

Masking of alphabetic character recognition by dynamic visual noise (DVN)¹

WILLIAM R. UTTAL^{2,3}
THE UNIVERSITY OF MICHIGAN

The recognition of alphabetic characters composed of dot patterns is interfered with by dynamic visual noise (DVN) composed of similar dots randomly plotted over the character field. In this study we examine the general ability of the visual system to detect test characters as the level of this DVN varies. Three experiments are performed. The first deals with the effect of the DVN when the characters are embedded in continuous bursts of the noise, while the second and third examine the effects of leading and trailing bursts of the DVN on character recognizability. Though the applicability of this paradigm to psychological questions is very broad, we concentrate on the problems of the channel capacity of the visual system and the persistence of the effects. The time course of the psychological instant is also considered.

A remarkably general experimental paradigm has made its appearance in recent years in many different experiments, each of which is presumably aimed at a different theoretical question. The paradigm typically involves the sequential presentation of two visual stimuli. Depending upon the interval between the two, a wide variety of different phenomena can be demonstrated.

This experimental paradigm has been used, for example, to study visual masking in a forward or backward direction. Such masking phenomena are characterized by interference with the recognition or detection of geometric forms by a preceding or following visual stimulus. The results of such experiments have typically shown that the masking effect can persist for prolonged periods. In general, when two stimuli activate the same receptors the masking occurs maximally at the shortest interstimulus intervals. Crawford's (1947) original studies of the backward masking phenomenon using flashes of light have been replicated many times. Generally the interference effects last for well over 100 msec. Schiller (1966) has shown further that the masking effect of letters on letters lasts longer than 200 msec, and that the effect is increased with an increase in the intensity of the masking stimulus. Schiller (1965), comparing monoptic and dichoptic masking, used a threshold test that apparently resulted in a less persistent masking effect lasting only for about 40 msec. But it should be noted that his masking stimulus never drove the probability of detection of the masked form to chance levels. Raab (1963) reviewed a large selection of older studies on retroactive masking—a phenomenon which has captured much of the attention in this area.

When the two visual stimuli do not fall on the same retinal areas there is also some interference. However, the result is quite different. Typically for the retroactive masking condition (metacounter) a U-shaped function obtains. Alpern (1953) found maximum dimming of a test stimulus when the masking stimulus followed about 120 msec after the test stimulus. Weisstein and Haber (1965) found a similar effect but with the maximum retroactive effect at 30 msec. Kolers and Rosner (1960) actually found dichoptic metacounter effects when the interstimulus interval was about 80-100 msec, and a decreasing effect as the interval increased. Kahneman (1967) has recently pointed out the intimate relation between the experimental results concerning metacounter and apparent motion and has reported peak metacounter effects at 80 msec.

It is not at all certain that the many different backward masking phenomena observed by the different investigators and the two different types of response we distinguish here are due solely to the conditions of overlap of the two stimuli. Kolers (1962) has considered the problem in detail and has also recognized the two different categories of response, which he named Type A and Type B, respectively. He lists a large number of other conditions such as the contrast and duration that affect the type of response one will obtain in any given experiment. Most interestingly, he also assumes the type of response can be affected by the response demanded of the S.

This same experimental paradigm has also been used to study the phenomena of short-term visual storage and the scanning processes used to inspect the stored image. Sperling (1963), following some of the older work of Baxt (1871), found that a static noise field completely obliterated the perception of an alphabetic character when the two were presented simultaneously. Furthermore, the noise terminated the scanning of a stored visual image when several characters were presented in the stimulus and the static noise appeared after the display. Interestingly enough he found that all Ss required at least 20 msec to see anything after leading static noise, when the intensities of the two were comparable. More intense leading noise fields could mask for as long as 100 msec.

Another application of this same paradigm is in the context of temporal acuity or perceptual simultaneity, or as it has sometimes been called—the psychological moment. The general notion here is that since the visual image is stored someplace in the visual system for a substantial period, two stimuli that impinge upon the eye within a certain minimum period should appear to be present simultaneously. Thus Piéron (1952) reports that the temporal resolution of the visual system is typically about 100 msec although he does mention that under certain stimulus conditions it can be reduced to 20 msec. In any event, for most practical situations such as the cinema, no break is seen in the action when individual frames are projected at about 40-msec intervals. Fraisse (1966) reported that characters appear to be simultaneous whenever the total duration of the two characters and the interval between them is less than 80-120 msec. Lichtenstein (1961) has obtained data showing an equivalent value of about 125 msec, and further showed that this result was independent of the interval pattern between the individual elements of a diamond-shaped stimulus array. (This latter paper is also of considerable interest as a guide to some of the older literature on this subject.)

The idea that is to be emphasized by pointing out the procedural similarities of all of these experiments is that their differing results may represent different aspects of a more general perceptual process. Hopefully, some future psychologist will be able to rationalize all these data in terms of some generalized ability of the visual system to deal with sequential stimuli.

It is, however, also probably true that there are several separate mechanisms that are mixed together in all these experiments. First, there is a simple signal-to-noise ratio consideration in which simultaneous stimulus and masking patterns interfere with each other to the degree that the stimuli are hidden or confused with visual noise of similar quality. This effect seems to be introduced

by either monoptic or fused dichoptic viewing. Schiller's (1965) investigations showed that masking with patterns was dichoptic, while masking with flashes of light was purely monoptic. These results suggest that this type of pattern or "signal-to-noise" masking is a process occurring at nervous levels higher than the retina. Second, there are also interacting processes due to the inertia of the transduction processes involved in vision. This latter effect would presumably be mediated by mixing due to simple photochemical inertia in the peripheral receptor layers of the retina.

