference between predicted and observed fixation duration or between predicted and observed probability and dividing it by the standard deviation of the observed value for that measure. (See the Appendix for a more complete description of the normalization procedure.) For E-Z Reader 1, we were fitting 10 independent data points with five parameters. However, $t(m)$ and $t(M)$ were not really free; they were set at convenient values (150 ms and 50 ms, respectively) that seemed consistent with earlier data on eye movement latencies. The results of the simulation in which the best fitting parameters were used indicate that E-Z Reader 1 gives a reasonably good fit of the data (see Table 1).

## E-Z Reader 2

E-Z Reader 1 was reasonably successful in predicting both gaze durations and skips even though it used only lexically

---

[3] Two points about our variability assumptions are in order. The first is that the total lexical access time is more variable than the familiarity check time (relative to the absolute time) because there are two random processes assumed: (a) variability assumed in generating $t(f_n)$ and (b) additional variability in producing $t(lc_n)$, which already has variability in it from sampling $t(f_n)$. This seemed psychologically reasonable to us, as one would expect the later stages of lexical access to be more variable, as they relate more to idiosyncratic properties of the word. The second is about our setting the standard deviation to be one third of the mean. This value was selected in all the gamma distributions so as to produce an overall variability of about 20% in the fixation durations; this seems consistent with the common observation that the variability in response times is roughly 20% of the mean. The reason that a greater variability than 20% in component processes is needed is that (a) when independent stages add, the standard deviation is less than the sum of the two components, and (b) a race between two components also decreases variability.

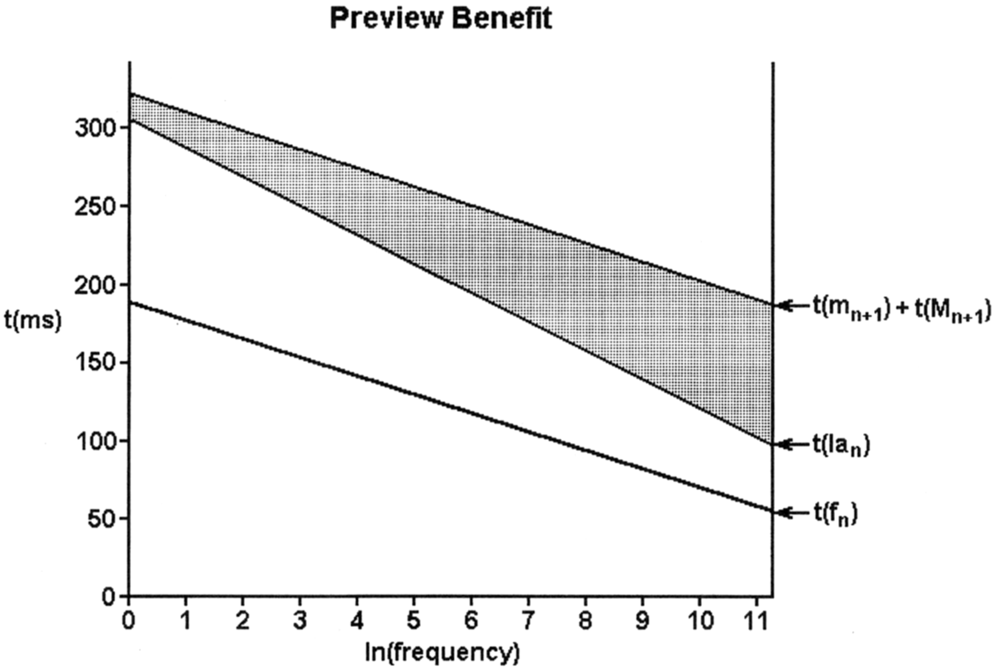


*Figure 6.* A diagram indicating how preview benefit is affected by word frequency. The bottom line represents the time to complete the familiarity check process, and the top line represents when the eye movement triggered by this familiarity check is executed. The middle line represents the time when lexical access is completed and attention moves to word $n + 1$. Thus, the vertical extent of the shaded region between the top two lines represents the time between when lexical access of word $n$ is complete and an eye movement is made to word $n + 1$, which is the time that the reader processes word $n + 1$ while still fixating word $n$.

based frequency measures. This raises the question of whether the performance of the model could be improved by incorporating the predictability of words into the model. As we indicated earlier, there is substantial evidence that the predictability of a word has effects on both the duration of fixations and on the probability of skipping a word. To the extent that these constraints facilitate lexical access (the engine driving eye movements in E-Z Reader), the conceptually driven processes underlying predictability need to be considered.

We attempted to incorporate the effects of top-down constraints on lexical access in E-Z Reader 2 in three different ways, each using the same empirically obtained predictability norms. These norms were gathered in an experiment separate from that in which the eye movement data were collected. The

Table 1
*Observed and Predicted Values of Gaze Durations and Probability of Skipping for E-Z Reader 1 for Five Frequency Classes of Words*

| Frequency class | Mean frequency | Frequency range | Gaze duration | | Probability of skipping | |
|---|---|---|---|---|---|---|
| | | | Observed | Predicted | Observed | Predicted |
| 1 | 3 | 1–10 | 293 | 304 | .10 | .09 |
| 2 | 45 | 11–100 | 272 | 279 | .13 | .19 |
| 3 | 347 | 101–1,000 | 256 | 257 | .22 | .30 |
| 4 | 4,889 | 1,001–10,000 | 234 | 236 | .55 | .46 |
| 5 | 40,700 | 10,001+ | 214 | 216 | .67 | .68 |

*Note.* The best fitting parameters for E-Z Reader 1 are $f_b = 254$ ms, $f_m = 22$ ms; $\Delta = 0.65$; $t(m) = 150$ ms; and $t(M) = 50$ ms. Root-mean-square deviation = 0.145. $f_b$ = the intercept of the mean completion time for the familiarity check stage; $f_m$ = the slope of the mean completion time for the familiarity check stage; $\Delta$ = the ratio of the length of the mean time for the lexical completion and the mean time for the familiarity check stage; $t(m)$ = the mean duration of the labile motor programming stage; $t(M)$ = the mean duration of the nonlabile motor programming stage.

20 participants in the norming experiment were presented with each sentence word by word, and after the presentation of each word, they were instructed to guess the next word in the sentence. The predictability values, $p_n$, in Equations 3–5 are the proportion of participants that guessed a particular word in a sentence when given the sentence up to that word.

We used three different rules for predictability. E-Z Reader 2a used the simplest rule. First, we assumed that predictability decremented the duration of the familiarity stage given in Equation 1 by an amount proportional to $p_n$ (see Equation 3). We assumed that $t(lc_n)$ is also decremented by $p_n$ by applying Equation 2 to the value of $t(f_n)$. It should be stressed that the $p_n$ values are not free parameters; they were not fit by using the eye movement data. Although this rule incorporates predictability into the model, it does so at the cost of making the model fairly implausible: The processing time for a completely predictable word is zero according to Equation 3. Moreover, the fit of E-Z Reader 2a, while still being reasonably good, was actually worse than the fit of E-Z Reader 1 (see Table 2). As a result, we constructed two other versions of E-Z Reader 2:

$$t(f_n) = \{f_b - [f_m \cdot \ln(\text{freq}_n)]\} \cdot (1 - p_n). \qquad (3)$$

In the simpler version (E-Z Reader 2b), we assumed that predictability had no effect on $t(f_n)$ but did have an effect on $t(lc_n)$. The rationale for this assumption was that predictability might only affect lexical identification after the initial stages of word encoding (roughly consistent with activation–verification models of lexical access such as Paap et al., 1982). Thus in E-Z Reader 2b, $t(f_n)$ is assumed to be unaffected by predictability and is the same as in E-Z Reader 1 (see Equation 1), but $t(lc_n)$ is given by Equation 4. That is, $t(lc_n)$, the difference between the lexical access time and the familiarity check time, which is equal to $\Delta \cdot t(f_n)$ in E-Z Reader 1, is decremented by $p_n$:

$$t(lc_n) = \Delta \cdot t(f_n) \cdot (1 - p_n). \qquad (4)$$

E-Z Reader 2b, while fitting the data better than E-Z Reader 1, has a serious problem: It does not account for either of the predictability effects in reading that we discussed earlier (i.e., that more predictable words are skipped more often and are fixated for less time). Specifically, by not having any predictability effect on the familiarity check stage, E-Z Reader 2b puts the whole effect of predictability after the reader has left the target word (because only the familiarity stage affects the decision to move from and/or skip the target word).

As a result, we created a third version (E-Z Reader 2c) in which predictability had an effect on both $t(f_n)$ and $t(lc_n)$, but in which the effect on $t(f_n)$ was attenuated. This version avoids the problems of E-Z Readers 2a and 2b with only a slight loss of simplicity. More specifically, in E-Z Reader 2c, $t(f_n)$ is given by Equation 5a, and $t(lc_n)$ is given by Equation 5b:

$$t(f_n) = \{f_b - [f_m \cdot \ln(\text{freq}_n)]\} \cdot (1 - \theta \cdot p_n) \qquad (5a)$$

and

$$t(lc_n) = \Delta \cdot \{f_b - [f_m \cdot \ln(\text{freq}_n)]\} \cdot (1 - p_n). \qquad (5b)$$

That is, $t(lc_n)$ is assumed to be the same as in E-Z Reader 2a, but the $\theta$ parameter "softens" the effect of predictability on $f_n$. If $\theta$ is 1, then Equation 5a reduces to Equation 3, and E-Z Reader 2c is the same as E-Z Reader 2a. If $\theta$ is zero, then there is no effect of predictability on $t(f_n)$, and E-Z Reader 2c reduces to E-Z Reader 2b. For intermediate values of $\theta$, the effect of predictability is less on $t(f_n)$ than on $t(lc_n)$. In this case, we did not search the whole parameter space but instead fixed $\theta$ at an intermediate value of 0.5 when fitting E-Z Reader 2c.

E-Z Readers 2a–c were fit to the same corpus of data by using multiple grid searches of the parameter space. The best fitting parameter values and the simulation results in which these parameters were used are presented in Table 2. As with E-Z Reader 1, there were 10 independent data points. For E-Z Readers 2a and 2b, there were five free parameters (although, as with E-Z Reader 1, $t(m)$ and $t(M)$ were not really free but were set to 150 ms and 50 ms, respectively). E-Z Reader 2c had one more free parameter ($\theta$), although this was set at a convenient value rather than all possible values being explored. As indicated earlier, the fit for E-Z Reader 2a was actually worse than for E-Z Reader 1, chiefly because it caused too much skipping of lower frequency words. However, E-Z Readers 2b and 2c both improved the fit over E-Z Reader 1 a bit. At first, it seems a bit curious that adding predictability improves the fits so little. However, that is likely due to the fact that predictability and frequency are highly correlated in our text (as is normal in discourse). Thus, E-Z Reader 1 makes up for not incorporating predictability by having a larger frequency effect in the lexical access process (i.e., a larger value for $f_m$) than the E-Z Reader 2 models. Of course, it cannot account for predictability effects in reading.

From our modeling efforts at this level, E-Z Reader 2c appears to be the clear winner. It produces the best fit to the data and is at least qualitatively consistent with what is known about the major effects in reading that were discussed earlier. Indeed, given the crude simplifying assumptions of the model (e.g., a linear relation between log frequency and processing time and a simple multiplicative relationship between predictability and processing time), it is unclear that any better fit could be expected from a model of this type than that given by E-Z Reader 2c. We could have attempted to "tweak" the model still further to produce better fits; however, we wanted to keep the model as simple as possible, and the agreement between predicted and observed values seems close enough that any attempt to improve the fit may be accounting for random error in the data. As a result, we decided to use E-Z Reader 2c as the framework in the models that follow, which attempt to give a more complete account of the eye movement record. For simplicity, we refer to Model 2c as *E-Z Reader 2* in what follows.

