Further Evidence Concerning Scanning and Sampling Assumptions of Visual Detection Models

by

George L. Wolford, David L. Wessel, and W. K. Estes

Technical Report No. 126

January 31 1968

Psychology Series

Institute for Mathematical Studies in the Social Sciences
Stanford University
Stanford, California

[1] Reported at the meeting of the Psychonomic Society, Chicago, Illinois, October 1967. This research was supported in part by Grant NGR-05-020-036 from the National Aeronautics and Space Administration, by Grant GB3878 from the National Science Foundation, and by Grant MH6154 from the United States Public Health Service.

Abstract

Following extended training in a visual detection task, functions were determined for individual subjects relating latency of detection responses to number of redundant signal elements embedded in tachistoscopic displays of letters and to distance between signal elements. Latency proved invariant with respect to number of redundant signals and varied nonmonotonically with distance. Of the several types of models considered in relation to previous studies, substantial support was forthcoming only for an independent stimulus sampling model. It is suggested that the detection method effects a relatively clear separation of the perceptual from the mnemonic aspects of the standard visual apprehension experiment, and that the sampling process may constitute only the first phase of a more general model which includes both parallel and serial information processing.

Introduction

A combination of theoretical and intuitive considerations has suggested that a model for the processing of information from tachistoscopic displays must involve both parallel and serial processes (see, e.g., Estes and Taylor, 1966; Sperling, 1967). During the tachistoscopic exposure, the elements of the display, or a sample therefrom, must be registered in some sense in the visual system. Further processing of the registered elements presumably depends in part upon task requirements. For an experiment conducted with a report procedure, Sperling (1967) proposes that the elements are scanned and transformed into auditory representations which may then be rehearsed during the interval between exposure and test. In a forced choice detection experiment, presumably the elements are scanned by some central mechanism until either the signal element is reached or all registered elements have been examined without encountering a signal (Estes and Taylor, 1966).

Evidence bearing relatively directly upon the assumed scanning process comes largely from the forced choice detection situation, which has the advantages of largely eliminating rehearsal and short term memory processes and permits accurate determinations of reaction times. In this situation Estes and Wessel (1966) found clear evidence of a systematic increase in reaction time as a function of display size, in accord with the serial scanning concept.

An independent source of evidence regarding the scanning process is available in the detection experiment by the incorporation of duplicate signal elements. In the basic experimental paradigm of

previous related studies, the displays are arrays of printed letters, usually filling randomly selected cells of a matrix. The subject is instructed that any one display contains exactly one of two possible signal elements, for example a B or an F. The remaining letters in the display are noise elements. The subject's task is to indicate by pressing one of two keys following each exposure which of the two signal elements was present in the given display. In the variation utilizing duplicate signal elements (Estes and Taylor, 1966) any one display would contain either two B's or two F's but the subject's task would still be simply to indicate which of the two signals was represented. Assuming that the scanning process on the trial continues only until a signal element has been detected, clearly one must predict that response time should be a decreasing function of the number of duplicate signal elements present. The first purpose of the present study is to evaluate this prediction. A second objective, having to do with other properties of the set of elements scanned, will be elucidated in a later section.

Method

Experiment I

Apparatus

The tachistoscope used in this experiment was identical to the one described by Estes and Taylor (1964, 1966), except that a viewing tunnel was added which placed the subject 24 in. from the stimulus field. A goggle frame was placed at the viewing end of the tunnel in order to obtain a standard visual orientation for all subjects. A fixation point

appeared in the center of the stimulus field and was kept on throughout the experiment. The pre- and post-exposure fields were otherwise dark. The stimulus exposure field had a brightness of 7 ft. candles. The subject responded by pressing one of two telegraph keys, placed to the right of the viewing tunnel. The keys required 13 grams of force to stop the timer.

Subjects

Four Stanford students served as subjects and were paid for their participation. The subjects were run for approximately 15 days each, with only the data of the last 6-8 days being used in the analysis.

Stimulus Materials

The stimulus materials were 4x4 arrays of English letters typed on white 8x5 in. cards with IBM Directory type. Each card contained either 1 or 2 signal elements with the rest of the array being filled with the remaining English consonants. The signal elements were either B's or F's. There were two identical sets of stimuli except that one set contained B's as the signal elements and the other set contained F's. One-half of the arrays contained one signal element per card and the other half contained two. The signal elements were placed randomly within the arrays. The 16 cell arrays subtended a visual angle of 2 degrees 10 min.