How can we separate the pattern or "signal-to-noise" type of masking from the retinal receptor effects? One way is to use Schiller's (1965) or Kolars's and Rosner's (1960) dichoptic observation technique. Another is to make use, as we have in the present experiment, of a special type of masking stimulus that is generated by the serial presentation of small dots at random locations. We have called this type of stimulus Dynamic Visual Noise (DVN). Dynamic visual noise was introduced in a slightly different form as a psychological research medium by MacKay (1967). He used random patterns in which a large number of points were presented simultaneously on sequential frames of a motion picture film strip. Each frame contained many dots but the pattern of dots was different from one frame to the next. Each frame was projected at ordinary cinema speeds. In the present experiments dots are plotted individually on the face of the oscilloscope with the interval following each dot individually controllable.

DVN of this sort has some important qualities. Since each retinal locus will be stimulated only infrequently on the average, it is unlikely, for practical noise levels, that retinal locations stimulated by signal dots will be simultaneously stimulated by noise dots, even with monoptic stimulation. Thus the retinal effects can be separated from the perceptual confusions of pattern or "signal-to-noise" effects. Masking with gross patterns or a bright flash of light delivered to the total retina, however, confounds masking due to signal and noise confusions and masking due to retinal lability and other complex interactive effects such as lateral summation or inhibition.

By using test patterns (alphabetic characters) composed of dots and dynamic visual noise (DVN) composed of equivalent dots, but ones located randomly in time and space, photoreceptive effects can thus be minimized, even in the monoptic or binocular cases. Because of the persistence of the visual image the characters, if plotted quickly enough, appear to be simultaneous displays of all constituent points. Similarly, the DVN appears to be composed of many dots at any given time, except for the very lowest interdot intervals. We thus are in a position to explore the nature of "signal-to-noise" or pattern masking per se, as well as study the time course of the decay of this pattern masking in a way that is independent of peripheral photochemical effects.

The particular experimental paradigm with which we have recently been concerned is aimed at measuring the effects of DVN on the recognition of these characters. Obviously, if the dots of the DVN are frequent enough, the character will be obscured simply because the individual dots of the character cannot be separated from the individual dots of the DVN. In the first experiment we establish the basic functional relationship between the DVN level and character recognizability when the DVN completely overlaps the character in time. In the second and third experiments reported here we examine the effects of DVN on character recognition when the DVN precedes or follows the presentation of the character. Specifically, the second experiment examines the effect on character recognition of leading noise separated from the character by a variable interval, while the third experiment examines a similar effect of trailing noise.

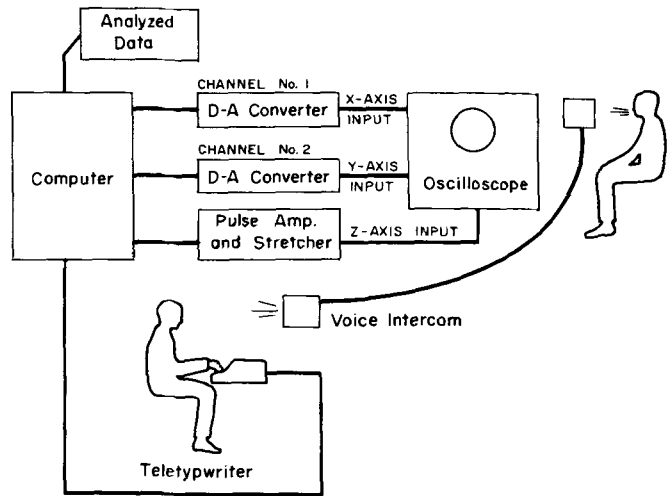


Fig. 1. A block diagram of the experimental apparatus used in this study.

METHOD

Subjects

Twenty undergraduate male and female Ss participated throughout the series of three experiments, although no more than eight were involved in any one experiment. Each data point in our graphs represents the averaged scores of six to eight Ss. While no S was arbitrarily discarded, the exigencies of academic life did not permit us to have a full complement of eight Ss for all cases or the same group of Ss throughout the entire course of each experiment. All Ss were given pretraining to familiarize themselves with the character set used in the experiments.

Apparatus

The S was seated in a darkened and acoustically insulated cubicle. A cathode ray tube (CRT) was positioned 14 in. in front of his eyes, measured at the bridge of his nose. A metal frame bearing against the S's forehead regulated the appropriate head position.

The face of the CRT was masked off, exposing a window $2\frac{1}{4} \times 1\frac{1}{2}$ in. The DVN described below filled this entire window when it was being presented. The character presented was slightly smaller; each character being a maximum of 1 in. high and $\frac{3}{4}$ in. wide. Thus the visual angle subtended by the entire display was about 9.2 deg vertically and 6.13 deg horizontally, while the characters subtended a visual angle of about 4.08 deg by 3.03 deg.

The CRT was driven directly by a digital-to-analog converter connected to the output of a medium-sized computer. The digital-to-analog converter used in our system is a 10-bit, two-channel system. One channel was connected to the x-axis input of the CRT, and the other to the y-axis input. A third channel was connected to the intensity control and determined when a point would actually be plotted. This third channel intensified the electron beam for 10 microseconds upon command. The system is thus acting as a point plotter. To form any pattern with such a point plotter, each and every point of that pattern must be individually calculated and plotted. Individual points could be plotted in about 20 microseconds, but the minimum interpoint interval actually depended upon the algorithms for coordinate specification for each character. Figure 1 shows the equipment arrangement used to accomplish these functions.