## E-Z Reader 3

E-Z Reader 2 successfully reproduced two basic aspects of eye movement behavior: gaze durations and skipping. The model's predictive power is limited, however, because it is restricted to molar-level measures of eye movements (i.e., gaze durations, interword saccades, and skipping). Finer grained measures such as the durations of first fixations on a word and the number of times individual words are fixated are beyond the scope of E-Z Reader 2 because such measures are logically dependent on intraword saccades (i.e., refixations). To remedy this problem,

Table 2

*Observed and Predicted Values of Gaze Durations and Probability of Skipping for E-Z Reader 2 for Five Frequency Classes of Words*

| | | Gaze duration | | | | Probability of skipping | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Predicted | | | | Predicted | | |
| Frequency class | Mean frequency | Observed | E-Z Reader 2a | E-Z Reader 2b | E-Z Reader 2c | Observed | E-Z Reader 2a | E-Z Reader 2b | E-Z Reader 2c |
| 1 | 3 | 293 | 300 | 298 | 292 | .10 | .13 | .10 | .12 |
| 2 | 45 | 272 | 273 | 276 | 272 | .13 | .20 | .20 | .20 |
| 3 | 347 | 256 | 255 | 256 | 255 | .22 | .27 | .30 | .27 |
| 4 | 4,889 | 234 | 225 | 236 | 232 | .55 | .49 | .47 | .46 |
| 5 | 40,700 | 214 | 203 | 212 | 210 | .67 | .65 | .68 | .63 |

*Note.* The best fitting parameters for Model 2a are $f_b$ = 255 ms, $f_m$ = 17 ms; Δ = 0.65; $t(m)$ = 150 ms; and $t(M)$ = 50 ms (root-mean-square deviation = 0.128). The best fitting parameters for Model 2b are $f_b$ = 248 ms, $f_m$ = 20 ms; Δ = 0.65; $t(m)$ = 150 ms; and $t(M)$ = 50 ms (root-mean-square deviation = 0.116). The best fitting parameters for Model 2c are $f_b$ = 242 ms, $f_m$ = 17 ms; Δ = 0.65; $t(m)$ = 150 ms; $t(M)$ = 50 ms; and $\theta$ = 0.5 (root-mean-square deviation = 0.104). $f_b$ = the intercept of the mean completion time for the familiarity check stage; $f_m$ = the slope of the mean completion time for the familiarity check stage; Δ = the ratio of the length of the mean time for the lexical completion and the mean time for the familiarity check stage; $t(m)$ = the mean duration of the labile motor programming stage; $t(M)$ = the mean duration of the nonlabile motor programming stage; $\theta$ = a parameter that attenuates the effects of predictability.

we augmented E-Z Reader 2 by adding a single mechanism to allow for multiple fixations on a word. The resulting model, E-Z Reader 3, thus maintains the basic structure of E-Z Reader 2 but allows for a more comprehensive account of eye movement behavior.

Staying with our minimalist approach, we adopted the working hypothesis that a single set of motor processes is responsible for programming and executing both interword and intraword saccades. The planning time for the labile stage of intraword saccades $t(r_n)$ was set equal to $t(m_n)$, and the time for nonlabile programming component, $t(R_n)$, was set equal to $t(M_n)$. The only difference between $r_n$ and $m_n$ is that they are initiated by different events: Although $m_n$ begins after $f_{n-1}$ has completed, $r_n$ starts as soon as a fixation on word $n$ begins. This difference is motivated by the following putative default mechanism that prevents the eyes from fixating any single location indefinitely: On moving the eyes to a particular viewing position within a word, immediately program a saccade to a second viewing position within the same word. This program to refixate a word is subject to cancellation by the completion of $f_n$—the signal that causes the motor module to begin planning a saccade to the next word. Thus, refixations are canceled by the same mechanism responsible for skipping: Programs for ensuing saccades can cancel earlier ones.

The motivation for this automatic refixation mechanism is as follows. Maintaining a single viewing position for identifying an object is unlikely to be optimal, especially in a task like reading in which the eyes move rapidly, and hence the location of the initial fixation on a word is likely to be somewhat variable. However, it is unlikely that a rapid refixation can be planned with a great deal of care. Hence, if a refixation saccade is programmed immediately to a location based on the physical characteristics of the word—two possibilities are the middle of the word or a location halfway between the current fixation location and the furthest end of the word—the refixation will generally place the reader in a location in which information that was difficult to extract on the initial fixation will be easier to extract on the subsequent fixation.

The primary advantage of the preceding conceptualization of refixations on words is parsimony: Refixations and interword saccades are both explained by the same theoretical constructs, without recourse to additional parameters. However, the distinction between $r_n$ and $m_n$ with respect to initiation caused an increased number of possible states for E-Z Reader 3 (relative to E-Z Reader 2). This fact is reflected in the OP diagram presented in Figure 7. However, other than the addition of two processes, E-Z Reader 3 is identical to E-Z Reader 2c (but with new best fitting parameters).

The corpus used with the previous simulations was expanded to include mean first-fixation durations, mean single-fixation durations, and the mean proportion of single and double fixations for each word in the five frequency classes. E-Z Reader 3 was then fitted to this corpus by using multiple grid searches to determine best fitting parameters. For E-Z Reader 3, there were 30 data points being fit. However, gaze duration is not completely independent of first-fixation duration and the refixation probabilities, and the refixation probabilities are not completely independent of each other. As indicated earlier, there were no free parameters added to E-Z Reader 2c. Hence, there were six parameters, but three were not really free: $t(m)$ was set to 150 ms, $t(M)$ was set to 50 ms, and $\theta$ was set to 0.5 as in E-Z Reader 2c. The simulation results are presented in Table 3.

As Table 3 indicates, the correspondence between observed and predicted values is quite good. First notice that the fits for gaze duration and probability of skipping are not quite as good as those of E-Z Reader 2, although the difference is not great.

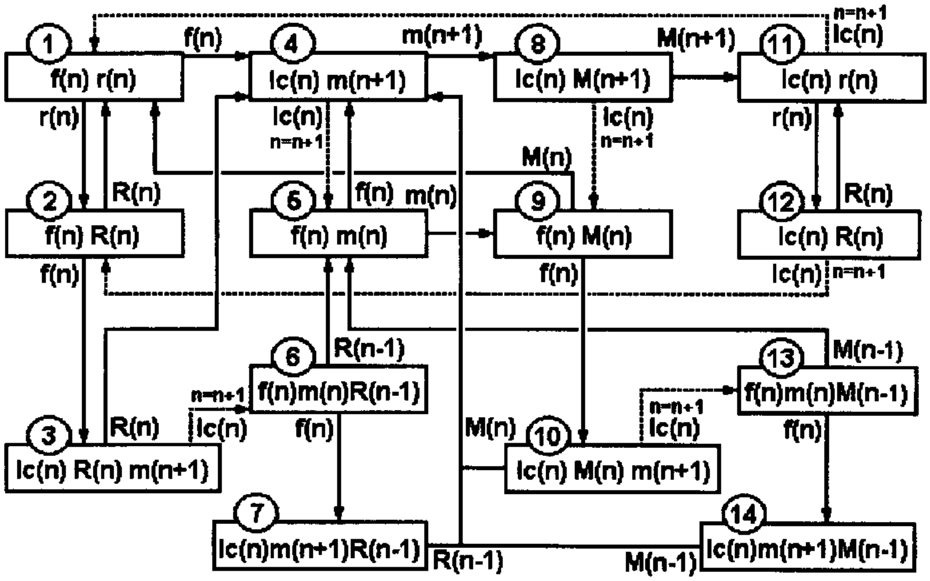


*Figure 7.* An order-of-processing diagram representation of E-Z Readers 3–5. The boxes are possible states that the model could be in, with the ongoing processes represented in the box. Each arrow is labeled by the process that has completed, and dotted arrows indicate that attention has shifted forward (indicated by $n = n + 1$ on the diagram). Note that $n$ indexes the attended word, not the fixated word. (The numbers given to the boxes are essentially arbitrary.) $f$ = familiarity check of the word; $lc$ = completion of the lexical access of the word; $m$ = a labile stage of saccade programming that can be canceled by a subsequent saccade; $M$ = a subsequent nonlabile stage of saccade programming. The additional states added are for planning and executing intraword saccades.

These discrepancies are induced by attempting to fit a substantial amount of additional data in Table 3 without adding any new parameters. In general, E-Z Reader 3 also makes reasonably good predictions for the first-fixation duration, single-fixation duration, and refixation data. The most serious discrepancies are (a) the predicted first-fixation and single-fixation durations for the lower frequency words are a bit long, (b) there is some nonmonotonicity in the predicted first-fixation durations with the predicted first-fixation duration for Class 2 larger than that for Class 1, and (c) the percentage of refixations for Classes 2 and 3 is underpredicted. (Parameter values can be changed to make the predicted durations closer to the observed values, but then the model seriously underpredicts the probability of refixating words.)

The cause of the nonmonotonicity in the first-fixation durations is complex because first-fixation durations are complex: They are a mixture of single-fixation durations and the durations of first fixations that are followed by refixations. As can be seen in Table 3, the locus of the problem is not the predictions for the single-fixation durations; they are monotonic and reasonably close to the observed values. Instead, the major cause of the nonmonotonicity of the predicted first-fixation durations is that the model is generating about the right number of refixations for Class 1 but too few for Classes 2 and 3. This causes nonmonotonicity because predicted (and observed) durations for first fixations followed by refixations are quite a bit shorter than single-fixation durations; thus, the percentage of these short durations for Class 1 is about right but is too small for Classes 2 and 3. This anomaly can be rectified by softening the assumption that the durations for the $m_n$ and $r_n$ processes are identical. (Additional simulations supported this conjecture.)

Two additional comments should be made about the constraints placed on the simulations of E-Z Reader 3 and the models to follow. First, somewhat better fits can be obtained by allowing the $t(m_n)$ parameter to be larger than 150 ms; however, we constrained it to be 150 ms so that the total motor programming time was not too large. Second, we initially assumed that all refixation eye movements would be directed to the fixated word, but this led to an unreasonable number of predicted refixations on high-frequency words. These occurred when an eye movement program had been executed from a lower frequency word to a high-frequency word, but with lexical access of the lower frequency word far from completed. This delayed the onset of processing of the high-frequency word and allowed refixations to take place even though the processing time for the high-frequency word, itself, was short. This led us to make the more reasonable assumption that was used in the simulations of E-Z Readers 3–5: The target of a "refixation" saccade is the attended word (word $n$) even when the attended word is not the fixated word and is to the left of fixation. As indicated earlier, the word to the left of fixation can be the attended word when the duration of the lexical completion stage is longer than the sum of the durations of the eye movement planning and execution stages (indicated in Figure 7 by path from State 1 to State 4 to State 8 to State 11). If a refixation program is also initiated before lexical completion (indicated by a transition to State 12 in Figure 7), a refixation will occur. (The refixation will be executed at the transition back to State 11.) By assuming that

Table 3

*Observed and Predicted Values of Gaze Durations and Individual Fixations, Probability of Skipping, Making a Single Fixation, and Making Two Fixations for E-Z Reader 3 for Five Frequency Classes of Words*

| Frequency class | Mean frequency | Gaze duration |  | First-fixation duration |  | Single-fixation duration |  | Probability of skipping |  | Probability of making a single fixation |  | Probability of making two fixations |  |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|  |  | Observed | Predicted | Observed | Predicted | Observed | Predicted | Observed | Predicted | Observed | Predicted | Observed | Predicted |
| 1 | 3 | 293 | 292 | 248 | 248 | 265 | 269 | .10 | .11 | .68 | .70 | .20 | .18 |
| 2 | 45 | 272 | 275 | 233 | 252 | 249 | 264 | .13 | .19 | .70 | .72 | .16 | .09 |
| 3 | 347 | 256 | 258 | 230 | 245 | 243 | 252 | .22 | .26 | .68 | .69 | .10 | .05 |
| 4 | 4,889 | 234 | 228 | 223 | 224 | 235 | 226 | .55 | .46 | .44 | .52 | .02 | .02 |
| 5 | 40,700 | 214 | 203 | 208 | 202 | 216 | 203 | .67 | .65 | .32 | .34 | .01 | .01 |

*Note.* The best fitting parameters for E-Z Reader 3 are $f_b = 250$ ms, $f_m = 16$ ms; $\Delta = 0.50$; $t(m) = 150$ ms; $t(M) = 50$ ms; and $\theta = 0.5$. The within-word motor programming parameters, $t(r)$ and $t(R)$, are assumed to have the same values as $t(m)$ and $t(M)$, respectively. Root-mean-square deviation = 0.194; this value is not strictly comparable to those of E-Z Readers 1 and 2. $f_b$ = the intercept of the mean completion time for the familiarity check stage; $f_m$ = the slope of the mean completion time for the familiarity check stage; $\Delta$ = the ratio of the length of the mean time for the lexical completion and the mean time for the familiarity check stage; $t(m)$ = the mean duration of the labile motor programming stage; $t(M)$ = the mean duration of the nonlabile motor programming stage; $\theta$ = a parameter that attenuates the effects of predictability.

the attended word is the target of a refixation, the "refixation" to the attended word to the left of fixation then becomes a regression back to the word to the left of fixation. Formally, this involves a reinterpretation of States 11 and 12 in Figure 7: We assumed (as shown in the figure) that the subscripts for $r$ and $R$ in States 11 and 12 are $n$ (the attended word) rather than $n + 1$ (the fixated word).