Procedure

Each subject began the day with 20 practice trials followed by 128 data collection trials. The stimulus cards were presented in a random order. The trials were subject paced. Each trial began with

the experimenter informing the subject that the stimulus card was in position. The subject then pressed a foot pedal when he felt he was ready, initiating a presentation of the stimulus card for a duration of 5 msec. The subject then responded by pressing either the B or F key and giving a confidence rating on a scale from 1 to 3 with 1 being a guess and 3 meaning certainty that the signal element had been detected. The subjects were given unlimited time for response. One day's session took approximately 50 min.

Experiment II

The second experiment was similar to the first with the exception that four new subjects were used and the stimulus materials were different. Stimulus cards contained 5x5 arrays of English letters and the number of signal elements per card varied from 0 to 3. A total of 500 cards was prepared, from which a sample of 128 was drawn for presentation to each subject each day, following the 20 trial warmup series.

In this experiment, the signal elements were either A's or T's. Ten percent of the stimulus cards contained no signal elements (blank trials) and 30 percent each of the cards contained 1, 2, or 3 signal elements. The remaining elements on any stimulus card were chosen from the 24 remaining English letters. The 5x5 arrays subtended a visual angle of 2 degrees 30 min. The subjects were run for about 15 days each, with data from the last 8-12 days being used in the analysis.

Results

Data of both experiments will be considered together in this section. Proportions of correct responses, given for each subject in Table 1, exhibit a uniformly increasing function of number of redundant signal elements, replicating the results of Estes and Taylor (1966). Although the analysis will not be presented here, it may be remarked that the fit of a fixed-sample-size, stimulus sampling model to the data of Table 1 is similar to that found in the previous study. The data from blank trials in Experiment II reveal a distinct response bias in three of the four subjects [P(A) = .37, .37, .33, .52 for subjects 5-8, respectively], which would have to be taken into account in a fully adequate model.

The results of primary interest, mean latencies for correct and incorrect response trials, are presented in Table 2. Error latencies are uniformly high and exhibit no significant variation as a function of number of signal elements, recalling the constancy over display size observed previously (Estes and Wessel, 1966). Correct response latencies are shorter than error latencies in all cases, and decrease uniformly as a function of number of signal elements.

The last result might be taken at first glance to support a serial processing model which assumes that elements of the display are scanned singly along some path, with the process terminating when a signal element is reached. However, it must be recognized that mean correct response times for all conditions undoubtedly represent mixtures of true detection trials and correct guess trials. Presumably the latencies

Table 1

Proportion of Correct Responses in Relation to Number of Redundant Signal Elements per Display

Experiment I No. Signal Elements			Experiment II No. Signal Elements			
Subject	1	2	Subject	1	2	3
1	.914	.977	5	.727	.881	.890
2	.793	.920	6	.755	.917	.921
3	.768	.886	7	.655	.831	.896
4	.705	.891	8	.628	.801	.881

Table 2

Mean Latencies of Correct (L_C) and Incorrect (L_I) Responses

Subject	Experiment I No. Signal Elements 1		2		Subject	Experiment II No. Signal Elements 1		2		3	
	L_C	L_I	L_C	L_I		L_C	L_I	L_C	L_I	L_C	L_I
1	617	687	582	721	5	914	946	891	955	879	966
2	1062	1787	880	1955	6	1139	1775	864	1673	871	1645
3	1160	1522	1034	1399	7	743	963	674	1017	613	1200
4	1338	1730	1150	1738	8	1491	2017	1394	2053	1204	2048
Mean	1044	1432	912	1453	Mean	1072	1425	956	1425	892	1465

of correct guesses should on the average be equal to those of errors, since both types of response arise when the subject fails to detect a signal element in the display and is forced to guess. If the proportion of true detections increases with number of signal elements, which can scarcely be doubted, then the changing mixture of true detections and guesses could yield the observed trend in mean latencies.

For our present theoretical interests, it is essential to estimate mean latencies on trials when true detections occur, since it is a trend in these that is implied by the serial processing model. Lacking any one, unexceptionable method of obtaining these estimates, we shall have to resort to two procedures which involve different auxiliary assumptions. The more direct of these approaches, perhaps, is the attempt to isolate a set of data representing only true detection trials by use of the confidence ratings. To this end, we have assembled in Table 3 mean correct response latencies for the subset of trials on which the highest of the three confidence ratings was given by the subjects. Our assumption that this procedure would yield data free of effects of guessing is borne out by the fact that the mean proportion correct was over .99 on high confidence trials.

