DOCUMENT RESUME

ED 234 356 CS 007 291

AUTHOR	Blanchard, Harry E.
TITLE	A Comparison of Some Processing Time Measures Based on Eye Movements. Technical Report No. 285.
INSTITUTION	Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.; National Inst. of Mental Health (DHHS), Rockville, Md.
PUB DATE	Sep 83
CONTRACT	400-76-0116
GRANT	NIMH-MH-32884; NIMH-MH-33408
NOTE	29p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Cognitive Processes; Comparative Analysis; *Eye Fixations; *Eye Movements; Higher Education; *Language Processing; Reading Comprehension; Reading Instruction; *Reading Research; *Research Methodology; Test Reliability; *Time; Word Recognition
IDENTIFIERS	*Gaze Duration

ABSTRACT

A study was conducted to provide a replication of the gaze duration algorithm proposed by M. A. Just and P. A. Carpenter using a different kind of passage, to compare the three gaze duration algorithms that have been proposed by other researchers, and to measure processing time in reading. Fifty-one college students read a passage while their eye movements were monitored. Five different measures of processing time in reading were each fit, using hierarchical multiple regression, to a model similar to that of Just and Carpenter. The processing time measures--Just and Carpenter's gaze durations, two modified gaze duration measures, number of fixations, and average fixation duration--were shown to be influenced by different independent variables. Also, some evidence was obtained that called into question one assumption of the gaze duration measure--that when increased processing time is needed, a trade-off occurs between fixation duration and the number of fixations on a word. The findings suggest that gaze duration measures should be considered indices of aspects of word processing during reading rather than as measures of actual processing time required by words. (Author/FL)

CENTER FOR THE STUDY OF READING

Technical Report No. 285

A COMPARISON OF SOME PROCESSING TIME MEASURES BASED ON EYE MOVEMENTS

Harry E. Blanchard
University of Illinois at Urbana-Champaign

September 1983

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02238

This work was supported by Grants MH 32884 and MH 33408 from the National Institute of Mental Health to George McConkie and National Institute of Education Contract HEW-NIE-C-400-76-0116 to the Center for the Study of Reading. I would like to thank Margaret Olson for her assistance in programming and data collection, Dr. George McConkie for his comments on this paper, and Dr. James Wardrop for his advice on the regression analysis.

Abstract

Fifty-one college students read a passage while their eye movements were being monitored. Five different measures of processing time in reading derived from the eye movement data were each fit, using hierarchical multiple regression, to a model similar to that of Just and Carpenter (1980). The processing time measures were Just and Carpenter's gaze durations, two modified gaze duration measures, number of fixations, and average fixation duration. The components of gaze duration, number of fixations and average fixation duration, were shown to be influenced by different independent variables. Also, some evidence was obtained which called into question one of the assumptions of the gaze duration measure: the assumption that there is a trade-off between increased fixation duration and making multiple fixations on a word when increased processing time is needed. It was suggested that gaze duration measures should be considered as indices of aspects of the processing of words during reading rather than as measures of the actual processing time required by those words.

A Comparison of Some Processing Time Measures Based on Eye Movements

Eye movements may provide a real time measure of mental processes during reading. Eye movement data indicate where in the text the reader fixated and how long the reader paused at each location. They do not, however, directly indicate what mental processes are responsible for how long the eye stays at each location or what units of text are to be allocated to each fixation. These two issues must eventually be answered by empirical investigation. Currently, a measure of processing time must be constructed by making assumptions on these issues. Using these assumptions, an algorithm can be defined which constructs an index of processing time from the raw eye movement data.

Just and Carpenter (1980) defined such an algorithm which produced what they called a gaze duration profile. The assumptions Just and Carpenter made were (1) on each fixation, a single word, the word that was directly fixated, was being processed (the eye-mind assumption), and (2) there was no processing lag: readers processed only what was being fixated; there were no temporal effects of processing previous words while fixated on the current word (the immediacy assumption). Seemingly consistent with these assumptions was the observation that readers averaged about 1.2 words per fixation. They further assumed that readers trade off fixation durations and number of fixations. For example, extended processing time could be realized as either one long fixation on a word or two or more shorter fixations. This will be referred to as the trade-off assumption. Using these assumptions, the gaze duration measure was constructed as follows: (1) for each fixation on a word, the fixation time was assigned to that word, (2) for words that received more than one fixation, the fixation times were summed, (3) words receiving no fixations were assigned a time of zero, and (4) data from fixations following regressive saccades or from rereadings were eliminated. Gaze durations from different subjects were averaged on a word by word basis, resulting in a mean gaze duration for each word in the passage. These mean gaze durations were then used as the dependent variable in a multiple regression analysis. The independent variables represented psycholinguistic factors which were believed to affect processing time. These variables explained a significant percentage of the variance in gaze durations. Since the psycholinguistic characteristics of immediately fixated words were shown to influence gaze duration, Just and Carpenter subsequently claim this as support for their assumptions.

