Reprinted from
15 October 1971, Volume 174, pp. 307-311

SCIENCE

Extremely Rapid Visual Search: The Maximum Rate of Scanning Letters for the Presence of a Numeral

George Sperling, Judy Budiansky, Jerry G. Spivak and Martha C. Johnson

Extremely Rapid Visual Search: The Maximum Rate of Scanning Letters for the Presence of a Numeral

Abstract. *Subjects searched a rapid sequence of computer-produced letter arrays for the presence of a numeral in one of the arrays. The subjects' scanning rates were computed from their precentage of correct detections of the location of the numeral. Scanning rates were very high and approximately the same for a wide variety of conditions; the highest scanning rates (125 and 75 letters per second for two subjects) occurred when there were 9 or 16 letters in each of the arrays and when new arrays were presented every 40 to 50 milliseconds. Giving the subject advance knowledge of the numeral to be presented made little difference in the scores.*

How fast can an array of letters be scanned to see whether it contains a numeral? This is a special case of a more general question: How quickly can a subject decide that an item from set A does not belong to set B? The answers to questions of this kind can give us basic data about the processes that underlie human pattern-recognition. In his classical experiments, Neisser (*1*) attempted to answer this kind of question by means of a procedure in which subjects searched a long list of items for the presence of a critical item. Because this procedure requires subjects to make eye movements, it is open to the criticism that the limiting factor in search is the rate of eye movements rather than the rate of information processing. Alternatively, search has been studied in exposures too brief to admit eye movements (*2*). Unfortunately, the effective visual duration of a brief stimulus is difficult to estimate unless post-exposure fields composed of visual noise (which looks like scattered bits and pieces of letters) are used to obliterate the visual persistence of the stimulus (*3*), and noise fields may introduce complications of their own. Some of these problems are overcome by using reaction-time rather than detection methods to study visual search (*4*). However, the interpretation of reaction-time experiments is exceptionally difficult unless the probability of a correct response is very high, and this is a serious limitation. With the evolution of computer systems for generating sequential visual displays [and their employment in closely related contexts (*5*)], the study of detection in complicated stimulus sequences has become technically feasible. The sequential search procedure (*6*) is the logical outcome.

In the sequential search procedure, subjects view sequences of arrays. In our experiments, each array—except one critical array—is composed strictly of letters. A critical array, which contains a numeral in a randomly chosen location, is embedded somewhere in the middle of the sequence. The task of the subject is to state the location of the numeral. When the subject is not told

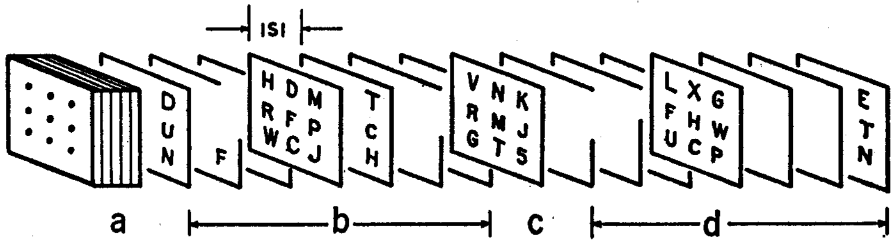


Fig. 1. Diagram of the stimulus sequence in the sequential search procedure. (a) Fixation field (1 second); (b) a randomly determined number (from 6 to 12) of letter arrays, each array having nine letters; (c) the critical array, in this instance it contains a numeral "5" in the bottom-right location; and (d) 12 more letter arrays. The interstimulus interval (ISI) is measured from onset-to-onset and is the same between all stimuli.

in advance which numeral will be presented, he must also identify the numeral.

Figure 1 illustrates a display sequence of nine-letter arrays. The correct response would be "5, bottom row, right." The arrays are generated by a DDP/24 computer (*7*), produced on a cathode-ray oscilloscope, and viewed with normal binocular vision (*8*). Subjects sit at the console, and, at the end of each display sequence, they type a number corresponding to the numeral and its location, making a response even when they are uncertain.

