VISUAL DETECTION WITH IMPERFECT RECOGNITION

by

John Holmgren

TECHNICAL REPORT NO. 128

March 29, 1968

PSYCHOLOGY SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

Abstract

A task involving both the detection and recognition of one of two possible critical elements embedded in a set of noise elements was investigated with the aid of a mathematical model. The model consists of three processes, these being detection, recognition and decision. The first of two experiments attempted to show the operation of two separate bias parameters in the decision process. While the results were in the right direction, the data did not unequivocally establish the necessity of two bias parameters. In the second study, it was found that while a subject's ability to detect a critical element in a display decreased as the number of noise elements in the display increased, his ability to recognize the critical element following detection remained constant. This finding was interpreted as supporting a strictly all-or-none view of detection.

Visual Detection with Imperfect Recognition[1]

John Holmgren

Stanford University

There are a wide variety of psychophysical procedures that may go under the general heading of visual detection tasks. In this paper we will be concerned with a type of detection task developed by Estes and Taylor (1964). This task is of special interest because of its potential use as a method for studying the processing of information from visual input. The task is of the following form. A visual display is presented to the subject tachistoscopically. The display consists of an array of elements (usually letters of the alphabet). A small number of elements (usually two) are designated as "critical elements." One member of this critical set is present in the display on each trial. The subject's task is to determine on each trial which member of the critical set was present in the display. The data from such experiments are frequently analyzed in terms of some type of high threshold detection model, i.e., it is assumed that the subject either detects the critical element, in which case he responds correctly with probability one, or he fails to detect, in which case he guesses, perhaps with some bias for one of the responses.

Yellott and Curnow (1967) have shown that a high threshold model is appropriate for such data. By using a critical set of three letters and allowing the subject a second guess following an error, they found that the subject's performance on the second guess could be adequately explained in terms of a high threshold assumption.

The high threshold models that have been applied to the above paradigm have the common assumption that if the subject detects a critical element in the display, then he recognizes which of the critical elements it is with probability one. It is likely, however, that in some situations the critical elements are similar enough so that the subject can detect the presence of a critical element in the display without being able to recognize which critical element it is. What are the implications of having imperfect recognition in a detection task? To help answer this question, we have formulated a high threshold model for the detection task with imperfect recognition. The model is an extension of one proposed earlier by Atkinson and Kinchla (1965).

The Atkinson-Kinchla model assumes two processes: an activation and a decision process. The activation process converts the visual stimulus into sensory information which is represented in the model by one of three possible sensory states. The decision process acts on the sensory information to determine a response by the subject. The model proposed here subdivides the activation process into two processes, detection and recognition. Each of the processes can be characterized by a transition matrix whose entries indicate the theoretical probabilities of entering the various states listed across the top of the matrix, given the initial conditions listed to the left of the matrix. Before presenting the matrix representation of the three processes, we introduce the following terminology:

S_i = an observable state indicating presentation on a given trial of critical element i ($i = 1$ or 2)

A_i = an observable state representing a response by the subject indicating that S_i was the critical element presented ($i = 1$ or 2)

d_0 = a hypothetical state indicating that the subject did not detect a critical element

d_i = a hypothetical state representing detection by the subject of critical element S_i ($i = 1$ or 2)

r_0 = a hypothetical state entered by the subject when he detects but fails to recognize the critical element

r_i = a hypothetical state indicating both detection and recognition of S_i ($i = 1$ or 2)

We can now present the three matrices characterizing the detection, recognition and decision processes.

$$
D = \begin{array}{c} \\ S_1 \\ S_2 \end{array}
\begin{array}{c}
\begin{array}{ccc} d_1 & d_2 & d_0 \end{array} \\
\begin{bmatrix} \sigma & 0 & 1-\sigma \\ 0 & \sigma & 1-\sigma \end{bmatrix}
\end{array}
$$

The detection matrix $\underset{\sim}{D}$ is identical to the matrix for the activation process in the Atkinson-Kinchla model.

$$
R = \begin{array}{c} \\ d_1 \\ d_2 \\ d_0 \end{array}
\begin{array}{c}
\begin{array}{cccc} r_1 & r_2 & r_0 & d_0 \end{array} \\
\begin{bmatrix} \delta & 0 & 1-\delta & 0 \\ 0 & \delta & 1-\delta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{array}
$$

The recognition matrix $\underset{\sim}{R}$ allows for the possibility that recognition following detection may not be perfect. The sensory states r_1, r_2

and d_0 are equivalent to the three sensory states in the Atkinson-Kinchla model. The addition of state r_0 requires elaboration of the decision process.

$$
G = \begin{array}{c c} & \begin{array}{cc} A_1 & A_2 \end{array} \\ \begin{array}{c} r_1 \\ r_2 \\ r_0 \\ d_0 \end{array} & \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ a & 1-a \\ b & 1-b \end{bmatrix} \end{array}
$$

Notice that the decision matrix contains two bias parameters, a and b, instead of one as in the Atkinson-Kinchla model. Guessing is governed by bias parameter a when a subject detects a critical element but does not recognize which of the critical elements was presented, and by b when the critical element is not detected. The first experiment described below attempts to show that there are conditions under which these two bias parameters are not equal.