Just and Carpenter's analysis has been criticized on several grounds. Hogaboam and McConkie (Note 1) argued that Just and Carpenter's assumptions were questionable and, in some cases, contradicted by current evidence. They cited evidence suggesting that information from words to the right of the center of fixation can be acquired on a fixation as well as from the immediately fixated word. Furthermore, the evidence to support the assumption that there is no processing lag is weak. In fact, there is more recent evidence demonstrating lagged effects for certain kinds of perceptual (Underwood & McConkie, Note 2) and psycholinguistic processing (Hogaboam, 1983). Also, to say that a reader averages 1.2 words per fixation does not mean that almost every word is fixated. In fact, 40% of the words in a passage may receive no fixations. Of the remaining 60%, many words receive more than one fixation.

Hogaboam and McConkie (Note 1) proposed an alternative algorithm which was consistent with the evidence that more than one word can be read on a single fixation. This algorithm, the RRG-1 (Read to the Right of Gaze) method, like Just and Carpenter's gaze duration profile, assumed no processing lag and a trade-off between fixation durations and number of fixations. However, it was assumed that for each fixation, more than one word could be processed. Words to the right of the fixation could be read on that fixation and then skipped over. So, RRG-1 was constructed by (1) for each fixation, the fixation time was equally distributed to the word being fixated and all words skipped by the following forward saccade, (2) for words receiving more than one fixation, times were summed, and (3) data from regressive eye movements and rereadings were excluded. The durations were to be used as the dependent measures in a multiple regression in the same way that gaze durations were used.

A second issue concerns the confounding of psycholinguistic and perceptual factors in Just and Carpenter's (1980) analyses (Kliegl, Olson, & Davidson, 1982). Just and Carpenter did not choose any predictors which represented perceptual variables. However, as Kliegl et al. pointed out, a primary influence on gaze duration might be coming from the effects which acuity limitations have on eye guidance. Longer words which extend further into the visual periphery are more likely to be fixated, while words which receive no fixations are usually short words which terminate closer to the fovea. Furthermore, a word could receive more than one fixation simply because on the first fixation the eyes were not in a "convenient viewing position" (O'Regan, 1981) from which the word could be easily identified.

When the initial fixation on a word is not near the center of it, the likelihood of a second fixation on that word is increased. Because fixation durations on a word were summed to produce gaze durations, perceptual factors such as these are confounded with any psycholinguistic factors which correlate with word length.

The best predictor in Just and Carpenter's (1980) model was, in fact, word length in syllables. This could have been due to a syllabic encoding process, or it could have been due to the eye guidance factors mentioned above. If the latter were the case, then word length in letters should be a superior predictor, for number of syllables does not correlate as well with actual visual length as number of letters does. Kliegl et al. (1982) replicated Just and Carpenter's analysis and then repeated the regression with word length in number of syllables replaced with length in number of letters. They found that number of letters accounts for whatever variance number of syllables accounts for as well as for additional variance not shared by number of syllables or other variables. Thus, perceptual factors might explain much of the variance predicted in Just and Carpenter's multiple regression model.

However, Kliegl et al. (1982) noted that any independent effects of number of letters and number of syllables could not be assessed, since these predictors were correlated. This limitation pertains to Just and Carpenter's (1980) entire set of predictors, which were all highly intercorrelated. There is no way to know what part of the shared variance is due to the influence of any particular predictor; it is impossible to assess independent contributions to variance when predictors are intercorrelated (Darlington, 1968; Kerlinger &

Pedhazur, 1973). This limits the theoretical interpretation of Just and Carpenter's psycholinguistic model as well as any attempt to separate perceptual and psycholinguistic factors by such an analysis.

