Some characteristics of selective attention in visual perception determined by vocal reaction time*

CHARLES W. ERIKSEN and JAMES E. HOFFMAN
University of Illinois, Champaign, Illinois 61820

Previous research had found that the accuracy with which an S could report an indicated letter or target in a briefly exposed multiletter display decreased as the number of irrelevant letters increased, and accuracy increased if the position of the target letter was indicated 150 msec before the display was presented. In the present experiment, these variables were reinvestigated using vocal reaction time as the dependent variable. An interpretation of the accuracy measures in terms of differential processing times was supported. The results of the two experiments were discussed in terms of a model of attentional selectivity.

An essential characteristic of a concept of selective attention is the ability to respond in a predetermined or preset way to only one or a small subset from a number of equally potent stimuli. The operation of selective attention in visual perception can be demonstrated in the following experimental arrangement: If a large number of letters or digits are simultaneously exposed for a brief duration, the human O is typically able to report only a small number, approximately four. Under the same conditions, if the O is told to report only the stimulus designated by a black bar or similar indicator, he can, by some process, select this particular stimulus and report it with perfect accuracy. A necessary condition for this demonstration is the overload of information provided the O. The existence of a limited channel capacity is implied by the concept itself. Without such a limit, there would be no necessity for selective attention, as all stimuli would then be processed with equal accuracy at all times.

Selective attention may be conceived as the programming by the O of which stimuli will be processed or encoded and in what order this will occur. Experimentally, the task becomes one of determining how the programming is carried out and the variables of which it is a function.

In prior experiments using multistimulus displays like those described above, we have determined some of the characteristics of the selection process. It requires time. On the order of 100 msec is required to respond to a black line indicator designating one of the display stimuli (Eriksen & Collins, 1969). We also

have found that the accuracy of the selectivity process is a function of the number of irrelevant stimuli in the display as well as their physical spacing (Eriksen & Rohrbaugh, 1970). Further, we have obtained evidence that when more than one stimulus is to be reported in such displays, the encoding of the separate stimuli appears to be serial in nature (Eriksen & Colegate, 1971).

In all these prior studies, the dependent variable was the accuracy with which the Ss were able to report the indicated display element or stimulus. The experimental technique was to present tachistoscopically the multielement displays at an energy level that led to above-chance accuracy but appreciably less than 100% correct. Thus, it was possible for the experimental variables to either improve or impair performance. While variations in accuracy permit inferences as to the speed of underlying perceptual processing, the relationship is not logically a necessary one. As we have pointed out elsewhere (Eriksen & Eriksen, 1971), to perceive more clearly or accurately does not necessarily require that the perception occurred more rapidly.

In the experiment reported in the present paper we have modified our experimental arrangement in order to use voicing latency as a dependent variable. Sternberg (1969) has shown most impressively the analytic power that can be obtained with latency or reaction time measures. It seemed highly desirable to determine if selective attention was amenable to this type of measurement.

Here we investigated several variables that had been found to influence the accuracy with which an indicated element could be processed from multielement displays exposed tachistoscopically. Again, a black line indicator was used to designate the display element S was to report, but,

instead of limiting the energy of the displays, they were presented and remained on until S reported. Instead of accuracy, which in the present experimental arrangement was in excess of 98%, Ss were instructed to speak aloud the designated element as quickly as possible, and the latency from display onset to initial voicing of the designated element was recorded. Previous research (Eriksen & Rohrbaugh, 1970) had shown that accuracy in reporting the indicated element decreased as the total number of elements simultaneously presented in a display increased from 4 to 12. On the other hand, accuracy was found to increase if the black line indicator preceded presentation of the display by as much as 150 msec. These, then, were the variables selected for reinvestigation in the present experiment.
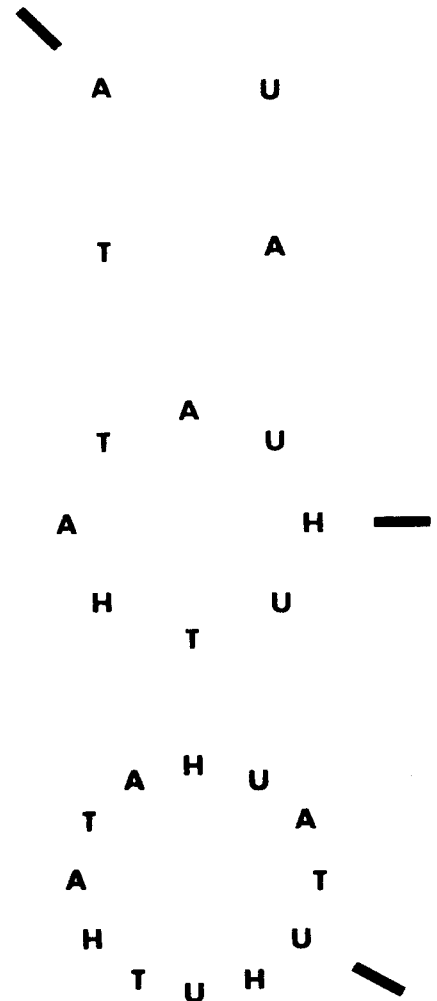

