Bulletin of the Psychonomic Society
1974, Vol. 2 (4B), 411-414

Visual detection accuracy and target-noise proximity*

WILLIAM P. BANKS†, DEBORAH BODINGER††, and MARTHA ILLIGE†††
Pomona College, Claremont, California 91711

This experiment demonstrates what we call a *proximity effect* in forced-choice visual detection: Detection accuracy improves as the distance is increased between the target and the noise items in the array that are confusable with it. The proximity effect is a natural prediction of Estes's theory that detection is mediated by feature-detecting receptive fields, but other recent models of visual detection do not predict it. However, the pattern of results seems best explained in terms of perceptual configurations in the array; when the target is grouped with confusable noise its visibility is less than when it is not.

In the forced-choice visual detection paradigm, the S is given brief presentations of arrays of characters containing one of a small prespecified set of targets and a number of nontarget (noise) items. He has the task of reporting which target was contained in each array. Estes (1972) has recently presented a model for performance in this and similar visual detection paradigms and applied it with some success to the major findings in the area. According to Estes's model, detection of the target element is mediated by feature-detecting receptive fields. The model assumes that the fields are fairly small, that there are a large number of receptive fields of any given type distributed over the retina, and that the fields are most densely packed at the fovea. The model also assumes mutual inhibitory interactions between receptive fields, and assumes further that the mutual inhibition decreases as distance between the fields increases.

Estes (1972) reviewed the three principal varieties of model for visual detection tasks (serial scanning models [e.g., Estes & Taylor, 1964]; parallel limited capacity models [Rumelhart, 1970]; and parallel unlimited capacity models [e.g., Gardner, 1973]), and he found them unable to account for a number of major findings that his feature-detection model can predict. While the three previous models differ in their ability to explain various effects, there is one, which we will call the *proximity effect*, that none but Estes's model seems able to handle.

The proximity effect is found when detectability of the target declines as confusable noise items (nontargets) are moved closer to it in the display. The proximity effect is not a matter of simple lateral masking. First, the effect refers specifically to reduction in visibility of the target caused by noise items that are confusable with the target: Its measure should be the difference between the effects of nonconfusable and confusable noise items on the visibility of the target. The nonconfusable noise item should create only lateral masking, while the confusable noise items will, in addition, inhibit feature-detecting fields shared with the target. Second, the proximity effect, where it has been found, seems to exist "at separations great enought to preclude retinal contour interactions (Estes, 1972, p. 283)." The proximity effect is a natural prediction of Estes's model; it results from interactions between spatially adjacent feature detectors. None of the other models have any provision for local interactions other than lateral masking among items, and they cannot explain the proximity effect without bringing in additional assumptions.

Important as the proximity effect is, the evidence for it comes from only a small number of studies, and in some cases the evidence is indirect. The present study is designed to demonstrate the proximity effect directly. Accuracy of forced-choice detection of the target (an F or a T) was examined as a function of the distance of confusable noise items from it in the display. A "hybrid F-T," which has been shown to be confusable with Fs and Ts (Shiffrin & Gardner, 1972), was used as the confusable noise, and when the F-T was not placed near the target a rectangular dot matrix was placed near it. The dot pattern was used to control for whatever lateral masking the confusable noise item might cause. To keep targets always at the same distance from S's fovea, circular displays were used, with the fixation point in the center (Egeth, Jonides, & Wall, 1972).

Figure 1 shows displays of the type used in the experimental condition that had two confusable noise items (not shown are the conditions with four confusable noise items or a control condition that had only dot patterns as noise items). As is seen in Fig. 1, the two F-T noise items were either next to the target or were one or two steps away from it on the circular display. These are Conditions 20, 21, and 22, seen Fig. 1A, B, and C, respectively. The conditions with four F-Ts placed them either next to the target (Condition 40) or one step away on either side

*This research was supported by a Pomona College Faculty Research Grant to W. P. Banks. The authors thank W. R. Garner and E. E. Smith for their valuable comments on this experiment.

†Requests for reprints should be sent to William P. Banks, Department of Psychology, Pomona College, Claremont, California 91711.
††Now at York University, Downsview, Ontario, Canada.
†††Now at the University of California Medical School, La Jolla, California.

(Condition 41). The conditions shown in Fig. 1, with two F–T's at various distances from the target, seek to establish the proximity effect. This effect will be observed if the percentage of correct identifications of the target increases as the F-Ts are moved further away from it. The conditions with four F-Ts were included to compare the effects of proximity and the number of noise items.

