identification process within E-Z Reader 7, and, if anything, understates the capabilities of the model. Here I discuss what I believe is a minimized strength of E-Z Reader 7, namely, that the model provides a natural framework for explaining how high-level cognitive processes influence eye-movement control. I begin by addressing the question: What is a high-level process? I then discuss how high-level processes might be explained within E-Z Reader 7.

What constitutes a high-level process is partly an issue of definition. High-level processes may be defined as processing based on information not contained within the lexical representation of a word. This is similar to the description of top-down processes in models of word processing (e.g., McClelland & Rummelhart 1981). Note that E-Z Reader 7 includes predictability as an element of word identification. Processes based on word predictability qualify as high-level processing in the sense that predictability accumulates across words and sentences. Monitoring predictability to enhance word identification appears to be a normal component of reading. I suspect that Reichle et al. would agree given that the predictability is a component of both the L_1 and L_2 stages of word identification in E-Z Reader 7. Based on the above definition, predictability represents an example of high-level information that directly influences eye-movement control during normal reading. Therefore, the conclusion that high-level processes influence eye movements only when "something is wrong" seems inconsistent (and unnecessary) with the structure of the model.

High-level processes may also be defined as based on later-occurring semantic processing, thereby excluding early-occurring visual processing. Determining whether high-level processes occur too late in the processing stream to influence eye movements becomes a critical issue. Although there is evidence for many high-level processes being slow, in the sense that they occur in late stages of word processing or even after a problematic word has been read (e.g., the garden path sentences used by Frazier & Rayner [1982]), there is growing evidence that high-level processes can influence early stages of word identification (Morris 1994; Sereno 1995; Wiley & Rayner 2000). This evidence again calls into question the necessity of the claim that high-level processes do not influence eye movements unless something goes wrong. My purpose here is not to resolve this definition issue but to suggest that Reichle et al. might be constraining their model unnecessarily. An untapped strength of E-Z Reader 7 is that it provides a transparent (i.e., definable) architecture for explaining how high-level processes influence eye movements (at least the decision of when to move the eyes). This contrasts with other models in which the architecture is not always transparent (such as, how hidden layers operate in connectionist models). Thus, my criticism of E-Z Reader 7 is that the architecture of the model is not fully utilized. Below I provide two examples of how the model may be applied.

Including a two-stage word identification system provides a natural architecture for separating the locus of low- and high-level processing influences on eye movements. Recent studies from my own lab support this conclusion. In one study (Raney et al. 2000), I recorded subjects' eye movements while they read a text once and then read either the same text a second time or a paraphrased version of the text. Paraphrases were created by replacing words with synonyms. For identically repeated target words, both first fixation duration and gaze duration were reduced during the second reading. For synonyms, only gaze duration was reduced during the second reading. For synonyms, early-occurring orthographic processing was not facilitated whereas later-occurring semantic processing was facilitated. This makes sense because no orthographic repetition occurs for synonyms, but semantic repetition does occur. In terms of the E-Z Reader model, the results for synonyms reflect no facilitation of the L_1 stage of word identification, but facilitation of the L_2 stage (a reduction in gaze duration, which reflects more later-occurring processes than first fixation duration).

In a similar study (Raney et al. 1996), fluent and nonfluent bilinguals read a text in one language and then reread either the same text or a translation. Embedded in the texts were cognate and noncognate target words. For fluent bilinguals, fixation durations were equivalent for cognates and noncognates during the second reading. For nonfluent bilinguals, fixation durations were shorter for cognates than for noncognates during the second reading. The low-level benefit of repeating the orthographic form (for cognates) interacted with high-level processes associated with comprehension level (fluency). These findings also map onto the model. Specifically, only semantic processes influenced fixation duration for fluent bilinguals (L_2), but both orthographic (L_1) and semantic processes (L_2) influenced fixation duration for nonfluent bilinguals.

To summarize, Reichle et al. describe E-Z Reader 7 as a processing model of eye-movement control. One constraint they impose on the model is that high-level cognitive processes do not influence eye movements unless normal reading processes are disturbed. This constraint makes the model conservative regarding what forms of information are allowed to influence eye movements. My own view is that there is enough evidence that high-level processes influence early and late stages of eye movements, for models of eye-movement control to incorporate these processes. E-Z Reader 7 provides a sensible architecture for explaining how high-level processes influence eye movements. Constraining the impact of high-level processes reduces the explanatory power of the model.