It was an interesting and unanticipated consequence of our model that it can predict a reasonable number of interword regressions (5% in some versions) even though it is driven by lexical access. Consistent with our elimination of the data from any sentence in the observed corpus of data when a reader made a regression back to a prior word, we also threw out any sentences in the simulation in which the simulation made a regression back to a prior word. We did not attempt to fit the interword regressions to the observed data because there are undoubtedly many other causes of interword regressions. However, in the General Discussion section we discuss some regression data we recently obtained that are nicely consistent with our modeling assumptions.

## E-Z Reader 4

E-Z Reader 3 was, in many ways, a success. With a small number of free parameters (six), it was able to predict mean first fixation, single fixation, and gaze duration and the mean proportions of single fixations, double fixations, and proportions of skips for the five frequency classes of words in our corpus. However, one aspect of the model remains troublesome: Parafoveal processing is assumed to occur with the same efficacy as foveal processing. This is clearly unreasonable because words are identified more rapidly in the fovea than in the parafovea (Rayner & Morrison, 1981). A manifestation of this problem is that the lexical-processing parameters appear to be unreasonably long. The intercept of $t(f) + t(lc)$ in E-Z Reader 3, $(1 + \Delta)*f_b$, is almost 400 ms, which would be the lexical access time of a word that is not predictable and has a frequency of one per million. If this time is considered as an average resulting from both parafoveal and foveal processing, it may not be that unreasonable, but it is implausibly long as an estimate of the mean foveal processing time for these words.

To counter this problem, we included in E-Z Reader 4 an additional parameter, $\epsilon$, that modulates the processing rates of $f$ and $l$ as a function of *eccentricity:* the distance of the word being processed from the word currently being fixated. The process durations for $f$ and $lc$ are adjusted by using

$$\text{duration}(x) = \text{duration}_0 \cdot \epsilon^x, \qquad (6)$$

where $\text{duration}_0$ is the normal duration of $f$ or $lc$ when word $n$ is fixated (i.e., word $n$ is in the fovea), and $x$ is an index of distance between the word currently being fixated and the word being processed (using the metric of word units). (When the word processed is the fixated word, $x$ is zero, when the word being processed is immediately to the right of the fixated word, $x$ is one, and so forth.) For example, if $\epsilon$ is 2 and if a word requires 200 ms to access when it is fixated, then the same word will take 400 ms to recognize if the prior word $(n - 1)$ is fixated and 800 ms if word $n - 2$ is fixated. In the model, these differences translate into different processing rates. That is, (using the same example), 100 ms of processing of word $n$ will cause lexical access to be half complete if it is fixated, one quarter complete if word $n - 1$ is fixated, and one eighth complete if word $n - 2$ is fixated.

Once again, best fitting parameters were found by iterative grid searches of the parameter space. Again, there are 30 (not completely independent) data points fit by one more parameter ($\epsilon$) than are fit in E-Z Reader 3. Thus, there were seven parameters, but as before, $t(m)$, $t(M)$, and $\theta$ were not really free parameters. The resulting parameter values and the outcome of a simulation based on these values are presented in Table 4. A quick glance at the predicted and observed values might lead one to conclude that the fit is about as good as in E-Z Reader 3. However, the global fit between predicted and observed values was substantially worse than in E-Z Reader 3, mainly because of the refixation data: E-Z Reader 4 seriously underpredicts the number of refixations on words. As with E-Z Reader 3, the predicted first-fixation and single-fixation durations for E-Z Reader 4 are both too long for the low-frequency words and too short for the high-frequency words. In addition, the gaze durations for the low-frequency words are a bit too short (a consequence of refixations being underpredicted).

What seems a bit more problematical, however, is that to achieve reasonably good fits, $\epsilon$ needed to be kept fairly small (1.30), which is probably a smaller drop-off in processing efficiency for lexical access than is reasonable. However, even with this relatively small decrement in processing as a function of eccentricity, the major problem with E-Z Reader 3 appears to be solved: The total lexical access time, $t(f_n) + t(lc_n)$, for unpredictable words with a frequency of one in a million is now a bit under 300 ms when those words are fixated.

## E-Z Reader 5

Even though E-Z Reader 4 seems basically reasonable and gives a reasonably good account of the data, we were somewhat concerned that the fit had deteriorated from that given by E-Z Reader 3. As a result, we made an adjustment of the eccentricity rule of E-Z Reader 4. Although this adds one more free parameter to the model, we feel that the modified rule is more psychologically reasonable than that of E-Z Reader 4. In E-Z Reader 5, the lexical completion process is slowed down more than the familiarity check process when a word is further from fixation. The motivation for this assumption is that the familiarity check process is a cruder one than lexical access. That is, partial identification of letters, letter features, or both, is likely to be sufficient for a familiarity judgment, whereas lexical access is likely to require full identification of most letters. Thus, the familiarity check process is likely to be degraded less by decreasing acuity than the lexical access process. As a result, we modified the eccentricity rule by introducing two $\epsilon$ parameters, $\epsilon_1$ and $\epsilon_2$: the first modifies $f_n$, and the second modifies $lc_n$. Each operates on its respective process by Equation 6. This adds one free parameter to those of E-Z Reader 4, making eight parameters (five free) fitting 30 data points.

As can be seen in Table 5, the overall fits of E-Z Reader 5 are substantially better than those of E-Z Reader 4 and are about the same as those of E-Z Reader 3. As with E-Z Reader 3, the

Table 4

*Observed and Predicted Values of Gaze Durations and Individual Fixations, Probability of Skipping, Making a Single Fixation, and Making Two Fixations for E-Z Reader 4 for Five Frequency Classes of Words*

| Frequency class | Mean frequency | Gaze duration | | First-fixation duration | | Single-fixation duration | | Probability of skipping | | Probability of making a single fixation | | Probability of making two fixations | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Observed | Predicted | Observed | Predicted | Observed | Predicted | Observed | Predicted | Observed | Predicted | Observed | Predicted |
| 1 | 3 | 293 | 282 | 248 | 259 | 265 | 273 | .10 | .09 | .68 | .80 | .20 | .11 |
| 2 | 45 | 272 | 266 | 234 | 257 | 249 | 263 | .13 | .16 | .70 | .79 | .16 | .04 |
| 3 | 347 | 256 | 250 | 228 | 245 | 243 | 248 | .22 | .25 | .68 | .73 | .10 | .02 |
| 4 | 4,889 | 234 | 221 | 223 | 220 | 235 | 220 | .55 | .47 | .44 | .53 | .02 | .00 |
| 5 | 40,700 | 214 | 199 | 208 | 198 | 216 | 199 | .67 | .68 | .32 | .31 | .01 | .00 |

*Note.* The best fitting parameters for E-Z Reader 4 are $f_b$ = 178 ms, $f_m$ = 12 ms; $\Delta$ = 0.65; $t(m)$ = 150 ms; $t(M)$ = 50 ms; $\epsilon$ = 1.30; and $\theta$ = 0.5. The within-word motor programming parameters, $t(r)$ and $t(R)$, are assumed to have the same values as $t(m)$ and $t(M)$, respectively. Root-mean-square deviation = 0.266; this value is comparable to that of E-Z Reader 3 but not to those of E-Z Readers 1 and 2. $f_b$ = the intercept of the mean completion time for the familiarity check stage; $f_m$ = the slope of the mean completion time for the familiarity check stage; $\Delta$ = the ratio of the length of the mean time for the lexical completion and the mean time for the familiarity check stage; $t(m)$ = the mean duration of the labile motor programming stage; $t(M)$ = the mean duration of the nonlabile motor programming stage; $\theta$ = a parameter that attenuates the effects of predictability; $\epsilon$ = a parameter that attenuates processing as a function of the eccentricity of the word from fixation.

first-fixation durations and single-fixation durations are a bit too long for the lower frequency words, and there is a slight nonmonotonicity in the predicted first-fixation durations (caused by the same underprediction of refixations as with E-Z Reader 3). However, other than that, it seems to be getting almost all of the other values about right. Moreover, all of the parameter values seem quite reasonable.

E-Z Readers 3–5 all do a reasonably good job of predicting the overall data; however, none is perfect. In some narrow modeling sense, E-Z Reader 3 might be viewed as the most satisfactory as it achieved almost as good a fit as E-Z Reader 5 by using two fewer free parameters. We feel, however, that E-Z Readers 4 and 5 are to be preferred because the $\epsilon$ parameters were added to make the model more psychologically reasonable by introducing decreasing visual acuity into the model, not to achieve a better fit. As a consequence, they also made the parameters for lexical processing time more psychologically reasonable.

As a result, E-Z Reader 5 appears to be our current ''state-of-the-art'' model. This is not to say that it could not be tinkered with; however, any more tinkering would entail adding more free parameters and be contrary to our minimalist intentions in this article. As a result, we now break our staff and end our modifications of E-Z Reader. (We discuss future directions later.) Instead, we turn to applying the E-Z Reader Models 3–5 to other phenomena.

## Additional Applications

The above-mentioned results indicate that E-Z Readers 3–5 predict the aggregate behavior of readers quite well. However, one can often be misled about a model's utility, psychological validity, or both, by simply evaluating aggregate properties (Hintzman, 1991). Consequently, we attempted to evaluate the models further by conducting some additional analyses.

### *Variability of Fixation Duration Measures*

An obvious starting point concerns the variability associated with the various fixation duration measures; E-Z Readers 3–5 predict the means of these measures about equally well, but do they also predict the appropriate pattern of variability? To answer that question, we constructed histograms of the observed first fixation and gaze durations in our corpus for each of the five frequency classes. The first-fixation duration histograms are presented in Figure 8, and the histograms for gaze durations are presented in Figure 9. In both cases, each datum represents the fixation duration on a single word for a single participant. A similar number of first fixation and gaze durations were then generated by simulating the data of 100 participants with E-Z Readers 3–5. The fixation durations predicted by the models are also presented in Figures 8 and 9 as histograms.