METHOD

Ss were tachistoscopically presented stimulus arrays that always had as a target either one F or one T, and they had to decide on every trial whether an F or a T was presented. The F or T was located at 1 of 7 positions in a circular stimulus array like that seen in Fig. 1. The other 6 positions had either 2 or 4 confusable noise items (the hybrid F-T) and enough nonconfusable noise items (a 1.25 x 1.25 cm dot matrix) to fill the circle.

The two independent variables of the experiment were the number of confusable noise items (0, 2, or 4) and the distance of the confusable items from the target. Distance was defined as the number of positions in the display between the target and the nearest confusable noise item. Since the confusable noise items were always symmetrically disposed on either side of the target, the steps of separation refer to the number of steps between target and confusable noise on either side. Figure 1 shows how spacing was varied for Conditions 20, 21, and 22. With four confusable noise items only two separations could be used, and confusable noise items were either adjacent to the target (Condition 40) or were one step away from it (Condition 41).

In addition to Conditions 20, 21, 22, 40, and 41, there was a condition with no confusable noise items at all. Since the target could be either an F or a T and could be in any of the 7 positions on the circle, there were a total of 84 different stimuli (6 x 2 x 7).

The circular display, as well as fixation and masking fields, were seen at a distance of 86 cm in a 3-field Iconix tachistoscope. The 7 positions of the display were equally spaced on a circle whose diameter subtended 2.54 deg of visual angle, and centers of adjacent positions on the circle were separated by 1 deg of visual angle. Targets and noise fields were centered on the 7 positions. The targets and confusable noise items were drawn in black ink with line width of approximately 8/10 mm, and they subtended .67 deg of visual angle horizontally and .83-deg vertically. The dot matrices subtended .83-deg horizontally and vertically and contained about 20 1-mm diam dots in an equally spaced matrix. While the center-to-center visual separation between adjacent elements was always1 deg, the shortest distance between the peripheries of adjacent elements varied with their positions on the circle. The distance between the edges of targets and the dot matrix ranged from 0 deg (physical contact) to .33 deg of visual angle. The distance between the closest parts of targets and confusable noise items when they were adjacent ranged from .33 deg of visual angle to .36 deg. All three fields were approximately square, subtending about 10 deg of visual angle, and had luminance of about 40 ml.

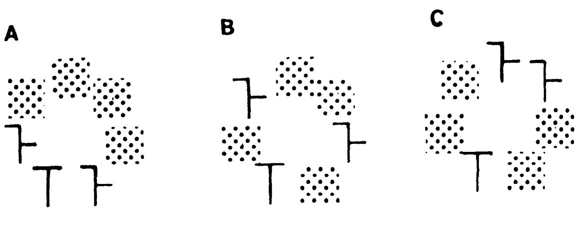

Fig. 1. Three of the 84 stimulus cards: A, B, and C are from Conditions 20, 21, and 22, respectively.

Procedure

On all trials S heard a verbal "ready" warning and then saw a fixation cross on a blank field for 2 sec before stimulus onset. The circular display then appeared for a short duration with its center where the fixation cross had been. Immediately after the stimulus disappeared it was replaced by a circular masking field with 7 positions just like the stimulus but with all 7 positions filled by the dot-matrix noise pattern. The masking field stayed in view until the next "ready" signal.

The S responded verbally "T" or "F" after each stimulus, speaking into a microphone placed about 2 in. from his mouth. The microphone was connected to a voice-operated switch, and the vocal response stopped a clock that had begun accumulating time at the offset of the stimulus display. The RT was recorded to the nearest millisecond. The Ss were told that their RTs were being recorded but that accuracy was more important than speed.

All Ss were first given a practice run with the 14 stimuli in the condition with no confusable noise items to familiarize them with the task. During this practice the stimulus duration was progressively reduced from 200 msec to about 50 msec. Then the stimulus duration was set separately for each S at a point at which they were approximately 60% correct (range was 57% to 64%) on stimuli in Condition 40. Only the stimuli in Condition 40 were used in the duration-setting trials, and most Ss saw each of the 14 stimuli in Condition 40 four times before data collection trials began.

Data trials were conducted in 3 separate blocks of 84, each block containing all the stimuli in a different random order. The first block was administered at the stimulus duration arrived at during practice. Before the second and third blocks more duration-setting trials with the 40 stimuli were given to readjust stimulus duration to the 60% criterion on these stimuli. Stimulus durations ranged from 20 to 60 msec.

The Ss were five students in the Claremont Colleges, paid $2/h for their participation. The experimental session, including rest breaks between blocks, lasted from 90 to 120 min.

