AF TECHNICAL REPORT NO. 5924

FACTORS DETERMINING THE LEGIBILITY OF LETTERS AND WORDS DERIVED FROM ELEMENTAL PRINTERS

The third of a series of reports
on
THE "INFOMAX" PRINCIPLE

Eugene R. Long
L. Sterling Reid
Ralph W. Queal

University of Virginia

August 1951

Aero Medical Laboratory
Contract No. W 33(038)-ac-21269
RDO No. 694-37

Wright Air Development Center
Air Research and Development Command
United States Air Force
Wright-Patterson Air Force Base, Ohio

McGregor & Werner, Inc.
Dayton, O. 150, 20 Nov. 52

FOREWORD

This report is the third in a series covering exploratory research on factors determining the legibility of letter and word patterns formed by an elemental printer (i.e., composed of individual elements or "dots"). This research on the "Informax principle" has been carried on by the University of Virginia for the Communication and Navigation Laboratory. These studies have been conducted under the general supervision of Dr. R. H. Henneman of the Psychology Department. Technical supervision was provided by the Aero Medical Laboratory, under Research and Development Order No. 694-37, Human Engineering Analysis of Aircraft Communication Techniques, with Mr. Julien M. Christensen as Project Engineer.

AD-A800032

ABSTRACT

This Technical Report covers the third and concluding experimental study concerned with the legibility of letter and word patterns formed by "dots" or individual elements. The first study was limited to letters, printed in black and white. The second study added letters printed in "gray scale" (i. e., printing signal and noise elements in brightnesses proportional to their respective intensities). The third study added four-letter words and four-letter "jumbles", printed in both gray scale and black and white. Other stimulus variables (matrix size, type of degradation, and degree of degradation) were the same as in the two earlier studies.

The same factors as in the previous studies were found to decrease legibility. Again legibility loss was minimized, by the use of a larger matrix and by printing in the gray scale. Four-letter words and single letters were equally legible (or illegible); four-letter jumbles were less legible under all experimental conditions. The less highly trained observers in this study demonstrated consistently lower and more variable performance than those of the earlier studies, using the same stimulus patterns. Tentative recommendations are made for the engineering design of elemental printers in order to achieve greatest legibility.

PUBLICATION APPROVAL

Manuscript copy of this report has been reviewed and found satisfactory for publication.

For the Commanding General:

Robert H Blount

ROBERT H. BLOUNT
Colonel, USAF(MC)
Chief, Aero Medical Laboratory
Research Division

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

INTRODUCTION

In two previous reports (2, 3), the advantages and disadvantages of a recently proposed pulse-activated visual presentation system were discussed together with descriptions of the preliminary research conducted to investigate it. More specifically, it was postulated that despite certain advantages of such a system, the signal would still be susceptible to various types of "noise" or distortion which might produce uncontrolled degrading changes in the elemental-formed visual message pattern, resulting in reductions of recognitions or legibility. It was further assumed that these signal distortions would result in such degradations of message patterns as (1) addition of spuriously active elements, (2) omissions of signal elements, and (3) simultaneous combinations of both addition and omission.

The research designed to ascertain the effects of these postulated types of degradation on the recognition of single letters, indicated that all of them reduced recognition, and approximately in proportion to their degree or amount. However, it was further found that these recognition losses might be reduced by either increasing the number of matrix elements or by allowing signal and noise to be printed in brightnesses proportional to their respective strengths, rather than in black and white alone.

This previous research involved only the recognition of single letters. It was decided, therefore, to re-investigate the effects of the same stimulus changes, employing letter groups as well as individual letters. This decision was made since it seemed reasonable to assume that a particular letter might be either more or less easily recognized when included in the context of four-letter groups than when presented alone. This question was investigated by presenting 11 different letters (under the types and degrees of degradation previously investigated) either individually or in one of the two following four-letter groups: (1) four-letter meaningful words (e.g., "BEAN"); (2) meaningless, random four-letter groups (e.g., "NBEA"). This study was designed to indicate whether or not such factors as meaning, grouping etc., played a significant part in the recognition of degraded, elemental-formed letter patterns.

A second feature of this experiment should also be pointed out. Since both of the two previous experiments utilized matched-group techniques, involving groups of four subjects, each of whom viewed all of the stimulus patterns, it was decided that a different design should be employed. This time a design employing many more subjects, each of whom viewed only a specific group rather than all of the slides, was utilized in order to determine to what extent the results of the first two experiments may have been contingent upon such factors as sampling, experimental design, etc.

With the exception of these two differences, the procedure, methodology, and rationale of this experiment were essentially the same as those of the two earlier experiments.

PLAN OF INVESTIGATION

The primary purpose of this third investigation was to learn the effects of the various types and degrees of stimulus change on the legibility of three types of letter patterns. As in the previous experiments, the following variables were manipulated; (1) number of matrix elements; (2) type of degradation; (3) degree of degradation; and (4) differential brightness of printed signal and noise. However, in this experiment a fifth variable type of letter pattern (single letters, meaningful four-letter words, and random four-letter groups), was also introduced. These variables were organized into the following 4 x 3 x 3 x 2 x 2 multifactorial design. (See Table I.)

Although the design of this experiment was basically similar to those previously employed (except for the inclusion of the variable stimulus type), its applications included several major differences. Instead of letting this design represent the distribution of stimulus changes made for all 26 letters of the alphabet, and instead of having each subject view all the slides, a different plan was adopted.

First, only 11, instead of all 26 letters, were included. These were viewed under all of the variable stimulus conditions, presented singly, and when included in both the meaningful and random letter groups. Thus, each combination of matrix size, degradation type, degree of degradation, and printed brightness was represented by 33 stimulus patterns --11 single letters, 11 meaningful four-letter words, and 11 random four-letter groups.

Secondly, 144 subjects were randomly distributed among the 48 different combinations of matrix size, degradation type, degree of degradation, and printed brightness. Further, the three subjects assigned to a given combination were to view only the 33 slides (11 single letters, 11 meaningful words, and 11 random four-letter groups) of that particular combination and no others. This was in contrast to the procedure previously employed, where each subject viewed all letter patterns under all combinations of stimulus change. It should be further pointed out that with this plan, the subjects were matched with respect to stimulus type, but were unmatched with respect to the other variables.

TABLE I
Experimental Design

Degradation Signal and Noise Printed in Black and White and Gray Scale	35-Element Matrix						140-Element Matrix					
	Omission		Addition		Omission + Addition		Omission		Addition		Omission + Addition	
	B&W	Gry	B&W	Gry	B&W	Gry	B&W	Gry	B&W	Gry	B&W	Gry
Single Letters	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	10%	10%	10%	10%	10%	10%	10%	10%	10%	10%	10%	10%
	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%
	60%	60%	60%	60%	60%	60%	60%	60%	60%	60%	60%	60%
Words	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	10%	10%	10%	10%	10%	10%	10%	10%	10%	10%	10%	10%
	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%
	60%	60%	60%	60%	60%	60%	60%	60%	60%	60%	60%	60%
Jumbles	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	10%	10%	10%	10%	10%	10%	10%	10%	10%	10%	10%	10%
	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%
	60%	60%	60%	60%	60%	60%	60%	60%	60%	60%	60%	60%

The first of the above alterations in procedure was carried out because it was assumed that the 11 randomly selected letters would yield a sample highly representative of all 26 letters, and thus reduce time and effort. This was particularly important here since the number of letter patterns was increased three-fold due to the introduction of the two letter-groups in addition to the single letters. The second alteration was introduced in order to ascertain to what extent the results of the first experiment were contingent upon a limited sampling of subjects, and the use of a particular matched group technique.