The precise manipulation of the duration of the intensity control would serve no useful function if the light emitted by the

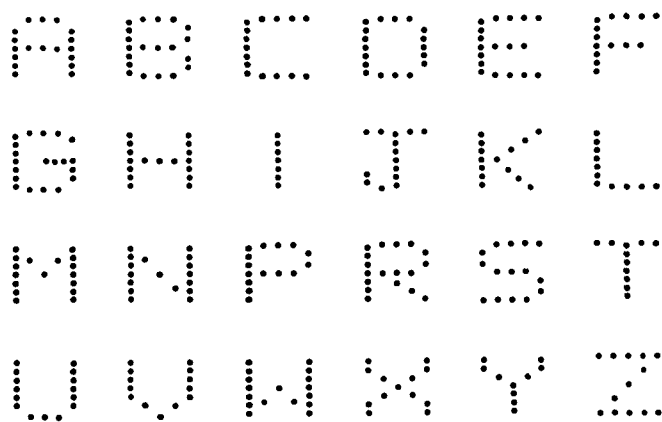


Fig. 2. The character set used in this study as test stimuli.

CRT phosphor persisted for any appreciable length of time. For our experiments a special ultra-short persistence phosphor (P-15) was used. Its light output decays to 0.1% of its original intensity in 50 microseconds. Thus the duration of each dot was of this order of magnitude. An entire alphabetic character was plotted in less than 500 microseconds. Therefore, all the dots of each character appeared to be plotted simultaneously. Figure 2 shows the family of alphabetic characters that were plotted by this method and used in our experiment. O and Q were not used, leaving a 24-character stimulus set. The DVN was controlled by the specification of a single constant in the computer program. This constant regulated the duration of the interval between sequential noise dots. The position of each dot was randomly selected by a two-dimensional random number generator. Figure 3 shows the general appearance of the DVN at three of the different interdot intervals.

As we have said, the visual system integrates, to a certain extent, dots presented within a given period so that the screen does appear to have many more than one dot present at all times except at the very lowest dot rates. Of course, the pattern appears to be constantly dancing from one configuration to another. However, as we shall see, the results of our experiments indicate that this integration is not complete and some separation of character dots and DVN dots is observed.

In addition to making the DVN field larger than the character field, we also attempted to control for pseudorecognition of characters in fixed positions by partial cues by jittering the position of the characters. Each character was presented in a location within the visual noise that varied in the vertical direction from trial to trial.

General Procedure

The entire experimental procedure was under the control of the computer. During each session the order of character presentation was randomized. However, the noise level remained constant during each 1-h session. Two Ss ran together in each 1-h session. Each served as the observer S for ½ h, and in the alternate ½ h transcribed the responses of his partner (who reported his observations verbally over an intercommunication system) into a special keyboard that entered the data into the computer. Both the stimulus character and the observer S's response were then displayed on a monitor television. The transcriber S gave feedback to the observer S by informing him of the correct character whenever an error was made. Ss were given a 2-min break at the end of 10 and the end of 20 min in the experimental cubicle.

Data were collected in raw form on a digital magnetic tape. At the end of each session, two analysis programs, DAP 1 and DAP 2

were used to summarize all results. DAP 1 reduced the data for each S to a summary table in which the number of presentations of each character, the number correctly identified, the percentage correctly identified, and the specific errors were tabulated. DAP 2 reduced these summary tables for each S to a group summary, identical in format to the individual ones, except that specific errors were not recorded because of space limits. Six to eight Ss were run each day and there were approximately 2500 responses for each condition indicated by a point on our graph. This number varied from 2000 to 3000 primarily depending upon the difficulty of the task and the amount of feedback given by the transcriber S.

In all instances the CRT displayed some sort of a ready signal. In the experiments with the leading noise, the noise itself was initiated to indicate to the observer S that the transcriber S had entered his last response and the system was ready for his next trial. The observer S held a pushbutton with which he controlled the emission of the stimulus, thus self-pacing the experimental sequence. In those situations without leading noise, a dim fixation point was plotted on the face of the CRT as a ready signal. It was extinguished when the button was pushed and, after a short pause, the character was presented. The trailing noise followed the character at the delay specified for each given condition.

In the first experiment the DVN completely overlapped the character in time and space. In other words, both leading and trailing noise were present. The duration of the DVN preceding the character was variable, depending upon the time the observer S took to release the character with the pushbutton. However, this period could be no less than 1 sec—a minimum duration that was programmed to prevent the S from shortening the duration of the leading noise to a brevity at which it might no longer be effective. The noise following the character lasted for a constant 1 sec. Since the display of the character could be accomplished by the computer in less than ½ msec, the noise at all interdot intervals appeared continuous, and the character appeared embedded in this continuous stream of dots. The noise levels used in this experiment corresponded to interdot intervals of 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, and 1 msec. Sampled data at longer interdot intervals showed no effect on character recognition scores. The maximum recognition score (group average) obtained under optimum conditions was about 95%. The 5% error margin was attributed to eye blinks, or transcription errors, and does not represent errors due to an inability to read the character through the DVN.

