

SESSION V GATHERING COMPLEX DATA

An automated eye movement laboratory for on-line electrooculography

DANA G. FISHER and ERNST Z. ROTHKOPF
Bell Laboratories, Murray Hill, New Jersey 07974

This paper describes an automated eye movement laboratory that uses electrooculography (EOG) to study people's eye movements while they read. An on-line minicomputer processes bioelectric potentials that correspond to saccadic eye movements. Horizontal saccades larger than 1.5 deg of visual angle are detected and analyzed in real-time as they occur. The laboratory is designed for prolonged yet unobtrusive observation of human eye movements during sustained reading periods of minutes or hours. All important functions regarding data collection and data reduction are performed automatically, according to simple procedures that can be applied uniformly and without bias to nearly all subjects that we study. Results from three experiments are cited in order to quantify the performance of the laboratory with respect to four criteria: saccade detection accuracy, measurement accuracy, sensitivity, and the uniformity of these measures over different subjects.

The cover story in the latest issue of *Scientific American* is about 7,500 words long. An average reader would take about 35 min to read the article and would make roughly 8,000 eye movements. Among the audience for that article, of course, different readers would bring different backgrounds to bear upon the text material. They would read for different purposes; they would require different amounts of time; they would learn different things; and in the process of reading they would show different patterns of eye movements.

Human saccadic eye movements are one of the few directly observable behavioral events that occur while people read. Eye movements attract the attention of behavioral scientists because eye movements may provide clues to the nature of the moment-to-moment mental activity that regulates the pace at which people acquire new information from written text (Just & Carpenter, 1980; Rayner, 1976; Rothkopf, 1978).

The laboratory described in this paper is designed to provide an efficient, general-purpose tool with which to study large numbers of eye movements that occur during sustained reading periods. The research issues that this laboratory is intended to explore include topics of persistence, individual differences in eye movement patterns and reading styles, and the relationships between

eye movements, text difficulty, and reading task requirements.

The procedures used for data collection and data reduction are in all important respects fully automated. Horizontal saccades are detected and analyzed by an on-line laboratory computer in real-time. We can apply without bias the identical criteria for the detection of saccades to nearly all subjects. The recording technique, electrooculography (EOG), is simple, unobtrusive, and allows us to record for periods of time up to several hours. Subjects sit comfortably and unrestrained in their normal postures. The large volume of data that is generated can be handled conveniently, during the experiment itself.

We measure four basic things: the magnitude, direction, and duration of each saccade and the duration of each fixation pause. Saccade magnitudes are measured in units that can be converted either into the number of character spaces traversed or into the degrees of visual angle traversed. The durations of saccades and fixations are measured in milliseconds. All events are located in time measured from the onset of the current text slide.

Unlike some eye movement laboratories, we do not map the eyes onto the text. EOG determines the position of the eyes relative to the subject's head, rather than their position relative to a point in space. It was not considered essential to determine, for each fixation pause, which word or words a subject fixated. Our approach emphasizes the study of aggregate eye movement patterns, rather than momentary states (Rothkopf, 1978). Moreover, recording techniques that allow a

The authors are grateful to Robert Palladino at DSI Company, who wrote the software for the on-site data collection and data reduction programs, and to Esther U. Coke and Dennis E. Egan for comments on an earlier draft of this paper. Requests for reprints should be directed to Dana G. Fisher, Bell Laboratories, 2D-449, 600 Mountain Avenue, Murray Hill, New Jersey 07974.

researcher to map the eyes onto the text require either that the subject's head be held stationary or that head movements be tracked automatically (Young & Sheena, 1975). The former method requires a bite board or a chinrest; the latter requires elaborate equipment that needs frequent recalibration. Neither method was consistent with our desire for prolonged, unobtrusive observation. To study the relationships between eye movements and text characteristics, we examine aggregate text units rather than word-by-word units. For example, we compare eye movements for a difficult paragraph with eye movements for a simpler paragraph. For circumstances in which a single text slide contains two different content regions, such as goal-relevant and background information, we can use an extrapolation procedure developed by Rothkopf and Billington (1979) that permits a comparison to be made between the eye movement characteristics attributed to each region.