The relationship between the observed and predicted histograms of both the first-fixation durations and gaze durations is reasonably close. The absolute ranges and the shapes of the observed distributions are in reasonably close agreement to those predicted by the three models. The major discrepancy is that the observed distributions are somewhat less variable than the predicted distributions. This is most likely due to our as-

Table 5

*Observed and Predicted Values of Gaze Durations and Individual Fixations, Probability of Skipping, Making a Single Fixation, and Making Two Fixations for E-Z Reader 5 for Five Frequency Classes of Words*

| Frequency class | Mean frequency | Gaze duration |  | First-fixation duration |  | Single-fixation duration |  | Probability of skipping |  | Probability of making a single fixation |  | Probability of making two fixations |  |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|  |  | Observed | Predicted | Observed | Predicted | Observed | Predicted | Observed | Predicted | Observed | Predicted | Observed | Predicted |
| 1 | 3 | 293 | 291 | 248 | 251 | 265 | 274 | .10 | .09 | .68 | .73 | .20 | .17 |
| 2 | 45 | 272 | 271 | 234 | 253 | 249 | 263 | .13 | .16 | .70 | .76 | .16 | .07 |
| 3 | 347 | 256 | 257 | 228 | 246 | 243 | 252 | .22 | .27 | .68 | .68 | .10 | .04 |
| 4 | 4,889 | 234 | 226 | 223 | 223 | 235 | 224 | .55 | .49 | .44 | .50 | .02 | .01 |
| 5 | 40,700 | 214 | 211 | 208 | 210 | 216 | 210 | .67 | .68 | .32 | .32 | .01 | .00 |

*Note.* The best fitting parameters for E-Z Reader 5 are $f_b$ = 195 ms, $f_m$ = 17 ms; $\Delta$ = 0.70; $t(m)$ = 150 ms; $t(M)$ = 50 ms; $\epsilon_1$ = 1.25; $\epsilon_2$ = 1.75; and $\theta$ = 0.5. The within-word motor programming parameters, $t(r)$ and $t(R)$, are assumed to have the same values as $t(m)$ and $t(M)$, respectively. Root-mean-square deviation = 0.198; this value is comparable to those of E-Z Readers 3 and 4, but not to those of E-Z Readers 1 and 2. $f_b$ = the intercept of the mean completion time for the familiarity check stage; $f_m$ = the slope of the mean completion time for the familiarity check stage; $\Delta$ = the ratio of the length of the mean time for the lexical completion and the mean time for the familiarity check stage; $t(m)$ = the mean duration of the labile motor programming stage; $t(M)$ = the mean duration of the nonlabile motor programming stage; $\theta$ = a parameter that attenuates the effects of predictability; $\epsilon$ = a parameter that attenuates processing as a function of the eccentricity of the word from fixation.

sumption that the standard deviations of the processes were equal to 0.33 of their means. Other simulations with smaller standard deviations produced better agreement with the observed histograms; however, these models fit the refixation and skipping data a bit less well.

### *Frequency Effects*

Our word corpus was taken from the sentences used by Schilling et al. (in press) that were constructed to examine the effects of word frequency on selected target words in reading, pronunciation, and lexical-decision tasks. One obvious question is whether E-Z Readers 3–5 would predict the frequency effects reported by Schilling et al. on their target words.

As already mentioned, Schilling et al. (in press) had readers read 48 sentences, half containing target words that were high in frequency of occurrence (over 46 per million, $M = 141$ by Francis & Kučera, 1982), and half containing targets that were low in frequency (less than 4 per million with a mean of 2). The mean first fixation, single fixation, and gaze durations were calculated for both the high- and low-frequency words, as were the frequency effects—the differences between the high- and low-frequency means for each measure (see Table 6).

Next, E-Z Readers 3–5 were applied to the Schilling et al. (in press) sentences and the predicted first fixation, single fixation, and gaze durations tabulated for each of those words that had been designated as a target by Schilling et al. The mean fixation durations for both high- and low-frequency words were then calculated, as were the mean frequency effects. These means are presented in Table 6. Note that the parameters in the models were not altered to fit these data: The parameter values displayed in Tables 4–6 were used. As Table 6 indicates, the models all tend to underpredict the frequency effects a bit. This is especially true of the first-fixation measures and is consistent with the overprediction of first-fixation durations and single-fixation durations in Frequency Classes 2 and 3 that produced relatively small predicted differences between Frequency Class 1 and Frequency Classes 2 and 3 in Tables 4–6. Model 5 makes slightly better predictions than the other two models both in terms of the absolute values and the sizes of the frequency effects.

One problem with this simulation is that the high- and low-frequency words were in different sentences; thus we are not sure whether the underprediction of frequency effects is a fundamental problem with the model or some accidental property of the particular sentence frames. As a result, we attempted a second simulation of frequency effects in which we modeled what would have happened if a high- and low-frequency word had appeared in each of the 48 sentence frames. This was accomplished by substituting the mean values of the Schilling et al. (in press) high- and low-frequency target words (141 and 2 per million, respectively) into each of the 48 designated target word locations. In this simulation, we not only tried to predict frequency differences in the gaze duration on word $n$ (the target word), but we also looked at frequency effects on spillover measures such as the gaze durations on word $n + 1$.

As the name of the phenomenon suggests, spillover is an effect of processing a given word that occurs after fixating that word. There are several alternate attempts to operationalize this term. One is to measure the duration of the first fixation after word $n$ (regardless of whether it lands on word $n + 1$ or $n + 2$). Another is to use the gaze duration on the first word fixated after word $n$. Typically, decreasing the frequency of word $n$ not only lengthens the gaze duration on that word but also lengthens the duration of both the subsequent fixation and the gaze duration on word $n + 1$ (Rayner & Duffy, 1986; Rayner et al., 1989). E-Z Reader naturally explains spillover effects in the following way: As the frequency of word $n$ decreases, the mean $lc_n$ process duration increases, thereby reducing any preview benefit that is gained in processing word $n + 1$ while word $n$ is fixated (see Figure 6). Less preview benefit for word $n + 1$ means that more time is necessary to process it, thus increasing fixation durations on it. Moreover, there are also possible chain reactions, whereby stealing preview time from $n + 1$ will affect fixation times on word $n + 2$.

The predictions made by E-Z Readers 3–5 are similar: The (spillover) effect of the frequency of word $n$ on the gaze duration of word $n + 1$ is approximately one half of the frequency effect on the gaze duration on word $n$, but virtually no spillover is predicted beyond word $n + 1$. The absolute size of the predicted frequency effect on the gaze duration on word $n$ was about the same as in the prior simulation: 40, 30, and 35 ms for E-Z Readers 3–5, respectively. These values are a bit smaller than the observed values but not that discrepant from the value observed by Schilling (in press) et al. or other prior studies (Rayner & Duffy, 1986; Rayner et al., 1989). Similarly, the predicted increases in gaze duration on word $n + 1$ (15, 15, and 22 ms for E-Z Readers 3–5, respectively) were all a bit smaller than the observed values in prior studies (which range from 30 ms to 50 ms).

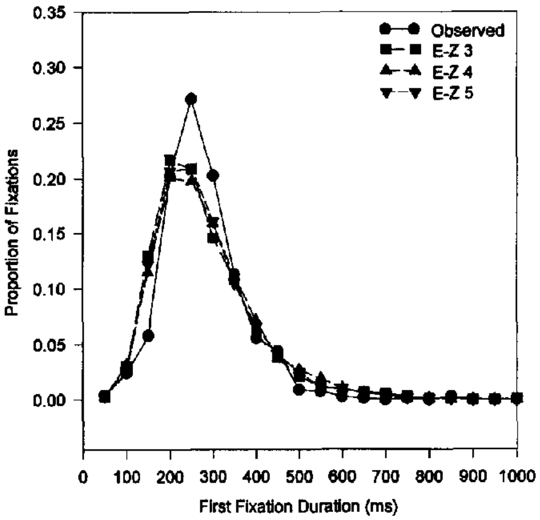
### *Sequential Effects*

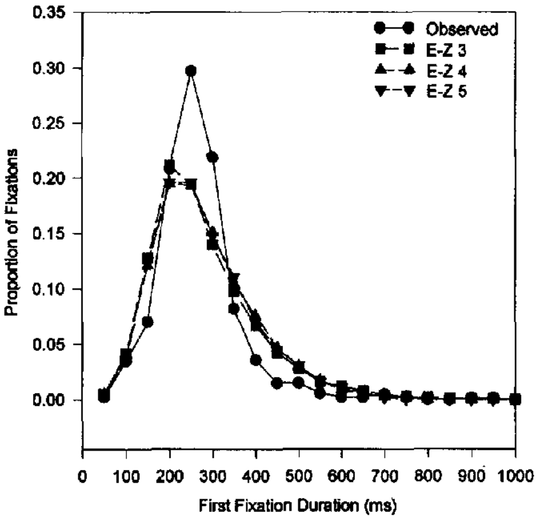
We also examined sequential effects in our corpus of data. This undertaking was not easy because there is not a large

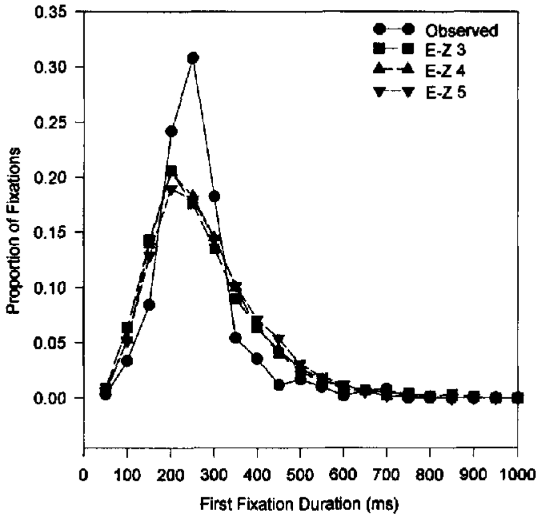
Table 6
*Observed and Predicted Frequency Effects (in Milliseconds) for Selected Target Words*

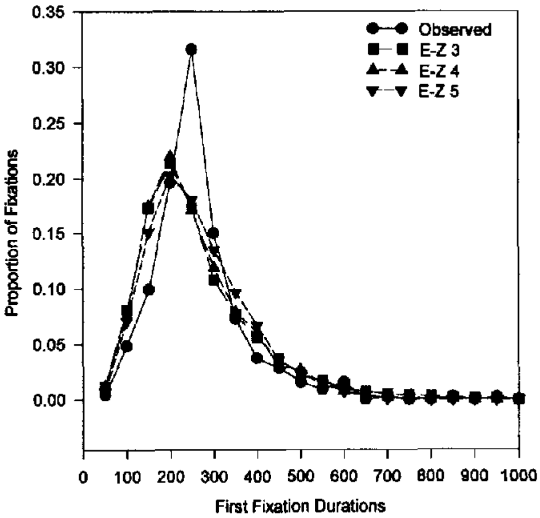
| | | Predicted value | | |
|---|---|---|---|---|
| Measure | Observed value[a] | EZ-Reader 3 | EZ-Reader 4 | EZ-Reader 5 |
| Gaze duration | | | | |
| High-frequency words | 248 | 267 | 262 | 260 |
| Low-frequency words | 298 | 294 | 296 | 298 |
| Difference | 50 | 27 | 34 | 38 |
| First-fixation duration | | | | |
| High-frequency words | 216 | 255 | 257 | 249 |
| Low-frequency words | 248 | 257 | 270 | 254 |
| Difference | 31 | 2 | 13 | 6 |
| Single-fixation duration | | | | |
| High-frequency words | 224 | 261 | 260 | 253 |
| Low-frequency words | 261 | 278 | 287 | 276 |
| Difference | 37 | 17 | 27 | 23 |

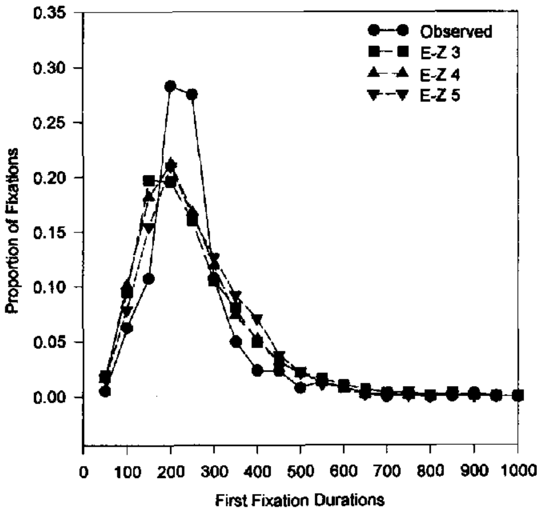
[a] The observed values reported here are not the overall values reported by Schilling, Rayner, & Chumbley (in press); our values were computed from the sentences used in the simulation (i.e., only from sentences that contained no interword regressions).


Frequency Class 1 First Fixation Durations


Frequency Class 2 First Fixation Durations


Frequency Class 3 First Fixation Durations


Frequency Class 4 First Fixation Durations


Frequency Class 5 First Fixation Durations

catalog of robust sequential effects found with eye movements during reading. For example, one prior attempt to examine the eye movement record from a large corpus of text (Rayner & McConkie, 1976) indicated that there was little or no sequential dependency for many global relationships (e.g., duration of a fixation given the duration of the prior fixation or the duration of a fixation given the length of the prior saccade).