METHOD

1. Design of Stimulus Patterns

In accordance with the methods of the two earlier experiments, single letters were composed of elements (either 35 or 140) in the type style similar to that proposed by the Control Instrument Company (1). These letter patterns, however, were composed of elements which were square instead of round, and which were immediately juxtaposed rather than separated.

The addition of spurious elements was made to fall randomly only on the unoccupied, non-figure (non-signal) areas; omission was similarly introduced but only on the occupied, figure (signal) areas; and simultaneous addition and omission was produced by simultaneously combining previous patterns of addition and omission. If the letter patterns were printed in black and white, undistorted figure (signal) elements were always black and undistorted non-figure (non-signal) elements were always white. In addition, figure (signal) elements which were omitted were always white, while added, spuriously active, non-figure (non-signal) elements were always black. If, on the other hand, the various signal and noise elements were printed in brightnesses corresponding to their assumed intensities (gray scale), omitted figure elements were printed in either white, light gray, or dark gray, but never black. Added, non-figure, spuriously active elements were printed in either black, dark gray, or light gray, but never white. Undegraded letter patterns were the same regardless of whether they were printed in gray scale or black and white--figure elements always being black; unoccupied, non-figure elements always being white.

The inclusion of the new "grouped" letter patterns involved no new assumptions as far as distribution and brightness of signal and noise elements were concerned. It simply involved the simultaneous combining of four previously determined individual letter patterns. Suppose, for example, the problem was to determine whether the letter "E" printed with the 140-element matrix in gray scale at 30% omission was more readily recognized when presented singly than when included in the context of a meaningful four-letter word such as "MOVE" or when included in the context of the same letters but rearranged in some random, meaningless sequence, such as "MEOV". The procedure for preparing these patterns involved simultaneously combining in one of two sequences the individual letter patterns of "E", "M", "O", or "V" --all printed with the 140-element matrix in gray scale and 30% omission.

The 11 letters used in carrying out this experiment were A, B, E, K, O, P, R, S, W, X, AND Z. The recognition of these letters was measured not only when presented alone, but also when included in both meaningful, as well as random, meaningless sequences. The letters and their grouped sequences were as follows:

Letter	Meaningful four-letter word	Meaningless, random four-letter sequence
A	Y(A)RD	RYD(A)
B	(B)EAN	N(B)EA
E	MOV(E)	M(E)OV
K	JO(K)E	J(K)E)
O	H(O)CK	(O)CHK
P	SLI(P)	L(P)SI
R	TU(R)F	UFT(R)
S	(S)TAG	G(S)AT
W	(W)ISH	H(W)SI
X	E(X)AM	ME(X)A
Z	QUI(Z)	IUQ(Z)

These stimuli were obtained by first randomly selecting 11 letters from the alphabet. Then four-letter words which included these letters were arbitrarily selected. Finally, after the letters composing each of these words were numbered from one to four, they were rearranged according to a table of random numbers. This was done because it was believed that the use of the two types of letter groups (one meaningful and one non-meaningful) would permit a more accurate independant estimate of such factors as meaningfulness, grouping, etc.

2. Preparation of Stimulus Material

After all types, degrees, locations, and brightnesses had been determined for the letters presented singly as well as those composing the letter groups, 35-millimeter transparencies were made for each of these patterns by photographing the corresponding patterns composed of wooden blocks, with sides painted black, white, light green and dark yellow.* The slides

*These two chromatic values were found necessary to yield photographed patterns of light gray and dark gray respectively.

for the individual letters were thus prepared as for the two earlier studies. The preparation of the slides of the letter groups simply involved the simultaneous photographing of four appropriate letter patterns instead of just one. Thus, these slides differed from the others in this study, and the previous studies, only in the sense that they contained four letters instead of one. (See Figures 1, 2, and 3 for representative samples.)

Other dimensions of the stimulus patterns remained unchanged from the previous studies. These include: (1) matrix area of individual letters-- 5" x 7"; (2) size of wooden blocks--either 1" x 1" or 1/2" x 1/2" depending on the matrix; and (3) size of letters on the slides 3/8" x 1/4" x 1/16" corresponding to visual angles of 1^{o} 30", 1^{o}, and 20" respectively when viewed at a constant distance of 14 inches. Exposure time, as in the earlier studies, was four seconds, while pre-and post-exposure times were six seconds in duration. The brightness of the exposure field, as well as of the pre and post-exposure fields was held constant at 87.5 millilamberts.

3. Procedure

The procedure followed in this third experiment was somewhat different from that of the two previous experiments. In each of the earlier experiments, only four observers were employed and each of these viewed all the slides.

In this experiment, 144 subjects were randomly assigned to one of the 48 cells (possible combinations of matrix, printed brightness, degradation type, and degree of degradation). Further, although a total of 1,584 slides was prepared, each observer viewed only 33--those representing the particular combination of matrix size, printed brightness, etc., to which he had been assigned. It should be further noted that with respect to the 33 slides which each observer viewed, 11 were single letters, 11 were meaningful four-letter words, and 11 were meaningless, random four-letter groups Thus, each subject viewed single letters as well as letter groups, but never at more than one combination of the variables mentioned. As was pointed out earlier, the subjects were, therefore, matched with respect to stimulus type, but were unmatched with respect to the other variables. This procedure was adopted in order to ascertain to what extent the results of the previous experiments may have been contingent upon a design and procedure which perhaps involved less independence through the possible operation of some complex sequence effect.