One of the most important considerations in the development of this laboratory was to ensure that the procedures for detecting saccades could be applied uniformly to all subjects without introducing bias. Our greatest concern in this regard was with the different levels of noise that exist between subjects in the analog eye movement signal. In EOG, noise can arise from extraneous muscle activity or from other electrical sources beneath the skin surface. It is essential to take precautions not only against misinterpreting noise as true saccades, but particularly against making any

systematic errors of identification between subjects. Dissimilar saccade detection rates between subjects, for example, or dissimilar false alarm rates could undercut efforts to investigate individual reading styles by introducing artificial or spurious differences. Our standards, therefore, could not be those that permitted the highest detection rates and lowest false alarm rates for each individual subject. Instead, we had to pursue the more egalitarian standard of uniform detection and false alarm rates over subjects who exhibited different noise levels.

The first section of this paper traces the signal path from its source at the subject's eyes to the experimenter's hard-copy summary statistics, in the context of a typical reading experiment. The focus is on the apparatus, procedures, and logic that have been developed to accomplish the task of automation. In the second section, the focus is on the saccade detection algorithm. Results are reported from three experiments that assess the performance of the algorithm with respect to four criteria: saccade detection accuracy, measurement accuracy, sensitivity, and the uniformity of these measures over different subjects. Since all four aspects of the laboratory's performance depend on how parameters are set in the algorithm, we explain how parameter settings affect performance, and how appropriate settings were chosen.

SIGNAL PATH

Figure 1 shows the overall architecture of the eye

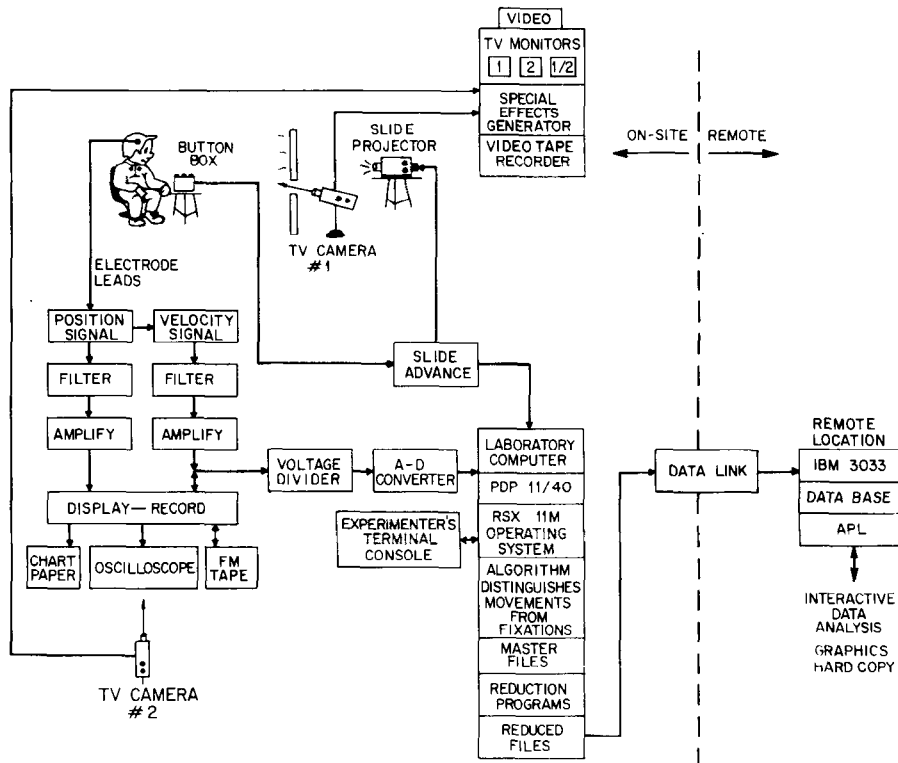


Figure 1. Schematic illustration of major devices, functions, and signal paths in the laboratory.

movement laboratory in schematic form. Figure 2 is a photograph that illustrates the setting of the laboratory during an experiment. A subject reads text that is displayed on a screen via rear projection. From the subject's position, a 60-character line subtends a visual angle of about 20 deg. Subjects read at their own pace, operating a button box connected to digital logic that advances a slide projector and reports that event to a computer program.

Usually, a subject is instructed to read for a particular purpose, such as to look for an answer to a certain question. The subject typically consults a notebook that contains instructions and reading objectives. While the screen is dark, the subject determines the target for the next slide and then looks up at the screen on which the text will appear. When the individual is ready to begin reading, he or she depresses a button that causes a text slide to appear and that initiates data collection. As soon as the subject either finds the target or reaches the end of the slide without finding it, he or she depresses the button again. This causes the screen to go dark and halts data collection. The subject then records a response and repeats the cycle in a self-paced fashion.