There does appear to be one robust relationship with respect to skipping: Gaze durations on word *n* are longer when word $n + 1$ is skipped than when word $n + 1$ is fixated (Hogaboam, 1983; Pollatsek, Rayner, & Balota, 1986). This fixation duration cost from skipping the next word is qualitatively predicted by the general E-Z Reader model because skipping is the result of a later eye movement program canceling an earlier one. This can best be understood by examining Figure 7 (see State 5). (Assume that word $n - 1$ is currently fixated and *n* is being processed.) In these cases, $m_n$, the labile program to move the eyes to word *n*, completed before $f_n$, the familiarity check on word *n*. Consequently, the model made a transition to State 9, in which $M_n$ completed (moving the eyes to word *n*). Gaze durations on word $n - 1$ are thus equal to the minimum time required to complete $m_n$ and $M_n$ (i.e., States 5 and 9). However, in cases in which word *n* is skipped, $f_n$ completed before $m_n$ (State 5), causing the labile program to word *n* to be canceled by a new program to move the eyes to word $n + 1$ (State 4). The gaze duration on word $n - 1$ thus reflects the minimum time necessary to complete $f_n$, $m_{n+1}$, and $M_{n+1}$ (States 5, 4, and 9).

Across our corpus, the fixation duration cost of skipping the next word was 38 ms. As expected, E-Z Readers 3, 4, and 5 all predicted fixation duration costs, but the sizes of the effect predicted, 100, 145, and 173 ms, respectively, were clearly too large. We are not sure whether this is a serious problem for the models or not. The effects predicted by the models were based on relatively small samples of words because a word contributed to the mean effect only if the successive word was both fixated and skipped by some proportion of the statistical subjects. As a result, the predictions may be biased in some way, may not be particularly reliable, or both.

In examining this conditional relationship, we uncovered a second dependency in our data—one that has not previously been reported in the literature: Gaze durations on word *n* were longer when word $n - 1$ was skipped than when word $n - 1$ was fixated. That is, there is also a fixation duration cost for skipping the prior word. There is no mechanism in E-Z Reader 3 that easily accounts for this effect. However, E-Z Readers 4 and 5, because the $\epsilon$ parameters adjust the processing rates of $f_n$ and $lc_n$ as a function of distance from the fovea, predict that the preview benefit on word *n* should be greater when the prior fixation is closer. As the prior fixation to word *n* is closer when word $n - 1$ is fixated than when it is skipped, E-Z Readers 4 and 5 should predict some fixation duration cost.

The mean fixation duration cost for skipping the prior word across our corpus was 50 ms. As expected, E-Z Reader 3 did not predict such an effect but instead predicted a mean benefit of 28 ms, whereas E-Z Readers 4 and 5 predicted fixation costs for skipping the prior word of 30 and 52 ms, respectively. One caveat is again necessary, the predicted effects are based on relatively small sample sizes and therefore may vary in terms of absolute size. Nonetheless, the close agreement between the observed successive fixation benefit and the benefit that was predicted by E-Z Reader 5 is encouraging.

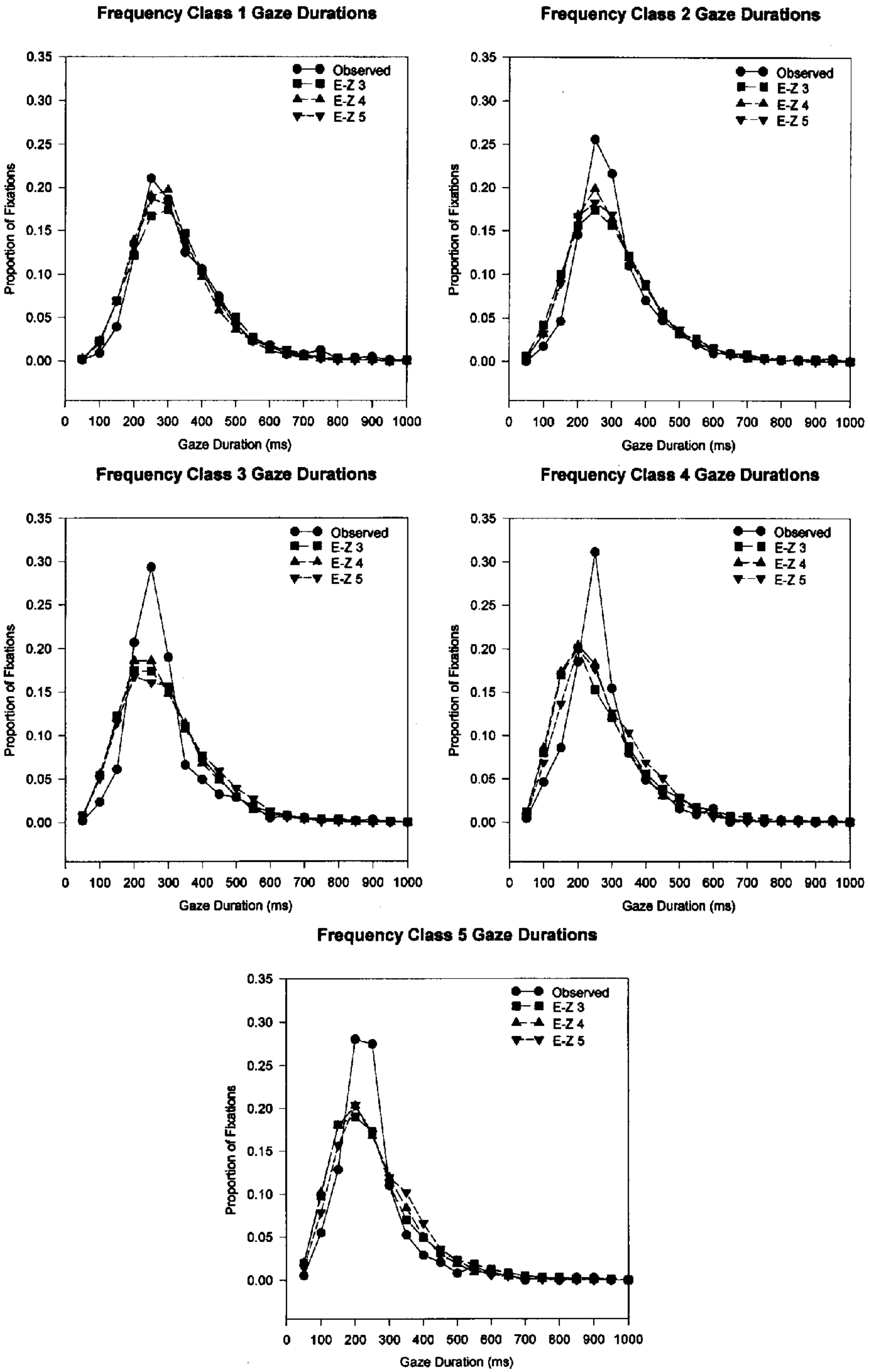
### *Preview Benefit*

Another aspect of reading that we simulated was the preview benefit from a parafoveal word. To do so, we simulated a hypothetical boundary experiment on the 48 target words used in the Schilling et al. (in press) experiment that were the focus of our frequency simulations. In this simulation, we compared two hypothetical conditions: one in which the target word appeared unchanged in the parafovea, and one in which the parafoveal preview was completely different from the target (such as random letters or a totally different word). We assumed in our simulation that processing essentially took an infinite amount of time in the parafovea when the preview was "wrong" and the normal amount of time (adjusted by an $\epsilon$ parameter, if appropriate) when the preview was "right." In fact, the predicted differences between the good and bad preview conditions (preview benefit) on gaze duration were predicted to be 88, 69, and 40 ms by E-Z Readers 3–5, respectively. The predictions for E-Z Readers 4 and 5 correspond quite well to observed values in the literature that range from around 40 to 60 ms. However, the preview benefit predicted by E-Z Reader 3 appears to be too large, consistent with the fact that it posits that processing in the parafovea is just as efficient as processing in the fovea.

### *Conclusions: E-Z Reader 3 Versus E-Z Reader 5*

Our chief motivation for going from E-Z Reader 3 to E-Z Reader 5 was to make the model more psychologically plausible by positing that lexical processing is slowed the further the word is from fixation (the $\epsilon$ parameters), even though the overall fits of the two models were approximately the same. There is certainly no question that an adequate model of reading has to assume some reduction in encoding efficiency the further the word being processed is from fixation; it remains an open question, however, as to whether the actual reduction in efficiency is consistent with the $\epsilon$ values that we fit in E-Z Reader 5. Classic psychophysical acuity functions are not much of a guide, as the dependent variable is accuracy (rather than time) and the stimuli are very different from words. Perhaps the most relevant study was by Rayner and Morrison (1981), which assessed loss of encoding efficiency for words as a function of eccentricity

*Figure 8 (opposite).* Observed and predicted frequency distributions of first-fixation durations. Separate distributions are presented for each frequency class. Each point represents the proportion of the first-fixation durations in a frequency class that occurred within a given 50-ms interval (e.g., the points above the abscissa labeled 200 represent the proportions of first-fixation durations between 150 and 200 ms that were observed in the sentence corpus and predicted by E-Z Readers 3–5).



Frequency Class 1 Gaze Durations

Frequency Class 2 Gaze Durations

Frequency Class 3 Gaze Durations

Frequency Class 4 Gaze Durations

Frequency Class 5 Gaze Durations

(while participants maintained fixation). They found that naming latency for high-frequency words increased by about 50 ms when the word began 1° to the right of fixation and by about 200 ms when the word began 3° to the right of fixation. In addition, accuracy decreased to about 95% at 1° and about 75% at 3°. In E-Z Reader 5, we assumed that processing efficiency for lexical access was decreased by 1.75 ($\epsilon_2$) for word $n + 1$ (which starts, on average, a bit over 1° from fixation) and by 3.06 ($\epsilon_2^2$) for word $n + 2$ (which starts, on average, about 3° from fixation). It is hard to align the parameter estimates with these data, especially because one does not know how to incorporate accuracy differences into an efficiency parameter, but they seem reasonably compatible. Perhaps we posited a bit too steep of a drop-off in efficiency for word $n + 1$ in the model; however, acuity might fall off more slowly in the Rayner and Morrison experiment than in reading because the word was presented in isolation.

In our subsidiary analyses, the models were also about equally good in predicting the data, although they had different strengths and weaknesses. First, E-Z Reader 5 predicted the fixation cost of skipping the prior word, whereas E-Z Reader 3 did not. This follows transparently from the fact that E-Z Reader 3 did not assume that words are harder to process the further they are from fixation; however, we are not sure why E-Z Reader 3 actually predicted a benefit rather than a cost. E-Z Reader 5 also predicted the size of the preview benefit better than did E-Z Reader 3 for similar reasons. In contrast, E-Z Reader 3 made better predictions on the fixation cost of skipping the succeeding word. The two models were about equally good at predicting the frequency and spillover effects, although E-Z Reader 5 did slightly better.

In spite of this apparent standoff, we believe that E-Z Reader 5 is our state-of-the-art model. We think that the eccentricity assumptions, rather than being free parameters to fit the data, are absolutely essential in a plausible account of reading. Moreover, it is clear that no minor change of E-Z Reader 3 will give an account of the fixation cost of skipping the prior word. In contrast, the failures of E-Z Reader 5 appear to be quantitative rather than qualitative and may be solved by tweaking some of the assumptions. However, we do not see much point in such an exercise. Instead, as we argue later, one of the goals of further modeling is to provide a more principled account of lexical access than our relatively crude functional equations. We think that the general framework of E-Z Reader 5 should be abandoned only if there is no reasonable account of lexical access that can reconcile its predictions with the observed data.

## General Discussion

We hope the success of the simulations of E-Z Reader Models 2–5 demonstrates that we have a plausible theory for how cognition drives the eye movement system in normal reading. Even E-Z Reader 5 is somewhat rough and incomplete, however. It is rough in two senses: (a) we have made oversimplified assumptions about processing, and (b) we have chiefly tried to predict differences among classes of words over a corpus of materials and have only made a limited number of finer grained analyses. It is also incomplete in two primary ways: (a) we have ignored the direct influence of supralexical comprehension processes on eye movements (such as syntactic garden path effects), and (b) we have ignored how the eye movement system selects a spatial target for an eye movement and how the movement is planned and executed.