In the present study, sequence effects with any group of 33 slides were minimized by numbering each slide and then placing them in random order. The groups of 33 slides were presented throughout the morning and afternoon, the appointed times for viewing a specific group being randomly assigned. The actual presentation of a group of slides required approximately five and one-half minutes, each slide being viewed for four seconds with six

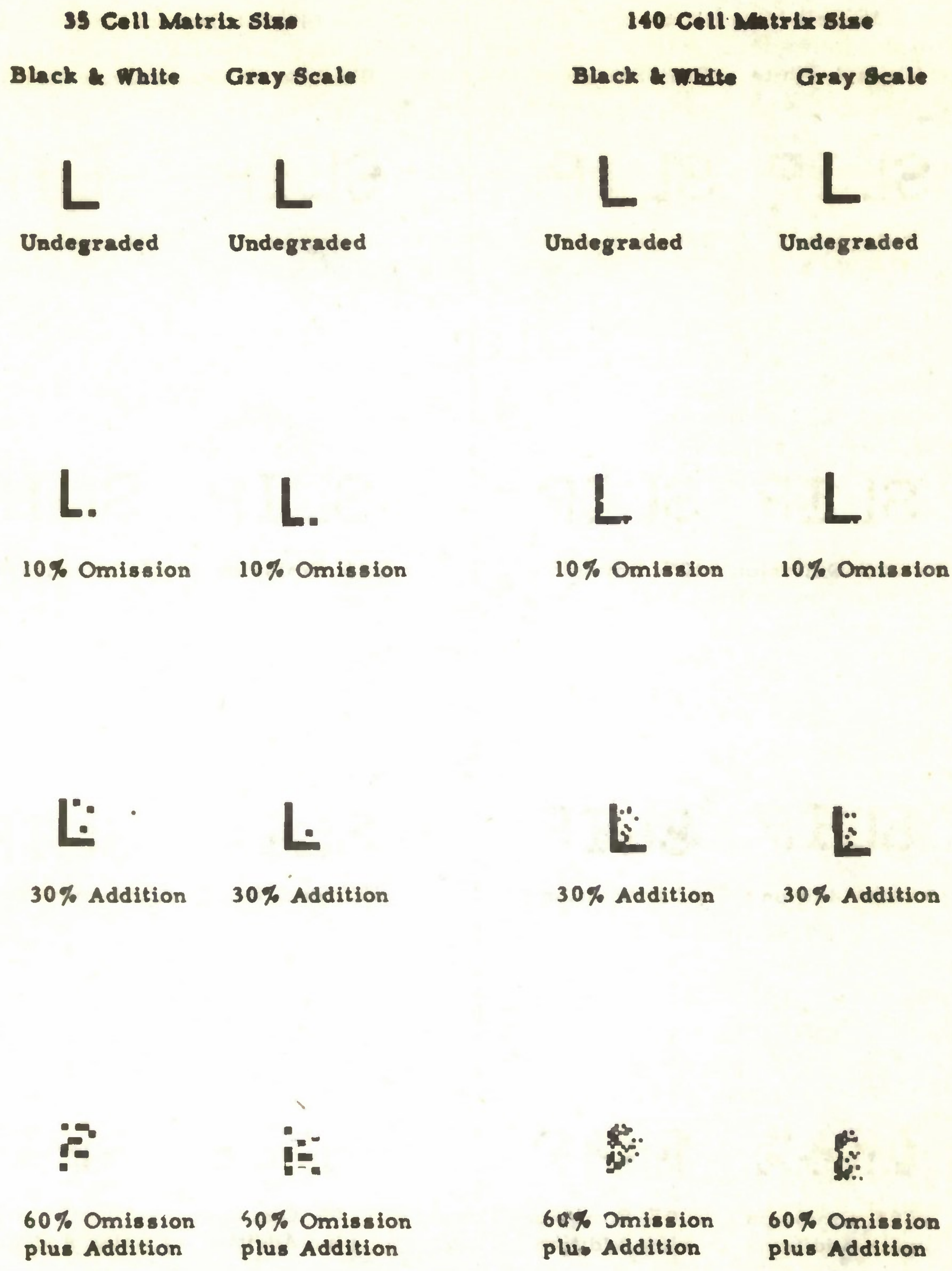

Figure 1. The above figure contains samples of the various stimuli used in this study.

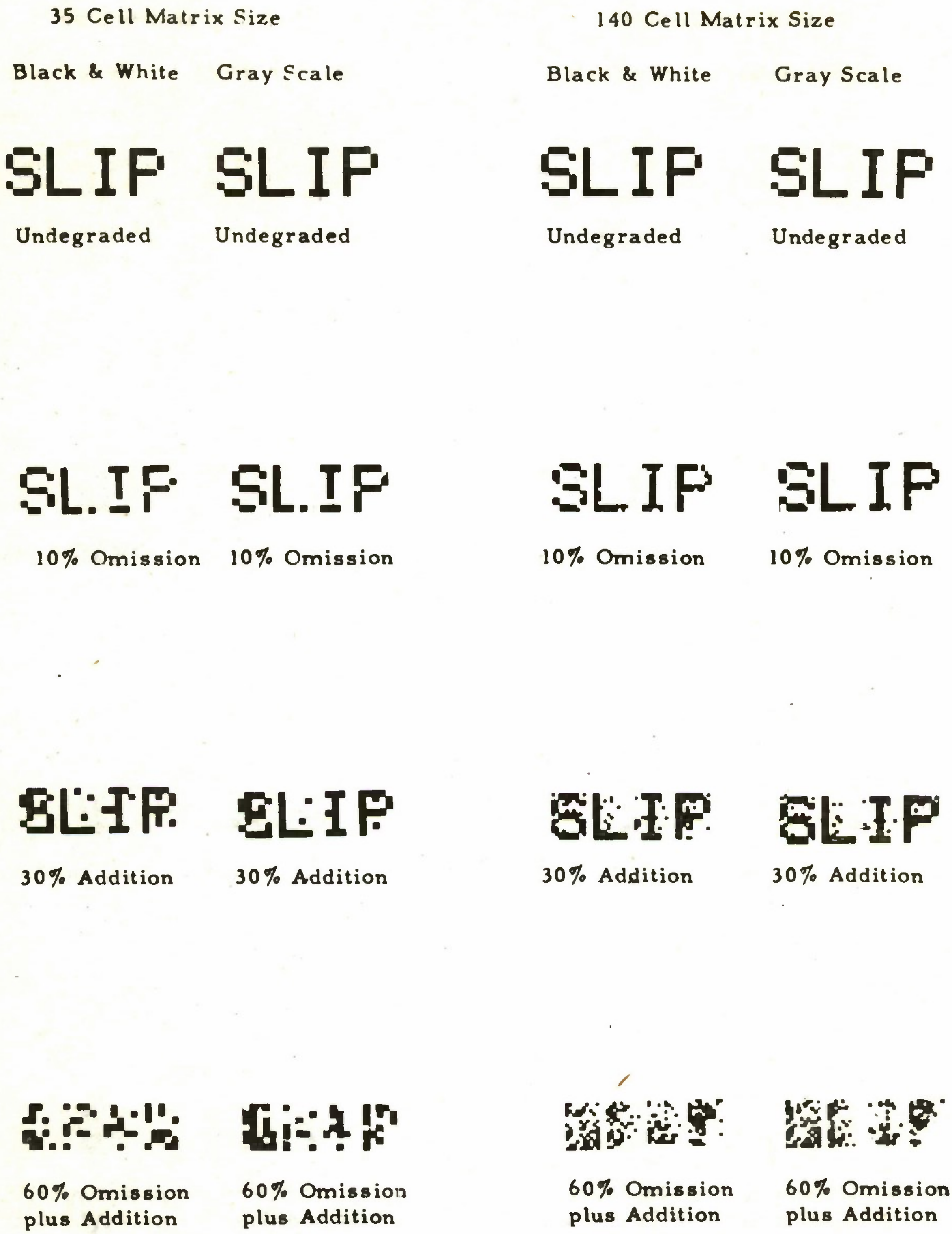

Figure 2 The above figure contains samples of the various stimuli used in this study.