A Beckman 16-mm Ag-AgCl surface electrode is taped about 20 mm from the external canthus of each eye, and an 11-mm reference electrode is taped to the dorsal surface of the right earlobe. These sensors record electrical changes that correspond to saccadic eye movements. This signal represents the position of the eyes relative to the subject's head. (For discussions of EOG, see Marg, 1951, and Shackel, 1967.) The signal from the electrodes is routed to a Beckman Type RM dynograph recorder, in which a Beckman 9859 direct nystagmus coupler splits the incoming signal into two paths. One path continues the original position signal. The other path uses a Beckman 9841 nystagmus velocity coupler with a rise time of 2 msec to differentiate the position signal with respect to time, yielding velocity information. We use ac amplification techniques with a 1-sec time constant in order to minimize the electrode drift that is known to occur with dc methods (Young & Sheena, 1975). The position and velocity signals undergo bandpass filtering in parallel, with a low-frequency cutoff at .152 Hz and a high-frequency cutoff at either 30 Hz or 150 Hz. Electrode drift is negligible. The position signal is normally amplified by a factor of about 10,000, and the velocity signal is amplified by a factor of about

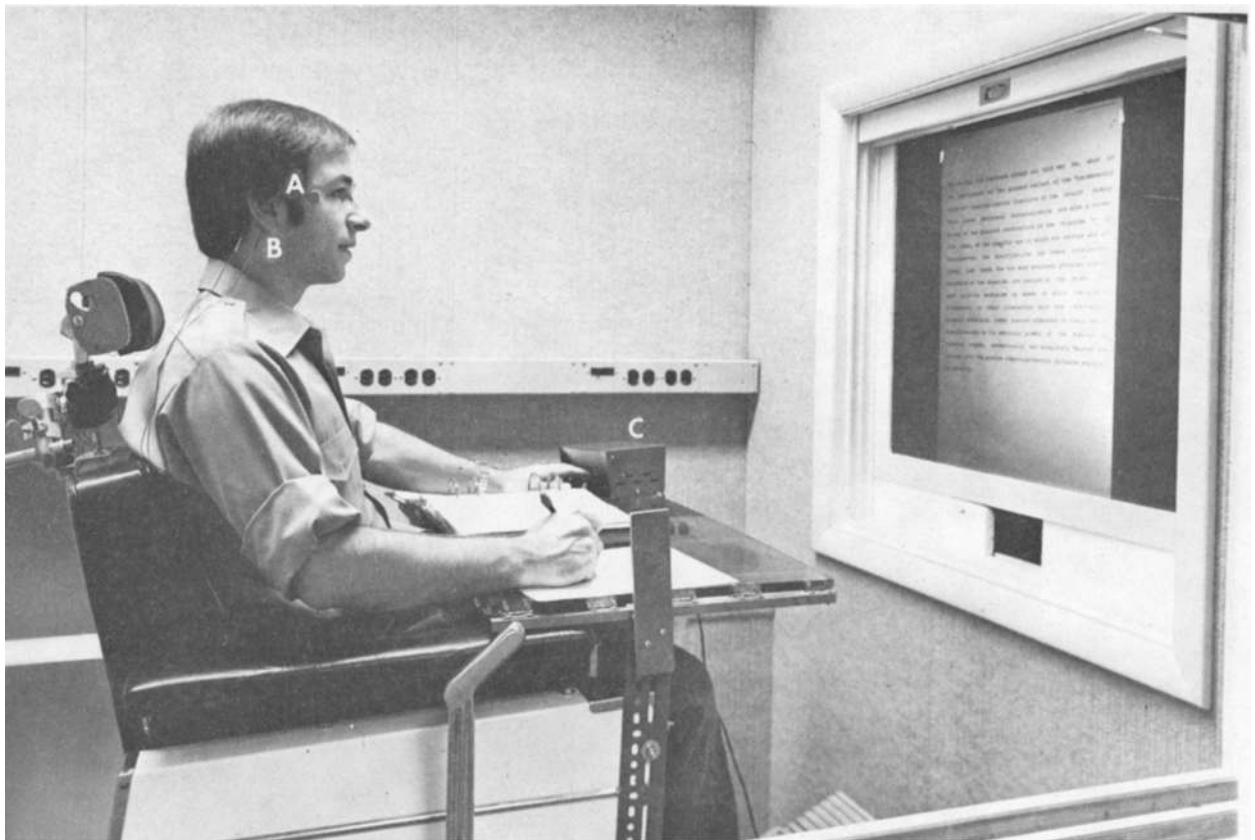


Figure 2. Photograph of a typical setting during a reading experiment. (a) Electrode attached with adhesive tape to the skin surface near the outer canthus of the subject's right eye. (b) Electrode attached to the dorsal surface of the subject's right ear. A third electrode, not shown, is attached near the outer canthus of the subject's left eye. (c) Button box allows the subject to advance a slide projector during self-paced reading experiments. TV Camera 1 is behind the rectangular opening at the bottom of the screen.

1,000. Individual subjects are calibrated by adjusting the amplification such that 30-character-space saccades consistently register 15-mm pen displacements on the velocity channel of the chart recorder. The amplification circuitry is identical for both position and velocity channels: A Beckman Type 481 preamplifier is followed by a Beckman Type 482M8 power amplifier.