Kliegl et al. (1982) also discussed another reason to be careful in the interpretation of the large word length effect in gaze duration analyses. The averaging procedure used could produce regression coefficients significantly different from zero even if fixations were distributed randomly across the text. With random placement of fixations, long words will be more likely to receive multiple fixations and short words will be more likely to receive no fixations at all. This by itself could produce a significant word length effect. Simply finding that word length accounts for variance in the gaze duration measure is not strong evidence that word length is determining processing time. In fact, for words on which there is a single fixation, there is no relation between word length and the duration of this fixation (Kliegl, Olson, & Davidson, 1983).

Finally, there have been concerns about the trade-off assumption and the loss of information involved in constructing gaze durations (see Hogaboam, 1983). The trade-off assumption states that extended processing time will be manifested as either a longer fixation duration or as multiple shorter fixations. Just and Carpenter (1978) based this assumption on Walker's (1933) finding that readers made either more fixations, longer fixations, or both when the difficulty of reading was increased. It is not clear, however, that this constitutes support for the trade-off assumption. In using the trade-off assumption to construct gaze durations, information is lost about any possible differential effects on fixation duration and number of fixations on a word. Furthermore, Kliegl et al. (1982), by repeating the gaze duration analysis using number of fixations as a covariate, found that most of the variance in gaze durations can be attributed to the number of fixations component. This result opened the possibility that Just and Carpenter's (1980) independent variables may have predicted gaze duration not because of a causal relationship between the two but as a result of gaze durations being correlated with another variable, number of fixations, which was actually causally related to the independent variables.

In a recent paper, Carpenter and Just (1983) introduced another measure, conditionalized gaze duration, which removed variation due to the probability of fixating a word. Conditionalized gaze durations were calculated in the same way as gaze durations, except that they were averaged only over those subjects who fixated a word for at least 50 msec. This eliminated observations of zero processing time and short durations probably due to measurement error. Note that although this removed variation due to fixating a word, it did not remove variation due to the number of fixations a word received, when the word was fixated.

The purpose of the present study was twofold. First, it provided another replication of the Just and Carpenter (1980) gaze duration analysis using a different kind of passage. Since there were some discrepancies between the results of Just and Carpenter and Kliegl et al. (1982), which the latter attributed to passage differences, another replication will be useful in investigating the basis for these discrepancies. Second, this study compared the three gaze duration algorithms which have been proposed and the two components of gaze, number of fixations and fixation duration. A model similar to the one used by Just and Carpenter was fit to each of these five measures of processing time. The measures can be compared by how well they fit the model and how each independent variable behaves. Further, number of fixations and fixation duration were compared to gaze duration to assess what their contribution to the gaze index might be.

The regression analyses used a set of six predictor variables: word length in number of letters, word frequency, beginning of line, part of speech, last word in sentence, and last word in paragraph. Except for word length and part of speech, these predictors were identical to those used by Just and Carpenter (1980). Word length was measured in number of letters rather than number of syllables because of Kliegl et al.'s (1982) findings. Part of speech was used in place of Just and Carpenter's case role factor. Part of speech and case role have a considerable degree of overlap; in fact, several categories are identical, but case role breaks down nouns and adjectives into their sentential functions rather than the formal proper-common distinction. An advantage of the part of speech classification is that it entailed a clear algorithm which other researchers could use in coding other texts (Carrithers and Bever, Note 3, criticized Just and Carpenter for not specifying the heuristics they used in assigning case roles).

In order to deal with the problem of intercorrelated predictors, hierarchical rather than simultaneous multiple regressions were done. Hierarchical multiple regression is the only way variance can be partitioned when predictors are correlated (Cohen & Cohen, 1975). In hierarchical multiple regression it is not the significance of the regression coefficients which is relevant, but rather the incremental proportion of variance a predictor adds and whether that proportion of variance is significant. The variables were ordered simply by causal priority: those variables which represented factors at an earlier level of processing were entered before variables which represented factors at later levels. Beginning of line was entered into the regression equation first, since the shortness of fixations first on a line is due to eye guidance factors. Similarly, word length, which also represented eye guidance factors, at least to some degree, was entered next. This was followed by word frequency, representing a factor involved in word identification, which was then followed by part of speech, representing syntactic processing. Last word in sentence followed by last word in paragraph, the interclause integration factors, were entered last.