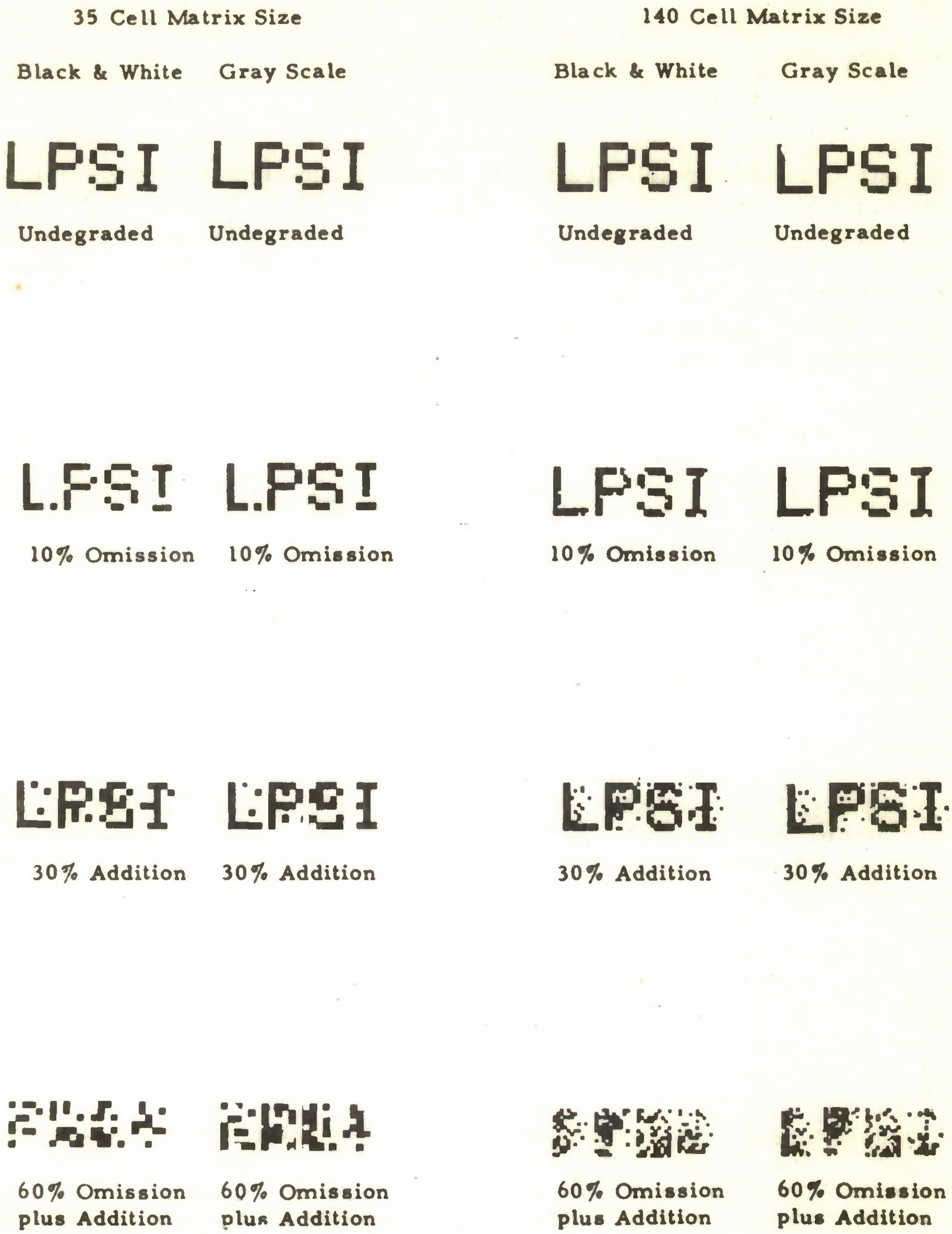

Figure 3. The above figure contains samples of the various stimuli in this study.

seconds separating successive presentations.

In summary, it should be pointed out that the plan and method of the present experiment differed from those previously followed in three ways: (1) only 11 randomly selected single letters of the alphabet were employed, instead of all 26; (2) these letters were presented not only singly, but also in the contexts of meaningful four-letter words and meaningless, rancom four-letter sequences; (3) the present experimental design utilized a much larger group of subjects, each viewing single letters as well as letter groups, but at only a single combination of matrix size, printed brightness, degradation type, and degradation degree.

RESULTS

1. Analysis of Results

a) The numbers of correct recognitions, together with the combinations and permutations of conditions under which they were obtained for each of the 144 subjects, are presented in Tables II, III, and IV. An analysis of these data (Table V) indicates that when the variances of each of the five variables is compared with that of the error term, all are found to be significantly larger. Thus, at this point, we may say that the same variables found to be significant within the limits of the second experiment are again found to be significant. These variables include matrix size, type of degradation, degree of degradation, and printed brightness. In addition to these, stimulus type (single letters vs. meaningful four-letter words vs. groups of four letters in random sequence) is also found to significantly influence numbers of correct recognitions.

b) When the variances of the second-order interactions are compared with that of the error term, two are found to be significantly larger--matrix size, type of degradation, and degree of degradation (A x B x C) and type of degration, degree of degradation, and gray scale (B x C x D). The variance interaction of matrix size, type of degradation, and printed brightness (A x B x D) approach significance but do not quite satisfy the five percent criterion level. It should also be noted that none of the interactions which include stimulus type is significant or even approaches significance, these variances ranging only from 0.1 to 1.2.

c) When the first-order interactions are compared with the error term, four are found to be significantly larger. These are matrix size and type of degradation (A x B), matrix size and degree of degradation (A x C), type of degradation and degree of degradation (B x C), and degree of degradation and printed brightness (C x D). Again, no interactions involving stimulus type approach significance.

TABLE II

Raw Scores for Letters*

Matrix Size	35 cell						140 cell					
Degradation Type	Omission		Addition		Omission plus Addition		Omission		Addition		Omission plus Addition	
Printed Brightness	B&W	Gray	B&W	Gray	B&W	Gray	B&W	Gray	B&W	Gray	B&W	Gray
	11	9	10	9	10	11	10	11	11	10	11	10
0%	11	9	9	10	11	11	11	11	10	10	11	11
	10	10	10	11	9	10	9	11	9	11	10	10
	10	9	9	9	7	9	10	8	10	10	8	10
10%	9	10	10	8	5	6	10	11	9	11	10	10
	10	9	10	8	8	9	10	9	10	10	8	10
	5	8	6	8	1	1	10	10	8	7	4	8
30%	2	9	9	9	0	1	8	11	5	6	1	8
	3	8	8	8	2	0	9	9	10	9	1	8
	0	6	3	5	0	0	10	11	8	7	1	2
60%	0	4	4	4	2	1	6	9	6	6	0	1
	2	6	6	6	0	0	9	11	5	8	0	3