To analyze the data, we consider the subject's report of the location because only this report will later enable us to compare the case of a known numeral and an unknown one-of-ten numerals. We suppose that a subject scans some average fraction p of the L locations in each array; thus he scans pL locations. The subject does not know when in the sequence the critical array will arrive nor in which location the numeral will occur. We assume that, when the numeral occurs in one of the pL scanned locations, it is reported correctly; if it occurs in one of the $(1 - p)$L other locations, its location must be guessed. To be conservative we assume that the subject uses the optimal guessing strategy; he scans a particular set of locations in all the arrays of a sequence. Therefore, he knows at least one of the locations he fails to scan so that when he must guess, he guesses only unscanned locations. Thus he needs to scan only L − 1 of the L locations to achieve a perfect score. The observed probability p_0 of being correct then is given by $p_0 = p + 1/L$, and the estimate $\hat{p}$ of p from the data is $\hat{p} = p_0 - 1/L$. Insofar as subjects cannot be quite so efficient, $\hat{p}$ will underestimate performance slightly (*9*).

The estimated *number* of locations per array scanned by the subject is pL. From $\hat{p}$L and ISI (the interstimulus interval between successive arrays) we obtain the estimated scanning time per letter τ, $\tau = \text{ISI}/(\hat{p}\text{L})$. The reciprocal of τ is the estimated scanning rate.

Data of two practiced subjects searching for a known numeral are illustrated in Fig. 2. Each data point is based on the average of about 60 trials, 30 trials in which the task was to detect a "2" and 30 in which the task was to detect a "5" (*10*). Figure 2, a and b, illustrates the letter scanning time versus the time-interval between arrays; Fig. 2, c and d, illustrates the estimated number of locations scanned ($\hat{p}$L). In these data we note the following.

1) Low and fairly constant scanning times (high rates) occur for a wide range of array sizes (from 9 to 25 letters) and for a wide range of ISI's (from about 40 to about 160 msec). The constant τ means that subjects scan either a few letters from each of many arrays (presented rapidly) or many letters from each of a few arrays (presented slowly) at an exactly equivalent scanning rate. For the faster subject (J.G.S.), the low scanning time of less than 10 msec per letter corresponds to a scanning rate of over 100 letters per second.

2) At long ISI's, search is most efficient—in the sense of number of locations scanned—for large arrays (16 and 25 letters). The number of locations scanned reaches its maximum value of about 14 locations in 160 msec (J.G.S.) and 10 locations in 200 msec (M.C.J.) (Fig. 2, c and d). In his best runs, subject J.G.S. was able to scan 16 locations. Here, as in all his runs, no further improvement resulted upon increasing the ISI from 160 to 320 msec.