Methods

Fifty-one University of Illinois undergraduates read a 417 word passage about the history of Alaska from a CRT while their eye movements were being monitored. The passage was non-technical prose which should have presented no problems for comprehension. The text was presented one line at a time on the CRT and the subject controlled when the next line of text appeared by means of a hand held button. So, the subject controlled the pace of reading but could not reread previous lines. Only horizontal eye movements were recorded. Subjects were given several short answer questions after they read the passage. The questions were straightforward and asked only about information explicitly stated in the passage. Subjects were informed beforehand that the purpose of the questions was to insure that they had read the passage, and that the answers would not actually be used for anything else.

The display unit used was a Digital Equipment Corporation Model VT-11, which has upper and lower case fonts. It was placed 48 cm away from the subject, which made one degree of visual angle equivalent to 4 character positions. Eye movements were monitored using an SRI Dual Purkinjie Image Eyetracker. The VT-11 and the eyetracker were interfaced with a Digital Equipment Corporation PDP-11/40 computer. The system sampled eye position every millisecond and was accurate to within an eighth of a degree of visual angle. Within typical subject variability, the system could record the position of the eye with an error range that was less than one character position. A bitebar and headrest were used to minimize head movements.

Data reduction and analysis

The eye movement data were first arranged into a fixation-based format. From the fixation data, gaze duration indexes were produced for each word in the passage in accordance with the gaze duration algorithm, and then the gaze duration indexes are averaged across subjects to produce a single profile for the passage. The profile contains 417 observations for the dependent variable, one observation for each word. However, eleven of the 417 words received no fixations by any subject (these were short words at the ends of lines). These words were eliminated from the analysis, so the profile actually contained 408 observations. The three gaze duration algorithms, gaze durations (Just & Carpenter, 1980), conditionalized gaze durations (Carpenter

& Just, 1983), and RRG-1 durations (Hogaboam & McConkie, Note 1), were implemented exactly as described in the introduction. The number-of-fixations index was simply the total number of times each subject fixated each word. The average fixation duration was calculated by taking the mean of all the fixations on each word.

Two of the independent variables, word length and word frequency, were continuous variables. Word frequency was log(f+1), where f was the cultural frequency of the word as determined by Kucera and Francis (1967). The rest of the predictors were coded (categorical) variables. Beginning of line, last word in sentence, and last word in paragraph were dichotomous variables coded as 1 or 0. Part of speech consisted of twelve groups: proper noun, common noun, proper adjective, adjective, verb, adverb, pronoun, preposition, article, conjunction, number, and quantifier. This factor was dummy variable coded with the quantifier group being the group coded by 0's, i.e. it was the group left out of the regression equation (Kerlinger & Pedhazur, 1973). Finally, one of Just and Carpenter's predictors, novel word, was not included in these analyses simply because the passage used contained no novel words. This passage contained no difficult vocabulary.

Results

Results from the multiple regression analyses are presented in Table 1. The incremental proportion of variance accounted for (squared part correlations) by each variable is reported, and those variables which add significantly to the explained variation are indicated. For the gaze duration profile, the multiple correlation (*R*) was .84. This is to be compared with

Just and Carpenter's (1980) reported multiple _R_ of .85 (R^2 = .72) and Kliegl et al.'s (1982) multiple _R_ of .67 (R^2 = .45). The exclusion of zero-gaze observations in conditionalized gaze durations weakens the fit of the model, lowering the multiple _R_ to .69. The RRG-1 durations are fitted by the model equally well, with the same multiple _R_ of .69. By definition, RRG-1 durations have no zero-gaze observations. The model is also fitted well by the number of fixations, multiple _R_ = .75. The weakest fit is provided by average fixation duration, multiple _R_ = .50. The good fit provided by number of fixations in contrast to the relatively poor fit provided by fixation duration reinforces Kliegl et al.'s conclusion that most of the variance accounted for in gaze durations is due to the relation between the independent variables and the number of fixations on each word.

Insert Table 1 about here

The incremental proportions of variance for the independent variables show a close similarity in the pattern of significant effects between the three gaze duration measures. The same variables show statistically significant effects for each measure. The principle difference lies in the proportion of variance accounted for by word length. For all the other variables, the difference in the proportion of variance accounted for compared across the multiple regressions does not exceed 7%. Despite the difference in the way the gaze and RRG-1 algorithms assign processing times to words, the fit of the independent variables is remarkably similar. The most likely reason for such similar results is that the specific differences between the algorithms do not come into play for most of the words in the data.