Fig. 1. Examples of 4-, 8-, and 12-element displays with indicator as they appeared to the S.

METHOD
Subjects

Four students, one male, served as paid Ss. All had normal or corrected-to-normal vision.

Apparatus and Stimuli

A Scientific Prototype Model GA three-field tachistoscope was employed with Sylvania FT45/CWX lamps. Luminances of the three fields were set at 6 mL, as measured with a Spectra spot photometer. A small black fixation cross was in Field 1, the adaptation field, which remained on except when the indicator and/or the stimulus displays were presented in Fields 2 and 3. Initiation of the trial started a Hunter Klockounter. A microphone located below the S's viewer triggered a Scientific Prototype No. 761-G voice relay, stopping the Klockounter. Voice reaction time was recorded in milliseconds.

The stimuli consisted of the capital letters A, T, U, and H (black Paratype No. 11316), which subtended .2 deg of visual angle in height. Target displays consisted of 4, 8, or 12 letters on a white plastic card. For all display sizes, letters were positioned around an imaginary circle, 2 deg of visual angle in diam, which was centered on the fixation cross.

In order to reduce materially the number of displays to be constructed, the following procedure was employed. For the 12-letter displays, the letters were positioned randomly in locations corresponding to the 12 clock positions, with 1 position remaining blank. Twenty-four such display cards were constructed, with the blank space occurring twice in each of the 12 clock positions. Next, clear plastic cards containing each of the 4 target letters in each of the 12 positions were constructed, and when one of these was placed in front of the corresponding 11-letter display card, a uniform 12-letter display resulted. A similar procedure was used in constructing the 8-letter displays. Here the letter positions corresponded to 12, 3, 6, and 9 o'clock, with the remaining 4 letters occurring midway between these positions. As before, 1 position was left blank, and clear plastic cards containing the target letter in the blank location were placed before the display when presented to S. Again, 24 display cards were constructed such that the blank position occurred in each of the 8 positions equally often. For the 4-letter displays, the letters occurred in positions midway between 1 and 2, 4 and 5, 7 and 8, and 10 and 11 o'clock, and again 24 display cards were constructed so that the blank position occurred in each of the 4 possible positions equally often. In the

12-letter displays, interletter distance was .57 deg of visual angle and in the 8- and 4-letter displays, .75 and 1.4 deg of visual angle, respectively.

The indicator was a black bar, .5 x .1 deg of visual angle, positioned outside the display circle, with its near end .58 deg of visual angle from the indicated letter. It was located on an extension of an imaginary radius from the center of the display through the indicated position. For the condition in the experiment where indicator and display occurred with simultaneous onset, the indicator was mounted as an integral part of the display card. For the leading indicator condition, where the indicator appeared 150 msec before onset of the display, the black line indicators were mounted on white plastic cards which were presented in Field 11 of the tachistoscope 150 msec before onset of the letter displays. Figure 1 shows typical 4- (top), 8- (middle), and 12-letter (bottom) displays with indicators as they appeared to the Ss.

Procedure

There was a total of six conditions: three display sizes and simultaneous or leading indicator. In each of the experimental sessions, Ss saw a block of 16 trials under each of the six conditions. The two indicator conditions were paired within a display size such that an 8-letter display with simultaneous indicator was followed by an 8-letter display with leading indicator. Order of presentation of display size and indicator conditions within display size were counterbalanced across sessions within Ss.