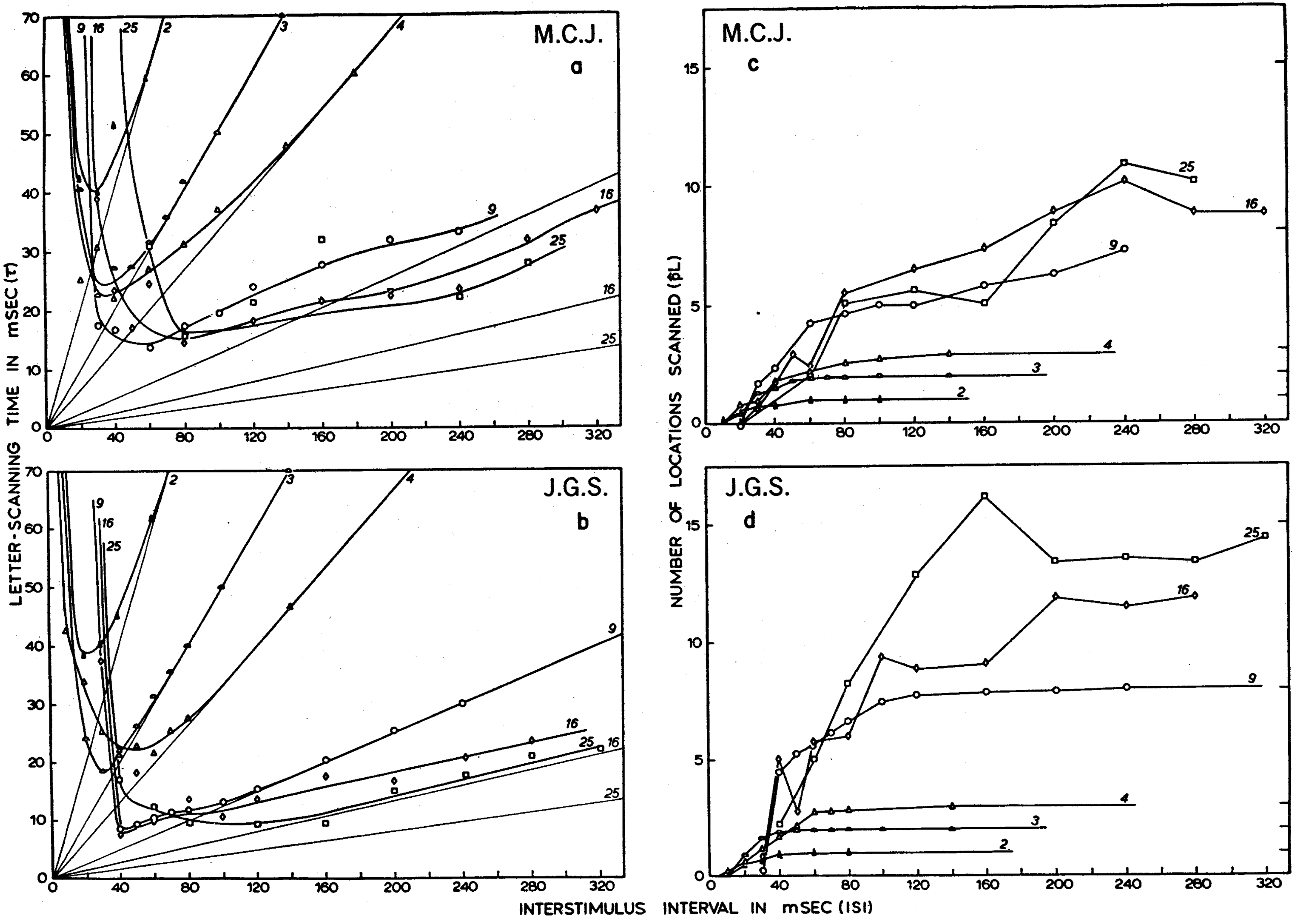
3) At short ISI's, small arrays (2, 3, and 4 letters) are searched most efficiently. The τ versus ISI curves for different sizes of arrays systematically cross at short ISI's (Fig. 2, a and b). That the curves cross (that is, are nonlaminar) indicates that the subject's search strategy changes (see below).

4) The consistently fastest scanning (minimum of minima of the τ versus ISI curves) occurs for arrays of 9 and 16 letters presented at ISI's of 40 msec (subject J.G.S.) and 60 msec (subject M.C.J.). The minima of the τ versus ISI curves are sensitive to the effects of practice, tending to decrease with practice. Thus, on her final runs, subject M.C.J.'s "minimum of minima" also occurred at 40 msec. Using data only from the final run to estimate the fastest scanning rates gives rates of about 1 letter per 8 msec (125 letters per second) and 1 letter per 13 msec (75 letters per second) for the two subjects, respectively. These rates occur with 9- or 16-letter arrays, or both, at ISI's of 40 to 50 msec and correspond to scanning of 5 to 6 and 3 to 4 locations, respectively, of the array (*11*).

We can exclude interpretations of the high scanning rates in terms of scanning of alternate arrays; a value of $\hat{p}$ greater than .5 suffices for this and $\hat{p} = .6$ (subject J.G.S.). In general, correct detections of the critical digit are clustered in a subset of scanned locations. This nonuniform error distribution is equivalent to an increase in $\hat{p}$ for a subset of the array, and establishes that approximately the same locations of every consecutive array are scanned.

To test whether the high scanning rate can be maintained when the subject must scan for an unknown one-of-ten numerals, subject M.C.J. was given extensive additional tests with the 9-letter stimulus and a 60-msec ISI (the condition that previously gave the maximum scanning rate). There were two blocks of 50 trials with each of the numerals 0 to 9, in which she was informed of the numeral in advance of the block of trials. There were six blocks of 50 trials in which the critical numeral was unknown to her (it could be any of the ten numerals). The 1300 trials were conducted in a counterbalanced series of 26 blocks with about 50 practice observations at the start of each block. Because the

Fig. 2. Estimated letter-scanning time (τ) and estimated number of locations scanned ($\hat{p}$L) as a functon of the time interval (ISI) between successive arrays. Data are shown for two subjects—a typical subject (M.C.J.) and a "rapid" scanner (J.G.S.). Data points obtained with arrays of the same size are connected; solid triangle, 2; half circle, 3; open triangle, 4; circle, 9; diamond, 16; and square, 25 letters, respectively. The theoretical limits set by perfect performance are indicated by the light straight lines (a and b) and by the scale marks on the right-hand ordinate (c and d).