In contrast to the gaze duration measures, number-of-fixations and average fixation duration show important differences from each other, as well as between themselves and the gaze duration measures. Viewing these measures as components of the gaze duration measure, we see that the word length effect in gaze duration comes almost entirely from the number-of-fixations, the word frequency effect almost entirely from fixation duration, and the part of speech effect from both components. It is clear that the independent variables do not predict these two components in the same way. There is also a small but significant effect of the last word in paragraph variable on number of fixations. The last word in a paragraph is also always the last word on a line, so it is not clear whether this is effect to due to the position of the word in the paragraph or the position of the word on the line. Considering the absence of an effect due to the other interclause integration variable, last word in sentence, which is not confounded with last word on the line, it is probably not appropriate to interpret this effect of last word in paragraph as due to interclause integration.

Two differences between the results of the present study and those of Just and Carpenter (1980) and of Kliegl et al. (1982) should be pointed out. In contrast to Just and Carpenter's analysis, the interclause integration factors, last word in sentence and last word in paragraph, do not account for a significant proportion of variance; indeed, they account for almost no variance at all. This finding is in agreement, however, with Kliegl et al.'s analysis, whose end of sentence variable accounted for merely 1% of their total variance. The incremental proportions of variance accounted for by word frequency and part of speech are much larger in the present analysis than the corresponding values in Kliegl et al.'s analysis. In contrast to the part of speech variable in the present analysis, Kliegl et al.'s syntactic processing factor, function-content word, coded much less information about the syntactic function of a word. This could explain the greater predictive ability of part of speech here.

A Test of the Trade-off Assumption

According to the trade-off assumption, an increase in processing time may be manifested as either an increase in fixation duration or as multiple short fixations. Therefore, the true indicator of processing would be total time on a word -- gaze duration. A simple interpretation of this assumption would imply that, because both number of fixations and average fixation duration represent processing time, the same factors should influence these measures in the same way. However, the results of the multiple regressions show some important differences. In particular, word length is a good predictor of number of fixations but not average fixation duration, whereas word frequency is a good predictor of fixation duration but not number of fixations. Thus the components of the gaze duration index do not reflect the same processes, as the trade-off assumption might predict.

A second prediction can also be made from a simple version of the trade-off assumption: in the long run, the processing time on a word should be the same whether the subject makes one long fixation or several shorter fixations.

That is, the mean gaze durations for words of a given category should be the same whether they received a single fixation or more than one fixation. The mean of the gaze durations on words that were nine or more letters long was calculated separately for cases where one fixation, two fixations, or three or more fixations were made on the word. Long words were selected in order to obtain a large sample of words with multiple fixations. The results, shown in Table 2, indicate that, contrary to the above predictions, the mean gaze duration nearly doubles when a second fixation was made and more than triples when three or more fixations were made. This is clear counterevidence for the simple trade-off assumption, as stated above.

Insert Table 2 about here

A more complex form of the trade-off assumption would state that multiple fixations are made whenever more processing time is needed than one fixation can provide. Multiple fixations on a word would only occur in instances where longer processing times are needed. Under this assumption, the increase in gaze duration shown when more than one fixation was made on a word can easily be explained: gaze duration increases with multiple fixations because more processing time is needed in those cases. However, if there is a trade-off occurring between making longer fixations versus more fixations on a word, then we would expect the mean gaze duration for words receiving two fixations to be considerably less than twice the mean for words receiving a single fixation. The actual data pattern, as seen in Table 2, indicates that the gaze duration for words fixated twice is almost twice that of words fixated only once, and words fixated three times have a gaze duration three times that of words fixated only once. This is not supportive of the trade-off assumption.

