Sperling, G., & Melchner, M. J. (1976). Visual search and visual attention. In V. D. Glezer (Ed.), *Proceedings of the Fourth Symposium of Sensory System Physiology: Information Processing in Visual System* (pp. 224 - 230). Leningrad: Academy of Sciences, Pavlov Institute of Physiology.

VISUAL SEARCH AND VISUAL ATTENTION

George Sperling
Bell Laboratories, Murray Hill, New Jersey
and
New York University, New York, New York
and
Melvin J. Melchner
Bell Laboratories, Murray Hill, New Jersey

In our laboratory we are studying the basic mechanisms of attention with particular reference to vision. The primary mechanisms of visual attention are of course overt: eye movements and body orientation. But even within a single eye fixation, attentive processes determine what particular kind of signals are analyzed and from what parts of the visual field they are accepted.

First we consider some examples from visual search. In normal visual search, the subject searches an array of objects (background objects) for a critical object (target) by moving his eyes over the array. While the pattern of eye movements is interesting in itself (Yarbus) it merely adds a complication (not under experimental control) to the analysis of attention. Therefore, in our experiments, we eliminate eye movements by having the subject keep his eyes fixated in the center of a display, and presenting new stimuli to him every t msec. This method gives the experimenter precise control over the flow of information to the visual system. When t is 240 msec, this display sequence simulates, approximately, the sequence which the eyes produce for themselves in natural visual search.

A typical paradigm is illustrated in Fig. 1. The subject sees first a fixation field, presented for one second, followed by a sequence of character arrays. Each character array is presented as a brief flash lasting a fraction of a millisecond. The arrays are clearly visible; in other experiments we have found no difference in performance between the brief flashes and arrays presented continuously for 200 msec. One of the character arrays, the critical array, contains the target characters. It is preceeded by a random number (from 7 to 12) of noncritical arrays and followed by at least 12 noncritical arrays. The subject does not know which array contains the target characters, nor what the particular target(s) will be, nor where in the array they are located. His task is to report the identity and location of target characters, and his degree of confidence in the correctness of his report. The data are analyzed in terms of *l* the average number of characters the subject scans in each array and τ, the average time it takes to scan one character.

In previous studies (Sperling, et al., 1971) we studied the ability of a subject to search for a single target numeral when the background characters were letters. Some salient results and conclusions were:

(1) Telling the subject in advance which particular numeral (e. g., «5») will be presented on each one of a long series of trials does not improve his ability to detect that numeral, relative to its detectability when the subject is searching for an unknown one-of-ten numerals. A related observation is that when the subject detects a target, he detects both its location and its identity. That is, he does not report seeing an unknown numeral in a particular location or seeing a particular number in an unknown location. Detection of a numeral amongst letters implies both: identity and location information.

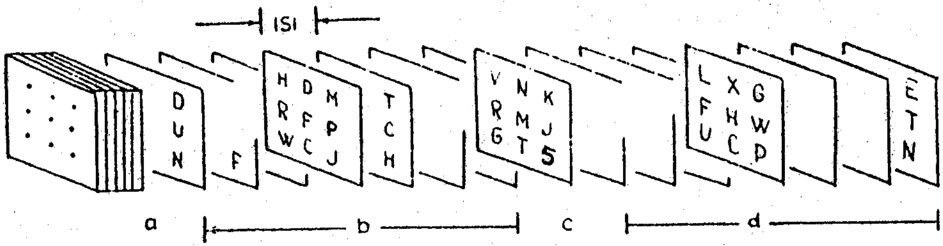

Fig. 1. Sequence of consecutive displays in the search experiment. (*a*) Fixation field; (*b*) 7 to 12 noncritical displays; (*c*) critical display with single target numeral; (*d*) 12 noncritical displays; (ISI) interstimulus interval, varied between sessions from 5 to 480 msec.

(2) The maximum number of characters a subject can scan in an array (max *l*, his «span») is about 15—25 characters. Increasing the number of characters in the array beyond a subject's span does not improve his performance.