By taking the product of the above three matrices, we obtain the performance matrix. The entries in this matrix are probabilities that can be directly estimated from the experimental data.

$$
P = D \cdot R \cdot G = \begin{array}{c c} & \begin{array}{cc} A_1 & A_2 \end{array} \\ \begin{array}{c} S_1 \\ S_2 \end{array} & \begin{bmatrix} \sigma\delta + \sigma(1-\delta)a + (1-\sigma)b & \sigma(1-\delta)(1-a) + (1-\sigma)(1-b) \\ \sigma(1-\delta)a + (1-\sigma)b & \sigma\delta + \sigma(1-\delta)(1-a) + (1-\sigma)(1-b) \end{bmatrix} \end{array}
$$

There are three conditions under which the extended performance matrix reduces to that of the Atkinson-Kinchla model: when $\sigma = 1$, when $\delta = 1$, or when $a = b$. If $\sigma = 1$, then the task involves only recognition, whereas if $\delta = 1$, the task involves only detection. Even if $\sigma < 1$ and $\delta < 1$, the extended model will, for the standard forced choice detection task, be indistinguishable from the simpler high threshold model when $a = b$. When $a = b$ the difference between the two models rests on the interpretation of the parameters. For the Atkinson-Kinchla model, the one parameter corresponding to $\sigma\delta$ represents the subject's ability to detect the presence of a critical element, whereas in our model $\sigma\delta$ represents a combination of detection and recognition.

Clearly, some method is needed to separate σ and δ. Data from the typical forced choice detection task are insufficient to this end. One way to accomplish the separation of σ and δ is to convert the two-choice task into one involving four responses. In the four-choice task, the subject's response would indicate not only which critical element he thought was presented, but also where in the display he thought it occurred. For instance, he might be required to decide whether the critical element was in the right or left half of the visual display. It is assumed that his performance on this portion of the task would be governed by only σ, the detection parameter. That is, given detection by the subject, it is assumed that he would respond perfectly regarding the side of the display on which the critical element had been presented. Thus σ could be estimated from that portion of the data. Using the estimate of σ, it would then be possible to obtain an estimate of δ.

Employing a procedure like the one described above, two experiments were run to study, within the framework of the above model, various aspects of a task involving detection with imperfect recognition. The first experiment involved an attempted manipulation of the bias parameters a and b, while the second investigated the relation between the sensitivity parameters σ and δ.

Experiment I

Can conditions be found under which the bias parameters a and b are unequal? The first experiment was an attempt to answer this question. Consider a situation in which the subject is given feedback at the end of each trial as to whether S_1 or S_2 had occurred on that trial. Let E_1 be the feedback if S_1 was presented and E_2 if S_2 was presented. Now if the probability of an E_1 when the subject is in sensory state d_0, $Pr(E_1|d_0)$, differs from the probability of an E_1 when the subject is in sensory state r_0, $Pr(E_1|r_0)$, then it seems reasonable to suspect that this might eventually result in two different guessing strategies, one employed when the subject is in state d_0, the other when in state r_0.

One way to create a difference between $Pr(E_1|d_0)$ and $Pr(E_1|r_0)$ is through the use of "blank" displays, that is, displays containing only noise elements. Since there is no critical element in such a display, according to the model the subject must go into state d_0. Now, if for blank displays, the probability of an E_1 differs from the probability of an E_1 for "signal" displays (i.e., displays containing a critical element), then $Pr(E_1|d_0)$ and $Pr(E_1|r_0)$ will differ. This first experiment employed blank and signal displays in a situation involving feedback to the subject. Because of the nature of the critical elements used in this study, it was possible to have a third type of display, called a "neutral" display. For these neutral displays, it was expected that subjects could go only into states d_0 or r_0. This will become clearer when the displays are described below. By choosing the probability of an E_1 for neutral displays to be different from the probability of an E_1 for both signal and blank displays, it was hoped that we could amplify the difference between a and b.

Method

The subjects for the first experiment were six Stanford undergraduates, four male and two female. All subjects reported having normal vision. They were each paid $12.25 for the seven sessions of the experiment.

The task, as described above, consisted of subjects judging, on each of a series of trials, which of the two critical elements had been presented in a display, and also on which side of the display the critical element had occurred. Each signal display consisted of one critical element and 15 noise elements. Figure 1 shows the two critical elements, S_1 and S_2, used in signal displays along with a noise element. All of the noise elements were identical. The tic in the critical elements was tilted 30° to the left (S_1) or right (S_2) of vertical. This type of display was used instead of displays made up of letters of the alphabet for two reasons. First, with this type of symbol, recognition could be easily controlled by simply varying the tilt of the tic in the critical elements. Second, the homogeneous background resulting from only one type of noise element increases the likelihood that the two critical elements will be equally detectable, that is, that the same value of σ will apply to both of them.

The 16 elements were arranged in a 4 x 4 matrix, with a vertical line dividing the display into two equal halves. Figure 2 shows a typical signal display. There were 32 signal displays consisting of two matching sets of 16 which differed only in the critical element, i.e., S_1 or S_2, which was presented. Each critical element appeared once in each of the 16 matrix positions.

There were 16 different neutral displays. These displays were identical to the signal displays except that instead of S_1 or S_2 as critical elements, they contained an S_0, i.e., a critical element in which the tic was vertical rather than tilted 45° right or left. S_0 is shown in Figure 1. The blank display was identical to the other two types except that there was no critical element; that is, it consisted of 16 noise elements.

The experimental apparatus employed was an automated two-field dual tachistoscope. The display terminus of the apparatus was located in a soundproof, air-conditioned room. It appeared as a wooden box 5 ft. 7 in. long, 4 ft. 1 in. wide, and 2 ft. 4 in. high, standing on four 8 in. legs on a table. At each of the two ends was a subject station, formed by a recess 8 in. wide and 10 in. deep, the full height of the box. In the recess was a panel containing a ground-glass rear-projection screen, 8 in. wide and 6.75 in. high, centered vertically. Behind the screen was a black metal plate bearing four lights and a large circular aperture, none of which was visible unless illuminated. The four lights were arranged in a horizontal row above the aperture, two to the left of the aperture and two to the right. A plastic eyepiece was mounted flush with the outer face of the box, 10 in. in front of the screen, aligned in height with the circular aperture. Below each observation station was suspended a response panel, at lap height, 12 in. wide and 10 in. deep. This bore two arrays of four adjacent rectangular buttons, each of which was 1 in. by 7/8 in.; one array was arranged vertically and one horizontally.