The data were also examined to determine whether individual words might show a pattern more consistent with the trade-off assumption, that is, whether some words would have the same gaze duration when fixated more than once as when they received one fixation. There was no evidence for such a pattern. Out of the 20 words which were nine or more letters in length, one received just single fixations; the gaze durations of the other words were composed of a combination of single and multiple fixation instances. In all but one of the latter cases, more of the gaze durations for each word came from single fixation instances than from multiple fixation instances (only two words were evenly divided between single and multiple fixation instances). In every case, multiple fixations resulted in a longer gaze duration than single fixations. For 70% of these words, gaze duration was increased by more than 150 msec when the word received two fixations. So, the individual words show the same pattern as the mean gaze durations in Table 2. Readers do not seem to be simply trading off between making longer versus making more fixations, with each providing an equivalent means of allowing more processing time. Rather, the making of longer fixations versus additional fixations is determined by different aspects of the reading process. This calls into question a basic assumption of the gaze duration and RRG-1 measures.

Discussion

The results of the regression analysis performed on the data from this study are very similar to those reported by Just and Carpenter (1980), though in the present study a hierarchical analysis approach was taken. The multiple correlation was almost identical to Just and Carpenter's; both of these studies yielded higher multiple correlations than those reported by Kliegl et al. (1982). This study also yielded higher proportions of incremental variance due to two psycholinguistic variables, word frequency and part of speech, when compared to Kliegl et al.'s hierarchical analysis with word length entered first. This may be due to differences in passages used, as Kliegl et al. suggested for their own results: Just and Carpenter's passages were more difficult, had more novel words, and more variance in word frequency than their passage. However, the passage used in the current study was more like that used by Kliegl et al. in posing no comprehension difficulties, so it would be difficult to explain the differences between this study and Kliegl et al.'s by invoking passage differences. It is likely that the difference reflects Kliegl et al.'s smaller sample size, which must have resulted in less stable gaze duration estimates for the words (as pointed out by these authors), and in their less fine-grained breakdown of language variables (only a function word/content word distinction, rather than classification by case role or part of speech).

One similarity between the results of this study and Kliegl e ..'s is the lack of an effect for the end of sentence variable, which is as d to represent interclause integration processes. In this study, the two interclause integration variables accounted for almost no variation in multiple regressions with any of the three gaze duration measures. Because Just and Carpenter did not do a hierarchical analysis and did not report incremental proportions of variance for their variables, it is not appropriate to compare the studies on this point. Effects due to the interclause integration factors are of considerable import to the theories of Just and Carpenter (1980) and Carrithers and Bever (Note 3), where they represent a special stage in which the preceding sentence or clause unit is semantically organized. Because both this study and Kliegl et al.'s study failed to find significant relationships between gaze duration and the interclause integration variables, the evidence for such a "wrap-up" stage must presently be regarded as, at best, equivocal.

In the current study, there is considerable similarity in the results obtained with the three gaze duration algorithms, presumably because they share so many common assumptions about the relation between the cognitive processes involved in reading, and the nature of eye movement control. There are two primary differences among them. First, they differ considerably in the amount of variability in the data. The total sum of squares, representing the total variability in the dependent variables, was 6,447,897 for the gaze duration, 2,267,004 for the conditionalized gaze duration, and 3,435,503 for the RRG-1 measure. Simply assigning zero's to unfixated words triples the variance in the gaze duration measure, as compared to the conditionalized gaze duration. Dividing fixation times among words, as the RRG-1 technique does, also increases variance as compared to the conditionalized gaze duration. Second, they differ substantially in the proportion of variance which can be accounted for on the basis of lengths of words. Since short words most often go unfixated, and hence get the zeroes or partial fixation times assigned to them in these techniques, the gaze duration and RRG-1 measures show a higher correlation with word length than does the conditionalized gaze duration. Thus, the added variance in these measures is strongly related to word length.

The eye-mind and immediacy assumptions, which underlie all the gaze duration measures, are increasingly coming under criticism with evidence that unfixated words are being processed and that there is frequently some processing lag. Furthermore, this study has provided good grounds to question the trade-off assumption as well. To the extent that the basic assumptions of the gaze duration measures are faulty, these measures are not accurate indicators of the actual processing times required by the different words in a passage (as assumed by Thibadeau, Just, & Carpenter, 1982). However, it is important to distinguish between the use of gaze durations as a true measure, as indicating the actual processing time, versus being an index of some aspect of the processing. A true measure of processing time, such as reaction time or reading time, is treated as a ratio scale of measurement. An index of processing time is treated as an interval scale or an ordinal scale. An index can be used to investigate whether certain processing events require more or less time than other events, perhaps with some suggestion of the magnitude of such differences, but it does not permit absolute statements about the actual amount of processing time required. It seems clear that, even if the gaze duration and RRG-1 measures fail as true measures of processing times for words, they can still serve as useful indices for comparing aspects of processing associated with different words during reading, because of their correlations with variables assumed to affect processing time.