In this section, we wish to discuss the road ahead. In doing so, we first discuss the omissions and whether it is productive in the near future to fill them in. We then discuss the key processes in the model in order to explore our central claims, and we indicate how some simplifying assumptions might be changed to better approximate the process of reading. We think that the modeling efforts in this article are a good framework that allows one both to see which directions of research are feasible in the foreseeable future and to be able to explore cognitive issues in reading on a firmer footing than before.

### *Omissions*

*Including parsing and text comprehension processes.* The most serious omission in the E-Z Reader model is a lack of influence of syntactic and text-based processing, except when they affect lexical access through predictability. In this case, it is not clear that there is much that can be done in the immediate future. As indicated earlier, models of parsing and on-line text processing are a long way from being able to make predictions at the level of time spent on individual words, let alone at the level of individual eye fixations. The current state-of-the-art models account for details of relatively local text comprehension processes and finesse questions of how sentences are parsed and how more global discourse structures are built (MacDonald et al., 1994), or sketch only certain components of a theory of more global processes (Frazier & Clifton, 1996; Frazier & Rayner, 1982), or are quantitative models at too crude a level to account for individual eye movements (Kintsch, 1988). Moreover, as we indicated earlier, even if one had a good quantitative model of text processing, modeling the effects of higher order processes on eye movements would be quite difficult; effects of disruption are complex, and the time course of many discourse processes is likely to be quite variable, and thus their effects would be difficult to localize. Thus, we see little likelihood in the near future that quantitative modeling of the effects of these processes on eye movements will help enlighten us much beyond what is already known at a qualitative level.

A somewhat deeper justification for omitting these processes

*Figure 9 (opposite).* Observed and predicted frequency distributions of gaze durations. Separate distributions are presented for each frequency class. Each point represents the proportion of gaze durations in a frequency class that occurred within a given 50-ms interval (e.g., the points above the abscissa labeled 200 represent the proportions of gaze durations between 150 and 200 ms that were observed in the sentence corpus and predicted by E-Z Readers 3–5).

in our model comes from a quasi-modular view of reading. That is, our view is that lexical access is largely performed by one module, and parsing and text comprehension are performed by other modules that are independent of lexical access in the sense of sharing only minimal processing resources. (We clearly do not hold a strict modular view, however, as we have allowed top-down influence on the speed of lexical access.) Moreover, because the higher order modules need input from the lexicon, their processing is ordinarily delayed relative to lexical access. Thus, it makes sense that these processes are usually too slow to be the usual signal to move forward and are better used as an occasional signal to stop lexical access and sort things out.

*Including more precise modeling of where the eye lands.* Our second omission is at the motor end. We have assumed that the motor programming system gets the signal "word $n + 1$" or "word $n + 2$" and somehow translates this either into a spatial target or into a distance for the eyes to move. Clearly, it would be desirable to understand the saccadic system more deeply by developing a model of how this translation process occurs. At our present level of understanding, however, it seems initially more feasible to grapple with the details of motor programming at a molar level. A good candidate for such a translation process is something like the general model proposed by McConkie et al. (1988), discussed earlier. It makes a specific claim about what the target for an initial fixation on a word is (the middle of a word) and induces general properties of the variability of eye movements from reading data that are consistent with what is known about motor movements.

We think a plausible next step would be to try to graft an eye-movement-location module (such as that proposed by McConkie et al., 1988) onto E-Z Reader 5 that would attempt to predict the exact locations of individual fixations. We would also need a quantitative model of where refixations go. Conceptually, this extension of E-Z Reader is not particularly difficult, but it clearly makes the modeling far more complex. The reason that it is not conceptually difficult to graft on such a module is that the basic focus of E-Z Reader is the attended word and not the fixated word. Earlier, we discussed one case in which there was a mismatch: when the reader had moved on to word $n + 1$ but was still attending to word $n$. In that case, we posited that the target for a refixation would be to the attended word; this would result in a regression back to the attended word if this saccadic program is not canceled. The way the model would handle saccadic overshoots and undershoots would be similar. It would be of interest to see whether "errors" (i.e., discrepancies between the word targeted for fixation and the word actually fixated) would substantially affect the fit of E-Z Reader 5 at the level of the predictions made in this article.

How such a composite model—E-Z Reader 5 combined with McConkie et al.'s (1988) spatial module—should be assessed at a finer grain of analysis, however, is not trivial. We think that it is likely that such a model would make good predictions about where readers land on words of a given length averaged over all words of that length in a corpus. This is because (a) McConkie's model is a good summary of a large set of data, and (b) there is evidence that the decision of exactly where to move is relatively independent of the decision of when to move (Rayner & McConkie, 1976; Rayner & Pollatsek, 1981). (The latter was the central concern of the E-Z Reader model.) However, it would be hard to claim that a good fit of landing positions was a big success of such a composite model, as the observed result would largely come from the properties of the spatial module, which has already been shown to fit landing data quite well. As a result, one would like to see the model pushed further by determining whether it could account for something like sequential patterns of eye fixations. The problem with predicting sequential patterns, as we mentioned earlier, is that there are few well-known sequential dependencies between individual fixation durations and saccade sizes, so that an interesting critical test of such a model is far from clear.

### Fleshing Out the Model

*Key assumptions.* There are two key assumptions of the model. The first is that the signal for an eye movement program is different from the one that shifts covert attention. The second is that eye movements can be programmed in parallel, and a later eye movement can cancel an earlier one. The latter assumption seems reasonable; it is a parsimonious explanation of saccadic eye movements in a simpler task (Becker & Jürgens, 1979), and there is little reason why such a mechanism would not operate in reading. Whether we have operationalized the eye movement location programming mechanism correctly in the present model by positing two discrete states, $m$ and $M$, however, is clearly an open question. Nonetheless, we think that our implementation may not be far from the truth. First of all, our positing the existence of the $M$ stage indicates that there is a "point of no return" where a saccadic program cannot be canceled. There is some indication that there may be essentially no such point for hand movements (e.g., Logan, 1982; Osman, Kornblum, & Meyer, 1990); however, the data of Becker and Jürgens (1979) and Abrams (1992) suggest otherwise for eye movements. In addition, Abrams's (1992) data support our assumption that there is neither any cost or benefit in reprogramming a saccade (i.e., $m$ and $M$ are the same for programs that cancel other saccades as for programs that initiate saccades). In his experiment, when participants needed to reprogram a saccade from a shorter to a longer saccade going in the same direction (i.e., as in our hypothesized reprogrammed saccades), the saccade latency was the same as for initially programmed saccades to the same location. (However, when saccades were reprogrammed to be shorter, the latency decreased.)

Our operationalization, which incorporates two discrete stages, is undoubtedly a bit oversimplified, however. Notably, it does not explain all of Becker and Jürgens's (1979) data. As indicated earlier, they observed three patterns of data when two targets were displayed in close succession: (a) two saccades are made (with a brief fixation intervening) when the (temporal) interval between the targets is relatively long, (b) a single saccade to the second target when the interval between the target is short, and (c) a single saccade to a "compromise" location when the interval between the targets is intermediate. Nothing in any of our models would account for the latter pattern of data. Clearly, the E-Z Reader model could be expanded to account for this phenomenon by dividing the $m$ state into two states: the first would be like the present $m$ state, and the later stage would be similar except that a later program does not merely cancel the first program but, instead, produces a compromise saccade (i.e., in which the target location is between the targets of the earlier and later programs). Because the current versions of the E-Z Reader model do not attempt to predict precise locations of fixations, such a model would be indistinguishable from the current model except that it would predict a few errors, mostly in the form of a refixation instead of a saccade to the next word. At the stage in which one is modeling the exact location of individual fixations (as sketched out in the previous section), however, it would be of interest to determine whether adding such a stage substantially improves the fit.

Our second key assumption, that there are separate signals for the eye movement program and the shift of covert attention, was primarily made because it seemed necessary to fit the reading data; it did not directly follow from any basic data on covert attention and eye movements. The physiological experiments that suggested a tight coupling (Wurtz et al., 1982) used an exogenous cue: Monkeys saw an abrupt change in the stimulus display that they were supposed to (a) respond to and (b) make an eye movement to. Wurtz et al. observed enhanced firing in the parietal lobe (interpreted as reflecting a shift of attention) followed by firing in other regions that are known to be involved in planning saccadic eye movements. However, we earlier argued that there was evidence that these processes could be decoupled. We also believe there is a good teleological argument for why covert attention and eye movements should be decoupled in continuous complex tasks such as reading. First, if the stimulus that is attended to needs to be encoded and incorporated into ongoing parsing and text comprehension processes, it seems functional not to move attention until this is done. Second, given that eye movement programs take a relatively long time to execute (about 150–200 ms), it is likely to be advantageous in many circumstances for the reader to be able to program an eye movement in advance of this attention shift. (In contrast, in experiments like those of Wurtz et al., in which attention is drawn by a sudden stimulus change, the situation is not conducive to such decoupling: There is no signal before the stimulus change that tells the eyes where to move.) If we are correct and such a decoupling of attention and eye movements is the norm for reading, we think it is unlikely to occur only in reading. If the organism is attempting to extract visual information rapidly and purposively in any task that involves a series of eye movements (e.g., extended visual search), it seems reasonable that a similar decoupling would happen.

The use of a separate signal to move the eyes is advantageous, of course, only if the signal to move the eyes is diagnostic that lexical access is likely to occur within 150–200 ms after the signal. Otherwise, frequent regressions would be necessary, and waiting for lexical access before deciding to move the eyes would then be a better strategy. This raises the possibility that the signal to move the eyes in the absence of an exogenous signal is not "hard-wired" but may depend on the visual task that is to be performed.[4]

We take this distinction between the two signals to be central to our model of reading. In our E-Z Reader models, we identified the first signal (to move the eyes) as a familiarity check stage and the second signal (to shift attention) as lexical access. In the actual modeling work, however, all that was assumed about these stages is that they were both linearly related to word frequency (which was confounded with word length) and that the slope of the lexical access function was steeper (see Figure 6). These assumptions allowed us to predict (a) frequency effects on the fixated word, (b) frequency effects on spillover, and (c) the foveal difficulty effect. Although the successful fits of the model suggest that such a distinction is valid, our assignment of psychological constructs to the two signals is clearly provisional.

We assumed (for the sake of parsimony) that there were linear fits of log frequency for the duration of both stages and that lexical completion time was a simple multiple of the familiarity check time. Second, we assumed in E-Z Readers 2c–5 that the familiarity check time and the lexical completion time were reduced by predictability, using two relatively simple multiplicative functions. We probably could have achieved better fits by relaxing either the linearity assumption or the multiplicative rules, but this seems like an empty exercise unless one has a deeper theory of the functional relationships one is positing. In addition, we assumed that frequency and predictability were the only two operative variables for either familiarity and lexical access. Obviously, a better characterization of the *f* and *lc* stages would arise from a more serious investigation of the variables that influence the duration of these stages and the functional relationship between these variables and the durations of *f* and *lc*.

*A better model of familiarity and lexical access.* We view the process of obtaining a better understanding of the familiarity check and lexical access processes as two pronged. One prong would be an empirical investigation of the effect of various variables on fixation times using the model as an analytical tool. The typical experiment would compare the reading of two sentences in which only a single target word was altered and would focus analysis on fixations on and near that target word, as in the analysis of frequency effects in spillover.