The position, velocity, and buttonpress signals can all be displayed simultaneously, on chart paper, an oscilloscope, or both. All three signals can be recorded on FM tape for long-term storage, via an Ampex PR-500 FM tape recorder. At 3.75 in./sec, the center carrier frequency is 6.75 kHz and the nominal data bandwidth is dc to 1.25 kHz. A 10-in. reel of .5 in. x 3,600 ft instrumentation tape can store about 3 h of recorded eye movements.

The voltage level of the velocity waveform is measured by a Digital Equipment Corporation (DEC) Model LPS-11 12-bit analog-to-digital converter at the rate of 1,000 samples/sec, under the control of a DEC PDP-11/40 minicomputer that runs a DEC RSX11M Version 3.2 operating system. During the sampling, which can take place either in real-time as the subject reads or during playback from an FM tape recorded at an earlier date, a computer program interprets the digitized velocity sample values and uses an algorithm to distinguish eye movements from eye fixations or extraneous noise.

This data collection program produces a master file on a disk that contains detailed information for each saccade detected. When it is convenient, these digital master files can be written to magnetic tape for long-term storage and/or they can be submitted to any one of several data reduction programs.

The on-site data reduction programs report summary statistics at various levels of detail, either at an operator's console or on a laboratory line printer. One reduction program creates a condensed file that consists of four key data points for each saccade: its starting time, its ending time, its direction, and the peak velocity it attained. (The peak velocity indicates the magnitude of the saccade, or how far the eye was displaced between fixations.) This file is transmitted via data links to a mainframe computer with a much larger capacity, in which a data base of eye movements can be assembled and manipulated as a unit for statistical analysis.

SACCADE DETECTION ALGORITHM

The saccade detection algorithm is illustrated in Figure 3. Panel a shows an analog velocity signal from a typical chart record of a subject's eye movements during reading. The center of the velocity channel represents zero velocity. The spikes in the chart record correspond to horizontal saccades. Spikes above the center of the velocity channel are rightward saccades. Spikes below