LETTER-SCANNING TIME IN mSEC (τ)
NUMBER OF LOCATIONS SCANNED (βL)
INTERSTIMULUS INTERVAL IN mSEC (ISI)
M.C.J.
a
J.G.S.
b
M.C.J.
c
J.G.S.
d

subject's performance was still improving slightly at the conclusion of this series, another series of 1000 trials was conducted with unknown one-of-ten numerals. In all these tasks, the analysis considers only reports of the location of the numeral (location scores) and not reports of the numeral itself (item scores). In fact, the location scores and the item scores are highly correlated.

In the counterbalanced series, which enables a direct comparison between numeral-known and numeral-unknown conditions, the subject scored slightly (but significantly) better when the numeral was known in advance ($\hat{p} = .46$ versus $\hat{p} = .38$). Performance improved slightly between series, and in the second series (in which only "unknown" numerals were tested) the final value of $\hat{p}$ is .47, equivalent to a scanning time of $\tau = 14$ msec. The numeral-by-numeral correlation between scores with known and unknown numerals is 0.97 (Table 1). We conclude that overall performance with a known numeral is slightly better than with an unknown one (although this advantage possibly may disappear with practice) and that the relative difficulty of each numeral individually is independent of whether the subject knows in advance which numeral will be presented.

An important result in Table 1 is the wide range of difficulties of individual numerals; $\hat{p}$ varies from .019 to .719. By selecting just one numeral to compare with an unknown one-of-ten numerals, one might conclude that advance knowledge of the numeral was helpful, of no importance, or harmful to performance, depending on which particular numeral happened to be chosen. To come to the correct conclusion (slightly helpful), it was necessary to compare all ten numerals in the numeral-known condition with those in the unknown one-often condition. Earlier, Fig. 2 showed data obtained with a particular known numeral, but we see from Table 1 that they were typical of the results with an unknown one-of-ten numerals; therefore, the results of Fig. 2 tentatively may be extrapolated to an unknown one-often numerals.

The scanning times of 8 to 14 msec per letter estimated from these search experiments are remarkably close to the scanning times of 10 to 15 msec per letter estimated from recall experiments in which a letter array is followed by an array of visual noise (*3*). We propose that the similarity of the letter scanning time in searching letters for a known

Table 1. Comparison of the estimated probability $\hat{p}$ of correctly detecting the location of (i) a known numeral and of (ii) an unknown one-of-ten numerals. Nine-letter arrays were presented at ISI's of 60 msec; approximately 100 trials were used to estimate each $\hat{p}$.

Numeral	Known	Unknown
0	.019	.011
1	.344	.296
2	.572	.636
3	.642	.691
4	.206	.293
5	.479	.449
6	.662	.646
7	.429	.406
8	.719	.648
9	.572	.633
Mean (series 1)	.464	(.383)*
Mean (series 2)		.469

* Data for these 300 unknown trials of series 1 are not included in the analysis.

numeral, in searching letters for an unknown one-of-ten numerals, and in the recall of letters is due to subjects' inefficient use of partial information in the search experiments. (Partial information consists of fewer features—for example, lines of a letter—than are required for identification, such as, only the vertical lines of H, N, and U.) Because subjects know neither the particular array nor the spatial location in which the numeral will appear, partial information gained from an incompletely analyzed numeral is immersed in the partial information from incompletely analyzed letters—not only letters from the same array but also from many other arrays. The information deluge would lead to multiple false detections (which are difficult to process) unless the subject sets a high criterion for his recognition response. Because the subject demands virtually complete information for a response, processing times for known and unknown numerals are similar, that is, the subject is doing the same complete analysis in both conditions. Complete analysis, of course, suffices even for recall—hence the similar scanning rate in recall tasks.