The somewhat unexpected result of this analysis is that the mean latencies in Table 3, presumably representing true detection trials under each condition, exhibit no trace whatever of a decreasing trend as a function of number of signal elements either in the group averages or in those of any individual subject.

The major possibility of an artifactual account of the horizontal

Table 3

Mean Correct Response Latencies for High Confidence Data

Subject	Experiment I No. Signal Elements 1	2	Subject	Experiment II No. Signal Elements 1	2	3
1	544	535	5	768	807	784
2	657	646	6	700	703	708
3	676	742	7	558	574	524
4	868	853	8	807	822	768
Mean	686	694	Mean	705	715	698

trends in Table 3 lies in the reduced range of variation in latency which may be entailed by restriction to a single confidence band. To appraise this factor, we computed standard deviations of all correct responses and of high confidence correct responses separately for each subject and condition in Experiment II. These statistics show that restriction to the high confidence band produced an across the board decrease of about 40 percent in the standard deviations, the overall values for 1, 2, and 3 redundant signal elements, respectively, being 284, 280, and 254 for pooled correct responses and 157, 165, and 156 for high confidence correct responses. Perhaps more significant, however, is the fact that even in the high confidence data the standard deviations are substantial in relation to the means, so the restriction has not eliminated the possibility of considerable variation between means. Further, we note that although the trends relative to number of signal elements disappear in the high confidence data, the pattern of individual differences is preserved.

For a second, relatively independent source of evidence on the redundant element function, we have corrected the mean correct latencies of Table 2 by the formula

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

where $P(+)$ denotes proportion of true detections, and L_+ mean latency of a true detection, with $P(+)=P(C)-P(I)$. The rationale for this analysis is given by Estes and Wessel (1966). The corrected group means are 852 and 850 for 1 and 2 signal elements, respectively,

in Experiment I, and 726, 856, and 823, for 1, 2, and 3 signal elements in Experiment II. Thus both approaches agree in supporting the conclusion that latencies of true detections are invariant with respect to number of redundant signal elements.

A second important aspect of the stimulus sampling process which may be illuminated by our data is a property which we term "connectedness." By a connected sample, we mean one all of whose elements come from some sub region of the display. According to one interpretation of a serial processing model, the subject is assumed to scan the displayed elements along some path (e.g., along rows of our matrix displays, from left to right in one row, right to left in the next, and so on), yielding a sample which is connected in the sense that all elements processed are adjacent along the scanning path. According to one interpretation of a fixed sample size model, the subject processes all elements falling in some portion of the display area (e.g., a circle around the fixation point). Contrariwise, a disconnected sample, implied by an independent sampling model, would in general contain elements coming from locations scattered in some fashion over the display area.

It is easy to see, without formal derivations, that if the sample is connected, then, up to some limit, probability of detection must be an increasing function of distance between two redundant signal elements. If, at one extreme, the two elements are adjacent, then the second adds little to the likelihood of detection, since usually either both or neither will be in the sample. As the distance increases, so also does the expected contribution of the second element, since it becomes

increasingly likely that the sample will include one or the other. If, however, the sample is unconnected, there need be no relation whatever between probability of detection and distance between signal elements.

To obtain evidence on this point, we have analyzed the data for probability correct in the case of two redundant signal elements for each experiment. Distance between the two signal elements in any display was measured in city block fashion. That is, if signal element 1 is in cell i,j of the matrix array and signal element 2 is in cell m,n, then the distance between them is taken to be $|i-m|+|j-n|$. Results of this analysis are summarized in Figure 1. In neither experiment is there any tendency apparent for probability correct to increase with distance.

Beyond taking this result as negative with regard to the hypothesis of connected sampling, we may inquire as to what process might give rise to the rather peculiar observed distance functions. The drop in probability correct at distances 4 and 5 in Experiment II must be taken seriously, for the numbers of observations are substantial (over 300 responses are represented at distance 4 and over 200 at 5), and the trend is very similar for all individual subjects.

