Reaction time in relation to display size and correctness of response in forced-choice visual signal detection¹

W. K. ESTES AND DAVID L. WESSEL

STANFORD UNIVERSITY

Response times were recorded in a two-alternative, forced choice visual detection situation. Stimulus displays, presented tachistoscopically, were randomly selected consonant letters distributed in random subsets of cells of a matrix. Display sizes in Experiment 1 were 8, 12, and 16 letters; in Experiment 2 — 1, 4, and 8 letters; on each trial S operated a key to indicate which member of a predesignated pair of letters (signal elements) was present in a given display. Correct response times, on the average, increased uniformly with display size. Incorrect response times were uniformly greater than correct response times and, except for a reduction in the case of one element displays, were constant over display size. These relationships appear to require a modification of one assumption in the earlier proposed serial processing model for tachistoscopic perception.

Previous studies of information processing from tachistoscopic displays by a forced-choice technique (Estes & Taylor; 1964, 1966) have provided evidence generally favorable to a dual-aspect model. This model assumes that, during an exposure, the elements of a tachistoscopic display are registered in the visual system, following which their central representations are scanned singly, the scanning continuing until either the signal element is detected or some random perturbation terminates the process. However, data in the experiments conducted previously have been limited to response frequencies and, from these, evidence regarding the hypothesized scanning process can be obtained only by a chain of inferences within the framework of a particular model. We wish now to explore a source of independent, and perhaps more direct, evidence regarding properties of the scanning process by utilizing data concerning reaction times.

Reaction time in this situation can be conceived as a sum of the time required to process the information from the elements of the stimulus display (t_s); the time required for selection of a response, i.e., decision time, (t_d); and the time required for execution of the response (t_r). It will be assumed that the third component, t_r , is constant, on the average, over all trials of an experiment regardless of characteristics of the stimulus display. For purposes of analysis it would be desirable also to assume t_d to be constant over all types of trials, but this simplification may not be justifiable. Thus we shall allow for the possibility that decision time has a different, presumably lower, value on trials when the signal element is perceived than on trials when the signal ele-

ment is not perceived and the subject is forced to guess. Finally, it will be assumed that stimulus processing time varies directly with the number of elements perceived (scanned, in terms of the dual-aspect model), and thus also, on the average, with the number of elements in the display. Our first experimental step has been to replicate a previous study of visual detection in relation to display size (Estes & Taylor, 1966, Experiment I) with provision for recording reaction times. Results of this first stage indicated the desirability of extending the range of display sizes and this has been done in the second of the experiments reported.

As a point of departure for theoretical analysis we have taken predictions regarding reaction times that can be derived from a serial processing model for visual detection (Estes & Taylor; 1964, 1966). Letting t_{sC} denote stimulus processing time on correct response trials and t_{sI} stimulus processing time on incorrect response trials the principal deductions are that

- (1) t_{sC} should be an increasing, negatively accelerated function of display size;
- (2) t_{sI} should be an increasing, negatively accelerated function of display size;
- (3) t_{sI} should be smaller than t_{sC} at all display sizes, although they should converge to a common asymptotic value as display size becomes very large.

We shall not give formal derivations of these predictions since they involve rather unwieldy algebra, but the bases for them can be seen fairly readily at an intuitive level from consideration of the hypothesized serial scanning process. The t_s values should be expected to increase as a function of display size simply because, assuming random placement of the signal element, the greater the number of elements present for processing on any trial, the more will be scanned on the average before the signal element is reached. The predicted inequality between t_{sI} and t_{sC} arises from the consideration that on correct response trials the scanning process must always continue from the starting point until the signal element is reached, whereas on incorrect trials the process may sometimes terminate prematurely as a result of some random perturbation (distraction or "lapse of attention").

Although stimulus processing times cannot be measured directly, predictions about them carry implications, under the relatively weak assumptions sketched above, regarding observed reaction times. Thus our plan is to compare functions obtained for reaction time

in relation to display size with those derivable from the serial processing model, hoping that, to the extent that the predicted pattern is not borne out in the data, the discrepancies may point the way to appropriate modifications of the original formulation.