ACKNOWLEDGMENTS
I wish to thank John Murray and Bette Bottoms for their helpful comments on drafts of this commentary.

Methodologies for comparing complex computational models of eye-movement control in reading: Just fitting the data is not enough

Ronan Reilly[a] and Ralph Radach[b]

[a]*Department of Computer Science, NUI Maynooth, Maynooth, Co. Kildare, Ireland;* [b]*Institute of Psychology, Technical University of Aachen, 52064 Aachen, Germany.* **ronan.reilly@may.ie** **http://cortex.cs.may.ie**
ralph.radach@psych.rwth-aachen.de
http://www.psycho.rwth-aachen.de

Abstract: As the number of computational models of eye-movement control in reading increases, so too will their coverage and complexity. This will make their comparison and testing increasingly challenging. We argue here that there is a need to develop a methodology for constructing and evaluating such models, and outline aspects of a possible methodology.

In recent years, research on eye movements in reading has made substantial progress. A key new development in the field is the emergence of computational models of eye-movement control during reading. The target article is a timely evaluation of this branch of reading research. The modeling principles and algorithms that different computational models embody reflect the theoretical viewpoints of their authors. In the case of E-Z Reader, sequential lexical processing is proposed as the obligatory trigger for the generation of all eye movements made in normal reading.

In contrast, Reilly and O'Regan (1998), following the theoretical framework developed by O'Regan (1990), demonstrated that a good account for the positioning of fixations in reading can be achieved by using a set of rather simple oculomotor heuristics. We believe that both of these positions have their merits and can account for important aspects of eye behaviour during reading. On the other hand, both approaches also have serious limitations. Therefore, the question of interest is not whether eye movements are determined by visuomotor factors or by linguistic processing,

https://doi.org/10.1017/S0140525X03450103 Published online by Cambridge University Press

but to what degree these two factors are involved and how they interact.

Recent modeling work by Engbert and Kliegl (2001) and Reilly and Radach (2003a) can be seen as attempts at reconciling these two views of the reading process. As things stand, however, all current computational models of eye-movement control in reading deal with the process at a relatively shallow level. As pointed out in the target article, one of the real challenges for the next generation of models will be to broaden their coverage to include cognitive and linguistic factors. Unfortunately, as models become more complex, their comparison will become more problematic. The main point of this commentary, therefore, is to make the case for the development of a methodology for the comparison of computational models of eye-movement control in reading.

Our methodological proposals fall under three headings: (1) the facilitation of the comparison of the structural and functional assumptions of competing models; (2) the grounding of models in the neuroscience of vision and language; and (3) the establishment of data sets for model comparison and benchmarking. With regard to the comparison of the structure and function of models, this could be facilitated by using a common implementation framework comprising a set of reusable software components (Schmidt & Fayad 1997). In software engineering terms, a framework is a reusable, "semicomplete" application that can be specialised to produce particular applications or, in this case, particular models. The components would need to be fine-grained enough to accommodate the range of model types and model instances described in the target article. If one could develop an acceptable and widely adopted modeling framework, it would be possible to establish a common basis on which to implement a variety of models. This would make the models more directly comparable not only in terms of their ability to account for data, but also in terms of their underlying theoretical assumptions. The modeling environment could provide a semi-formal language with which a model's structures and processes function could both be unambiguously articulated. This would aid the task of both designing the models and communicating the design to other researchers.

Functionalist computational models, of which E-Z Reader is an excellent example, are inherently underdetermined in terms of their relationship to the brain mechanisms that underlie them. For example, one could envisage a family of E-Z Reader-like models with quite different combinations of parameters and/or parameter values that would be capable of providing an equally good fit to the empirical data (e.g., Engbert & Kliegl 2001). One way to reduce this lack of determinism is to invoke a criterion of biological plausibility when comparing models. We agree with the authors that there is an increasingly rich set of data emerging from the field of cognitive neuroscience which could be used to augment the traditional behavioural sources of constraint on computational models. We believe that models of reading can no longer avoid scrutiny from this perspective. Another, not unrelated, factor in assessing competing models is to take due account of the evolutionary context in which our vision system evolved. Because it evolved for purposes quite different from reading, we need to beware of too-easy recourse to arguments of parsimony, particularly when they are couched solely in terms of the reading process itself. A model with the minimum of modifiable parameters may be parsimonious on its own terms but fail the test of biological realism when compared with, say, a model that comprises an artificial neural network with many hundreds of adjustable parameters. While evolution is parsimonious in the large, when we look at brain subsystems in isolation, such as those involved in reading, we need to be careful how we wield Occam's razor.