An interpretation which is compatible with the results on connectedness and which might also account for the distance functions is a stimulus sampling model with independent sampling probabilities for the elements (Estes, 1959). The principal assumptions would be that the probability of sampling, that is processing, a signal element in any cell of the

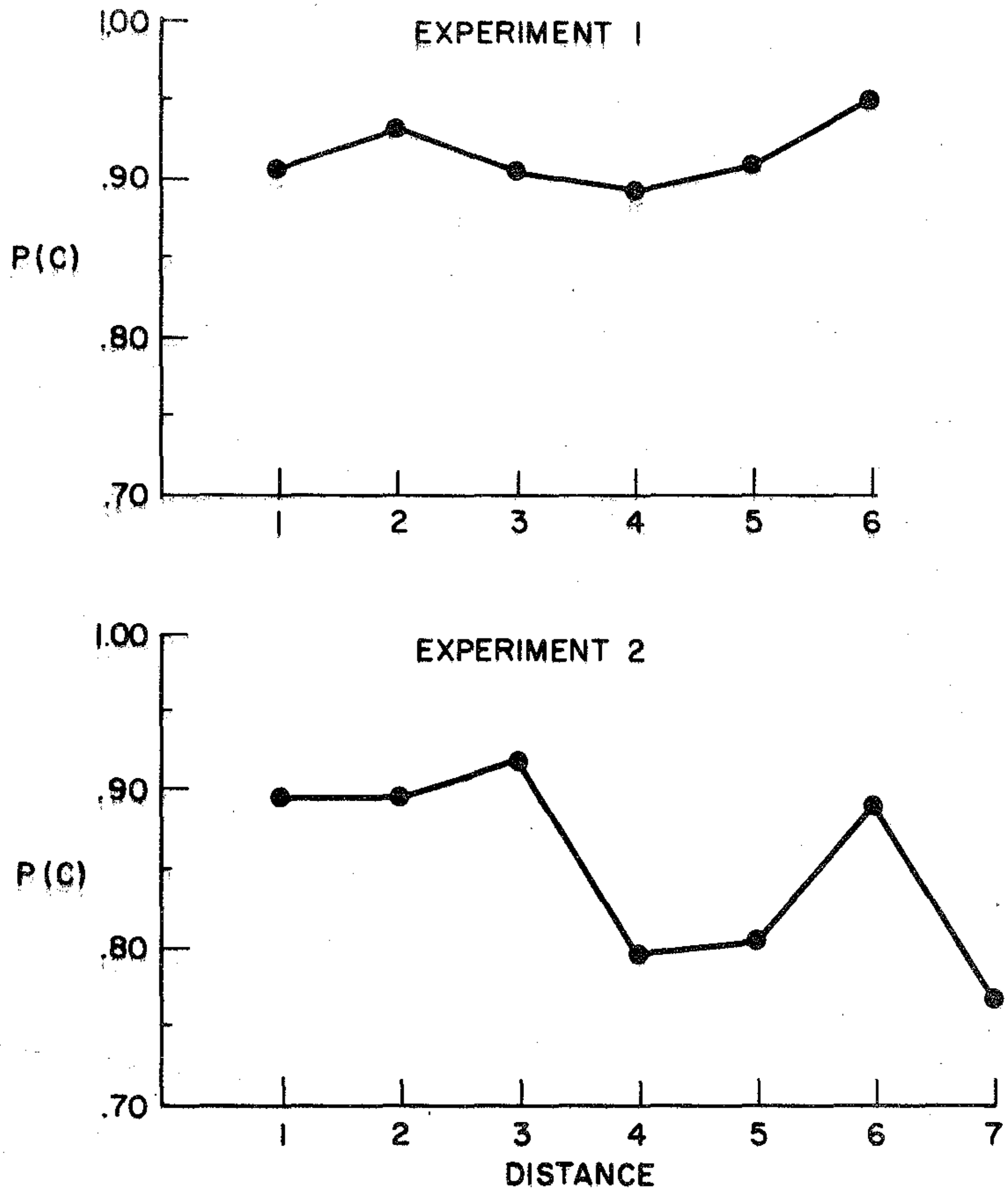
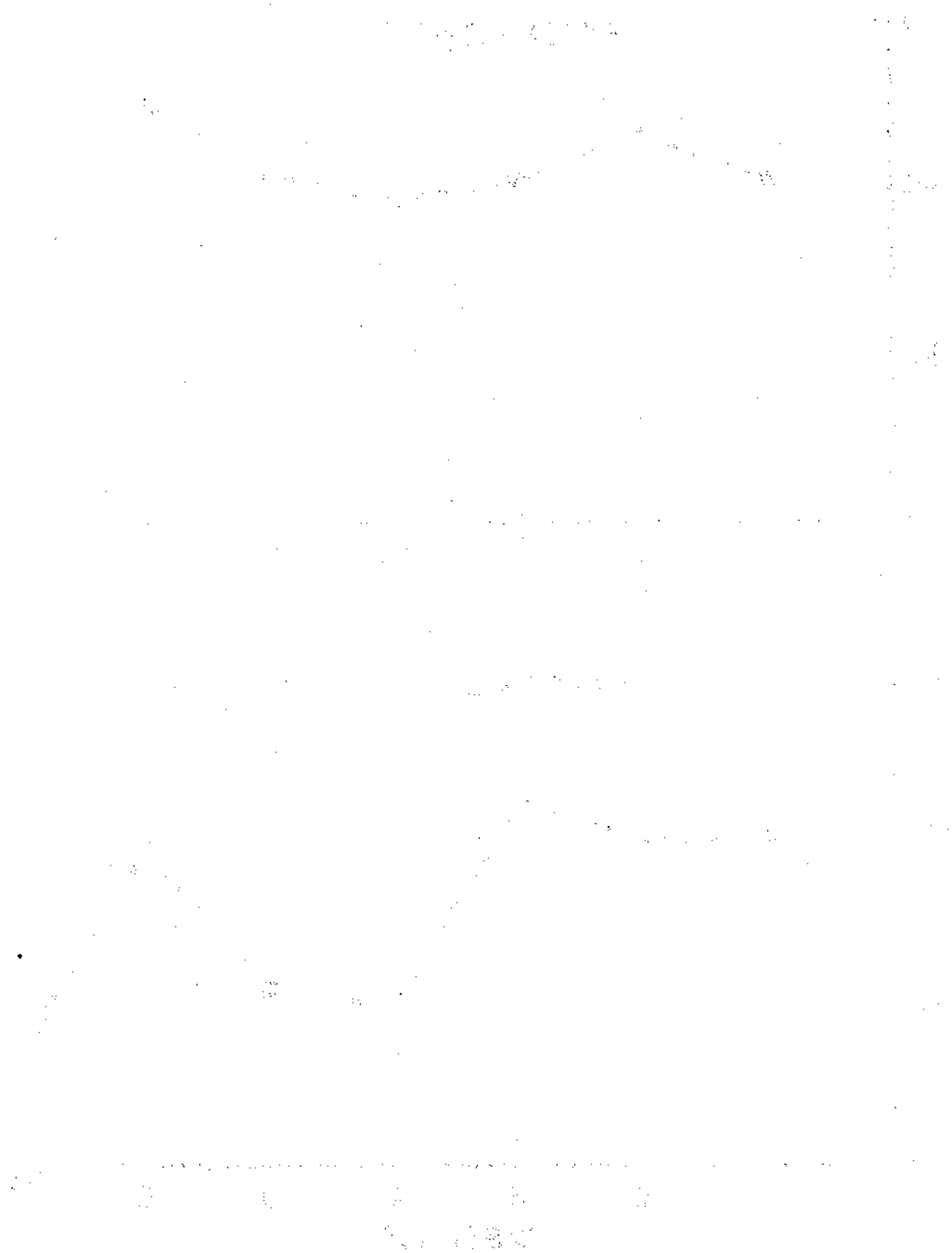