EXPERIMENT 1

Method

Apparatus. The tachistoscope was identical to that used in previous experiments (Estes & Taylor; 1964, 1966), except that the fixation point on the pre-exposure field was eliminated. Pre-exposure and post-exposure fields were dark.

Since it was desired to obtain response latency measures, two response keys were placed directly in front of the subject as he was seated in the viewing position. The distance between the keys was 3 in. They were of the telegraph type and required 13 gm of force to close the contacts, the key moving over a distance of approximately 3/32 in. A timer was started simultaneously with the onset of the stimulus display on each trial and was stopped by the operation of the response key, permitting recording of the subject's reaction time in units of .001 sec.

Stimulus materials. The stimulus materials were 8, 12, and 16 element displays, precisely those used in an earlier experiment (Estes & Taylor, 1966). Randomly drawn sets of consonant letters were entered in cells of a hypothetical 4 by 4 matrix, each row and column containing 2, 3, or 4 letters for the three display sizes, respectively. Each display contained exactly one "signal" element (the two signal letters, B and F, appearing equally often) plus the appropriate number of "noise" elements.

Procedure. Subjects were 20 undergraduate Stanford students enrolled in Introductory Psychology. The basis for their participation was a course requirement.

The experimental room was illuminated by a 150-W bulb encased in a white globe mounted on the ceiling; this was adjusted to about 1/3 normal brightness. The subjects were given approximately 5 min. to adapt to the room brightness. The subject was seated at the apparatus and shown a few sample displays during the reading of instructions, which were as follows:

"This piece of apparatus is a tachistoscope, a device for presenting visual displays for very brief periods of time. The display duration in this experiment will be 1/20 of a second.

"You will be shown various arrays of consonant letters which appear like this (demonstration). Each display will contain either the letter "B" or the letter "F". Never both and always at least one of them. Your task is to tell me which of the letters "B" or "F" appears in a given display by pressing one of the keys before you. You are to respond as rapidly as you comfortably can. Since the display is very short you will be given a warning, a click, which sounds like this (demonstration). This click will precede each display by 1.5 sec. and will

serve to help you attend to the display. If you are unsure of how to respond, please guess. Please use two hands, one on each key, and remember to respond as rapidly as you are able. Any questions? We will begin by running some practice trials."

During preliminary practice, the subject was then presented 10 trials each of 8, 12, and 16 element arrays in that order. Then each subject received 16-trial blocks of 8, 12, and 16 element arrays in the same order. This sequence was repeated again after a 3 min. rest yielding a total of 96 experimental trials from which data were collected. All subjects received the same random presentation order. The intertrial interval was approximately 14 sec.

Results

Differences in correct response proportions between the first and second replications were small and inconsistent. For the two replications pooled, the proportions were .745, .731, and .692 for display sizes of 8, 12, and 16, respectively. Following the procedure of Estes and Taylor (1964, 1966), we can, for each display size, take the observed proportion of correct responses as an estimate of $P(C)$, the probability of a correct response, in the formula

$$P(C) = \frac{P}{D} + (1 - \frac{P}{D}) \frac{1}{2},$$

where D denotes display size and P the number of elements effectively processed, or perceived, by the subject from the display on a given trial, and solve for P . The values obtained from the present data are 3.92, 5.54, and 6.14 elements, for the display sizes 8, 12, and 16, respectively, in fair correspondence with the values of 4.43, 5.33, and 5.57 reported by Estes and Taylor (1966).