*Note: Each subject here is the same subject in the corresponding cell for words and jumbles stimulus type in Table III and IV

TABLE III

Raw Scores for Words*

Matrix Size	35 cell						140 cell					
Degradation Type	Omission		Addition		Omission plus Addition		Omission		Addition		Omission plus Addition	
Printed Brightness	B&W	Gray	B&W	Gray	B&W	Gray	B&W	Gray	B&W	Gray	B&W	Gray
0%	10	11	11	10	9	11	11	11	11	11	10	11
	10	9	10	10	11	11	11	11	10	10	11	10
	10	10	10	10	7	9	9	11	9	10	9	10
10%	11	11	8	9	6	9	11	11	9	11	9	9
	11	10	7	10	8	8	10	11	10	10	9	10
	11	7	6	10	6	6	11	10	10	11	11	9
30%	5	7	10	5	0	3	10	11	9	8	4	9
	1	8	6	6	0	1	8	11	7	9	1	9
	1	8	5	8	2	2	2	10	10	10	3	6
60%	0	4	1	5	0	0	6	11	7	6	0	0
	0	4	4	4	0	0	5	10	6	7	2	2
	1	6	4	8	0	1	9	11	0	8	1	1

* Note: Each subject here is the same subject in the corresponding cell for letters and jumbles stimulus type in Table II and IV

TABLE IV
Raw Scores for Jumbles*

Matrix Size	35 cell						140 cell					
Degradation Type	Omission		Addition		Omission plus Addition		Omission		Addition		Omission plus Addition	
Printed Brightness	B&W	Gray	B&W	Gray	B&W	Gray	B&W	Gray	B&W	Gray	B&W	Gray
0%	10	11	10	10	10	9	10	11	11	11	11	9
	10	8	10	10	11	11	11	11	10	10	11	10
	10	7	9	10	7	9	9	11	10	11	10	10
10%	9	11	7	9	7	8	10	8	9	9	7	9
	8	9	7	10	6	4	10	11	7	10	10	11
	9	8	6	9	5	7	11	9	8	8	9	10
30%	5	4	5	6	1	0	10	11	8	8	2	7
	2	8	8	6	0	1	6	10	5	6	1	6
	5	8	6	5	0	2	10	9	9	9	2	5
60%	0	3	1	3	0	0	6	11	6	7	0	0
	0	6	2	3	0	1	5	10	6	3	3	1
	1	4	4	7	0	1	8	10	3	6	0	2

*Note: Each subject here is the same subject in the corresponding cell for words and letters in Table II and III

TABLE V --- The Analysis of Variance of Correct Recognition Scores

	Source of Variation	Sum of Squares	df	Mean Squares
A.	Matrix size	412.4	1	412.3
B.	Type of Distortion	695.2	2	347.6
C.	Degree of Distortion	2686.8	3	895.6
D.	Printed Brightness	133.4	1	133.4
E.	Stimulus Type	29.9	2	15.0
	AxB	62.5	2	31.2
	AxC	116.9	3	38.9
	AxD	0.0	1	0.0
	AxE	3.4	2	1.7
	BxC	397.9	6	66.3
	BxD	12.4	2	6.2
	BxE	3.1	4	0.8
	CxD	67.2	3	22.4
	CxE	8.7	6	1.4
	DxE	3.3	2	1.6
	AxBxC	163.6	6	27.3
	AxBxD	10.6	2	5.3
	AxBxE	0.4	4	0.1
	AxCxD	4.8	3	1.6
	AxCxE	7.3	6	1.2
	AxDxE	1.0	2	0.5
	BxCxD	82.2	6	13.7
	BxCxE	14.1	12	1.2
	BxDxE	3.8	4	1.0
	CxDxE	5.0	6	0.8
	Error	587.9	288	2.0
	Total	5513.7	431	

	Source of Variation	F Ratio	df	P
1.	Matrix Size (A)/Error	206.1	1;288	.01>P
2.	Type of Distortion (B)/ Error	173.8	2;288	.01>P
3.	Degree of Distortion (C)/Error	447.7	3;288	.01>P
4.	Printed Brightness (D)/Error	66.7	1;288	.01>P
5.	Stimulus Type (E)/Error	7.5	2;288	.01>P
	AxB/Error	15.6	2;288	.01>P
	AxC/Error	19.4	2;288	.01>P
	AxD/Error	0.0	1;288	.05<P
	AxE/Error	0.8	2;288	.05<P
	BxC/Error	33.1	6;288	.01>P

Source of Variation	F Ratio	df	P
BxD/Error	3.1	2;288	.05<P
BxE/Error	0.4	4;288	.05<P
CxD/Error	11.2	3;288	.01>P
CxE/Error	0.7	6;288	.05<P
DxE/Error	0.8	2;288	.05<P
AxBxC/Error	13.6	6;288	.01>P
AxBxD/Error	2.6	2;288	.05<P
AxBxE/Error	0.05	4;288	.05<P
AxCxD/Error	0.8	3;288	.05<P
AxCxE/Error	0.6	6;288	.05<P
AxDxE/Error	0.25	2;288	.05<P
BxCxD/Error	6.8	6;288	.01>P
BxCxE/Error	0.6	12;288	.05<P
BxDxE/Error	0.5	4;288	.05<P
CxDxE/Error	0.4	6;288	.05<P
AxB/AxBxC	1.1	2;288	.05<P
AxC/AxBxC	1.4	3;6	.05<P
BxC/AxBxC	2.4	6;6	.05<P
BxC/BxCxD	4.8	6;6	.05>P
BxD/BxCxD	-	2;6	-
CxD/BxCxD	1.6	3;6	.05<P
A/AxBxC	15.1	1;6	.01>P
B/AxBxC	12.7	2;6	.01>P
B/BxCxD	25.4	2;6	.01>P
B/BxC	5.3	2;6	.05>P
C/AxBxC	32.8	3;6	.01>P
C/BxCxD	65.4	3;6	.01>P
C/BxC	13.5	3;6	.01>P
D/BxCxD	9.7	1;6	.05>P

TABLE V cont'd.--The Analysis of Variance of Correct Recognition Scores

d) After these comparisons are made, the significant first-order interactions may be further evaluated by comparing their variances with those of the appropriate significant second-order interactions. When this is done, only the interaction of type of degradation and degree of degradation (B x C) is found to be significant, and then only when compared with the interaction of type of degradation, degree of degradation, and printed brightness (B x C x D).

e) A re-evaluation of the manipulated variables, involving a comparison of them with their significant first-and second-order interactions, shows that all of the variables are significantly larger. This suggests that their effects are not restricted to the particular values of the variables employed in this experiment, and therefore permits greater generalization.

2. Interpretation of Results.

If one compares the results of the analysis of variance of this experiment with those of the first or second studies, certain marked differences may be seen. The most apparent difference is that concerned with the independence of effects of the variables. In the third experiment the effects, while related to, are not restricted by, any particular values of the other variables. This is somewhat different from the results of the first two experiments where generalizations concerning effects were related, to but in addition restricted by, the particular values of the other variables. A case in point is the finding in Experiments I and II that the increases in numbers of correct recognitions produced by increases in number of matrix elements was not only related to, but was also dependent upon the degree of degradation present. It should be added that this condition of dependency was found to hold not only for matrix size, but for other variables also.