Displays were projected onto the screen through the large aperture, providing an illuminated circle 2-1/16 in. in diameter. Stimuli were displayed by a random-access slide projector (Spindler & Sauppe model SLX-750) modified to mount a special light source (Sylvania electronic tube #R1131C) characterized by rise time within 0.05 msec. and decay time within 0.025 msec. A second projector, optically identical to the first but holding a single slide, served to illuminate the screen between stimulus exposures. Both projectors were concealed within the display box. Brightness determinations, made with a Macbeth Illuminometer, for the primary (stimulus) and secondary fields were 2.3 and 3.6 apparent foot-candles for observation station #1, and 1.6 and 1.4 apparent foot-candles for station #2. A given trial consisted of the following sequence of events:

1. A pre-exposure light came on for 3 sec.

2. The pre-exposure light went off and the display came on. A warning click over the intercom preceded the onset of the display by .5 sec.

3. The display went off and the post-exposure light came on. During the first 2 sec. following offset of the display the subject made his response.

4. One of four reinforcement lights came on. This light remained on for 3 sec.

5. The reinforcement light and the post-exposure light went off. The screen remained dark for .1 sec. while the next trial was read in by the tape reader.

The subject's response was made by pressing one of the four buttons in the horizontal row on the response panel. Numbering the buttons 1 to 4 from left to right, button 1 indicated an S_1 on the left side of the line, button 2 an S_1 on the right side, button 3 an S_2 on the left side, and button 4 an S_2 on the right side. The response panel contained a black vertical line between buttons 2 and 3. Buttons 1 and 3 were labeled with critical element S_1, while 2 and 4 were labeled with S_2. Numbering the reinforcement lights 1 to 4 from left to right, the reinforcement light that came on indicated reinforcement of that response having the corresponding number.

Before describing the structure of the experimental sessions, let us introduce the following additional notation:

D_s = presentation of a signal display (noise elements plus an S_1 or S_2 element)

D_n = presentation of a neutral display (noise elements plus an S_0 element)

D_b = presentation of a blank display (only noise elements)

S_L = presentation of a critical element (S_1, S_2, or S_0) to the left of the dividing line

S_R = presentation of a critical element to the right of the dividing line

A_L = a response by the subject indicating presentation of S_L

A_R = a response by the subject indicating presentation of S_R

E_L = feedback indicating presentation of S_L

E_R = feedback indicating presentation of S_R

An experimental session consisted of 240 trials. There was an equal number of signal, neutral and blank trials. For signal trials, S_1 and S_2 were presented equally often. The critical element was equally likely to be on the right or left of the vertical line. The reinforcement light for a signal trial indicated which critical element was in the display and on which side of the dividing line it had occurred.

For neutral trials, S_L and S_R were equally likely to be presented. The reinforcement light correctly informed the subject as to which side of the dividing line the critical element was on, however, the subject was falsely informed that either S_1 or S_2 had occurred. The probability of an E_1 equalled .1, while the probability of an E_2 equalled .9.

On the blank trials, subjects were falsely informed that a critical element had been presented in either the right or left half of the display. E_L and E_R were presented equally often, while the probability of E_1 equalled .9 and the probability of E_2 equalled .1. Figure 3 summarizes the probabilities of the various events on a given trial.

Each session lasted about 35 minutes. Subjects were run for a total of seven sessions. The first two sessions consisted of only signal trials; that is, trials with displays containing S_1 or S_2. During the first session the exposure time of the stimulus was decreased after every block of 48 trials. The exposure time ranged from 100 msec. to 15 msec. Based on his performance on the first day, an exposure time was chosen for each subject for the second day so that all subjects had about the same level of performance. Sessions three through seven constituted the experimental sessions during which displays were presented and reinforcements given according to the schedule described above.

Results

On each trial the subject indicates both where in the display he thought the critical element had occurred and which of the two critical elements he thought had been presented. We will refer to the former as the "detection response" and the latter as the "recognition response." Using this terminology, we now introduce the following notation:

C_d = a correct detection response

I_d = an incorrect detection response

C_r = a correct recognition response

I_r = an incorrect recognition response

As stated above, it is assumed that the detection response is a function of σ only. This would not be true if it were possible for the subject to detect a critical element without being able to decide whether it was to the right or left of the dividing line. To test the validity of this assumption, we look at $Pr(C_r|I_d)$, the probability of a correct recognition response given an incorrect detection response. Of course this probability is defined only for signal displays. If the assumption is wrong, then on some proportion of the trials involving an error in detection, the subject will actually have seen the critical element and will recognize it with probability δ. Therefore $Pr(C_r|I_d)$ would be greater than chance; that is, greater than .5. If, however, the assumption is correct, a detection error always indicates a failure to detect; thus the probability of a correct recognition response should be .5. The observed $Pr(C_r|I_d)$ over the last four sessions for the six subjects is as follows: .55, .55, .54, .57, .58, .71. It appears that our assumption is not strictly correct. However, except for one subject, the deviations from chance are not great enough to cause rejection of the assumption. The one extreme subject made only 18 detection errors on signal trials over the last four experimental sessions. Therefore his observed $Pr(C_r | I_d)$ is not a very reliable estimate of the true value.