For example, we indicated in the introduction that there was some evidence that the presence of lexical neighbors may have an effect on lexical access even when word frequency was controlled. For example, there is evidence from the lexical-decision task that the presence of higher frequency neighbors (e.g., *space* is a higher frequency neighbor of *spice*) has an inhibitory effect on lexical access (Grainger & Jacobs, 1996). This inhibition is likely to show up mainly in the latter stages of lexical access (e.g., it would be predicted to occur in the second stage of models such as the activation–verification model of Paap et al., 1982). If so, E-Z Readers 3–5 predict that such inhibitory effects would not be observed on gaze durations on the target word (because that is controlled by the duration of *f*) but instead would be observed in spillover effects such as increased fixation duration on the word following the target word or regressions back to the target word. Note that E-Z Reader in fact predicts regressions back to the prior word when the duration of the *lc* stage is long (as when verification takes a long time). In fact,

---

[4] Indeed, some recent experiments in our lab (Rayner & Fischer, 1996; Rayner & Raney, 1996) have demonstrated that when participants are instructed to search through text for a target word, there is no frequency effect: participants look no longer at lower frequency words than at higher frequency words. Thus, the decision to move the eyes in such a search task is presumably made on the basis of a judgment like "Does the fixated word orthographically match the target word?"

this pattern of data was observed by Perea and Pollatsek (in press). In contrast, effects of increasing the number of neighbors (which is often confounded with the frequency of bigrams and trigrams) might be expected to have primarily a facilitative effect early in lexical access (as words with more neighbors would plausibly produce greater overall facilitation in the lexicon), and thus this variable would be expected to influence first-fixation durations and gaze durations. Results consistent with this prediction were obtained by Lima and Inhoff (1985).

Similarly, one could investigate the independent contributions of variables such as word length, presence of lower frequency lexical neighbors, and the effects of various types of predictability manipulations. Such a program of research would allow one to obtain a better functional relationship between various experimental variables and the durations of *f* and *lc*. This would provide a better characterization of what these two stages are and whether our preliminary identification of them is correct. A particularly interesting (and difficult) domain for an expanded model would be an investigation of lexical ambiguity effects. There are now several studies that indicate that lexical and phonological ambiguity affect fixation times and, furthermore, that these effects are modulated by prior sentence context (Binder & Morris, 1995; Duffy, Morris, & Rayner, 1988; Rayner & Duffy, 1986; Rayner & Frazier, 1989; Rayner, Pacht, & Duffy, 1994). As all of these studies have found that lexical ambiguity affects the gaze duration on the ambiguous word, our model would place at least part of these effects relatively early in lexical access (in the familiarity check stage). If that conclusion proves to be implausible given a reasonable theory of lexical access, then the interpretation of *f* and *lc* may have to be changed. For example, *f* might be the process of achieving lexical access in the narrow sense of the term—making contact with an orthographic, and/or phonological lexicon—whereas *lc* might be the process of extracting the meaning of the word.

The second prong of such a research program would be to develop a "deeper" model of lexical processing (e.g., similar to Seidenberg & McClelland, 1989) so that one could account for the influence of various variables in a process model rather than merely assuming some function relating processing times to various variables. We see these efforts as complementary with theory influencing experiment, and vice versa.

*Accounting for refixations.* Another likely oversimplification of our model is the assumption that refixations are always driven by the signal that a new fixation has begun. It seems quite plausible that at least some refixations may be driven by more cognitive signals such as the processing of subword units such as morphemes. Beauvillain (1996), for example, has shown that refixations on isolated words are in fact influenced by the morphemic structure of the word. If these findings extend to reading, our assumptions about refixations need to be complicated. One possibility, of course, is that refixations are sometimes driven by identifying a subword unit such as a morpheme.[5] Extending the model in this way, however, is not simple. First, one would want to ascertain with some certainty which subword units were capable of influencing eye movements. Second, it is not clear whether it is plausible that the process of identifying a morpheme (or other subword unit) is rapid enough in order to affect refixations in silent reading of text. Third is the question of the target for such intelligent refixations. For example, if the reader has processed the first morpheme and now wants to process the second, can the eye movement system plausibly know the spatial location of the second morpheme?

In summary, we think that modifications of the refixation process should be made only after it is clear that the simple mechanism posited in E-Z Readers 3–5 is wrong and one has a relatively clear idea of what is driving refixations. We should point out that our present dumb refixation mechanism does exhibit some emergent intelligent behavior. It predicts that the duration of the first of two fixations and the probability of a refixation are both influenced by variables that affect the duration of *f*. In addition, an assumption that the reader programs a refixation to a location near the furthest boundary of a word would allow the reader to refixate reasonably intelligently by fixating a different part of the word.

*Accounting for predictability effects.* Here again, our functional assumptions about how predictability influences lexical access are undoubtedly a crude approximation to how top-down and bottom-up influences interact. Our proportional assumptions capture, to some extent, the plausible idea that predictability will have a bigger impact (in absolute terms) when the bottom-up processing is taking a longer time. However, there are at least two ways in which the way our models use predictability is bound to be only a rough approximation to reality. Remember that our predictability norms were generated in the usual manner: by having readers predict the next word (a) with little time pressure and (b) in the absence of any stimulus input from the next word.

With respect to point a, in various studies of reading, quite different effects of predictability have been found, ranging from small effects on fixation durations on the target word and no effect on skipping rate (Zola, 1984) to modest differences in skipping rate (about 5%) combined with somewhat larger fixation effects (Balota et al., 1985) to relatively large differences in fixation duration and skipping rate (Ehrlich & Rayner, 1981). All studies generated predictability norms in roughly the same manner, and the predictable words were approximately equally predictable, but, intuitively, the likely efficacy of the predictability seems different in the experiments. In the Zola study, the predictability usually relies strongly on the word immediately before the target word, whereas in the Ehrlich and Rayner study, the target word was, in some sense, "set up" by several sentences before the sentence that contains the target word. (The predictability in the Balota et al. study was usually set up earlier than in the Zola study, but in the same sentence.) This indicates that a closer approximation to modeling the effects of predictability should probably consider the time at which the context becomes active. This, of course, is a difficult and contentious issue in the current literature as there are conflicting claims as to which aspects of the prior text are really doing the facilitation

---

[5] A recent experiment by Hyönä and Pollatsek (in press) indicated that the pattern of fixation durations and location of refixations on long compound words (averaging 12 letters) in Finnish were in fact affected by the morphemic components of these words. This, however, might not generalize to English because Finnish (like German) has a very productive system for forming compounds (e.g., "snowball fight field" would be a single word in Finnish).

(Hess, Foss, & Carroll, 1995; Morris, 1994; Schustack, Ehrlich, & Rayner, 1987).

Point b suggests a different shortcoming in our modeling. Our predictability norms indicate how much support top-down influences have in the absence of any stimulus information. However, reading is a more interactive process, and there may be many situations in which a word will not be predictable in the absence of any information but quite predictable given minimal information such as approximate word length and the first letter (Haber, Haber, & Furlin, 1983). Thus, predictability in the absence of any input is likely to be an imperfect measure of the actual strength of the top-down influences on word identification in reading.

Incorporating refinements into the model to address either point, of course, would require a serious model of how top-down and bottom-up processes interact. We see the incorporation of such a model into the present framework as an important, though difficult, direction of future research. However, we think the present version of the model can serve as a useful ''null hypothesis'' against which to test more detailed predictions.

### *Some Speculation on Learning to Read and Individual Differences*

A major finding in the psychology of reading is that lexical access processes are a major factor in both learning to read and in individual differences in reading (see Rayner & Pollatsek, 1989). An example of the latter is the study from which our materials were taken (Schilling et al., in press). In that study, naming and lexical-decision times for critical target words correlated with the gaze durations; moreover, individual differences on these word identification tasks predicted individual differences in gaze duration quite well. These facts, of course, can easily be handled by our model by adjusting the $f$ and $lc$ durations for individual readers.

Our model raises a subtler, and possibly more interesting, question: Is the linkage between cognition and eye movements that we have proposed a learned skill? Also, are there individual differences in this skill? Perhaps this issue can be framed by outlining a few simple alternatives. One is that reading is a relatively unique skill and that in most visual tasks, one programs both saccades and shifts of attention when lexical access (or its equivalent for object recognition) is complete. Thus, in reading, the ''cheat'' of programming an eye movement by a preliminary process such as $f$ must be learned. If so, then one might expect beginning readers' eye movements to be able to be modeled by a simpler model with only one stage (in which the duration of $lc$ is set equal to zero). On the other hand, the decoupling of a prior familiarity check process from lexical or object identification may either be hard wired into the visual motor system or may be learned by ages 5 or 6 in other visual tasks so that there is no visual skill to be learned in reading. In the latter view, the only skills to be learned in reading are visual word decoding and general linguistic skills.

More generally, this discussion raises the question of whether there are important individual differences among readers that are not simply attributable to linguistic processing. These differences could either be in how cognitive processes are hooked up to the eye movement system (as we just discussed) or in the durations of the motor programming times. These ideas can be tested by separately assessing individual differences in lexical processing (using word identification tasks) and motor programming skills (using simple eye movement tasks) and then using the model to predict how these individual differences would affect the pattern of eye movements in reading.

## Summary

As we have indicated, our model is not a complete model of reading. It is neither a deep model of cognition nor a deep model of motor programming. Instead, its focus is on how cognition ''talks to'' the eye movement system. As we indicated in the prior section, the model is provisional in many ways. We do feel, however, that it is an important first step. First, our successful simulations indicate that such a model is both possible and plausible. Second, its simplicity is a major virtue. It is interpretable enough that when it fails, it is reasonably clear why it fails, and the assessment of modifications (in some cases) is relatively straightforward. Furthermore, it is the only extant model of eye movement control that can simultaneously account for fixation time on a word and whether or not a word is skipped, as well as a number of other eye movement phenomena. Even though there is clearly more work to be done, we feel that the E-Z Reader model provides a useful framework for the work ahead.

## References

Abrams, R. A. (1992). Planning and producing saccadic eye movements. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 66–88). New York: Springer-Verlag.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 340–357.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17,* 364–390.

Beauvillain, C. (1996). The integration of morphological structure and whole-word form information during eye fixations on prefixed and suffixed words. *Journal of Memory and Language, 35,* 801–820.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double-step stimuli. *Vision Research, 19,* 967–983.

Binder, K. S., & Morris, R. K. (1995). Eye movements and lexical ambiguity resolution: Effects of prior encounter and discourse topic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1186–1196.

Blanchard, H. E. (1985). A comparison of some processing time measures based on eye movements. *Acta Psychologica, 58,* 1–15.

Bouma, H., & deVoogd, A. H. (1974). On the control of eye saccades in reading. *Vision Research, 14,* 273–284.

Carpenter, P. A., & Just, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275–307). New York: Academic Press.

Chumbley, J. I., & Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Memory & Cognition, 12,* 590–606.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27,* 429–446.

Ehrlich, K. (1983). Eye movements in pronoun assignment: A study of sentence integration. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 253–268). New York: Academic Press.

Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior, 22,* 75–87.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20,* 641–655.

Eich, J. M. (1985). Levels of processing, encoding specificity, elaboration, and CHARM. *Psychological Review, 92,* 1–38.

Fisher, D. L., & Goldstein, W. (1983). Stochastic PERT networks as models of cognition: Derivation of the mean, variance, and distribution of reaction time using order-of-processing (OP) diagrams. *Journal of Mathematical Psychology, 27,* 121–151.

Fodor, J. A. (1983). *Modularity of mind.* Cambridge, MA: MIT Press.

Francis, W. N., & Kuĉera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Frazier, L., & Clifton, C. (1996). *Construal.* Cambridge, MA: MIT Press.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14,* 178–210.

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

Haber, L. R., Haber, R. N., & Furlin, K. R. (1983). Word length and word shape as sources of information in reading. *Reading Research Quarterly, 28,* 165–189.

Henderson, J. M., Dixon, P., Petersen, A., Twilley, L. C., & Ferreira, F. (1995). Evidence for the use of phonological representations during transsaccadic word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 82–97.

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

Hess, D. J., Foss, D. J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General, 124,* 62–82.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace model. *Psychological Review, 95,* 528–551.

Hintzman, D. L. (1991). Why are formal models useful in psychology? In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 175–194). Hillsdale, NJ: Erlbaum.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics, 57,* 787–795.

Hogaboam, T. W. (1983). Reading patterns in eye movement data. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 309–332). New York: Academic Press.

Huey, E. B. (1908). *The psychology and pedagogy of reading.* New York: Macmillan.

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review, 96,* 208–233.

Hyönä, J., & Pollatsek, A. (in press). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance.*

Inhoff, A. W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behavior, 23,* 612–624.