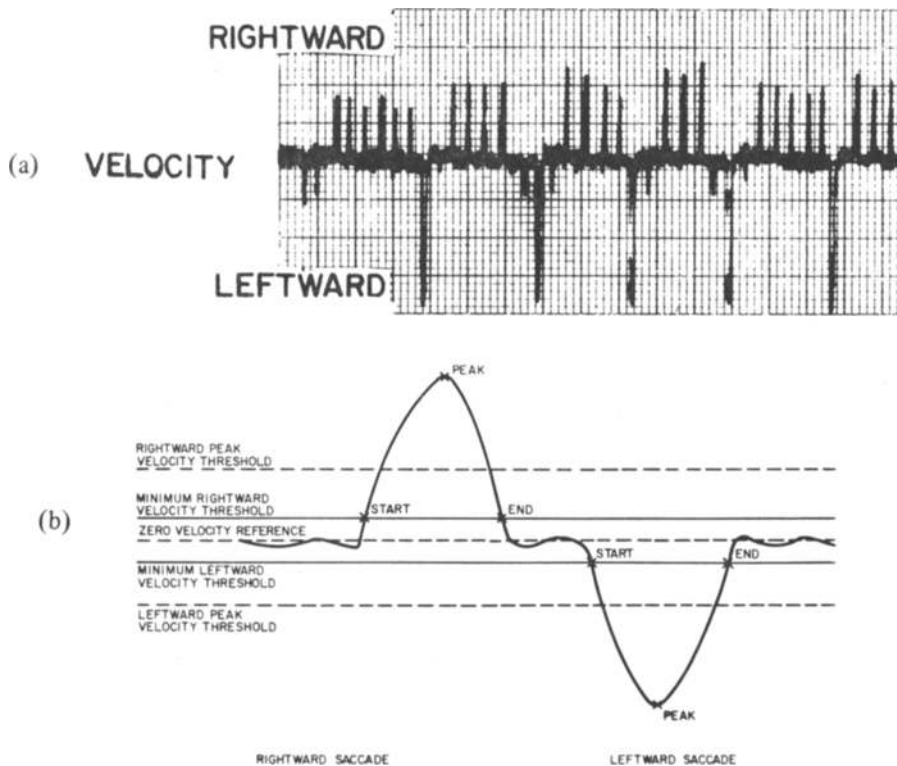


Figure 3. (a) Analog velocity signal from a typical chart record of a subject's eye movements during reading. (b) Schematic illustration of the saccade detection algorithm. The velocity waveform here is intended to represent spikes associated with rightward and leftward saccades.

the center of the channel are leftward saccades. The height of each spike indicates the peak velocity attained, which is known to be correlated with the magnitude of the saccadic eye movement. Saccade durations are about 20-70 msec, and they are also correlated with saccade magnitude. The portions of the chart record between spikes, in the center of the velocity channel, correspond to eye fixations, which are periods of about 200-250 msec during which the eyes have little or no velocity.

The elements of the saccade detection algorithm are shown schematically in Panel b of Figure 3, to illustrate how the algorithm would detect a rightward and a leftward saccade. Five parameters are indicated: a zero-velocity reference value, minimum rightward and leftward velocity thresholds, and rightward and leftward peak velocity thresholds. These five parameters are related in a straightforward way to the major features in the chart record. The zero-velocity reference value is approximately the midpoint of the velocity channel. The band between the minimum rightward and leftward velocity thresholds corresponds roughly to the noise in the center of the velocity channel that is noticeable during eye fixations. The rightward and leftward peak velocity thresholds are criteria with regard to the height of each spike.

The zero-velocity reference value is determined independently for each subject in a calibration procedure. The remaining four values are set at the same fixed distance from the zero reference point for all subjects. In the digitized velocity signal, a rightward saccade, for example, is defined as a string of at least 15 consecutive sample values, each of which exceeds the minimum rightward velocity threshold, and at least 1 of which exceeds the rightward peak velocity threshold. The duration of a saccade is measured from the first superthreshold velocity value to the last superthreshold value.

The main reason for analyzing velocity rather than position is that the velocity signal is less sensitive to head movements. To illustrate, if a subject makes a head movement while maintaining a fixation, the position of the eyes relative to the head changes and the change is reflected in the position signal. But since head movements do not generate the velocities that saccadic eye movements do, very little change is reflected in the velocity signal, and the head movement probably will be overlooked by the saccade detection algorithm.

PERFORMANCE EVALUATION

The issue of selecting optimal velocity threshold and duration criteria is threefold: to identify quantitative measures of the laboratory's performance, to investigate how various parameter settings would affect performance, and to select criterion parameter values that optimize performance.

We identify four principal aspects of performance: detection accuracy, measurement accuracy, sensitivity,

and the uniformity of these measures over different subjects. Detection accuracy refers to the ability to distinguish true saccades from background noise, and it is quantified in terms of hit rates and false alarm rates. Measurement accuracy refers to the ability to determine correctly the durations and magnitudes of saccades in milliseconds and character spaces. Sensitivity refers to the smallest saccades that the laboratory is capable of detecting. The uniformity issue, which arises because different subjects exhibit different levels of background noise, is concerned with minimizing the differences between subjects in each of the preceding three measures of performance.

The existence of different noise levels between subjects complicates our efforts to identify optimal parameter values by introducing tradeoff relationships between performance criteria. For example, lenient duration and peak velocity thresholds might achieve a high degree of sensitivity for low-noise subjects without compromising detection accuracy; but the same thresholds applied to high-noise subjects would tend to cause higher false alarm rates, which would have an adverse effect not only on detection accuracy, but also on uniformity between subjects. In fact, any difference at all between the duration and peak velocity thresholds applied to different subjects would tend to affect some aspect of uniformity adversely. Therefore, we sought thresholds that could be applied identically to all subjects and that still achieve acceptable levels of performance in terms of sensitivity, detection accuracy, and measurement accuracy.