It was found that four of the subjects were more likely to make a correct detection response when the critical element was to the left of the vertical line than when it was to the right; one subject showed the reverse, while one showed no difference. Because of the differences between the two sides, it was decided that a separate σ must be estimated for each side of the display. To obtain separate estimates of σ for right and left it is necessary to estimate any bias that might exist for a right or left detection response. Let us therefore define the following terms:

σ_L = the probability of detection given presentation of S_L

σ_R = the probability of detection given presentation of S_R

p = the probability of A_L given failure to detect

The equations below give the theoretical expressions for various probabilities observed in the data. From these equations estimates of all of the parameters in the model can be obtained.

$$Pr(A_L | D_b) = p$$

$$Pr(A_L | D_s \;\&\; S_L \;\text{ or }\; D_n \;\&\; S_L) = \sigma_L + (1 - \sigma_L)p$$

$$Pr(A_L | D_s \;\&\; S_R \;\text{ or }\; D_n \;\&\; S_R) = (1 - \sigma_R)p$$

$$Pr(A_l | D_b) = b$$

$$\Pr(A_1|D_n) = \sigma a + (1 - \sigma)b \quad \text{where} \quad \sigma = \frac{\sigma_L + \sigma_R}{2}$$

$$\Pr(C_r|D_s) = \sigma\delta + (1 - \sigma\delta)\frac{1}{2}$$

Solving the above expressions for the various parameters, we obtain the following estimates:

$$\hat{p} = \Pr(A_1|D_b)$$

$$\hat{\sigma}_L = \frac{\Pr(A_L|D_s \;\&\; S_L \;\text{ or }\; D_n \;\&\; S_L) - \hat{p}}{1 - \hat{p}}$$

$$\hat{\sigma}_R = 1 - \frac{\Pr(A_L|D_s \;\&\; S_R \;\text{ or }\; D_n \;\&\; S_R)}{\hat{p}}$$

$$\hat{b} = \Pr(A_1|D_b)$$

$$\hat{a} = \frac{\Pr(A_1|D_n) - (1 - \sigma)\hat{b}}{\hat{\sigma}} \quad \text{where} \quad \hat{\sigma} = \frac{\hat{\sigma}_L + \hat{\sigma}_R}{2}$$

$$\hat{\delta} = \frac{2\Pr(C_r|D_s) - 1}{\hat{\sigma}}$$

Table 1 gives the estimates of the six parameters for each of the six subjects. It is clear from the values of $\hat{\delta}$ that this task involves both detection and recognition. If recognition were perfect following detection, $\hat{\delta}$ would be equal to 1. It could be argued that $\hat{\delta}$ was less than 1 for the reason that having to make both a detection and a recognition response impairs the subject's performance on the recognition phase of the task. A preliminary study showed that performance on the recognition task was the same regardless of whether or not the subject made a detection response.[2]

It is interesting to note that there appears to be no correlation between the value of $\hat{\sigma}$ for a subject and his value of $\hat{\delta}$. In other words, at least for this type of display, a subject's ability to detect the critical element in the display does not seem to be related to his ability to recognize that element following detection.

For all subjects except subject 6, $\hat{b}$ is greater than $\hat{a}$. If it is assumed that subjects are trying to use information from feedback to increase their proportion of correct responses, then this difference is in the expected direction -- i.e., for the feedback schedule used, $Pr(E_1|d_0)$ is greater than $Pr(E_1|r_0)$. Even though $\hat{b}$ is less than $\hat{a}$ for subject 6, the fact that there is a difference between $\hat{a}$ and $\hat{b}$ supports the hypothesis that two bias parameters are operating in this situation rather than one. It is clear, however, that for most subjects the difference is rather small. For subjects 1, 3 and 4, this difference was not significant at the .05 level.[3]

Discussion

It could be argued that if the subjects had been run for a number of sessions beyond the five in this study, that they would have eventually shown large differences between $\hat{a}$ and $\hat{b}$. There is some evidence that the differences were still increasing at the end of this study. The mean value of $\hat{a}$ over the first three sessions was .54 and for $\hat{b}$, .51. Over the last two sessions $\hat{a} = .46$ and $\hat{b} = .54$. If subject 6 is excluded from the average, then over the first three sessions $\hat{a} = .54$ and $\hat{b} = .54$; while for the last two sessions $\hat{a} = .45$ and $\hat{b} = .57$.

There is another possible explanation for the lack of large differences. It was assumed that for neutral displays, the subject went into either state d_0 or state r_0. Doubtlessly, on some neutral trials subjects went into r_1 or r_2. In fact, it could be argued that subjects never go into state r_0 -- that all recognition errors following a detection are due to confusion of one critical element with the other. If this were true, then $\hat{a}$ would not be an estimate of the subject's bias, but of his probability of confusing the critical element in a neutral display with an S_1, in which case manipulation of the reinforcement schedule would not necessarily affect it. The use of confidence ratings would help to resolve this issue. Experiment II employed such ratings.

Experiment II

The purpose of this second experiment was to investigate how the sensitivity parameters σ and δ changed with changing display size. In the usual two-choice detection task, the probability of a correct response decreases as the number of noise elements increases (Estes and Taylor, 1965). Is this decrease due to a decrease in σ only, or a decrease in both σ and δ ? If detection is truly all or none, then one would expect δ to remain constant over display sizes. On the other hand, if the all-or-none assumption is incorrect, it seems reasonable to assume that the larger the number of noise elements in the display, the smaller the amount of information the subject will obtain about the critical element and therefore the smaller δ will be.

Method

The subjects for this second experiment were eight Stanford undergraduates, five male and three female. All subjects reported having normal vision, corrected or uncorrected. They were paid $16.25 for the nine sessions of the study.

The elements making up the displays for this study were of the same type as those used for the signal displays in the first experiment. There were six types of displays, differing as to the number of noise elements they contained. Potential 16-cell matrices, four elements wide and four elements high, were filled with 1, 2, 4, 6, 8, or 16 elements. For each display size, there were 32 displays. These were two matching sets of 16 displays which differed only in the critical element (S_1 or S_2) that was presented. Each critical element appeared once in each of the possible 16 cells of the 4 x 4 matrix. A vertical line divided each display in half, bisecting the potential 16-cell matrix. For all but the one element displays, there was an equal number of elements on each side of the vertical line.