Prior to the experimental sessions, all Ss were run two practice sessions to introduce them to the task. In each experimental session, each S was given 5 min to dark adapt followed by eight practice trials of the first condition to be run that day. Ss were instructed to initiate the trial when they had the fixation cross in clear focus and to vocalize the indicated letter as quickly as possible while maintaining a low error rate. Error rates averaged less than 2% and varied from 1% under the four-element simultaneous condition

to 3% under the eight-element leading indicator condition.

Since Ss initiated trials, a control was needed for the 150-msec difference in display onset between the leading and simultaneous indicator conditions. To render the delay comparable under the simultaneous condition, a white card containing a black dot, .22 deg of visual angle and located in the center, occurred in Field 2 for 150 msec before onset of the letter display.

RESULTS

In Table 1, mean voicing reaction times are shown for each of the three display sizes under the leading and simultaneous indicator conditions. A three-way analysis of variance (display size, indicator condition, and Ss) showed that the main effects of display size, $F(2,6) = 98.8$, $p < .001$, indicator condition, $F(1,3) = 40.3$, $p < .01$, and Ss, $F(3,6) = 255$, $p < .001$, were significant. In addition, the Display by Indicator, $F(2,6) = 9.8$, $p < .05$, and Indicator by Ss, $F(3,6) = 9.9$, $p < .01$, interactions were also significant.

As seen in the table, reaction time increases progressively as display size increases from 4 to 12 elements. Also, for all three display sizes, reaction time is shorter under the leading indicator condition. If the indicator precedes the display by 150 msec, reaction time is 81 msec faster for an 8-element display and 82 msec faster for a 12-element display than is obtained under the simultaneous indicator condition. However, the leading indicator results in a gain of only 48 msec for the 4-element display. The smaller gain with the 4-element display accounts for the significant Display by Indicator interaction. Further investigation of the Indicator by Ss interaction revealed that all Ss had shorter reaction times with leading indicators, but the amount of facilitation the leading indicator produced varied between Ss.

DISCUSSION

The above results constitute a complete confirmation of the Eriksen and Rohrbaugh (1970) experiment. In that study, the displays were but briefly presented and the dependent variable was accuracy of target identification as a function of display size and leading or simultaneous indicator conditions. By showing that the effects of these variables are also apparent using reaction time as the dependent variable, the present experiment helps to bridge the gap between tachistoscopic studies of attention and informatioprocessing and the line of experimental investigation that has employed

reaction time as the measure of these processes. Whether all tachistoscopic studies of information processing are amenable to the more powerful reaction time approach still needs to be determined. It would be interesting, for example, if visual masking could be detected by means of differential reaction times.

It was assumed in the Eriksen and Rohrbaugh experiment that the decrease in accuracy that occurred with increasing number of display elements reflected an increase in time required to select the target element. The assumption was that the decreased accuracy resulted from a decay in legibility of the icon over time while the selection was occurring. Similarly, the increased accuracy that had occurred with a leading indicator was attributed to the location of display position having occurred before the display was presented, and thus encoding of the target element occurred while its iconic representation was still highly legible. The ability to replicate the results in the present experiment with direct measures of latency lends credence to the assumption that variation in accuracy of the previous experiment reflected the time required for such attentional selectivity to occur.

In the simultaneous indicator condition of the present experiment, the increase in voicing reaction time with an increase in display elements would suggest that the programming of which stimulus to encode is not completely effective in eliminating extraneous or irrelevant stimuli. Voicing time increased 43 msec as the display size increased from 4 to 8 elements, and an additional 27 msec was required when the display contained 12 elements. Over this limited range of display sizes the increase in reaction time does not appear to be linear. Other data (Sperling, 1960; Eriksen & Lappin, 1967) have indicated that the processing of 4 or fewer elements such as letters or digits involves a different processing rate than when a larger number of letters or digits are involved. The suggestion of nonlinearity of reaction time with display size may reflect this shift between the way a smaller number of elements is handled as opposed to larger numbers.