Figure 1. Proportion of correct responses on a function of distance between duplicate signal elements.



matrix display has some value θ_{ij} which depends upon the position of the cell, and probably also upon the density of neighboring noise elements which might exert masking effects. Then, if there is only one signal element in a display and it is located in cell i,j, the probability of a true detection on any trial is simply θ_{ij}. If there are two redundant signal elements, one in cell i,j and one in cell m,n, then the probability of a true detection is $\theta_{ij}+\theta_{mn}-\theta_{ij}\theta_{mn}$, and so on. With regard to the last analysis, distance per se would be irrelevant to detection probabilities, yet some distances would be expected to yield lower detection probabilities than others simply because they happened in our situation to be associated with pairs of cells involving higher θ values.

In order to provide an independent test of this interpretation, we have obtained estimates of the θ value for each cell of each of our matrix arrays from the observed probability correct obtained under the condition in which only one signal element was present. The observed probabilities correct by position for the 5x5 matrix of Experiment II are shown in Table 4. Some of the correct responses represented in each cell are, however, attributable to guessing. Thus, to obtain estimates of the true detection probabilities, we let

$$\theta_{ij} = 2[P_{ij}(C)- \frac{1}{2}],$$

where $P_{ij}(C)$ denotes observed proportion correct in cell i,j and θ_{ij} the estimated sampling probability for an element in that cell. To obtain theoretical detection probabilities for any given distance between two redundant critical elements, we simply take all pairs of

Table 4

Probability Correct by Position for 5x5 Matrix of Experiment II

.82	.76	.82	.82	.72
.82	.86	.82	.84	.77
.78	.94	.88	.74	.69
.56	.59	.57	.60	.62
.54	.50	.59	.50	.59

Table 5

Correct Response Probability as a Function of Distance between two Redundant Signal Elements: Experiment I, 4x4 Display

Distance	Subject No. 1 Pred.	1 Obs.	2 Pred.	2 Obs.	3 Pred.	3 Obs.	4 Pred.	4 Obs.	Overall Pred.	Overall Obs.
1	.98	1.00	.83	.88	.88	.93	.79	.86	.87	.91
2	.99	.98	.86	.93	.88	.90	.86	.91	.90	.93
3	.99	.94	.88	.88	.90	.88	.86	.91	.91	.91
4	.99	1.00	.90	.95	.87	.73	.81	.89	.89	.89
5	.99	.94	.99	.95	.87	.89	.75	.86	.90	.91
6	1.00	1.00	1.00	1.00	.88	1.00	.67	.80	.89	.95

Table 6

Correct Response Probability as a Function of Distance between two Redundant Signal Elements: Experiment II, 5x5 Display

	Subject No.									
	1		2		3		4		Overall	
Distance	Pred.	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.	Obs.
1	.93	.87	.97	.98	.89	.82	.75	.87	.89	.89
2	.82	.90	.89	.90	.89	.89	.77	.86	.84	.89
3	.96	.96	.98	.97	.88	.90	.79	.85	.90	.92
4	.84	.81	.85	.82	.74	.78	.66	.73	.77	.79
5	.86	.82	.94	.87	.78	.79	.67	.74	.81	.81
6	.87	.97	1.00	1.00	.84	.77	.78	.79	.87	.88
7	.78	.71	.99	.82	.74	.82	.60	.71	.78	.77

cells separated by the given distance, compute the theoretical detection probability for each on the basis of the appropriate pair of estimated θ values, and compute the average for the set. The theoretical values so obtained are compared with observed values for all subjects of Experiment I in Table 5 and of Experiment II in Table 6. On the whole the theoretical account appears quite satisfactory, especially in the case of Experiment II, where the complex relation between accuracy and distance is found to be well predicted by the theoretical values.

Discussion

The salient finding of this study is that the proportion of correct responses increases systematically with number of redundant signal elements per display, whereas latency of correct response on true detection trials, however estimated, is invariant with respect to this variable. The latter aspect seems incompatible with any model assuming a serial processing of the elements in the display which terminates when a signal element is reached. The first principal question at issue is how this result can be reconciled with previously published evidence apparently implying a serial scanning process of precisely this type.

Some of the apparently most cogent evidence pointing to a process of serial scanning which terminates upon detection of a signal came from a study by Estes and Wessel (1966) which showed that latencies of correct detection responses increase systematically with the size of the display. The increasing trend in that study held over a wide range of display sizes, even when correct guesses were partialed out,

yielding estimates of mean latencies on true detection trials. The one loose end was that latencies of incorrect responses were constant over a wide range of display sizes, a fact not easy to fit into the serial scanning conception.

If we reject the idea of a serial process terminating upon detection, how can we account for the display size functions in the data of Estes and Wessel? One possibility is that the processing is serial but that it continues until all elements have been processed regardless of how many signal elements are encountered. Although a logical possibility, this interpretation seems remote to us. Such a strategy would not fit in with the instructions or training given the subjects nor with their accounts of their own behavior. Neither would it help to explain why correct detection latencies but not error latencies vary with display size.

An alternative suggestion is that detection latency is related to the amount of processing required to discriminate a signal element from the background, and especially from neighboring noise elements sharing common properties with it. Increasing the display size would lead to increased processing times simply because it entails an increase in the number of noise elements, and usually also a decrease in the average distance between these and signal elements. In this interpretation, the reason why error latencies do not vary with display size is that errors occur on trials when no signal element is included in the set sampled by the subject; in these instances the total processing time is determined not by the number of elements on display, but

rather by the decay rate of the representation in the visual system.