The apparatus for this experiment was the same as for the first study. The sequence of trial events was also the same as for Experiment I, with the following exceptions. The pre-exposure light was on for 5 seconds instead of 3. The subject had 3 seconds to respond instead of 2. Following the response, the subject gave one of three confidence ratings: 1 if he was sure which critical element was presented and also which side it was on; 2 if he thought he knew which side it was on but not which one it was; and 3 if he did not see a critical element at all.

The confidence rating was made by pressing one of three buttons in a vertical row on the response panel.

Each subject was run for a total of nine days. On day 1 the exposure time of the stimulus was set at 100 msec. and was decreased by steps down to 15 msec. Based on the results of day 1, the subjects were run at tentative exposure times on day 2. As in Experiment I, an attempt was made to obtain about the same level of performance from all subjects. On the third day any final necessary adjustments in the exposure times were made. The exposure time for each subject remained constant over the remaining six days. The exposure times used were 20 msec. for four of the subjects, 25 msec. for two, and 30 msec. for the remaining two.

Results

The observed quantities used to estimate parameters were the proportions of correct and incorrect detection responses along with the proportions of correct detection responses followed by either a correct or an incorrect recognition response. The theoretical expressions for these quantities are the following:

$$Pr(C_d) = \sigma + (1 - \sigma)\frac{1}{2}$$

$$Pr(I_d) = (1 - \sigma)\frac{1}{2}$$

$$Pr(C_d \,\&\, C_r) = \sigma\delta + \sigma(1 - \delta)\frac{1}{2} + (1 - \sigma)\frac{1}{4}$$

$$Pr(C_d \,\&\, I_r) = \sigma(1 - \delta)\frac{1}{2} + (1 - \sigma)\frac{1}{4}$$

Solving the above system of equations for σ and δ yields the following estimates:

$$\hat{\sigma} = Pr(C_d) - Pr(I_d)$$

$$\hat{\delta} = \frac{Pr(C_d \ \& \ C_r) - Pr(C_d \ \& \ I_r)}{\hat{\sigma}}$$

It can easily be shown that these are maximum likelihood estimates.

It was not necessary to estimate separate σ's for the right and left half of the displays. The probability of an A_L given a confidence rating of 3 serves as an estimate of p, the bias parameter for a detection response. For no subject did this estimate differ significantly from .5, therefore the formula given above provides an accurate estimate of the overall value of σ. Considering the results of the first study, it is a little surprising that there was no detection bias in this experiment.

Table 2 shows the estimates of σ and δ for each subject at each display size. Figure 1 shows a plot of $\hat{\sigma}$ and $\hat{\delta}$ against display size for the average over seven of the eight subjects. Subject 3 was discarded due to his abnormally low probability of a correct detection response for one element displays. Also, since his $\hat{\delta}$ for four of the display sizes was 0, entering his values into the average would give the plot of $\hat{\delta}$ a false look of stability over the last four points. It is clear from Figure 1 that while $\hat{\sigma}$ shows a steady decrease with increasing display size, $\hat{\delta}$ shows no such trend.

Table 3 presents for each display size the 16 basic probabilities averaged over seven subjects; that is, the probability of each of the four possible responses given each of the four possible stimuli. The 16 corresponding theoretical probabilities are also given in the table. The theoretical probabilities are estimated from seven parameters -- a $\hat{\sigma}$ for each display size plus one overall $\hat{\delta}$. Most subjects had a higher probability of a correct recognition response given the stimulus combinations S_R, S_1 or S_L, S_2 than given either of the other two stimulus combinations. This was a major source of differences between the theoretical and empirical probabilities. The reason for this is not clear, particularly when one notes that the probability of a correct detection response is approximately constant across the four stimulus combinations. The observed results could be caused by an unusual response bias employed by subjects when they are in state r_0, however the data show no evidence of any such bias. Since confidence rating 2 is the rating given when subjects feel that they have detected but not recognized the critical element, the probabilities of the various responses given a rating of 2 should serve as an estimate of any bias when in state r_0. Given a confidence rating of 2, subjects made all four responses with approximately equal frequency.

The first study brought up the problem of how often the two critical elements are confused with one another. The data from the confidence ratings in this study can help to answer that question. If we make the assumption that when the subject gives a 1 as his confidence rating he is in either state r_1 or r_2, then his probability of an error on the recognition response provides an estimate of the number of times he confuses S_1 and S_2. Averaged over all subjects and display sizes, the probability of an incorrect recognition response given a confidence rating of 1 equals .11. Averaging over display sizes is justified since this probability showed no change with increasing display size. While this estimate of the probability of confusion is based on the somewhat tenuous assumption given above, it indicates that confusions are not the major cause of recognition errors following a detection. Therefore the recognition process as formulated in this paper seems adequate for this data.

Discussion

Neisser (1967) cites considerable evidence supporting the existence of a visual trace lasting as much as five seconds after tachistoscopic presentation of a stimulus. Subjects can gain information about the stimulus by, in some sense, scanning this trace. Neisser refers to the trace as the "icon" in order to differentiate it from other related constructs. This icon is generally considered to fade gradually following presentation of the stimulus. While gradual fading is an intuitively appealing supposition, the data from this experiment do not appear to support it. If the icon were a gradually fading trace, then the larger the display size, the more on the average the image would have faded by the time the subject located the critical element. Greater fading should make recognition more difficult, thus giving rise to smaller values of δ for larger display sizes. The observed constancy of $\hat{\delta}$ over display size supports the alternative assumption that elements in the trace drop out in an all-or-none manner. However, the support given by the data to that assumption is somewhat indirect. If the all-or-none assumption is really correct, then δ should remain constant over variation in the duration of the icon. The icon's duration is affected by variables such as the brightness of the stimulus field and the exposure time of the display. Constancy of $\hat{\delta}$ over changes in these variables would provide much more convincing evidence for the all-or-none assumption than that given here. Experiments involving manipulation of these variables are planned for the near future.