RESULTS

Table 1 shows how detection accuracy improved as the F-Ts were moved away from the target in both Conditions 20, 21, 22 and Conditions 40 and 41. All five Ss showed such a proximity effect in both sets of conditions. Separate analyses of variance of the two sets of conditions (using the arcsine transformation of error scores) gave $F(2,8) = 15.04$ ($p < .01$) for the proximity effect when there were 2 F-Ts in the field and $F(1,4) = 29.91$ ($p < .01$) for it when there were 4 F-Ts. The analyses included blocks and targets (F vs T) as variables, and neither of these variables had any significant main effects or interactions with any of the other variables.

As is seen in Table 1, the reaction times (RTs) for correct respones show the same trends as the accuracy scores. Correct RTs and error proportions covary quite closely over the other variables of the experiment that are not reported in Table 1. The RTs associated with errorful responses were virtually constant for Conditions 20, 21, and 22, being 740, 739, and 738 msec, respectively. The RTs for errors in Condition 40 averaged 784 msec, and in Condition 41, 875 msec. There were no errors on data-collection trials in the condition with no F-Ts, and this condition gave a mean RT of 483 msec.

One peculiar result seen in Table 1 is the superiority in accuracy of Condition 41 over 21. To test the reliability of this result an analysis of variance (of arcsine of errors) for Conditions 20, 21, and 40, 41 was performed. This analysis included number of F-Ts (2 or 4), distance from the target (0 or 1 space), blocks (1, 2, or 3), and target (F or T) as orthogonal independent variables. The interaction between Proximity and Number of F-Ts was reliable [$F(1,4) = 17.44$, $p < .025$], and this interaction is clearly the result of the fact that Conditions 20 and 40 are equal ($t < 1.0$) and Conditions 21 and 41 differ ($t(4) = 5.36$, $p < .01$). This interaction was consistent over blocks and had the same form for both F and T targets (both $Fs < 1.0$). The analysis also showed a reliable proximity effect with $F(1,4) = 16.24$, $p < .025$.

Table 1
Accuracy and Speed of Detection as a Function of the Number of Confusable Distractors and Their Distance from the Target

Number of Confusable Distractors		Spaces Between Target and Confusable Distractor: 0	1	2
2	Percent Correct	73.8	78.6	97.2
	RT (msec)	655	560	544
4	Percent Correct	73.0	92.4	
	RT (msec)	655	562	

DISCUSSION

The results demonstrate the existence of a proximity effect as defined in this paper. A given number of confusable noise items interferes more with detection of the target as they are moved closer to it in the display. The possibility of lateral masking accounting for the proximity effect was eliminated by the procedure of placing nonconfusable noise, the dot matrix, near the target in the conditions where the F-T was not near it.

The results also show that increasing the number of confusable noise elements does not impair detection performance as much as moving the confusable noise closer to the target does. Comparison between Conditions 20 and 40 shows that the additional two F-Ts in Condition 40 have no effect on accuracy. Conditions 21 and 41 seem to show that additional noise items can even improve performance in some circumstances, or at least that proximity relations in the display can be more potent determiners of performance than the sheer number of noise items.

These results pose a difficulty for models of detection performance that have no provisions for local interactions among items. Parallel models, whether they postulate limited or unlimited capacity, could, however, easily incorporate such a provision by assuming that feature extractors or channels have mutually inhibitory relations with their neighbors. These models might not need to be changed in any other respect, but they would no longer be strictly parallel models. Serial scanning models, on the other hand, seem less able to accommodate the proximity effect. Since serial models make a number of other predictions at variance with effects observed in the detection paradigm, there seems to be no reason to look for ways to modify them to take account of this effect.

The surprising finding that performance in Condition 41 was better than performance in Condition 21 needs some explanation. There are, unfortunately, several possible explanations of this finding, and all of them cannot be completely eliminated on the basis of the data at hand. A speed-accuracy tradeoff, at least, seems unlikely as an explanation of the results, since correct responses took only 2 msec longer in Condition 41 than in Condition 21. Also, mutual inhibition of inhibition by neighboring F-Ts in Condition 41 is implausible. If such a process reduced the confusing effect of the F-Ts in Condition 41 it should also work in Condition 40 to some extent, but the extra F-Ts in Condition 40 do not improve performance over that in Condition 20.