To further analyze the data, it is useful to distinguish three levels of mental processing: P1, processing of a single character; P2, processing of characters in the same array; and P3, processing of characters from consecutive arrays. We can ask to what extent each of these processes, individually, is composed of successive operations (serial processing) or simultaneous operations (parallel processing). For example, in an entirely serial model, the processing of a single character (P1) might consist of the serial comparison of the representation of the unknown character with the memory representation of each of the ten numerals (*12*). This process ends when the unknown character is identified as a particular numeral or as a nonnumeral, and processing of another character from the same array begins (P2). Once the next array appears, processing of characters from the original array is terminated, and processing of the first new character begins (P3).

The data of the present experiments, while they do not suffice to exclude or to prove any particular model, can most plausibly be accounted for by assuming that both P1 and P2 are parallel processes (*13*). With respect to P1, the great similarity of results in the numeral-known and numeral-unknown conditions (Table 1) implies that the same processing occurs in both conditions. This similarity is predicted naturally by a parallel processing model of P1 (all memory representations are compared simultaneously with the unknown character) but only with considerable complications by serial processing models.

With respect to P2, simple serial models predict that when either array size or ISI is increased, the number of locations scanned increases. (Giving the subject more opportunities for successful scanning—more letters to scan or more time to scan them—should improve his performance.) In fact, when array size is increased, the nonlaminar τ versus ISI curves (Fig. 2) imply that the subject does not simply do more scanning, rather, increasing array size is detrimental to performance when ISI is small and helpful when ISI is large. When ISI is increased, the steep negative slopes of the τ versus ISI curves imply that the subject does not simply do more scanning, instead he derives a special benefit when the ISI is sufficiently long (≥ 40 msec) to enable him to scan 3 to 4 letters. A plausible interpretation of these data is that P2 is a parallel process, and that the two subjects, respectively, scan at least 3 to 4 and 5 to 6 locations in parallel from the same display. Scanning efficiency is impaired when either (i) fewer letters are presented than the smallest number the subject normally scans in parallel or (ii) the array is presented for less than the 40 msec needed to scan this number of letters. The data suggest that the remaining locations (up to a total of 10 to 14) also are scanned in parallel but less efficiently. However, scanning does not proceed independently at each location. If it did, $\hat{p}L$ would increase in proportion to array size—which it obviously does not do.

Sternberg and Scarborough (*6*) showed that, in arrays of size one, consecutively occurring numerals are processed at overlapping times, that is, that P3 is partly parallel and not purely serial. Taken together with our results, this means that all component operations of scanning overlap each other in time. Scanning may be concentrated on either a few locations or spread out over many; for a considerable range of conditions the overall scanning rate is approximately constant (*13*). For the two subjects, the maximum overall rate varies from about 75 to 125 letters per second (equivalent to the discrimination of 1 letter from a numeral every 8 to 14 msec).

What is the role of eye movements in visual search? In visual search, as in nearly all visual tasks, the eyes make quick saccadic movements at rates not exceeding about 4 or 5 movements per second, and thereby the eyes effectively transform the visual input into a sequence of stimuli with ISI's of 200 to 250 msec (*14*). Even these minimum times between eye movements (for example, 200 to 250 msec) are five times longer than the ISI's that produce the fastest scanning rates in the present experiment (40 to 50 msec). When arrays are presented at ISI's comparable to those between eye movements, scanning rates generally are much lower than the maximal rates—in fact, the data imply that most scanning occurs during the first half of the ISI (Fig. 2, c and d). Therefore, in the simple search for a numeral among letters, the rate of eye movements is a factor that significantly limits the rate of search, and a method like the sequential search paradigm is needed to estimate a scanning rate that is unconfounded by eye movements. Can replacing the sequence of stimuli generated by eye movements with a sequence of stimuli generated by a computer result in a comparable improvement in performance in complex visual tasks such as reading? That is, can eye movements gainfully be replaced by stimulus movements, even when the rate of processing rather than the rate of eye movements is the limiting factor? These are important practical questions, and the sequential presentation procedures offer the means to answer them.