The problem which remains, then, is to determine which of the available indices is best to use. Because the results from the three gaze duration measures studied were so similar, which one is selected is not likely to drastically change the outcomes of a study. On the other hand, there are some differences. In particular, it would seem best to choose the measure that is least influenced by eye guidance factors and which shows the best prediction by those independent variables which are most clearly related to psycholinguistic factors. The gaze duration measure is probably the least appropriate, because of the way it treats words that are not fixated and because of the heavy influence from word lengths. Conditionalized gaze durations and RRG-1 durations are quite similar; however, conditionalized gaze durations show the smallest influence from word lengths and the best prediction by word frequency and part of speech. However, a final possibility, which may be the most justifiable, is to use both number of fixations and mean fixation durations as indices, rather than trying to combine them into a single index. This seems most appropriate, since they appear to reflect somewhat different aspects of the processing taking place. Determining which is influenced by the variables under study provides more information than using a combined measure, and avoids problems associated with the trade-off assumption.

We do not yet understand the nature of the relationship between cognitive processes and eye movement control well enough to permit deriving an accurate measure of processing time from eye movement records. At the present time we must settle for indicators of aspects of processing rather than true processing time measures for words in the on-going reading situation. It is important that we strive to understand the nature of these indices, and what aspects of processing they are reflecting, if they are to be of the most use to us in our study of language processing during reading.

Reference Notes

1. Hogaboam, T. W., & McConkie, G. W. The rocky road from eye fixations to comprehension (Tech. Rep. No. 207). Champaign, Ill.: University of Illinois, Center for the Study of Reading, 1981.

2. Underwood, N. R., & McConkie, G. W. Perceptual span for letter distinctions during reading (Tech. Rep. No. 272). Champaign, Ill.: University of Illinois, Center for the Study of Reading, 1983.

3. Carrithers, C., & Bever, T. G. Eye-movement patterns confirm theories of language comprehension. Unpublished manuscript, Columbia University, 1982.

References

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

Cohen, J., & Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1975.

Darlington, R. B. Multiple regression in psychological research and practice. Psychological Bulletin, 1968, 69, 161-182.

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

Just, M. A., & Carpenter, P. A. Inference processes during reading: Reflections from eye fixations. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), Eye movements and the higher psychological functions. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978.

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

Kerlinger, F. N., & Pedhazur, E. J. Multiple regression in behavior research. New York: Holt, Rinehart and Winston, 1973.

Kliegl, R., Olson, R. K., & Davidson, B. J. Regression analysis as a tool for studying reading processes: Comment on Just and Carpenter's eye fixation theory. Memory and Cognition, 1982, 10, 287-296.

Kliegl, R., Olson, R. K., & Davidson, B. J. On problems of unconfounding perceptual and language processes. In K. Rayner (Ed.), Eye movements in reading: Perceptual and language processes. New York: Academic Press, 1983.

O'Regan, K. The "convenient viewing position" hypothesis. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), Eye movements: Cognition and visual perception. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1981.

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

Walker, R. Y. The eye movements of good readers. Psychological Monographs, 1933, 44 (3, Whole No. 199), 95-117.

Table 1

Incremental Proportions of Variance Accounted for by Each Independent Variable in Multiple Regressions on Five Processing Time Measures

	Dependent Variable				
Factor	Gaze Duration	Conditionalized Gaze Duration	RRG-1 Duration	Number of Fixations	Average Fixation Duration
Beginning of Line	.01**	.01**	.03**	0	0
Word Length (letters)	.45**	.16**	.23**	.39**	.06**
Word Frequency	.15**	.15**	.10**	.05**	.11**
Part of Speech	.08**	.15**	.12**	.08**	.08**
Last Word in Sentence	0	0	0	0	0
Last Word in Paragraph	0	0	0	.04**	.01
Total R^2	.69	.48	.47	.56	.25

** F value $p < .01$

Table 2

Gaze Durations on Words Nine

or More Letters in Length

Number of Fixations	*n*	Mean	*SD*
1	521	276	132
2	133	488	148
3 or more	23	949	261