Another explanation of this finding is based on search strategies. It assumes that the S searches the side of the display opposite to the massed F-Ts for the target in the fading iconic image. Performance would thus be poor when the target is imbedded in the F-Ts and good when it is on the opposite side of the display circle from them. Condition 41 gives better performance than Condition 21 by this account because the group of four F-Ts would be a better cue than the group of two. There are, however, a number of difficulties with this account. First, the extensive practice trials with Condition 40 can only induce a strategy exactly opposite to that assumed; for correct detection in Condition 40, the S must search in the midst of the F-Ts. Second, such a strategy will aid performance on only 1/3 of the trials (Conditions 22 and 41) and it seems unlikely that Ss would deliberately settle on a strategy that aids performance so infrequently. Third, this strategy should depress performance in Condition 40 just as it boosts it in Condition 41 (relative to 20 and 21, respectively) because the larger cluster of F-Ts in Condition 40 than 20 should cause the S to search the wrong area of the display more often in 40 than 20. As seen in Table 1, this prediction does not hold up, and there is no evidence of a floor effect that could obscure the predicted difference. Finally, Howard Egeth (personal communication) has obtained similar proximity effects in several experiments in which there was no possibility for the S to base search strategies on the configuration of elements in the display. Thus, scanning strategies are probably not responsible for the effect.

A final explanation of the paradoxical difference between Conditions 21 and 41 is able to explain the other results of the experiment as well. According to this explanation, processing of elements in the display is determined by the way in which they are grouped together by the perceptual system. Grouping, in turn, is determined by proximity and similarity of elements, good form and whatever other Gestalt principles operate to create perceptual configurations in the display. When the target is grouped together with the noise elements its visibility is not as great as when it can be processed as a single item. Thus, in Conditions 40 and 20 identification of the target is poor because it is grouped with and processed with the adjacent F-Ts. In Condition 41 the F-Ts form their own group and are processed separately from the target. Performance is much better in 41 than in 21 because the extra two F-Ts in 41 cause the confusable noise elements to cluster together as a "good" form separate from the target, while in 21 the noise items are more likely to form a group with the target. In Condition 22, where the F-Ts are furthest removed from the target, the target is almost never grouped with the F-Ts and is consequently identified very accurately. Finally, in the conditions with no F-Ts at all, the dot patterns, bearing almost no similarity to the target, group entirely with each other.

The possible importance of configuration in the visual detection paradigm has recently been pointed out by Kinchla (1974), and a number of researchers have found potent effects of stimulus configuration in visual information processing tasks where the elements, not the configuration they form, are the units of analysis required (see e.g., Clement & Weiman, 1970; Pomerantz & Garner, 1973). It remains to be seen just how many results of visual detection experiments can be explained in terms of perceptual grouping of elements, but a number of candidates can be suggested. The facilitation of detection by use of redundant noise elements (McIntyre, Fox, & Neale, 1970) could result from the redundant noise elements being grouped separately from the target more often than the varied ones because of similarity grouping. The effect of variations of target-noise similarity (e.g., Gardner, 1973; McIntyre, et al, 1970) could also be explained by similarity grouping. Possibly, too, the reduction in target visibility with increases in the number of noise items, when the display has a constant areal extent, may be partially due to increased grouping of targets with noise because of proximity. In any event, proximity and similarity between target and noise may jointly reduce visibility of the target only because they increase the likelihood of the target being grouped with the noise. Precisely why such grouping has this effect and how these variables operate to create perceptual clusters are questions that deserve further study.

REFERENCES

Clement, D. E., & Weiman, C. F. R. Instructions, strategies and pattern uncertainty in a visual discrimination task. Perception & Psychophysics, 1970, 7, 333-336.

Egeth, H., Jonides, J., & Wall, S. Parallel processing of multielement displays. Cognitive Psychology, 1972, 3, 674-698.

Estes, W. K. Interactions of signal and background variables in visual processing. Perception & Psychophysics, 1972, 12, 278-286.

Estes, W. K. & Taylor, H. A. A detection method and probabilistic models for assessing information processing from brief visual displays. Proceedings of the National Academy of Science, 1964, 52(2), 446-454.

Gardner, G. T. Evidence for independent parallel channels in tachistoscopic perception. Cognitive Psychology, 1973, 4, 130-155.

Kinchla, R. A. Detecting target elements in multielement arrays: A confusability model. Perception & Psychophysics, 1974, 15, 149-158.

McIntyre, C., Fox, R., & Neale, J. Effects of noise similarity and redundancy on the information processed from brief visual arrays. Perceiption & Psychophysics, 1970, 7, 328-332.

Pomerantz, J. R., & Garner, W. R. Stimulus configuration in selective attention tasks. Perception & Psychophysics, 1973, 14, 565-569.

Rumelhart, D. E. A multicomponent theory of the perception of briefly exposed visual displays. Journal of Mathematical Psychology, 1970, 7, 191-218.

Shiffrin, R. M., & Gardner, G. T. Visual processing capacity and attentional control. Journal of Experimental Psychology, 1972, 93, 72-82.

(Received for publication July 31, 1974.)