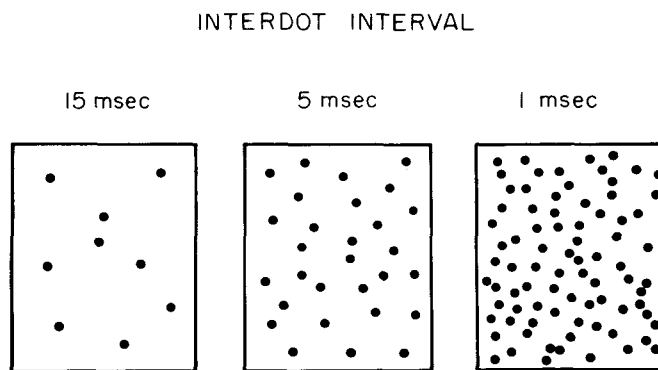


Fig. 3. A visual impression of the dynamic visual noise (DVN) at the three interdot intervals indicated. These are not copies from a photograph, but an artist's impressions of the apparent dot densities. The characters, composed of dots of the same size and brightness, appear physically placed within the limits of this DVN.

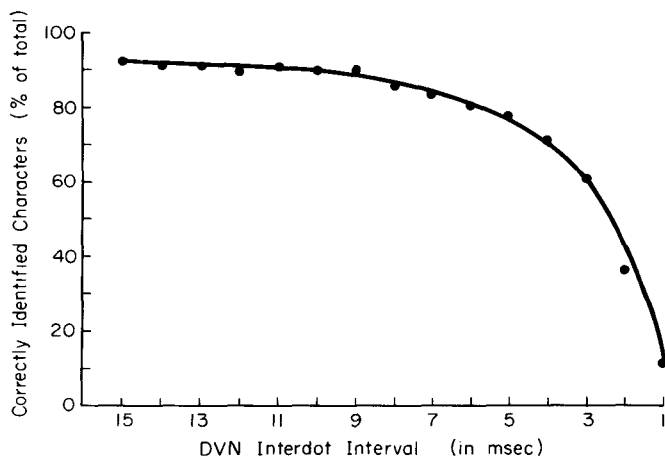


Fig. 4. The results of Experiment 1 showing the decrease in character recognizability as the interdot interval of the DVN decreases.

The observer S was signaled that the previous response had been acquired and that a new stimulus was available by the reappearance of the DVN on the CRT. All Ss were instructed to respond with phonetic alphabet names (for example, "A as in ALPHA") to help reduce transcription errors.

In the second experiment, the noise following the character was deleted and the DVN preceded the character by a variable delay (the independent variable) that changed from day to day. The values of the DVN used in this experiment are indicated in Fig. 4. While we originally started out collecting data at rather great densities, the stability of the data (a typical result in a computer-controlled experiment) led us to reduce progressively the number of data points collected in this and the third experiment. Data were obtained for two values of DVN interdot interval, 3.0 msec and 1.0 msec.

In the third experiment the leading noise was deleted and DVN was presented only following the character. A variable delay, changed from day to day was interposed between the character and the DVN. The values of the delay used for each of the two conditions of DVN used (3.0 msec and 1.0 msec) are indicated in Fig. 5.

RESULTS

In all three experiments the percentage of the total number of characters presented that were correctly identified was the main dependent variable. Figure 4 presents the results of Experiment 1. The data are plotted with the interdot interval of the DVN on the abscissa and the percentage of correctly identified characters on the ordinate. There is a gradual but extremely regular decrease of the ability of our Ss to identify the characters as the interdot interval decreases. It should be appreciated that it is not possible to use the DVN metric of interdot interval directly to specify the noise level, for there is another important variable confounding the actual noise level. The eye, because of inertia of the photochemical effects does integrate over time and, as shown in Fig. 3, there are apparently many more than a single dot present on the surface of the CRT at any given perceptual instant. In fact, the eye, as a first approximation, seems to be sliding an integrating temporal window along the time axis. The width of this temporal window is associated with the persistence of visual images among other factors. As we shall see, however, there is an enormous discrepancy between this simplistic notion of an integrating temporal window and the separability of the characters from the DVN in Experiments 2 and 3.

In Experiment 1 a significant decrease in the recognizability of the characters begins at a DVN interdot interval of about

10 msec. The downward trend of the curve, as can be seen, gradually accelerates, decreasing to a level only slightly above chance (4.1% is the chance performance level for our 24-character alphabet) when the DVN interdot interval is 1 msec.

However, there are irregularities implicit in this smooth curve that are not immediately obvious. For example, the characters do not lose their ease of recognition equally. This is evidenced by the data summarized in Table 1 in which the data for all Ss for DVN with an interdot interval of 3 msec are tabulated. This table is the equivalent of the single data point for 3 msec in Fig. 4. The recognizability of the characters at this DVN interdot interval varies from 87.6% for the character I, down to 32.3% for the character R. Two features of character recognizability emerge from an examination of this table and the related confusion matrix data. First, straight lines are seen with a great deal of ease compared to curved lines in the DVN. Second, some character groups like P, R, K, and F, and A, H, W, M, N, and U have large confusion scores. Why they should do so is obvious if one examines the font shown in Fig. 2. The distinction between each of the members of these two sets of characters depends primarily upon a small number of critical distinguishing dots which are extremely susceptible to DVN obscuration. Confusion data for the entire set of characters used in this experiment are shown in Table 2. In this table all the data for all the different values of interdot interval have been pooled. Errors for approximately 37,000 trials are tabulated with the exception that no error scores less than 10 have been tabulated to avoid cluttering up the table. All error scores greater than 50 have been underlined for emphasis.