Inhoff, A. W., Pollatsek, A., Posner, M. I., & Rayner, K. (1989). Covert attention and eye movements during reading. *Quarterly Journal of Experimental Psychology, 41A,* 63–89.

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

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329–354.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension.* Boston: Allyn & Bacon.

Kennison, S. M., & Clifton, C. (1995). Determinants of parafoveal preview benefit in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 68–81.

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction–integration model. *Psychological Review, 95,* 163–182.

Klein, R. M., & Pontrefact, A. (1994). Does oculomotor readiness mediate cognitive control of visual attention? Revisited! In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 333–350). Cambridge, MA: MIT Press.

Kolers, P. A. (1976). Buswell's discoveries. In R. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes* (pp. 373–396). Hillsdale, NJ: Erlbaum.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35,* 1897–1916.

Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997). Mr. Chips: An ideal-observer model of reading. *Psychological Review, 104,* 524–553.

Lima, S. D., & Inhoff, A. W. (1985). Lexical access during eye fixations in reading: Effects of word-initial letter sequences. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 272–285.

Logan, G. D. (1982). On the ability to inhibit complex movement: A stop-signal study of typewriting. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 778–792.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101,* 676–703.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial fixations on words. *Vision Research, 28,* 1107–1118.

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

McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics, 25,* 221–224.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99,* 440–466.

Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 92–103.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 667–682.

Murdock, B. B. (1993). TODAM2: A model for the storage and retrieval of item, associative, and serial-order information. *Psychological Review, 100,* 183–203.

O'Regan, J. K. (1979). Eye guidance in reading: Evidence for the linguistic control hypothesis. *Perception & Psychophysics, 25,* 501–509.

O'Regan, J. K. (1981). The convenient viewing hypothesis. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 289–298). Hillsdale, NJ: Erlbaum.

O'Regan, J. K. (1990). Eye movements and reading. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 395–453). Amsterdam: Elsevier.

O'Regan, J. K. (1992). Optimal viewing position in words and the strategy–tactics theory of eye movements in reading. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 333–354). New York: Springer-Verlag.

Osman, A., Kornblum, S., & Meyer, D. E. (1990). Does motor programming necessitate response execution? *Journal of Experimental Psychology: Human Perception and Performance, 16,* 183–198.

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

Perea, M., & Pollatsek, A. (in press). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance.*

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 148–162.

Pollatsek, A., & Rayner, K. (1990). Eye movements and lexical access in reading. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 143–163). Hillsdale, NJ: Erlbaum.

Pollatsek, A., Rayner, K., & Balota, D. A. (1986). Inferences about eye movement control from the perceptual span in reading. *Perception & Psychophysics, 40,* 123–130.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88,* 93–134.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Ratcliff, R. (1988). Continuous versus discrete information processing: Modelling accumulation of partial information. *Psychological Review, 95,* 238–255.

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

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85,* 618–660.

Rayner, K. (1979). Eye guidance in reading: Fixation locations in words. *Perception, 8,* 21–30.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41,* 211–236.

Rayner, K. (1995). Eye movements and cognitive processes in reading, visual search, and scene perception. In J. M. Findlay, R. W. Kentridge, & R. Walker (Eds.), *Eye movement research: Mechanisms, processes and applications* (pp. 3–22). Amsterdam: North-Holland.

Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology, 40,* 473–483.

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

Rayner, K., & Duffy, S. A. (1988). On-line comprehension processes and eye movements during reading. In M. Daneman, G. E. MacKinnon, & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (Vol. 6, pp. 13–66). Hillsdale, NJ: Erlbaum.

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics, 58,* 734–747.

Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 779–790.

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

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology, 12,* 206–226.

Rayner, K., & Morris, R. K. (1992). Eye movement control in reading: Evidence against semantic preprocessing. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 163–172.

Rayner, K., & Morrison, R. E. (1981). Eye movements and identifying words in parafoveal vision. *Bulletin of the Psychonomic Society, 17,* 135–138.

Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of prior encounter and discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory and Language, 33,* 527–544.

Rayner, K., & Pollatsek, A. (1981). Eye movement control in reading: Evidence for direct control. *Quarterly Journal of Experimental Psychology, 33A,* 351–373.

Rayner, K., & Pollatsek, A. (1987). Eye movements in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 327–362). Hillsdale, NJ: Erlbaum.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice Hall.

Rayner, K., & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin and Review, 3,* 245–248.

Rayner, K., & Sereno, S. C. (1994). Eye movements in reading: Psycholinguistic studies. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 57–81). San Diego, CA: Academic Press.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes [Special issue]. *Language and Cognitive Processes, 4,* 21–50.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1188–1200.

Rayner, K., Slowiaczek, M. L., Clifton, C., & Bertera, J. H. (1983). Latency of sequential eye movements: Implications for reading. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 912–922.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin and Review, 3,* 504–509.

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (in press). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition.*

Schustack, M. W., Ehrlich, S. F., & Rayner, K. (1987). The complexity of contextual information in reading: Local and global influences. *Journal of Memory and Language, 26,* 322–340.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Stelmach, L. B., Campsall, J. M., & Herdman, C. M. (1997). Attentional and ocular movements. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 823–844.

Suppes, P. (1990). Eye-movement models for arithmetic and reading performance. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 455–478). Amsterdam: Elsevier.

Thibadeau, R., Just, M. A., & Carpenter, P. A. (1982). A model of the time course and content of human reading. *Cognitive Science, 6,* 101–155.

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review, 76,* 282–299.

Vitu, F. (1991). The influence of parafoveal processing and linguistic control on the optimal landing position effect. *Perception & Psychophysics, 50,* 58–75.

Wolverton, G. S., & Zola, D. A. (1983). The temporal characteristics of visual information extraction during reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 41–51). New York: Academic Press.

Wurtz, R. H., Goldberg, M. E., & Robinson, D. L. (1982). Brain mechanisms of visual attention. *Scientific American, 246,* 124–135.

Zola, D. (1984). The effect of redundancy on the perception of words in reading. *Perception & Psychophysics, 36,* 277–284.

# Appendix

## Some Details of the Simulations

The programs for the E-Z Reader models are written in C++ and can be run on any IBM-compatible PC. (They are available from Erik D. Reichle on request.) There are two core programs: The first implements E-Z Readers 1 and 2 and corresponds to the state space of the Order-of-Processing or OP diagram that is presented in Figure 5. The second program implements E-Z Readers 3–5 and corresponds to the state space of the OP diagram in Figure 7. Because the programs are very similar, the following discussion is limited to the second program and, more specifically, to (a) how E-Z Reader 5 operates, (b) how the model's performance was evaluated, and (c) how the best fitting parameter values were found.

To provide a sense for how E-Z Reader 5 works, consider what happens when the model "reads" a single sentence. The familiarity check, $t(f_n)$, and lexical completion, $t(lc_n)$, process durations are first calculated (see Equations 5a and 5b, respectively) for each word in the sentence by using: (a) each word's frequency of occurrence (as tabulated in the norms of Francis & Kučera, 1982), (b) each word's mean predictability (as obtained in the predictability norming study), and (c) parameter values obtained by iterative grid searches of the model's parameter space (this is discussed later). The values of $t(f_n)$ and $t(lc_n)$ are then used as the means of gamma distributions, from which the actual process durations, $T(f_n)$ and $T(lc_n)$, are sampled. In a similar manner, the labile, $t(m_n)$, and nonlabile, $t(M_n)$, saccadic programming parameters are used as the means of gamma distributions, from which the predicted process durations of both the interword—$T(m_n)$ and $T(M_n)$—and intraword—$T(r_n)$ and $T(R_n)$—saccadic programming times are sampled. Again, $t(m_n)$ and $t(M_n)$ are free parameters that could have, in principle, been estimated by searching the parameter space; however, in our simulations, they were fixed at 150 ms and 50 ms, respectively.

For each sentence, 1,000 statistical subjects were run, with the sampling from the gamma distributions being done independently for each statistical subject. Gamma distributions were used for convenience and because response latencies are often approximately distributed as gamma distributions. In our simulations, component gamma distributions were constructed by convolving nine exponential distributions, which means that the histograms of the gamma distributions are unimodal and approximately normal, but with a positive skew.

After the process durations for a particular sentence are sampled, the model starts to read the sentence from State 1 (see Figure 7), where the first word of the sentence is fixated. First, the durations of the two ongoing processes, the familiarity check and labile programming component of an intraword saccade (i.e., $T(f_1)$ and $T(r_1)$, respectively), are compared. Then, the process with the shortest duration "completes": Its duration is subtracted from the duration of the other process to simulate the amount of processing that occurred in the unfinished process during the completion of the sorter process. Finally, the model makes the appropriate state transition.

For example, if $T(f_1)$ equals 150 ms and $T(r_1)$ equals 100 ms, then the labile programming for an intraword saccade completes, and the process duration of the familiarity check is decremented by 100 ms (i.e., $T(f_1)$ now equals 50 ms). The model then moves to State 2, where the familiarity check continues, and the nonlabile programming component for an intraword saccade begins. However, if $T(f_1)$ equals 100 ms and $T(r_1)$ equals 150 ms, then the familiarity check on the first word completes, the labile program to make an intraword saccade is canceled, and the model moves to State 4. In State 4, the completion of lexical access and the labile programming of an interword saccade (to word 2) begin.

The cycle of (a) comparing the durations of the active processes, (b) completing the process with the shortest duration, and (c) making the appropriate state transition continues until lexical access of the last word in the sentence has been completed. Two points need to mentioned with respect to saccades: First, saccades cannot be programmed to locations beyond the last word. Second, saccades require time (25 ms) to execute; this time is subtracted from any processes that happen to be ongoing when the saccade is made.

On each pass through the sentence (there is one pass per statistical subject), the fixation durations and number of fixations are tabulated for each word. These tabulated values are then used to calculate the mean first-fixation duration, single-fixation duration, gaze duration, and probabilities of skipping, fixating once, and fixating twice (i.e., refixating) for each word in the sentence (excluding the first and last words of each sentence). Finally, the means for the individual words are used to calculate the means for the five frequency classes (e.g., see Table 6).

The model's overall performance was measured by using the root mean square of the normalized difference scores (errors) between the observed and predicted means of the five frequency classes for each of the dependent measures. The normalization process allowed the errors to be evaluated on a common scale (i.e., milliseconds and probabilities were converted to unitless scores). The normalization process that we used was to square the difference between the observed and predicted values and then to divide this difference by the standard deviation of the observed values. For example, for Frequency Class 3, the error score would be the squared difference between the observed and predicted first-fixation durations divided by the standard deviation of the observed first-fixation durations for that frequency class. For the probabilities, the standard deviation was just the square root of $p(1 - p)$, where $p$ was the observed probability for a given frequency class. The single-fixation duration and refixation means were not included in this measure because their values are largely redundant with the other measures.

In all of the simulations reported in this article, the best fitting parameters were found by completing iterative grid searches of the parameter space; in other words, the value of each parameter was systematically varied orthogonal to the other parameters. Each combination of parameter values was evaluated by using the goodness-of-fit measure that was discussed in the previous paragraph. In doing these grid searches, the estimated values of $f_b$, $f_m$, and $\Delta$ were subject to a plausibility constraint: The values of these parameters had to be such that lexical access occurred within a particular time window (i.e., between 100–300 ms). The values of $t(m_n)$, $t(M_n)$, $t(r_n)$, and $t(R_n)$ were also subject to a plausibility constraint in that it typically takes 200–250 ms to program and execute a saccade. Consequently, these saccadic programming parameters and the time required to execute a saccade were held constant. The values of the remaining parameters (i.e., $\theta$, $\epsilon_1$, and $\epsilon_2$) were chosen somewhat arbitrarily: The value of $\theta$ was selected so as to provide an intermediate effect of predictability on the time required to perform a familiarity check, whereas $\epsilon_1$ and $\epsilon_2$ were selected so as to modulate the parafoveal processing by an amount that is more or less consistent with results obtained from word recognition studies (Rayner & Morrison, 1981) and with our intuitions about the difference between the familiarity check and full lexical access processes (see text).

Received August 5, 1996
Revision received June 21, 1997
Accepted June 23, 1997 ■