GEORGE SPERLING, JUDY BUDIANSKY
JERRY G. SPIVAK, MARTHA C. JOHNSON
Bell Telephone Laboratories,
Murray Hill, New Jersey 07974,
and *New York University,*
Washington Square,
New York 10003

References and Notes

1. For a summary, see U. Neisser, *Cognitive Psychology* (Appleton-Century-Crofts, New York, 1967).
2. V. D. Glezer and A. A. Nevskaia, *Dokl. Acad. Sci. USSR* **155**, 711 (1964); W. K. Estes and H. A. Taylor, *Proc. Nat. Acad. Sci. U.S.* **52**, 446 (1964); W. K. Estes and H. A. Taylor, *Percept. Psychophys.* **1**, 9 (1966); P. Shaw, *ibid.* **6**, 257 (1969).
3. G. Sperling, *Hum. Factors* **5**, 19 (1963); *Acta Psychol.* **27**, 285 (1967); P. Liss, *Percept. Psychophys.* **4**, 328 (1968); D. A. Allport, *Psychonom. Sci.* **12**, 231 (1968).
4. W. K. Estes and D. L. Wessel, *Percept. Psychophys.* **1**, 369 (1966); R. S. Nickerson, *J. Exp. Psychol.* **72**, 761 (1966); S. Sternberg, paper presented at the Eastern Psychological Association meeting in Boston, Mass. (1967); R. C. Atkinson, J. E. Holmgren, J. F. Juola, *Percept. Psychophys.* **6**, 321 (1969).
5. M. S. Mayzner, M. E. Tresselt, N. Tabenkins, R. Didner, M. S. Helfer, *Percept. Psychophys.* **5**, 297 (1969); R. N. Haber and L. S. Nathanson, *ibid.*, p. 359; V. Zinchenko and G. Vuchetich, *Science in Progress* (Novosti Press Agency), issue 12 (Dec. 1969), p. 25.
6. G. Sperling, in *Early Experience and Visual Information Processing in Perceptual and Reading Disorders*, F. A. Young and D. B. Lindsely, Eds. (National Academy of Sciences, Washington, D.C., 1970), p. 198. An alternative term is "simulated search," proposed by S. Sternberg and D. L. Scarborough, *Visual Information Processing and Control of Motor Activity* (Bulgarian Academy of Sciences, Sofia, in press).
7. J. Budiansky and G. Sperling, *Internal Memorandum, Bell Telephone Laboratories* (1969), p. 153.
8. The viewing distance was 40 inches (1 inch = 2.54 cm); subjects maintained fixation at the center of the array. Letters were 0.4 inch high and spaced apart (center-to-center) 0.65 inch horizontally and vertically. The letters O, B, I, S, Z, and Q were not used. Letters were composed of 22 points (on the average) and illuminated briefly to a luminous directional energy of 1.1 candle-microseconds per point upon a uniform background of 1.6 footlambert (1 footlambert = 1.076 mlam) [see G. Sperling, *Behav. Res. Methods Instrum.* **3**, 148 (1971)].
9. The difference in the estimated number of letters scanned ($\hat{p}L$) by assuming the most efficient or the most inefficient guessing strategy is $\hat{p}/(L-1)$, which is small for large arrays.
10. Within a daily session, array size was held constant and ISI was varied. Sessions and ISI's within sessions were conducted in a counterbalanced order; all of the sessions for detection of "2" occurred before any of the sessions for detection of "5." After every trial, subjects were given knowledge of results.
11. C. W. Erickson and T. Spencer [*J. Exp. Psychol. Monogr.* **79**, 1 (1969)] in formally similar experiments obtained totally different results. In their experiments, a critical item remained detectable even at exeedingly high rates of presentation. In our experiments, such effects occur only when insufficient precautions have been taken to ensure that the characters that follow a critical item overlap every part of the critical item.
12. S. Sternberg, *Acta Psychol.* **30**, 276 (1969).
13. D. Rumelhart, *J. Math. Psychol.* **7**, 191 (1970); see also Sperling (*6*).
14. B. Erdmann and R. Dodge, *Psychologische Untersuchungen über das Lesen auf experimenteller Grundlage* (Niemeyer, Halle, 1898); R. S. Woodworth and H. Schlosberg, *Experimental Psychology* (Holt, New York, 1954).
15. We acknowledge an award of the John Simon Guggenheim Memorial Foundation to G. Sperling. Individual responsibility for the research was divided as follows. Most of the computer programing was done by Mrs. Judy Budiansky; Mrs. Martha C. Johnson tabulated the data and served as a subject; J. G. Spivak ran some preliminary experiments and served as a subject; G. Sperling did the rest.

17 March 1971