A second difference that may be seen concerns the relative efficacy of printed brightness as compared to matrix size. In the second experiment, changing from black-and-white to the gray scale increased the number of correct recognitions far more than did increasing the number of matrix elements from 35 to 140. In Experiment III, however, the reverse appears to be true. In view of these and other important findings, it would appear that all treatments and their interactions should be carefully scrutinized so that these differences may be explained or resolved.

a) Matrix size: The conclusion that increases in matrix size produce increases in numbers of correct recognitions is suggested by the fact that its variance is significantly larger than that of the error term. That this increase is related to the effects of other variables is indicated by the fact that the second-order interaction of matrix size, degradation, type, and degree of degradation (A x B x C) is significantly larger than the error term, and is also clearly shown in Fig. 4. To this extent all of the experiments agree. However, they do not agree when the variance of the first-order interaction of matrix size and degree of degradation (A x C) is compared with the second-order interaction of matrix size, type of degradation and degree of degradation (A x B x C). In the first two experiments, (A x C) was significantly larger than (A x B x C); in this experiment it is not. A reason for this difference in Experiment III is suggested by Fig. 4. Here it can be seen that the interaction of matrix size and degree of degradation (A x C) is obviously dependent upon the type of degradation, even at zero degree of degradation. In the earlier experiments, a similar dependency upon type of distortion existed, but not to the same degree. The problem which arises is how to explain this differential effect of degradation type on the interaction of matrix size and degree of degradation. Although only suggestive evidence is available, it appears that sampling might in part account for it. In the present study, only 11 of the 26 letters were viewed . It seems quite possible that the 11 letters selected for Experiment III might not have been completely representative of the whole alphabet. This appears to be even more probable when one considers that for a given letter presented with a particular degree and type of degradation only one of an infinitely large population of patterns was employed. Thus inadequate sampling of either letters or patterns per se might account for this discrepancy.

In addition to recognizing the possible operation of this factor it should be further noted that different groups of subjects viewed these letters under the various combinations of conditions. Perhaps more important, these subjects were not previously trained to a uniform level of proficiency in making the observations, thus perhaps accounting for differences in performance even at the zero degree of degradation. If subsequent research proves this reason to be valid, then perhaps the failure to demonstrate an interaction between matrix size and degree of degradation independent from type of degradation has been in part resolved.

The differences between the findings of the experiments may be further resolved if it is recalled that in the earlier experiments matrix size (A) was found to be significantly larger than the second-order interaction of matrix size, type of degradation, and degree of degradation (A x B x C). Thus, in the earlier experiments, if the interaction of matrix size and degree of degradation (A x C) had not been independent from degradation type (C) so that it could be used as an error term, then there would have been no differences between the experimental results as far as matrix size is concerned, because in both Experiments II and III the effect produced by increases in matrix size are independent from the interaction of matrix size, type of degradation, and degree of degradation (A x B x C).

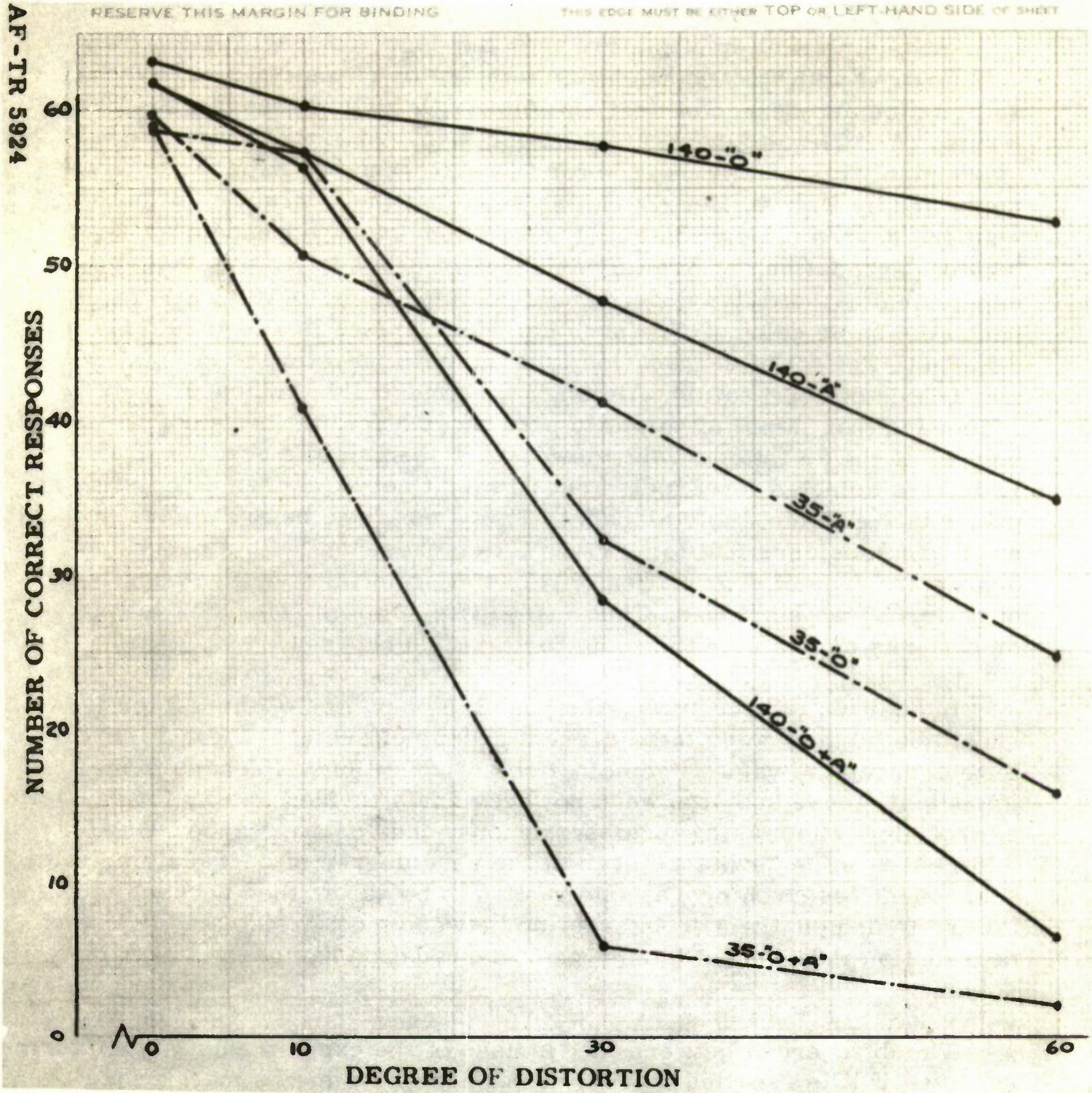

Figure 4. Graph showing the interaction between matrix size, type and degree of degradation.

b) Type of degradation: For the most part in this experiment changes in type of degradation produced changes in recognition quite comparable to those of the earlier experiments. One exception is the fact that degradation type did not significantly interact with matrix size and printed brightness (A x B x D). That some interaction took place is indicated by the magnitude of the mean square of this term and may also be readily seen in Fig. 5. The reason for the failure to attain significance may be again explained perhaps in terms of the heterogeneity of the subjects placed in each cell. It is highly probable that any, or all of several procedures would have yielded a significant interaction among these variables. These might include: (1) increase in number of subjects; (2) better equating of subjects; (3) pre-experimental training of subjects to a higher and more uniform level of performance.