(3) Subjects approach their asymptotic performance when arrays are presented every 120 msec; increasing t to 240 msec improves performance only slightly; increases in t beyond 240 msec are of no benefit whatsoever. A corollary conclusion is that when a subject searches large arrays naturally by means of eye movements, (i. e., corresponding to a new input about every 240 msec.) his processing capacity is unused for about half of the time (between 120 and 240 msec.).

(4) The most efficient search (min [τ]) occurs when new arrays occur every 40 to 50 msec (corresponding to 20 to 25 fresh

arrays per second). In these presentations, most subjects achieve a τ of less than 10 msec, corresponding to scan rates in excess of 100 characters per second. When arrays are more closely spaced than every 40 msec, performance deteriorates rapidly.

(5) The conclusion from these and related studies is that subject search in parallel for 15—25 characters in an array. There is some sharing of processing capacity between nearby locations of the same array (so that enlarging an array by adding new characters does not improve performance as much as would be expected if the new characters were processed entirely independently.)

In an attempt to produce an optimum array of characters for visual search, we constructed an array in which the number of characters in each local area was matched to the density of information processing capacity of that area. This array consisted of very small characters in the center, surrounded by larger and larger characters in each successively larger ring around the center. To our surprise we found that performance was not optimal but awful, and this led us to investigate the question of whether subjects can search in parallel for targets of different sizes.

The experimental paradigm was the same as before except that the critical array now contained two target numerals, a large one and a small one, chosen independently, and the subject's task was to report both numerals, both locations, and both confidences. The spatial arrangement of the small characters in the **inside** and the large characters in the **outside** of an array is illustrated in Fig. 2. The size and number of the **inside** and **outside** numerals were chosen so that in control experiments, in which the task of the subject was to report only **inside** or only **outside** numerals, the probability of a correct report was approximately matched in the **inside** and the **outside**.

Two other conditions were investigated, **Noise**, and **Reversal**. In the noise condition, the **inside** was composed of large characters (the same size as the **outside**) but detection of a critical numeral was made comparably difficult by superimposing a randomly-chosen, squiggly line segment («noise» segment) to each **inside** character. (In the **reversal** condition, background characters in the **inside** were taken to be numbers and the target was a letter.)

In all conditions, the **outside** was the same. The background characters were all the letters of English except that B, S, Z, Q, O, I were omitted because of their similarity to the numerals 8, 5, 2, 0, 0, 1. In the reversal condition, which was studied separately after the others, the numerals 0 and 1 also were omitted.

Some data from a typical subject viewing the display with **Small** characters are illustrated in Fig. 2a. The ordinate indicates the probability of correctly identifying the small target numeral (from the **inside**); the abscissa indicates the probability of cor-