The uniformity issue is particularly important because a key objective of this laboratory is to investigate individual differences in eye movement patterns. Any lack of uniformity could cause artificial or spurious differences to appear between subjects when none actually exist. The emphasis on uniformity between subjects might be contrasted with an approach that seeks to maximize sensitivity for each individual subject.

Three experiments were conducted to investigate the adequacy of the saccade detection algorithm and to collect the data we required in order to select optimal parameter settings.

Experiment 1

This experiment used simulated saccades produced by an electronic signal generator to assess the measurement accuracy of laboratory hardware and software in the absence of noise, under conditions in which the input signal was known, regular, and easily described.

Method

We used sine waves to simulate saccades between 4 and 20 character spaces. In our laboratory, a 60-character-space line subtends a visual angle of about 20 deg. Saccades of different magnitudes were simulated by adjusting the voltage amplitude and frequency of the sine waves according to known guidelines

regarding the voltages to expect in EOG signals (Shackel, 1967) and the durations of saccades (Fuchs, 1971, 1976). Test signals were fed through a voltage divider to produce amplitudes in microvolts comparable to those of saccades. These signals entered our apparatus at the junction where electrode leads are normally plugged in. From there, they underwent the identical signal processing as normal saccades.

Results and Discussion

The saccade detection algorithm correctly detected all simulated saccades and made no false alarms. Simulated saccades were detected accurately at rates up to 40 per second, which is about an order of magnitude higher than the rate of normal saccades during reading. Thus, the speed of software execution is more than adequate for our purposes.

The correlation between the measured output voltage in the amplified velocity signal and the mean digitized peak velocity was $r = .999$ over the 30 test signals. This finding confirms that our sampling rate of 1,000 samples/sec is frequent enough to permit accurate estimates of peak velocity across the range of voltages we anticipate.

The duration and peak velocity data were orderly and unexceptional. We observed well behaved relationships between simulated saccade magnitude and digitized peak velocity and between predicted and observed duration.

We conclude in general from this experiment that our equipment and procedures are adequate for detecting and measuring signals similar to those we encounter in normal reading.

Experiment 2

The purpose of this experiment was to collect the data required to establish the relationships between objective magnitudes of normal saccades and their durations and peak velocities as we measure them. Subjects were instructed to make saccadic eye movements back and forth between targets separated by a known horizontal distance. Experimental conditions were designed to permit us to assess effects both of different low-pass filter settings and of time elapsed during the experiment on the durations and peak velocities of saccades.

Method

Each stimulus slide contained two Xs that served as fixation targets. The set of 15 stimulus slides contained targets that were from 2 to 30 horizontal character spaces apart, in increments of 2 character spaces. Subjects were instructed to alternate their gaze between the two targets at their own pace, making 10 round trips per slide. A block of slides consisted of a random order of the 15 stimulus slides. Each subject viewed a practice block followed by six test blocks. Eight subjects were tested individually in sessions that lasted about 2 h.

Results and Discussion

Figure 4 shows the mean digitized peak velocities, averaged over subjects, plotted as a function of target separation for rightward and leftward saccades. We

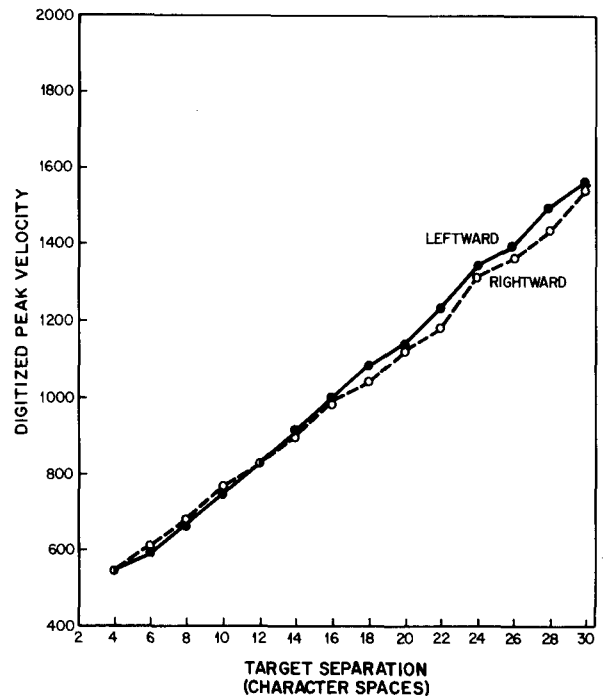


Figure 4. Mean digitized peak velocities of saccades as a function of target separation for leftward and rightward saccades in Experiment 2.