Turning to our analyses having to do with distance between redundant signal elements, it seems clear that detection latency does not increase as a function of distance in our situation, implying that the set of elements sampled on the trial is not connected, as we have defined this term. These results would not necessarily be incompatible with a serial processing model in which the scanner skips over some elements, thus generating an unconnected sample, but as indicated above, such a model finds little positive support in other aspects of detection data. A fixed sample size model in which the sample comprises those elements falling in some particular sub area of the visual field would seem to be incompatible with the obtained distance functions. The principal remaining interpretation, and one which has some independent support, is that the elements of the display are sampled independently with the sampling probabilities varying over the field.

The model which receives positive support from our data assumes such independent sampling, with the sampling probabilities, θ values, varying as a function of position of the element in the display. We hasten to add, however, that this description holds only for variation in a particular display. We do not mean to imply that θ values are determined solely by position. Rather, it seems quite likely that the θ value associated with an element in any particular location depends upon the density of neighboring elements which might generate masking. (It was overlooking of the latter possibility that led to the inference that the functions for probability versus display size

obtained by Estes and Taylor (1966) were incompatible with an independent sampling model.) It should be added that our results concerning independent processing of elements in different locations in the visual field agree well with the results of Ericksen and Lappin (1967) obtained in a different situation and with somewhat different procedures, as well as with those of Schlosberg (1948) with regard to apprehension span for dot patterns. For inter-element distances of the order of magnitude of those involved in most current work on visual information processing, including all of the studies from our laboratory, the assumption of independent sampling appears quite well established.

With respect to one general qualitative issue that has been prominent in the recent literature, we might note that the type of model favored by our results involves parallel rather than serial processing. However, we do not mean to imply that this is the whole story. Sperling (1967) interprets results of his series of studies as pointing to a model including elements of both parallel and serial processing, with the parallel aspect being evidently sufficient for "letter recognition." The latter aspect perhaps agrees with our own conclusion regarding detection in the forced choice situation. Since the forced choice detection technique has been developed expressly for the purpose of separating perceptual and mnemonic aspects of the visual apprehension task better than preceding methods, it seems quite possible that parallel and serial processing mechanisms have to do, respectively, with these two aspects.

It appears that perception of the elements of the display,

involving processing to the point of detecting differences between signal and noise elements, can be quite fully handled by a parallel processing model of the general type proposed above. However, it may be that, once the sampling has occurred, a serial processing mechanism takes over when the results of initial stimulus sampling are being coded, or transferred to short term memory.

This conclusion should be qualified in one respect. The inferences drawn by Sperling (1967) and ourselves regarding parallel versus serial processing depend on the assumption that, if the elements of a display are scanned serially, the time required to process each element is measurable by present techniques. An alternative possibility is that both stages involve serial processing, with the scanning rate of the first stage being so fast that variations are not reflected in response times. Although not distinguishable on presently available data from the conception of an initial phase of parallel processing, this possibility may be testable by experiments involving suitable manipulations of the arrangement of signal and noise elements within the display.

References

Eriksen, C. W., and Lappin, J. S. Independence in the perception of simultaneously presented forms at brief durations. J. exp. Psychol., 1967, 73, 468-472.

Estes, W. K. Component and pattern models with Markovian interpretations. In R. R. Bush and W. K. Estes (Eds.), Studies in Mathematical Learning Theory. Stanford, Calif.: Stanford Univer. Press, 1959, 9-53.

Estes, W. K., and Taylor, H. A. A detection method and probabilistic models for assessing information processing from brief visual displays. Proc. of the National Academy of Sciences, 1964, 52, No. 2, 446-454.

Estes, W. K., and Taylor, H. A. Visual detection in relation to display size and redundancy of critical elements. Perception and Psychophysics, 1966, 1, 9-16.

Estes, W. K., and Wessel, D. L. Reaction time in relation to display size and correctness of response in forced-choice visual signal detection. Perception and Psychophysics, 1966, 1, 369-373.

Schlosberg, H. A probability formulation of the Hunter-Sigler effect. J. exp. Psychol., 1948, 38, 155-167.

Sperling, G. Successive approximations to a model for short term memory. Acta Psychologica, 1967, 27, 285-292.