How does number of irrelevant display elements affect the processing time for the target element? One can conceive of a number of different ways that this could occur. In the Eriksen and Rohrbaugh experiment, errors in identifying a target letter were found to match an adjacent letter more often than chance would indicate. This would suggest some kind of successive approximation or

focusing in as the characteristic of the location process.

We can conceive of the selection as a focusing or narrowing in on the target element that is gradual and requires time. Simultaneously with this focusing in, elements in the gradually narrowing focus field are given initial processing. Items may cease to be processed as they drop from the focus field with increased narrowing. The more narrow the focus field has to become to encompass only the target item, the longer will be the time required for the process to occur. Thus, a longer time would be required to focus down to 1 out of 12 positions of the visual field than to focus down to 1 out of 4 positions. With the increased time, the probability increases that a nontarget item in the narrowing focus field will complete processing through to the point of encoding before the target item itself reaches this stage. This would tend to produce the kind of errors that would have been correct responses to adjacent stimuli.

In this model of gradual focusing in on the target item, we have assumed that initially all items in the field are given a low level of processing. Some items will be processed to more advanced stages because they remain longer in the gradually narrowing focus field. It assumes that parallel processing occurs at these low stages, even though the actual encoding, a process by which an item is transferred into a short-term storage system, may itself be serial. It does fit the observation that Ss can generally report the approximate number of elements that were presented in a visual display, even though they may not be able to identify more than the target element.

The possibility of a time-consuming focusing-in process is consistent with the results obtained in the leading indicator condition. If an indicator is presented sufficiently long before the display, the indicator information should be capable of being processed and the precise position in the visual field focused in upon before the display occurs. Under this circumstance, reaction time in vocalizing the target letter should be independent of the number of noise or irrelevant letters in the display. Reaction time to all display sizes would then be equivalent.

There is a suggestion of such an effect occurring, indicated by the significant interaction of display size with simultaneous vs leading indicators. Presenting the indicator 150 msec before the display tended to eliminate the difference in reaction time between 4- and 8-element displays. This would be commensurate with an interpretation that 150 msec is

sufficiently long to encompass the latencies involved in a focusing process narrowing down to the equivalent of about one-fourth of the display and also long enough for a large proportion of the latencies necessary for a precision of one-eighth of the display size. This explanation would account for the reaction-time difference between 4- and 8-element displays being reduced from 43 msec to only 10 msec under the leading indicator condition and also why the difference between 8- and 12-element displays remains at approximately 27 msec under the leading indicator condition.

We are assuming here that the focusing-in process is actually spatial in the visual field. In the present experiment, the number of display elements is confounded with the spacing between elements. Thus, as the number of elements in the display increases, the number of elements occupying a given space in the display also increases. If 150 msec is sufficient to locate one quadrant of the display, there would be only one element to be encoded and reported in that quarter of the field. A reduction to one-eighth of the field would be required to retain only one element in the focus for an 8-element display and a reduction of 1/12 to narrow down for 1 element of the 12-element displays.

The above description of the selectivity process is admittedly highly speculative, and various other interpretations would be commensurate with the data. It does, however, provide a rough formulation around which further investigations of attentional selectivity can be organized.

REFERENCES

ERIKSEN, C. E., & COLEGATE, R. L. Selective attention and serial processing in briefly presented visual displays. Perception & Psychophysics, 1971, 10, 321-326.

ERIKSEN, C. W., & COLLINS, J. F. Temporal course of selective attention. Journal of Experimental Psychology, 1969, 80, 254-261.

ERIKSEN, C. W., & ERIKSEN, B. A. Visual perceptual processing rates and backward and forward masking. Journal of Experimental Psychology, 1971, 89, 306-313.

ERIKSEN, C. W., & LAPPIN, J. S. Independence in the perception of simultaneously presented forms at brief durations. Journal of Experimental Psychology, 1967, 73, 468-472.

ERIKSEN, C. W., & ROHRBAUGH, J. W. Some factors determining efficiency of selective attention. The American Journal of Psychology, 1970, 83, 330-342.

SPERLING, G. The information available in brief visual presentations. Psychological Monographs, 1960, 74(11, Whole No. 498).

STERNBERG, S. Memory scanning: Mental processes revealed by reaction-time experiments. American Scientist, 1969, 57, 421-457.