Response time data are given in Table 1 in terms of means and standard errors of the latencies in seconds for correct and incorrect responses separately at each display size. Although proportions of correct responses did not differ significantly between the first and second replications of the experiment, a practice effect is manifest in the consistently lower response times for the second replication. However, the trends over display size are the same for both replications. Correct re-

Table 1. Response Time in Seconds as a Function of Display Size

	Experiment 1					
	Correct Responses			Incorrect Responses		
	8	12	16	8	12	16
First Replication						
M	1.324	1.304	1.381	1.680	1.558	1.587
SE _M	.041	.038	.043	.075	.084	.073
Second Replication						
M	1.120	1.203	1.292	1.603	1.402	1.610
SE _M	.029	.033	.039	.097	.067	.067
Pooled Data						
M	1.220	1.253	1.337	1.644	1.482	1.599
SE _M	.026	.025	.029	.060	.054	.049

sponse times increase as a function of display size, as anticipated on the basis of the serial processing concept, the function being clearly concave upward. Incorrect response latencies are greater than correct response latencies at all display sizes and are virtually constant over display size, at any rate showing no sign whatever of an increasing trend.

The sharply different pictures for correct and incorrect response times as a function of display size seem likely to be of considerable theoretical import. However, we shall defer discussion of this result until we report Experiment 2 in which the range of display sizes is extended to lower values.

EXPERIMENT 2

Not only the virtual constancy of incorrect response times over display size but also the very small difference in correct response time between displays of 8 and 12 elements in Experiment 1 seemed to call for an investigation of smaller display sizes. Thus in this experiment the range of this variable is from the minimal display, one-element, to eight elements, included to provide a basis for comparison with previous results.

Also a number of measures were taken in order to reduce the standard errors of measurement associated with our determinations of response time. First, the subjects were run for three sessions, with the data of the last two being used for analysis, in order to avoid the practice effect observed between the first and second replications in Experiment 1. Second, the subjects were urged to respond as rapidly as they could without making errors. Third, the foreperiod was reduced from 1.5 sec. to .5 sec., which might be expected to reduce the mean and variance of response times (Karlin, 1959).

Finally, we wished to evaluate a procedure which would permit the subject to initiate a stimulus display. In the previous experiment, and in those of Estes and Taylor (1964, 1966), the display on each trial had been controlled by a timer, the stimulus onset coming at a fixed interval following a ready signal regardless of the subject's behavior. There is some reason to believe that subjects perform more uniformly in this type of situation if they are permitted to initiate the stimulus when they are ready. Thus the apparatus was modified by the addition of a foot switch to permit the subject to initiate stimulus onset. However, in order to determine the effects of this modification for future reference, half of the data in this experiment were collected with the subject-paced procedure and half with the previously standard experimenter-paced procedure.

Method

Apparatus. The apparatus was identical to that of Experiment 1, with the addition of a foot switch that could be used by either the experimenter or the subject to initiate a trial.

Stimulus materials. The displays were 1, 4, and 8 element arrays made with the same materials and con-

straints as those of Experiment 1, and in fact the 8-element displays were identical to those used in Experiment 1. For each display size a complete set of displays consisted of 32 cards; these were two parallel sets of 16, differing only in respect to the signal element, B or F. Each of the 1-element displays included only one or the other of the signal elements, each cell of the 4 by 4 matrix being used equally often.

Procedure. Subjects were 12 undergraduate and graduate Stanford students who were paid for their participation. At least 5 min. of dark adaptation was given each subject before each session.

Each subject was run for three consecutive days. The first day was used solely as a practice session. Two different types of trials were used equally often for all subjects, trials initiated by the experimenter, EP trials, and trials initiated by the subject, SP trials. Blocks of 16 EP or SP trials were alternated throughout the experiment. Within trial blocks each subject received a different random presentation order in which all three display sizes were mixed together. On an experimental day the subject received one block each of SP and EP trials as a warm-up. These were followed by 6 alternating EP and SP blocks making a total of 96 trials per condition for each experimental day. Exposure duration was 1 msec. on all trials.²

On Day 1 the subject was given a set of instructions similar to those in Experiment 1 with the following additions. The subject was told of the sequence of events on each type of trial. On experimenter paced trials the experimenter spoke the word "ready" and immediately thereafter operated the foot switch; after .5 sec. the display was presented for 1 msec. On subject paced trials the experimenter spoke the word "ready", after which the subject operated the foot switch when he was indeed ready and the 1 msec. stimulus display appeared after a .5 sec. delay.