Again, type of degradation interacted significantly with degree of degradation (A x B), this interaction being related to, but not completely dependent upon, matrix size (A). This may also be seen in Fig. 4. One result, however, which does not entirely follow the results of Experiment II is the finding that the degradation type variance is significantly larger than that of the interaction of degradation type and degree of degradation (B x C). In Experiment II B was not significantly larger than B x C. It would appear that this finding, like that concerning matrix size, does not constitute a refutation of previous experimental results but merely again indicates the role played by such factors as experimental design, sampling, and training of subjects. Actually, this finding means that the results of this experiment permit greater generalization than was permissible in the earlier experiments.

c) Degree of degradation: Little needs to be said about the effects of degree of degradation. Again, as in the earlier experiments, it interacts significantly with the other variables. The principal exception in this experiment is the non-significance of the first-order interaction of printed brightness and degree of degradation (B x D). Since the second-order interaction of type of degradation, degree of degradation, and printed brightness (B x C x D) is significant, it appears that the interaction of degree of degradation and printed brightness is dependent upon the type of degradation present. This is clearly shown in Fig. 6.

No better explanation for this suggests itself than those offered to explain the non-significance of the interaction between matrix size and degree of degradation. Again, it appears that the dependency of this first-order interaction (C x D) like (A x C) on type of stimulus (B) is, in part, due to letter sampling. The implicit assumption is that if all 26 instead of only 11 letters of the alphabet had been used, the results might have more closely resembled those of Experiment II.

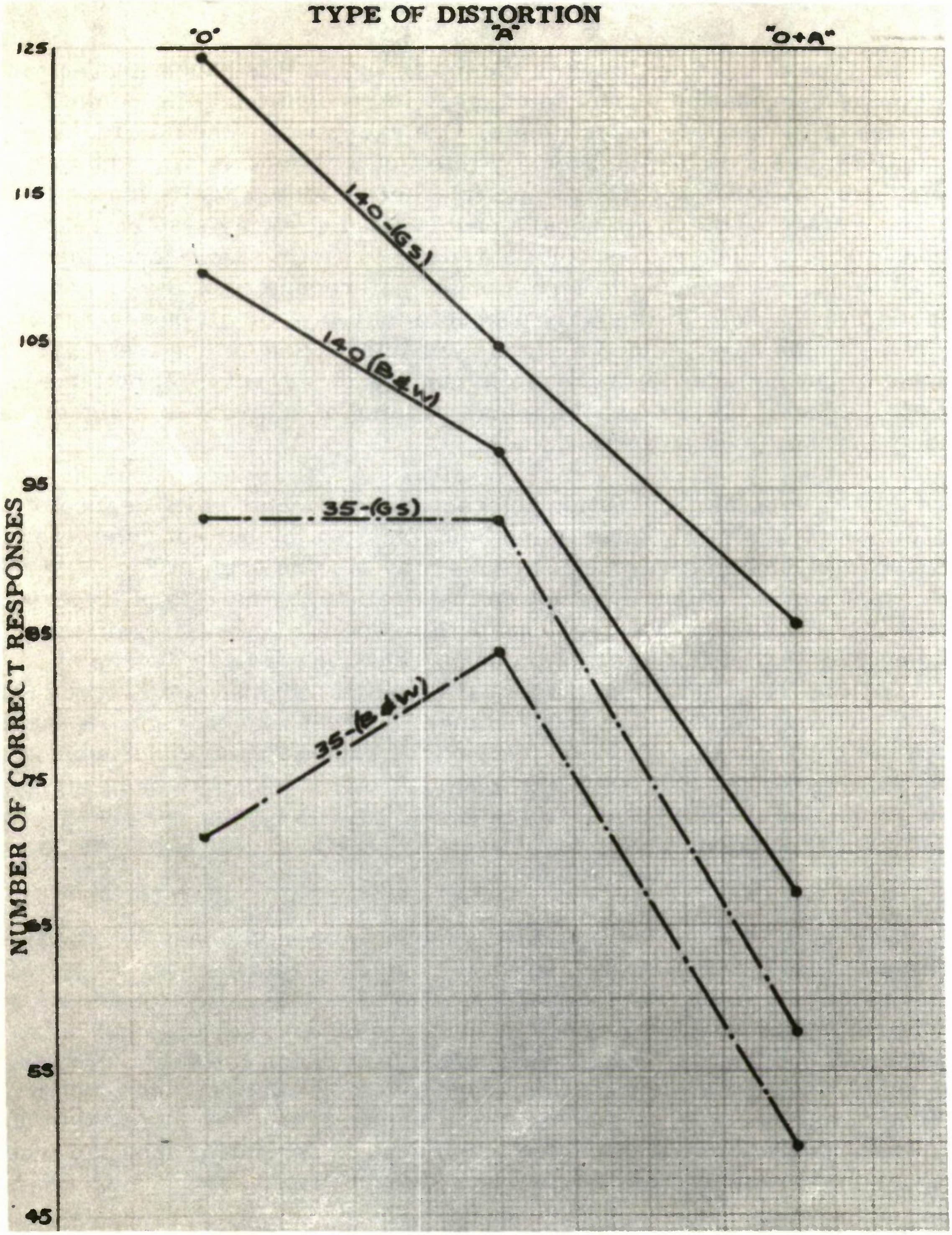

Figure 5. Graph showing the interaction between matrix size, printed brightness and degradation type.