While the evidence supports an all-or-none model of the detection process, the data in this paper give little direct information as to whether or not the recognition process following detection is also best characterized by an all-or-none model. In detection, the subject must find the critical element in the display; this is not difficult to visualize as an all-or-none process. However, for the displays in these studies, recognition involves a determination by the subject of the orientation of the tic in a critical element. It does not seem nearly as natural to view this as an all-or-none process. A second guess study similar to that of Yellott and Curnow (1967) should help to determine the nature of the recognition process.

One could argue that the data presented both in this and the previous experiment are dependent on the special nature of the stimuli used. Perhaps for other types of displays, detection and recognition cannot be so neatly separated. Consider, for example, displays with only one type of noise element, such as those used here. Assume that one of the critical elements is differentiated from the noise elements by its value on some dimension, say shape. Now let the other critical element be differentiated from the noise elements by its value on some other dimension, color for instance. Clearly, for this type of display, detection implies recognition, therefore no separate recognition process is involved in the task. However, whenever the means used to differentiate noise elements from critical elements do not also allow complete discrimination between critical elements, a model combining detection and recognition processes should be applicable.

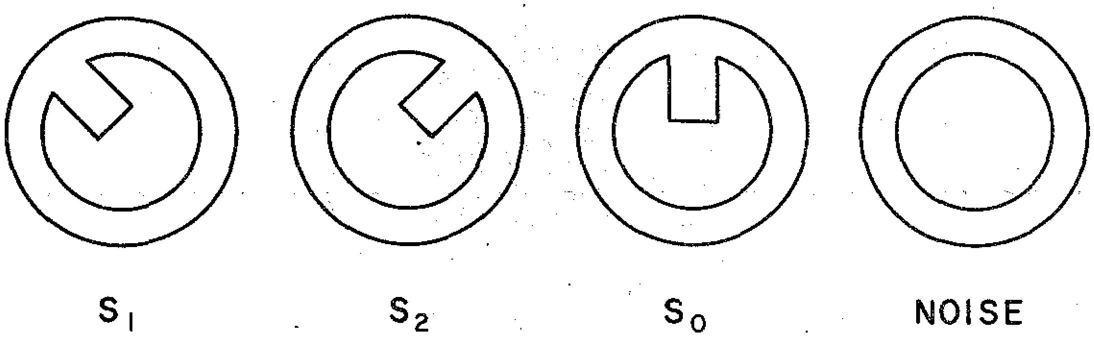

Fig. 1. The four symbols used in the displays for Experiment I.