Results

Proportions of correct responses for each display size and condition are given in Table 2. Somewhat unexpectedly, the experimenter paced and subject paced conditions showed very little difference in accuracy of response and certainly there was no advantage for the subject paced procedure. Pooling the data for the two conditions at each display size, we obtain estimates of P , the mean number of elements perceived per trial, of .948, 3.392, and 4.384, for the display sizes of 1, 4, and 8, respectively. The last of these is quite close to

Table 2.
Correct Response Proportion by Display Size and Condition

	Experiment 2		
	Display Size		
	1	4	8
Experimenter Paced	.976	.948	.791
Subject Paced	.972	.901	.758

Table 3.
Response Time in Seconds by Display Size and Condition
Experiment 2

	Correct Responses			Incorrect Responses		
	1	4	8	1	4	8
Experimenter Paced						
M	.587	.644	.775	.734	.933	.878
SE _M	.014	.011	.018	.105	.078	.035
Subject Paced						
M	.549	.644	.764	.686	.894	.886
SE _M	.009	.013	.015	.100	.067	.034

the value obtained for display size 8 in Experiment 1, indicating that the additional practice given subjects in the present experiment had little effect on accuracy of detection.

Response time data for this experiment, presented in Table 3, again show very little difference between the experimenter paced and subject paced conditions. In contrast to the stability of the frequency data, however, is the large overall reduction in mean response time, of the order of 1/2 sec., from the level observed in Experiment 1. Presumably this reduction reflects not only the additional practice but also the changes in instructions and in foreperiod duration.

As in Experiment 1 correct response times are uniformly lower than incorrect response times and increase with display size. The curve relating correct response time to display size is again concave upward, though not as markedly so as in Experiment 1, and in this case there is clearly a significant increase in correct response time with each increase in display size from 1 to 4 to 8. Thus it appears that the very small difference in correct response time between the display sizes of 8 and 12 in Experiment 1 must be attributable to some idiosyncrasy of the particular displays used in that experiment. There would seem to be no doubt that correct response time increases monotonically with display size over the range we have studied even for well practiced subjects.

The situation is somewhat different for incorrect response times. In the present experiment there is an increase from display size 1 to 4, but this should perhaps not be given great weight since there were very few observations at display size 1 (only 9 errors in the experimenter paced condition and 11 in the subject paced condition, in contrast to the *N*s of 367, and 381, respectively for correct responses). Between display sizes of 4 and 8, where there are sufficient observations for reliability, there is again no trend toward increase in incorrect response time with display size. Thus, incorrect response time is evidently constant over the range of display sizes we have studied, except for a possible display deviation at size 1.

DISCUSSION

The relative constancy of incorrect response time

over most of the range of display sizes indicates that one aspect of the original serial processing model is certainly incorrect. Incorrect responses do not arise primarily from instances in which the scanning process terminates before the subject reaches the signal element in a display on a given trial. The simplest interpretation of this result would seem to be that errors occur on trials when the critical element is not sampled at all, i.e., is not registered in the subject's visual system, and that when the signal element is not included in the set registered, scanning continues for a fixed interval of time which is independent of display size.

More important, perhaps, for theoretical purposes are the correct response time data; for these, suitably analyzed, should provide information concerning stimulus processing time on trials which terminate with detection of the signal element. However, the simple statistics of correct response times, for example the mean at each display size, cannot be used directly for this purpose since they must represent mixtures of trials on which the signal element was detected and trials on which the correct response occurred by guessing. Thus, mean correct response time at any display size must itself be a weighted mean of response times on "true detection" trials and mean response times on "correct guess" trials. We can, however, decompose this mixture analytically, for the mean response time on correct guess trials should be equal to the mean time on incorrect response trials. We shall denote mean reaction time on correct response trials, true detection trials, and incorrect response trials, respectively, by L_C , L_+ , and L_I . In terms of the notation introduced above for components of reaction time,

$$L_+ = t_{s+} + t_{d+} + t_r$$

and

$$L_I = t_{s-} + t_{d-} + t_r$$

Here the + and - subscripts distinguish stimulus processing times and decision times on true detection trials from corresponding times on trials when the signal element is not detected and the subject has to guess. According to our assumptions the set of correct response times arising in a series of experimental trials is a mixture of reaction times from true detection trials, whose mean is L_+ , and reaction times from correct guess trials, whose mean is equal to L_I .