The results of Experiments 2 and 3 have been plotted together in Fig. 5 to emphasize the symmetry of the obtained data. The data for the two DVN conditions for leading and trailing DVN can be seen to produce nearly equivalent results. There is but the most moderate decrease in the ability of the S to recognize the character in the 3-msec DVN as the delay between the noise and the character decreases for both the leading and trailing noise. The lowest recognition scores for the 3-msec noise (86.5% and 82.5%) occur when the noise immediately follows or precedes the character with no interposed delay. This, however, is a considerably higher recognizability score compared to the situation in which the characters are embedded in equivalent DVN levels as in Experiment 1. The comparable value for the embedded condition is 61.6%. Similarly, the decrease in recognizability when the leading or trailing noise is at the 1-msec level is considerably less than the masking produced by

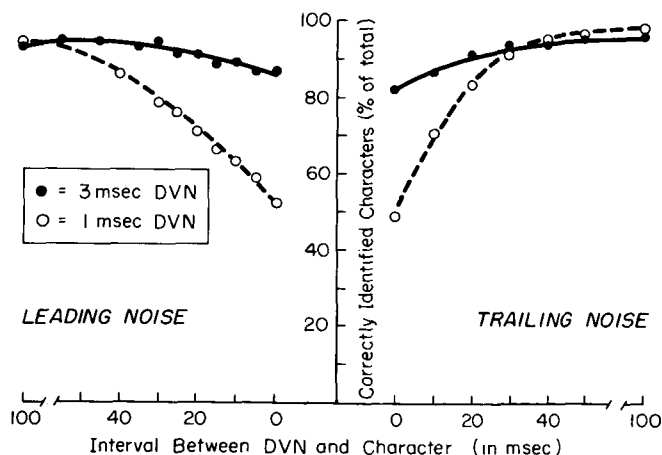


Fig. 5. The results of Experiments 2 and 3 showing the effects of bringing the character closer to either leading or trailing DVN for two values of interdot interval.

Table 1

The data for the 3 msec interdot interval condition of Experiment 1, showing the range of recognizability scores for the character set at this noise level. This table is one of those summaries of data from only six Ss and thus the total N is somewhat less than the 2500 typically obtained.

Symbol	N	NC	%C
A	74	42	56.7
B	66	36	54.5
C	74	44	59.4
D	72	46	63.8
E	73	55	75.3
F	65	49	75.3
G	80	53	66.2
H	66	37	56.0
I	81	71	87.6
J	73	46	63.0
K	67	53	79.1
L	61	46	75.4
M	75	29	38.6
N	72	30	41.6
P	86	42	48.8
R	71	23	32.3
S	67	40	59.7
T	64	40	62.5
U	68	53	77.9
V	59	32	54.2
W	70	28	40.0
X	75	52	69.3
Y	69	48	69.5
Z	94	67	71.2

($NT = 1722$, $NCT = 1062$, $\%C = 61.6$)

continuous noise with an interdot interval of 1 msec. The values in this case are 52.6% for leading noise immediately preceding the character, 49.1% for trailing noise immediately following the character, and 12.5% for continuous noise.

The important result here is that, even though an initial consideration of the appearance of the DVN suggests the existence of a temporal integrating window, the differences reported above suggest that the S is not limited by this integration when he has the task of separating the character signals from the DVN. He is, in this case, able to separate the characters and the DVN from each other even when the delay between the two is virtually nonexistent. The older notions of the psychological moment or of short-term visual storage that implied that any two stimuli which occurred within a brief interval were dealt with as if they were simultaneously present are strongly challenged by these results.

Figure 5 also shows that even the moderate decrease in recognizability only occurs with delays of about 30 or 40 msec even under the most effective masking condition (DVN interdot interval of 1 msec). These results indicate a temporal discrimination capability of the eye for these viewing conditions that is considerably more acute than might have been predicted from the reports cited in the introduction to this paper. As we shall discuss below, we believe this to be mainly due to the fact that we have separate retinal photoreceptive lability effects from pattern discrimination effects. The effects examined here fit mainly in this latter class.

DISCUSSION

The general experimental paradigm we have used in this experiment, characterized by the sequential presentation of a signal and DVN, has wide applicability. As we mentioned in the introduction, the findings of this experiment are relevant to current interests in short-term visual storage, backward and forward masking, notions of the psychological moment, character recognition, and signal detection. The paradigm also provides insight into the neurophysiological foundations of the visual process by separating peripheral photoreceptor effects from the

pattern discrimination effects presumably mediated by the central nervous system.

Nevertheless, in the present experiment we have focused on two main issues. First, we have concerned ourselves with the recognizability of the dot characters in varying densities of DVN. Second, we have measured the persistence of the masking effect when the character and the DVN are not simultaneous, thus considering the nature of the temporal window that is defined by these transient persistencies. Our results in the latter case show that our situation fits in the Type A response category (monotonically decreasing masking effects with increasing interstimulus interval) suggested by Kolers (1962). This is in spite of the fact that the contrast, luminance, and size of the individual dots are identical and the overall patterns are similar—conditions that Kolers suggests should lead to a Type B response (a U-shaped response curve with maximum masking at some intermediate interstimulus interval). However, our task and stimulus conditions are so different from the ring and disc type of stimulus used in many of the other studies that the two situations may not be truly comparable.

Experiment 1, as we have said, was designed simply to measure recognition performance as a function of the DVN level. However, in the spatio-temporal universe of these visual experiments we cannot directly use the usual single dimensional measures of, say, an acoustic signal, to specify the information content of either the test character or the noise. Some approximation to the information content of the character is possible because we have essentially collapsed time to a single instant when we plotted the points so closely together. However, a similar notion of the information content of the DVN in terms of the number of noise dots that occur in a given interval is not immediately available. This is so both because of the random spatial distribution of the noise dots and because of the temporal integrating properties of the visual system.