Finally, the issue of appropriate data sets with which to test and compare computational models of eye-movement control needs closer attention than has been given to date. The Schilling et al. (1998) data set used to parameterise and test E-Z Reader and several other models discussed in the target article is not particularly extensive. A good case can be made for establishing a range of publicly accessible data sets against which any proposed model can be tested. This would be similar to what has been done, for example, in machine learning, in data mining, and, most notably, in the field of language acquisition (MacWhinney 1995). Furthermore, the corpus of benchmark data should be extended to include a variety of languages, alphabets, and scripts. The more successful models will be those that can readily generalise beyond just one language and one writing system.

Eye-movement control in reading: Models and predictions

Eyal M. Reingold

Department of Psychology, University of Toronto, Toronto, Ontario M5S 3G3, Canada. **reingold@psych.utoronto.ca**
http://psych.utoronto.ca/~reingold/

Abstract: It is argued here that a critical prediction of the E-Z Reader model is that experimental manipulations that disrupt early encoding of visual and orthographic features of the fixated word without affecting subsequent lexical processing should influence the processing difficulty of the fixated word without producing any processing effect on the next word. This prediction is explained and illustrated.

In the target article, Reichle et al. introduce a comprehensive framework for evaluating models of eye-movement control during reading. The authors also provide an updated version of the E-Z Reader model (Reichle et al. 1998; 1999) and argue that the qualitative and quantitative predictions that follow from this model closely match empirical findings concerning a wide range of reading phenomena. Consequently, they contend that the new version of their model, E-Z Reader 7, constitutes the best available computational framework for modeling eye-movement control during reading. The purpose of this commentary is to derive and illustrate a critical and as yet untested prediction that is unique to the E-Z Reader model. The proposed empirical strategy is illustrated in Figure 1 and will be outlined below.

As illustrated in Figure 1, three core aspects of the E-Z Reader model are central to the present proposal: (1) The E-Z Reader model introduces a distinction between two stages of lexical processing: an early lexical processing stage corresponding to the extraction and identification of the orthographic form of the word (L_1), and a late stage involving access to the phonological and semantic forms (L_2); (2) the programming of a saccade to the next word (word_{n+1}) is initiated following the completion of L_1 of word_n; and (3) parafoveal preview of word_{n+1} begins following the completion of L_2 of word_n. Therefore, according to the E-Z Reader model, variation in the duration of L_2 of word_n – $t(L_2)$ – critically determines the duration of parafoveal preview of word_{n+1}.

As shown in Figure 1, the duration of the parafoveal preview of word_{n+1} equals the duration of the interval between the initiation and execution of the saccade to word_{n+1} minus $t(L_2)$ of word_n. In the current implementation of the E-Z Reader model, two variables, word frequency and contextual constraint or predictability, influence the duration L_2 of word_n and consequently should also control the duration of the parafoveal preview of word_{n+1} and the magnitude of any benefit when word_{n+1} is later fixated (e.g., shorter fixations on word_{n+1}, greater probability of skipping word_{n+1}). Consistent with this prediction, greater parafoveal preview benefit on word_{n+1} has been demonstrated when word_n is a high frequency word (Henderson & Ferreira 1990; Kennison & Clifton 1995) and when word_n is highly predictable from the preceding text (Balota et al. 1985). These findings are typically taken to suggest that as the difficulty of foveal processing increases, parafoveal preview benefit decreases.

However, the E-Z Reader model dictates more precise inferences concerning any effects of experimental manipulations of the characteristics of word_n on the subsequent processing of word_{n+1}

https://doi.org/10.1017/S0140525X03450103 Published online by Cambridge University Press