To estimate L_+ from the data at each display size, we proceed as follows. In the expression for L_C as a weighted mean of L_+ and L_I ,

$$L_C = P\left(\frac{1}{C}\right) \cdot [P(+)L_+ + P(I)L_I],$$

we can insert observed values of mean correct and incorrect response latencies, L_C and L_I ; of proportions of correct and incorrect responses, $P(C)$ and $P(I)$; and of the proportion of true detections, $P(+)$, estimated from the relation

$$P(C) - P(I) = P(+);$$

and finally solve for L_+ ,

$$L_+ = \frac{P(C)L_C - P(I)L_I}{P(+)}$$

to obtain the desired estimate. Application of this procedure to the data of Experiment 1 yields L_+ estimates of .999, 1.120, and 1.127 sec. for display sizes of 8, 12, and 16, respectively. For Experiment 2 the estimates are .564, .620, and .724, for display sizes of 1, 4, and 8, respectively.

Evidently we may conclude that, on the average, response time on true detection trials increases with display size, as required by the conception of a serial scanning process. It would not be judicious to go further and attempt to infer the precise form of the relationship between L_+ and the display size from our data, for our estimator of L_+ involves a ratio of random variables and we can assume neither that the estimator is unbiased nor that the bias is constant over display size. This problem might be circumvented in further work either by determinations of L_+ at various display sizes for individual subjects run over long sequences of trials, or perhaps by use of confidence ratings which might enable one experimentally to segregate true detection trials from correct guess trials.

Taking results of this study together with those of Estes and Taylor (1966) it seems reasonable to conclude that of the three principal concepts involved in the original serial processing model for tachistoscopic perception, those assuming that on each trial a random sample of the elements displayed is registered in the visual system and then scanned sequentially until the signal element is reached have been supported by several independent lines of evidence. The third concept, that the scanning process is terminated by an exponentially distributed random event, has proved incompatible both with estimates of variability in number of elements

scanned per trial (Estes & Taylor, 1966; Estes, 1965) and with the present results concerning incorrect response times, and thus must be modified. It now seems more likely that whenever the signal element is included in the set sampled on a trial, the scanning process always continues until the signal element is processed; that errors occur only on trials on which the signal element is not sampled; and that on trials of the latter type the scanning process continues until each element of the sample has been scanned at least once, ultimately being terminated by some event which does not depend on sample size. Little is known about the generality of these conclusions, however, for the few relevant studies have involved a very limited range of conditions, especially with regard to stimulus parameters.

References

- Estes, W. K. A technique for assessing variability of perceptual span. *Proc. Nat. Acad. Sci.*, 1965, 4, No. 2, 403-407.
- Estes, W. K., & Taylor, H. A. A detection method and probabilistic models for assessing information processing from brief visual displays. *Proc. Nat. Acad. Sci.*, 1964, 52, No. 2, 446-454.
- Estes, W. K., & Taylor, H. A. Visual detection in relation to display size and redundancy of critical elements. *Percept. & Psychophys.*, 1966, 1, 9-16.
- Karlin, L. Reaction time as a function of foreperiod duration and variability. *J. exp. Psychol.*, 1958, 59, 185-191.

Notes

1. This research was supported in part by Grant NGR-05-020-036 from the National Aeronautics and Space Administration and in part by Grant MH-6154 from the United States Public Health Service. Reproduction in whole or in part is permitted for any purpose of the United States Government.
2. That is, the time interval control was set for 1 msec. Approximately 0.3 msec. is required for light intensity to build up to maximum on the exposure field, so effective exposure duration was somewhat less than 1 msec.

(Accepted for publication September 3, 1966.)