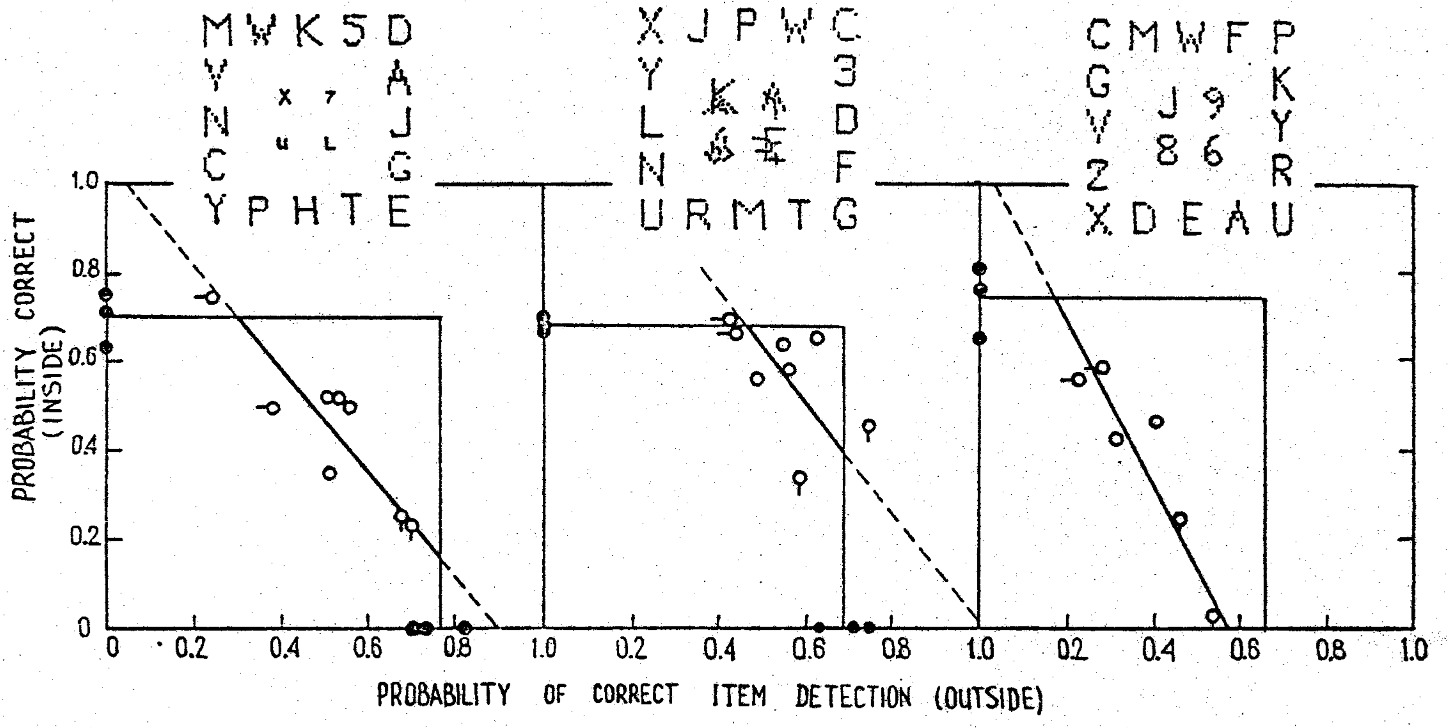

Fig. 2. Data from divided attention experiments. Photographs indicate the stimuli. Each open circle represents data from one session (30—60 trials); each filled circle represents data from one control session. The diagonal lines are best fits to the data, the darkened portion represents the Attention Operating Characteristic (AOC).

rectly identifying the large target numeral (from the **outside**). In these experiments both numerals always occurred in the same array.

Each point in Figure 2 represents data from a different session. In some sessions the subject was instructed to give most of his attention to the large **outside** characters, in other sessions to the small **inside** characters, and in still others he was told to pay equal attention to both. Figure 2 indicates that indeed he was able to follow these instructions, and that he was able to «trade-off» performance on one class of targets against the others. The range of performances of which he is capable, as he varies his attention from being devoted entirely to the small targets to entirely to the large targets, defines his «attention — operating — characteristic» that is, his AOC for this task. In this task, the AOC is approximately a straight line with slope of -1, indicating that the subject can exchange a certain amount of probability on one task (Δp_1) for an equal amount on the other (Δp_2).

In control conditions, the subject was told to report just one kind of target (e. g., the **outside** target) for the entire session. These data are graphed directly on the axes of Fig. 2. A vertical line is drawn through the mean of the **outside** control data, and a horizontal line through the **inside** mean. The intersection of these two points defines the «independence point», the point at which the subject would operate if he could perform both search tasks simultaneously without any interference, i. e., independently of each other. Insofar as the AOC lies inside the independence point, it represents some degree of interference between the two tasks.

Interference between two search tasks does not occur because of any memory deficit. To prove this the display was altered so that the targets remained the same but each background character was replaced with just a single dot. In this case, subjects gave errorless reports of both targets. Thus, the subject's inability to report both targets in the experimental condition is due to the mutual interference of the two search tasks.