However, we can specify precisely the functional relationship between percentage character recognizability and the interdot interval of the DVN on the basis of our experimental results. This function will serve the same purpose as a more conventional plot of signal-to-noise effects on recognizability, since the information content of the character set—the signal—whatever it is, is constant from one condition to the next.

Let us assume the integrating time window of the visual system to be of the order of 100 msec, for example, and therefore that the number of dots simultaneously perceived during any given 100-msec window is a direct function of the interdot interval (IDI). Therefore, we can now replot the data of Fig. 4 with the new abscissa N , where $N = 100/\text{IDI}$. This has been done, and the data of Experiment 1 are replotted in Fig. 6. The straight line so generated intercepts the vertical axis at about 103%. This mathematical fiction reflects the facts that some of our data are outside the region within which the addition of DVN dots has any effect, and that 100 msec is not exactly correct as an estimate of the integrating period. We may assume that this value should be 100%. Therefore, the line has a slope of $(-100)/76 = -1.31$ where 76 is the intercept on the abscissa. The equation of the straight line using the slope-intercept formula ($y = mx + b$; where m = the slope and b = the y intercept) is:

$$\% \text{Correct} = -1.31 (\text{number of DVN dots}) + 100 \quad (1)$$

Since the number of DVN dots is equal to $100 \text{ msec}/\text{IDI}$, the analytical expression for our original nonlinearized recognizability function is

$$\% \text{Correct} = -1.31 \left(\frac{100}{\text{IDI}} \right) + 100 \quad (2)$$

Table 2

Confusion matrices for all data of Experiment 1. Errors for about 37,500 trials are tabulated but only error scores greater than 10 have been entered to avoid cluttering up the table. Each error score greater than 50 has been underlined for emphasis.

		RESPONSE CHARACTER																									
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	P	R	S	T	U	V	W	X	Y	Z		
STIMULUS CHARACTER	A	x	10						19					42	47		13					11	16				
	B	15	x		23		20	21	14	12		11	34			29	22				21		14				
	C			x	40	11	17	71		18		13	<u>63</u>							11			35				
	D		27	13	x			12		11		12	13	14	13					<u>91</u>		14	17				
	E		40		12	x	<u>54</u>	12				18	30		11								17				
	F					12	x					44				42	23						32				
	G	21	20	25	15	14		x	21				15	10						15	10		23				
	H	20							x					50	<u>80</u>					11		48					
	I									x	15	11								24			17	13			
	J									<u>60</u>	x	15								<u>120</u>			25				
	K						10			<u>75</u>		x	29		27		15						18				
	L			12						27		34	x								10		23				
	M								41			12		x	<u>219</u>					37	12	46	23				
	N								<u>80</u>			12		<u>281</u>	x					29	16	<u>146</u>	10				
	P	12	12	11		10	<u>119</u>		13	13	30			10		x	<u>81</u>						23				
	R		21				<u>67</u>			11		<u>53</u>			15	<u>320</u>	x					20					
	S		11		13	16		17										x					<u>66</u>	12	37		
	T									<u>131</u>	<u>81</u>													18 *			
	U												13	20	30					x		11					
	V		35		13								11	45	42					11	x	12	<u>62</u>	11			
	W												46	<u>168</u>							<u>72</u>	x	42	17			
	X								34					<u>14</u>			12	15			<u>36</u>	13	x	15			
	Y			13						13		36												x	28	17	
	Z		35			11				<u>67</u>	16									15	15		16	<u>73</u>			
											11						12	21	21		11		<u>79</u>		x		

or

$$\% \text{Correct} = 100 - \frac{131}{\text{IDI in msec}} \quad (3)$$

The only data point that deviates from the assumption of linearity is that one on Fig. 4 that corresponds to an IDI of 1 msec. We believe this deviance is almost entirely due to the continued recognizability of the characters I, K, L, and (surprisingly) X. The data for X are thought to be artifactual, for several Ss reported that they gave an X-response when they saw nothing, since X had been the first character to disappear completely as the DVN level increased. This disappearance is most probably due to the absence of any long straight line segments in the particular X-pattern we used in the experiments.

With some trepidation we can also point out that the intercept of the straight line on the abscissa is a rough estimate of the maximum information capacity of the visual system, for it is at this intercept [actually the intercept with the chance guessing value (4.04%) would be more appropriate] that theoretically no further information processing capacity is available for character recognition. Our hesitancy to do so is primarily based upon the fact that certain characters are indeed still recognizable. Nevertheless, 76 dots in a 100-msec integrating period, or converted to our original metric—a 1.31-msec interdot interval, can be considered a preliminary estimate of the information-processing capacity of the total visual system in the current context. Though this value is somewhat higher than earlier estimates, it is for a situation which is somewhat more comparable to the real visual processing tasks of everyday life.

Now that we have established a formal relation for the effect of DVN on character recognizability, we can consider the results of Experiments 2 and 3. A major outcome of these experiments is the difference that obtains between the results for a character embedded (as in Experiment 1) in 1- or 3-msec DVN, and the equivalent points of Experiments 2 and 3. In general, we see that placing the character at either the beginning or the end of the DVN allows a much higher level of character recognizability. The most salient difference is for those points for which there is no delay between the character and the DVN. (These are the four

points in Fig. 5 that have abscissae of zero.)