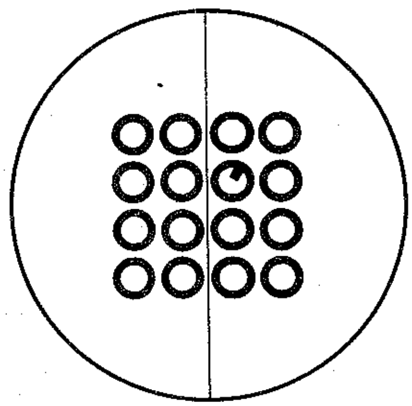

Figure 2. A typical display; actual size

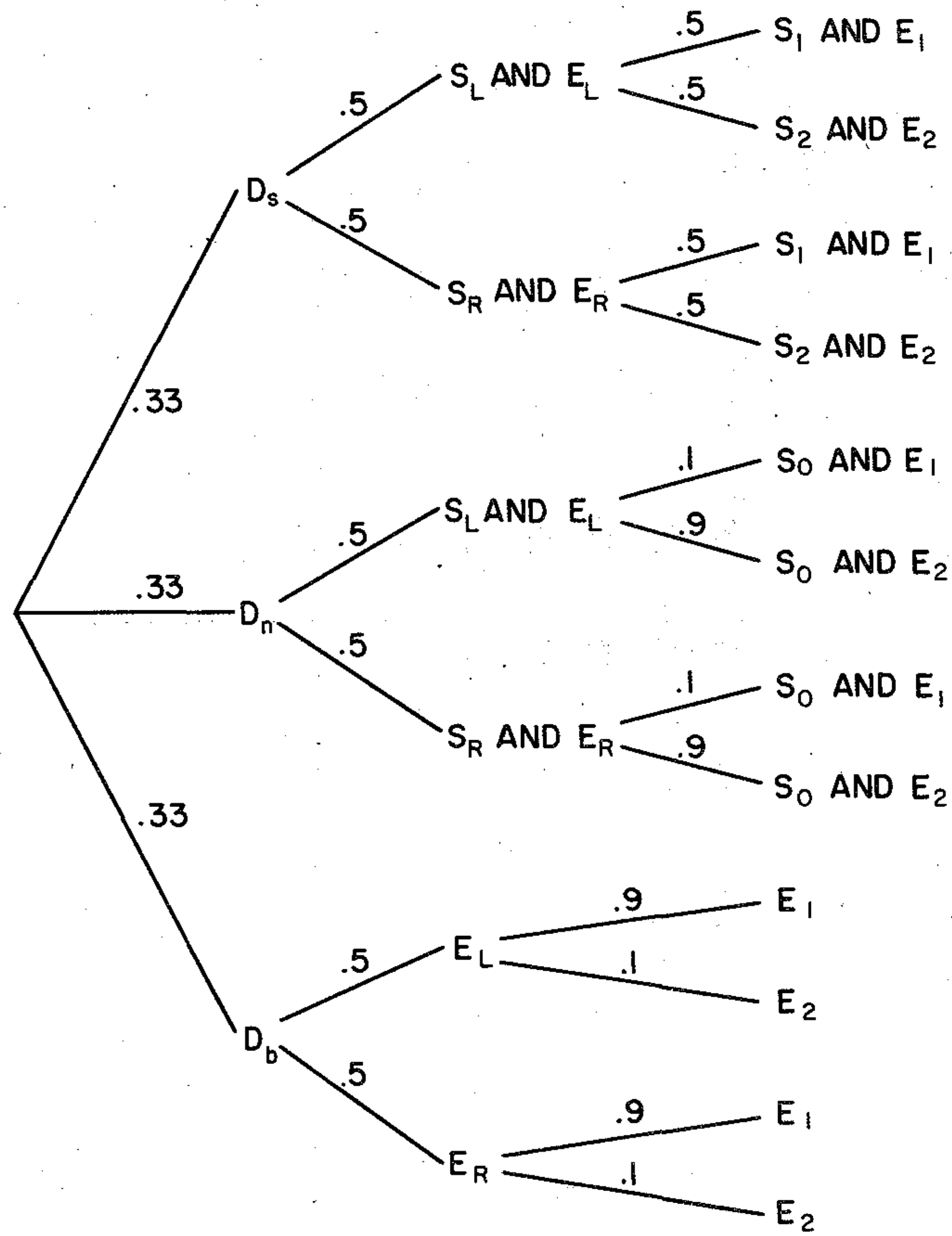

Fig. 3. Probabilities of various events for a given trial of Experiment I.

Subjects	$\hat{\sigma}_L$	$\hat{\sigma}_R$	$\hat{\delta}$	$\hat{p}$	$\hat{b}$	$\hat{a}$	$\hat{b}-\hat{a}$
1	.55	.55	.74	.50	.53	.47	.06
2	.97	.73	.74	.48	.55	.45	.10
3	.57	.67	.75	.54	.50	.46	.04
4	.79	.67	.66	.62	.61	.57	.04
5	.75	.41	.61	.37	.66	.30	.36
6	.93	.88	.68	.38	.35	.49	-.14

Table 1. Parameter estimates for Experiment I.

		Display Size 1		2		4		6		8		16	
		$\hat{\sigma}$	$\hat{\delta}$	$\hat{\sigma}$	$\hat{\delta}$	$\hat{\sigma}$	$\hat{\delta}$	$\hat{\sigma}$	$\hat{\delta}$	$\hat{\sigma}$	$\hat{\delta}$	$\hat{\sigma}$	$\hat{\delta}$
Subjects	1	0.98	0.70	0.88	0.60	0.69	0.61	0.63	0.42	0.56	0.74	0.37	0.56
	2	0.93	0.89	0.96	0.58	0.81	0.77	0.77	0.71	0.59	0.65	0.50	0.83
	3	0.71	0.15	0.38	0.25	0.35	0.00	0.30	0.00	0.11	0.00	0.00	0.00
	4	0.96	0.89	0.99	0.73	0.91	0.76	0.95	0.73	0.83	0.74	0.75	0.62
	5	0.98	0.10	0.70	0.24	0.52	0.10	0.40	0.34	0.33	0.11	0.13	0.46
	6	1.00	0.63	0.92	0.67	0.78	0.73	0.75	0.72	0.68	0.78	0.56	0.42
	7	0.97	0.68	0.70	0.80	0.49	0.72	0.55	0.80	0.51	0.74	0.28	0.87
	8	0.99	0.89	0.95	0.74	0.87	0.82	0.84	0.81	0.79	0.81	0.71	0.70

Table 2. Estimates of σ and δ for each subject at each display size.