observe a similar, although slightly weaker, relationship between target separation and saccade duration. These relationships, however, which correspond to well-known facts about the dynamics of saccades, are orderly for all subjects tested. Since the differences between subjects and the differences associated with repeated saccades within slides are both considerably smaller for peak velocity than for saccade duration, we expect to use the peak velocity measure to infer the magnitudes of saccades made during reading experiments.

Differences between 30-Hz and 150-Hz low-pass filter conditions were negligible, with regard both to duration and to peak velocity. There was a small but measurable tendency toward smaller peak velocities in later blocks. We can prevent this signal attenuation in long reading experiments by recalibrating the amplification for each subject about every 30 min or so.

In general, data from this experiment establish orderly relationships between the magnitudes of saccades and their durations and peak velocities. These data provide the empirical foundation required to infer the magnitudes of saccades in character spaces from what we measure about them.

Experiment 3

This experiment was designed to study the different noise levels that exist between subjects in the EOG signal and to collect the information required both to

assess the effects of different algorithm parameters on laboratory performance and to identify optimal parameter settings. We regard optimal settings as those that permit the best possible levels of detection accuracy, measurement accuracy, and sensitivity, while still achieving uniform levels of performance between subjects whose noise levels are different. In this study, we quantified the noise levels of different subjects and selected subjects who represented low, medium, and high noise levels. We then analyzed sample reading records in detail, graphically, for the purpose of distinguishing true saccades from imposters. Finally, we examined the consequences on performance of various parameter settings and made decisions regarding optimal settings.

Method

Noise levels were quantified on the basis of brief samples of reading records from each of 24 subjects. For each subject, we collected 20,000 digitized velocity samples over a period of 20 sec and computed a frequency distribution of these values around each subject's zero-velocity reference value, which was defined as the median of the distribution. Subjects were rank ordered according to the standard deviations of their distributions. We selected the lowest four, middle four, and highest four subjects to represent low, medium, and high noise levels.

For each of the 12 test subjects, we made a detailed study of a continuous 40-sec eye movement record. For the purpose of this analysis, we used exceptionally lenient criteria in the saccade detection algorithm, in order to detect all true saccades plus an appreciable number of imposters. For each subject, we consulted both the analog chart record and the corresponding digital output, and we distinguished true saccades from imposters by inspection. Altogether, a total of about 1,800 true saccades and about 1,000 imposters were identified.

Results and Discussion

We observed that a value of 128 units for the minimum rightward and leftward velocity thresholds would exclude about half the imposters, without excluding any true saccades. We judged that the remaining imposters could be excluded by the other saccade detection criteria, duration and peak velocity.

Figure 5 shows the false alarm rates and miss rates we would obtain in this sample as a function of various duration and peak velocity criteria. Both error rates are expressed as a percentage of only those true saccades whose peak velocities were greater than the prevailing peak criterion at each point, on the grounds that the specification of a peak criterion implies that no attempt is being made to detect saccades whose peaks are less than the criterion itself.

A tradeoff relationship exists between detection accuracy and sensitivity: As peak velocity criteria increase, lower error rates are achieved only at the cost of decreased sensitivity. At the duration criterion of 14 msec, however, the false alarm and miss rates both exhibit shallow slopes over the range of peak criteria considered. We reasoned, therefore, that 14 msec was the best of the three proposed duration criteria. Based on the empirical relationships developed in Experi-