The fact that characters can be separated from noise if they are at the beginning or the end of the noise is an extraordinary result. It indicates first of all, that the simplistic notion of a temporal window in which all stimuli are integrated must be modified. In other words, it raises the question: How can this separability occur if our perceptual world is made up of a string of psychological moments strung together in a fixed way like a string of sausages? Our initial interpretation of this result is that the integrating moments or windows are not arranged end to end like those metaphorical sausages, but are more closely approximated by a sliding temporal window that moves along the real temporal dimension. Thus the leading characters can be seen and interpreted rather well because they occur in the last position

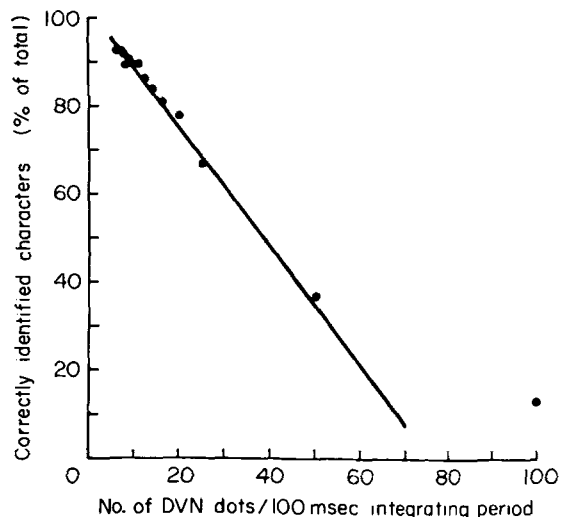


Fig. 6. A replot of the data shown in Fig. 4 assuming a 100-msec integration period. This graph shows the excellent fit by a straight line except for the one point at the 1-msec interdot interval. This single deviant point is almost totally due to the continued high recognizability of the characters I, K, L, and X.

of the sliding temporal window before the leading edge of that window enters the noise. The trailing characters can be seen because they also are present in a window that does not include the noise, i.e., the window whose trailing edge is just emerging from the DVN as the character is emitted. The notion of a sliding temporal window and the resultant separability of DVN and character holds, of course, even if the masking due to trailing noise and the masking due to leading noise are, as is suggested by our data, additive in either a linear or in some weighted fashion.

How then can we accommodate this notion to the tracking data of Stroud (1955), who has shown that stimulus information delivered within a 100-msec period cannot effectively produce a separate response? Similarly, how can we make this result compatible with the notion of perceptual simultaneity of stimuli occurring within 100 msec of each other? We would like to propose that the psychological moment of those experiments might indeed be a "receptor moment" or a "response moment," rather than a "central psychological moment." A receptor moment would result from those interactive phenomena that are due to the integrating properties of the retina. Such effects seem to have very long persistence and seem capable of overriding the otherwise more sensitive temporal discrimination abilities of the central nervous system. Thus, those experiments that involve sequential stimulation of the same retinal locations typically show an inertia greater than that produced by experiments which stimulate separate loci. (The one exception to this generality is the disc and annular ring typically used in the metacontrast experiments. But that situation has several special features that suggest other mechanisms may be involved.)

Therefore, the results obtained by an experiment such as the present one in which separate loci are stimulated should provide more valid estimates of the temporal discriminatory powers of the higher levels of the nervous system, unconfounded by retinal effects and the resulting "receptor moment." In fact, from a certain point of view we can consider stimulation at the same retinal loci to be misleading in the same sense that we would be misled if we ignored the decay time of a particularly persistent CRT phosphor.

Stroud's data might possibly also be explained in terms of the physical inertia of the response mechanism or some other motor programming limitation. While in both cases the receptor and effector obviously do limit the total performance of the organism, they are in no sense indicative of a psychological moment, if by a psychological moment we are referring to that indivisibly small quantum of perceptual and decision-making time.

The nature and the shape of the visual integrating period and the psychological moment shall be considered in detail in a subsequent paper. For the present, let us simply mention several possible alternatives. In addition to the sausage-like string of psychological moments, we have already suggested the possibility of a sliding temporal window. There are many different "shapes" the temporal window might take in either case. The window could be one that is heavily weighted to emphasize the information at its center. Thus, rather than being rectangular in its weighting and equally open throughout its course, it is possible that the window might symmetrically decrease its degree of "openness" the further a stimulus is from its center. The overall weighting distribution could be triangular, a normal distribution, or even some more complex shape. Some of our data in Fig. 5 suggests that it is indeed nonsymmetrical. This is suggested by the fact that the persistence of the effect of noise as the interval between the character and noise is varied is not identical for both forward and backward masking. Leading noise appears to have

some effect even when the interval is as great as 40 msec. Trailing noise, on the other hand, seems to have an effect only when the interval is shorter than 20 msec. This suggests that the leading edge of the temporal window protrudes considerably further ahead in time than its trailing edge drags behind.

The important general conclusion to which we have come is that the many points of dissimilarity between data of this sort and the previous studies mentioned are due to the fact that we are dealing with only a portion of those processes that can produce persistent masking effects. Our experiments, because of the statistical properties of DVN, examine those processes that are specifically characteristic of the central pattern recognition processes and are therefore unconfounded estimates of visual temporal resolution or the psychological moment. While at some later time others may be able to subdivide the pattern effects we have dealt with to an even finer degree, we believe that these measures give more meaningful estimates of man's ability to process sequential visual signals than the older studies. A strong cognitive component dependent upon the nature of the task is suggested. Thus a host of new questions is opened up by this technique, some of which we shall deal with in forthcoming reports.