Figure 2b illustrates performance in the **Noise** condition. The **outside** search task was the same as in the **Small** condition; the **inside** search task was matched to be of equal difficulty. Nonetheless, we see here that the AOC curve is closer to the independence point. This subject can carry out these two search tasks (with targets of equal size) with very little mutual interference.

Figure 2 c illustrates performance in the **Reversal** condition, where the subject searches for a letter target among numerals on the **inside**, and for a numeral target among letters on the **outside**. The data show the mutual incompatability of these two search tasks is nearly total.

In order to examine the mechanism by which the subject moves along the AOC curve, that is, the mechanism by which attention is shifted from one search task to the other we must examine the 2×2 contingency table (Fig. 3) which gives the joint occurrences of correct reports on the two tasks. This table has three degrees of freedom: two of these, the marginals, P_1 and P_2 are used to make the AOC. The third degree of freedom (correlation) provides information about the mechanism of attention. We consider here two (of many) possible mechanisms: **sharing** and **switching**. The mechanisms are outlined below without formal derivations.

In pure sharing, attention is assumed to be divided between the two tasks in some fixed proportion, which does not vary from trial to trial. Insofar as there is less attention devoted to each task than in a control condition, performance suffers relative to the control. However, the 2×2 contingency table is assumed to show statistical independence.

		Task 1		
		Wrong	Right	
Task 2	Wrong	$(1-p_1)(1-p_2)-c$	$p_1(1-p_2)+c$	$(1-p_2)$
	Right	$(1-p_1)p_2+c$	p_1p_2-c	p_2
		$(1-p_1)$	p_1	

Fig. 3. Contingency table for the joint probability of correct responses in a divided attention experiment. The algebraic expressions represent the predictions of pure «shared» attention (statistical independence of the two tasks) when c is zero, and represent some degree of «switching» for $c>0$.

In attention switching, two different attention states (S_1, S_2) are assumed to occur randomly, from trial to trial, in the search task. (Ideally, S_1 and S_2 are states of pure shared attention, but this is not a necessary assumption.) To move along the AOC curve, the subject varies the proportion of times he is in S_1. Two interesting properties of attention switching are (1) mixtures of S_1 and S_2 produce a straight line AOC curve connecting S_1 and S_2, and (2) the contingency table for a mixed state is the mixture of the two separate contingency tables (S_1 and S_2). From (2) it can be shown that, under the conditions of the experiment, any contingency table produced by switching between states with purely independent tables has a negative correlation and shows nonindependence by the Chi-square test. With the aid of some additional assumptions, one can estimate the particular two states between which attention is being switched.

When this analysis is applied to the data described above, we discover that the major mechanism of altering attention is switching i. e., altering the proportion of times the subjects is in S_1 or S_2. On the other hand, we can also reject the hypothesis, for at least some of the data, that there are only two attentional states (i. e., the states determined by the intersections of the AOC curve with the control condition lines). Thus attention sharing is not merely a mechanism for generating the end points of the AOC but is also a mechanism for moving along an AOC.

In conclusion, we see that subjects cannot simultaneously search for a large and a small target as well as they can search for equal-sized targets. The instruction to give equal attention to the search for two different-sized targets causes the subject to switch his attention from trial-to-trial between searching primarily for large and searching primarily for small targets.

In a more general vein, we remark that the AOC is a general way of studying attention and particularly of describing the compatibility of two tasks. A pair of tasks to be performed simultaneously determines an AOC. To compare two pairs of tasks, one cannot just use one condition of attention for each pair, as this would be comparing one point from each of two curves and not comparing two curves. (An analogous problem occurs in signal detection theory with ROC curves.) The mechanism by which a subject varies his performance along an AOC curve in the tasks we studied was primarily by switching attention between extreme states, but some sharing of attention also occurred.

References

1. Sperling G., Budiansky, Judy, Spivak, Jerry G. and Johnson, Martha C. Extremely Rapid Visual Search: The Maximum Rate of Scanning Letters for the Presence of a Numeral. Science, 1971, 174, 307—311.
2. Yarbus A. L. Eye Movements and Vision. (Translated by B. Haigh). New York: Plenum Press, 1967.