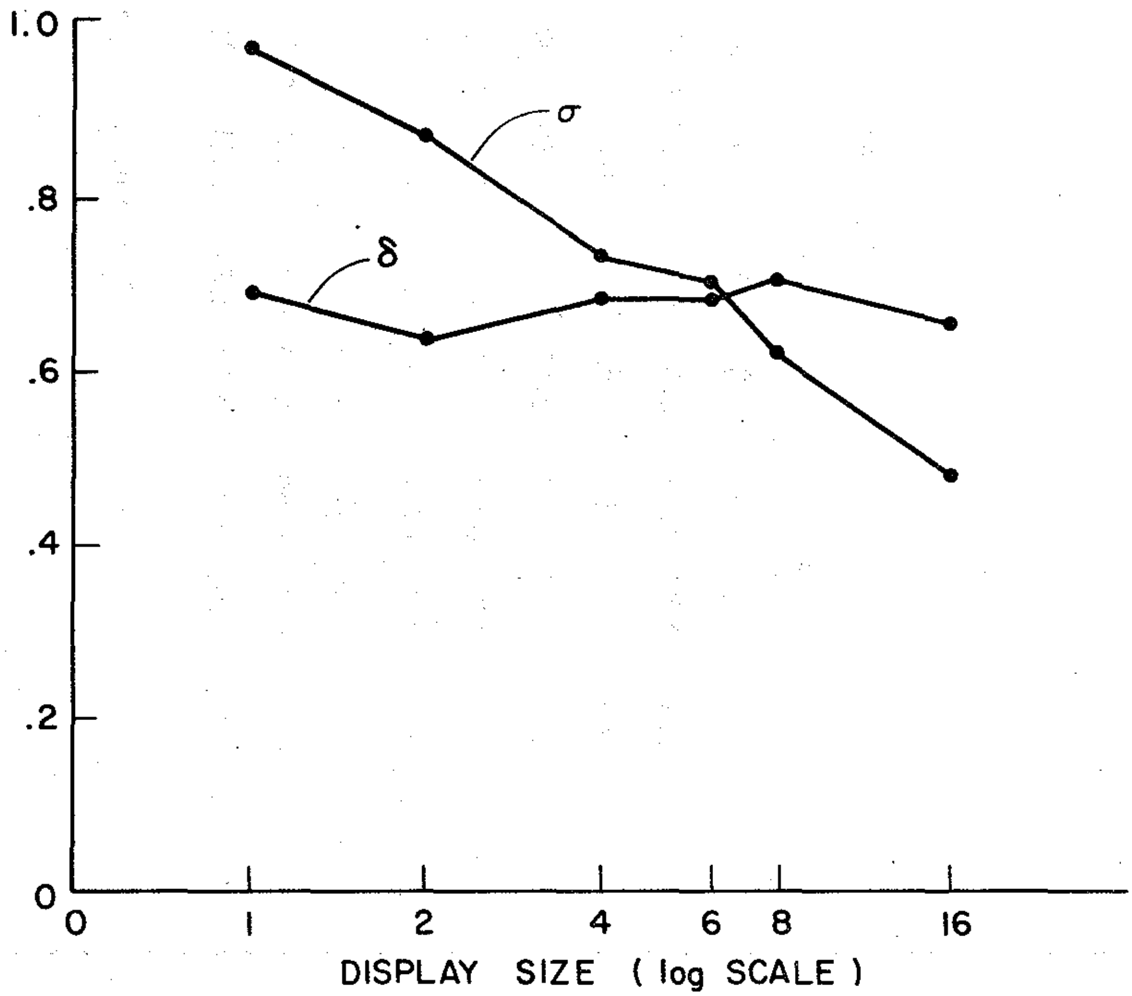

Fig. 4. Estimated values of σ and δ for Experiment II

Display Size		S_L								S_R							
		S_1				S_2				S_1				S_2			
		A_L		A_R		A_L		A_R		A_L		A_R		A_L		A_R	
		A_1	A_2	A_1	A_2	A_1	A_2	A_1	A_2	A_1	A_2	A_1	A_2	A_1	A_2	A_1	A_2
1	obs.	.799	.195	.006	.000	.141	.838	.012	.009	.009	.000	.903	.088	.000	.021	.226	.753
	th.	.817	.169	.007	.007	.169	.817	.007	.007	.007	.007	.817	.169	.007	.007	.169	.817
2	obs.	.686	.242	.039	.033	.120	.790	.054	.036	.027	.039	.787	.147	.003	.033	.265	.699
	th.	.757	.177	.033	.033	.177	.757	.033	.033	.033	.033	.757	.177	.033	.033	.177	.757
4	obs.	.630	.236	.081	.054	.159	.715	.069	.057	.039	.098	.710	.153	.040	.114	.196	.650
	th.	.673	.189	.069	.069	.189	.673	.069	.069	.069	.069	.673	.189	.069	.069	.189	.673
6	obs.	.614	.248	.088	.051	.147	.697	.102	.054	.054	.102	.730	.114	.065	.093	.247	.595
	th.	.656	.192	.076	.076	.192	.656	.076	.076	.076	.076	.656	.192	.076	.076	.192	.656
8	obs.	.503	.276	.149	.073	.131	.698	.097	.074	.071	.120	.628	.180	.069	.123	.169	.639
	th.	.607	.199	.097	.097	.199	.607	.097	.097	.097	.097	.607	.199	.097	.097	.199	.607
16	obs.	.414	.274	.140	.173	.145	.605	.123	.127	.075	.172	.536	.217	.055	.193	.228	.524
	th.	.525	.210	.132	.132	.210	.525	.132	.132	.132	.132	.525	.210	.132	.132	.210	.525

Table 3. Observed and theoretical probabilities of each response given each stimulus

Footnotes

[1]This research was supported by Grant NGR-05-020-036 from the National Aeronautics and Space Administration.

[2]A study was run in which a series of displays were presented tachistoscopically to five subjects. There were 32 displays, identical to the signal displays defined in Experiment I. On each trial subjects were required to respond in one of three ways. They made a detection response only, a recognition response only, or both detection and recognition responses. An experimental session was divided into thirds, a single type of response being given in each third. Subjects were run for eight days, 192 trials per day. The first two days were practice sessions; data was gathered over the remaining six sessions. The order in which the three response types were used by the subjects was unique on each day, all possible orders being given over the six experimental sessions. The probability of a correct recognition response, averaged over subjects, was .74 when both detection and recognition responses were given, while for recognition alone it was .73. For no subject did the difference between these two probabilities reach the .05 level of significance.

[3]The significance of $\hat{b} - \hat{a}$ was tested by assuming that $\hat{a}n_1$ and $\hat{b}n_2$ are binomially distributed where

n_1 = the number of neutral trials on which the subject goes into state r_0,

n_2 = the number of blank trials on which the subject goes into state d_0.

It is assumed that n_1 equals $\hat{\sigma}$ times the total number of neutral trials; n_2 is taken to be the total number of blank trials. We can then use the following statistic to test the significance of $\hat{b} - \hat{a}$:

$$z = \frac{\hat{b} - \hat{a}}{\sqrt{\frac{\hat{a}+\hat{b}}{2}\left(1 - \frac{\hat{a}+\hat{b}}{2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

References

Atkinson, R. C. and Kinchla, R. A. A learning model for forced-choice detection experiments. _British Journal of Mathematical and Statistical Psychology_, 1965, _18_, 183-206.

Estes, W. K. and Taylor, H. A. A detection method and probabilistic models for assessing information processing from brief visual displays. _Proceedings 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 & Psychophysics_, 1965, _1_, 9-16.

Luce, R. D. Detection and Recognition. In R. D. Luce, R. R. Bush and E. Galanter (Eds.), _Handbook of Mathematical Psychology_, Vol. 1. New York: Wiley, 1963, pp. 103-189.

Neisser, U. _Cognitive Psychology_. New York: Appleton-Century-Crofts, 1967.

Yellott, J. I., Jr. and Curnow, P. F. Second choices in a visual span of apprehension task. _Perception & Psychophysics_, 1967, _7_, No. 2, 307-311.