REFERENCES

- ALPERN, M. Metacontrast. *Journal of the Optical Society of America*, 1953, 43, 648-657.
- BAXT, N. Über die Zeit welche nötig ist, damit ein Gesichtseindruck zum Bewusstsein kommt und über die Grösse (Extension) der bewussten Wahrnehmung bei einem Gesichtseindrucke von gegenbener Dauer. *Pflüger's Archiv für die gesamte Physiologie*, 1871, 4, 325-336.
- CRAWFORD, B. H. Visual adaptation in relation to brief conditioning stimuli. *Proceedings of the Royal Society*, 1947, 134B, 283-302.
- FRAISSE, P. Visual perceptive simultaneity and masking of letters successively presented. *Perception & Psychophysics*, 1966, 1, 285-287.
- KAHNEMAN, D. An onset-onset law for one case of apparent motion and metacontrast. *Perception & Psychophysics*, 1967, 2, 577-584.
- KOLERS, P. A. Intensity and contour effects in visual masking. *Vision Research*, 1962, 2, 277-294.
- KOLERS, P. A., & ROSNER, B. S. On visual masking (metacontrast) dichoptic observation. *American Journal of Psychology*, 1960, 73, 2-71.
- LICHTENSTEIN, M. Phenomenal simultaneity with irregular timing of components of the visual stimulus. *Perceptual & Motor Skills*, 1961, 12, 47-66.
- MACKAY, D. M. Ways of looking at perception. In W. Wathen-Dunn (Ed.), *Models for the perception of speech and visual form*. Cambridge: The M.I.T. Press, 1967. Pp. 25-43.
- PIÉRON, H. *The sensations: Their functions, processes and mechanisms*. New Haven: Yale University Press, 1952.
- RAAB, D. H. Backward masking. *Psychological Bulletin*, 1963, 60, 118-129.
- SCHILLER, P. H. Monoptic and dichoptic visual masking by patterns and flashes. *Journal of Experimental Psychology*, 1965, 69, 193-199.
- SCHILLER, P. H. Forward and backward masking as a function of relative overlap and intensity of test and masking stimuli. *Perception & Psychophysics*, 1966, 1, 161-164.
- SPERLING, G. A model for visual memory tasks. *Human Factors*, February 1963, 19-31.
- STROUD, J. The fine structure of psychological time. In H. Quastler (Ed.), *Information theory in psychology*. New York: Free Press, 1955.
- WEISSTEIN, N., & HABER, R. N. A U-shaped backward masking function in vision. *Psychonomic Science*, 1965, 2, 75-76.

NOTES

1. This project was supported in part by NSF Grant GJ114.
2. The author wishes to express his appreciation to Mr. Jon Baron and to the members of the Faculty Psychophysics Seminar at the University of Michigan whose conversations contributed to the development of many of the ideas expressed in this paper. The competent experimental assistance and editorial advice of Miss Anne Byrnes are also gratefully acknowledged.
3. Address: Department of Psychology, The University of Michigan, Ann Arbor, Michigan 48103.

(Accepted for publication January 20, 1969.)

NOTES ADDED IN PROOF

1. It has been pointed out to me that the explanation I presented for the data of this experiment—the sliding temporal window—is not unequivocal. An equally strong case can be made for a model in which randomly placed moments lead to a lower than average number of dots for a trailing or leading character, than an embedded one. Since our data are pooled, it is difficult to distinguish between these two models. In retrospect, it appears that this criticism is valid, not only for our paper, but for all experiments in which data for many trials are pooled. To the best of our knowledge the only report that deals experimentally and specifically with the issue of a sliding versus a discrete moment, is a recent study by Allport (1968). His conclusion is, however, in agreement with the less strong argument of this paper, namely that the moment or integrating period is more likely to be sliding continuously in time and thus be synchronized to any stimulus event, than it is to be discrete and independent of a stimulus such as our character.

2. A continuing literature survey has also turned up a number of studies overlooked in the discussion sections of this paper.

An earlier paper dealing with the recognition of dot geometric forms in static dot noise has been encountered. French (1954) discovered that increasing the number of dots in the noise background decreased the recognizability of geometric forms, while increasing the number of dots in the form increased their recognizability. His study did not deal with the temporal dimensions

emphasized in this paper, but deserves acknowledgment as an antecedent and relevant study.

Kinsbourne and Warrington (1962a, 1962b) also studied the effects of static noise patterns on alphabetic character recognition. Their studies used filled figures, but found many of the same features of visual masking that we report in this paper.

The work of Mayzner, Tresselt, and Helfer (1967) also seems to have many areas of overlap in both technique and interpretation with our own studies. In their case the masking stimuli are other alphabetic characters and the order or presentation is critical. However, they also find much shorter periods of time typical of the integrating period than is found in the ring and annuli experiments.

REFERENCES ADDED IN PROOF

- ALLPORT, D. A. Phenomenal simultaneity and the perceptual moment hypothesis. *British Journal of Psychology*, 1968, 59, 395-406.
- FRENCH, R. S. Pattern recognition in the presence of visual noise. *Journal of Experimental Psychology*, 1954, 47, 27-31.
- KINSBOURNE, M., & WARRINGTON, E. K. The effect of an aftercoming random pattern on the perception of brief visual stimuli. *Quarterly Journal of Experimental Psychology*, 1962, 14, 223-234.
- KINSBOURNE, M., & WARRINGTON, E. K. Further studies on the masking of brief visual stimuli by a random pattern. *Quarterly Journal of Experimental Psychology*, 1962, 14, 235-245.
- MAYZNER, M. S., TRESSELT, M. E., & HELFER, M. S. A provisional model of visual information processing with sequential inputs. *Psychonomic Monograph Supplements*, 1967, 2, 91-108 (Whole No. 23).