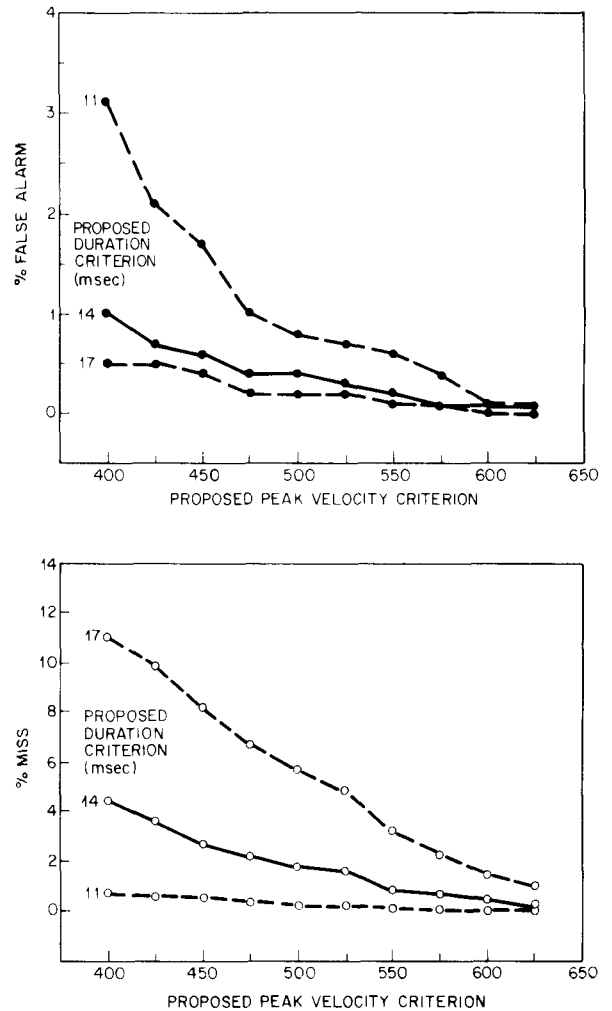


Figure 5. False alarm and miss rates in Experiment 3 that would result from various duration and peak velocity criteria. Both error rates are expressed as a percentage of all true saccades identified by hand whose peak velocities are greater than or equal to that of the peak criterion at each point.

ment 2, we judged that a peak criterion of about 450 units would detect virtually all saccades equal to or larger than four character spaces.

These duration and peak velocity criteria result in a false alarm rate of .6% and a miss rate of 2.7%. At these rates, there are too few false alarms to indicate reliable differences between noise groups. Since identical duration and peak velocity criteria are applied to all subjects, there is no a priori reason to expect that miss rates, sensitivity, or measurement accuracy should vary according to noise groups.

GENERAL CONCLUSIONS

This laboratory meets our requirements for prolonged yet unobtrusive observation of horizontal saccadic eye movements during sustained reading periods of minutes or hours. All important functions regarding data collec-

tion and data reduction are performed automatically, according to simple procedures that can be applied uniformly and without bias to nearly all subjects that we study. We anticipate that the principal benefits of this laboratory will follow from the ability to collect and to study large numbers of eye movements efficiently and conveniently. This ability will serve a number of research objectives, including the study of individual differences in eye movement patterns and reading styles, studies of persistence in reading, and the investigation of relationships between eye movements, text difficulty, and reading task requirements.

REFERENCES

- FUCHS, A. F. The saccadic system. In P. Bach-y-Rita & C. C. Collins (Eds.), *The control of eye movements*. New York: Academic Press, 1971.
- FUCHS, A. F. The neurophysiology of saccades. In R. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes*. Hillsdale, N.J.: Erlbaum, 1976.
- JUST, M., & CARPENTER, P. A. A theory of reading: From eye fixations to comprehension. *Psychological Review*, 1980, 4, 329-354.
- MARG, E. Development of electro-oculography. *Archives of Ophthalmology*, 1951, 45, 169-185.
- RAYNER, K. What guides a reader's eye movements? *Vision Research*, 1976, 16, 829-837.
- ROTHKOPF, E. Z. Analyzing eye movements to infer processing styles during learning from text. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and the higher psychological functions*. Hillsdale, N.J.: Erlbaum, 1978.
- ROTHKOPF, E. Z., & BILLINGTON, M. J. Goal guided learning from text: Inferring a descriptive processing model from inspection times and eye movements. *Journal of Educational Psychology*, 1979, 71, 310-327.
- SHACKEL, B. Eye movement recording by electro-oculography. In P. H. Venables & I. Martin (Eds.), *Manual of psychophysiological methods*. Amsterdam: North-Holland, 1967.
- YOUNG, L. R., & SHEENA, D. Survey of eye movement recording methods. *Behavior Research Methods & Instrumentation*, 1975, 7, 397-429.