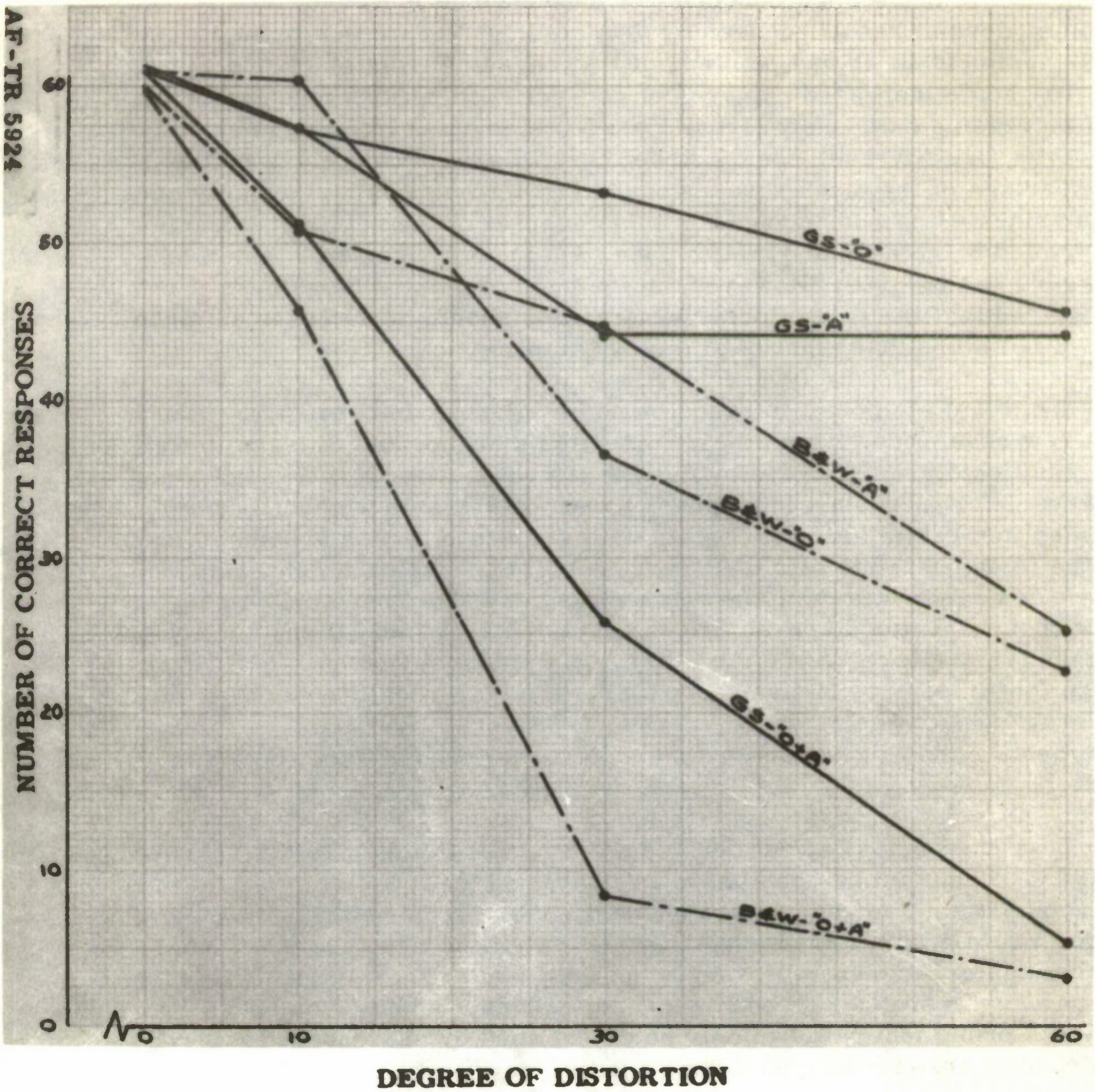

Figure 6. Graph showing the interaction between printed brightness, type and degree of degradation.

d) Printed brightness: Printed brightness has already been discussed in part in the section concerned with degree of degradation. However, another deviant result concerning printed brightness needs to be covered. In Experiment II, it was pointed out that changing from black and white to gray scale increased numbers of correct recognitions much more than did increasing matrix size from 35 to 140 elements. In Experiment III, however, the reverse appears to be true, this being suggested by the comparison of the mean squares of the two terms. Again, the same question arises -- how to account for this reversal of the relative efficacies of printed brightness and matrix size. Again, only the same tentative answers can be given. However, in the case of printed brightness it is suggested that the role of observer learning should be stressed as much as, if not more than, that of stimulus sampling.

This last statement is based on some informal introspective reports of a few observers. It was their contention that they had to learn to utilize the additional cues afforded by the use of gray scale. This was not true, however, for increases in number of matrix elements which apparently yielded immediate aid in recognition. Unfortunately, this hypothesis cannot be tested with the data of Experiment III because each subject made only a few observations in the course of a very short period of time. At first thought it would appear that the data of Experiment II should yield suggestive evidence, since the observers made many observations over much longer periods of time. However, it will be recalled that when several types of performance curves were informally calculated for Experiments I and II, they showed no systematic trends, thus suggesting little or no learning.

It should be pointed out here, however, that this last rather informal analysis should not be considered as indicating that absolutely no learning took place. The very nature of the distribution or grouping of the slides would prohibit any very sensitive test of the hypothesis. In other words, the experiment was simply not designed to show the role of learning *per se* in such a psychophysical experiment. Whether or not it played a prominent role can only be determined by subsequent research designed to answer the question.

The same holds true for the hypothesis concerned with the sampling of letter patterns. If the recognition of some letters is improved more by increasing matrix size than by shifting from black-and-white to gray scale, then these individual differences among letters should be determined by subsequent research, since the discovery of such differences might have important implications for future engineering use. Thus far it appears that for the entire alphabet, changing to gray scale improves recognition more than does increasing matrix size. However, for the 11 specific letters used in Experiment III the reverse seems to be true, although both produce significant increases in recognition.

e) Stimulus type: The fifth and final factor to be evaluated is stimulus type -- do the procedures of presenting letters individually, in the context of four-letter meaningful words, or in random four-letter groups (jumbles) alter the number of correct recognitions? The analysis of the data (Table V, Figs. 7, 8, 9, 10, and 11) indicates that letters presented individually are as frequently recognized as they are when included in meaningful four-letter words. They are significantly less frequently recognized than either of these, however, when included in a random four-letter group.

Another important finding is that the variable, stimulus type, did not significantly interact with any of the other variables. This is indeed a characteristic which differentiates it from the other variables. Although the effects of the other variables were not found to be dependent upon each other, they were nevertheless related, as was indicated by the significant interactions. This was not true for stimulus type.

3. Summary of Results:

a) Numbers of correct recognitions were again found to be increasingly reduced by all types of degradation as the degree of degradation was increased. Greatest loss was again produced by simultaneously combined addition and omission.

b) Both increasing number of matrix elements from 35 to 140, as well as changing from black-and-white to gray scale, increased the number of correct recognitions, increasing number of matrix elements being the more effective. Again, however, greatest increases were produced when both were utilized simultaneously.

c) Although all four of the above variables were related, their independent effects were significantly larger than, and thus not limited by, these interactions. In this respect, the results of the third experiment differ from those of the first two, where generalizations of the effects of all of the variables except degree of degradation were limited to the variable values investigated.

d) The fifth variable, stimulus pattern, was also found to significantly affect numbers of correct recognitions. Letters were as frequently recognized when presented individually as when included in meaningful four-letter words. Numbers of correct recognitions were reduced, however, when the same letters were included in random four-letter groups (jumbles). In addition to these differential effects of the different stimulus patterns, it was found that they did not interact with the other variables, but were completely independent from them.

e) Incidental observations strongly suggest that the proficiency level of the observers plays a prominent role in the recognition of degraded

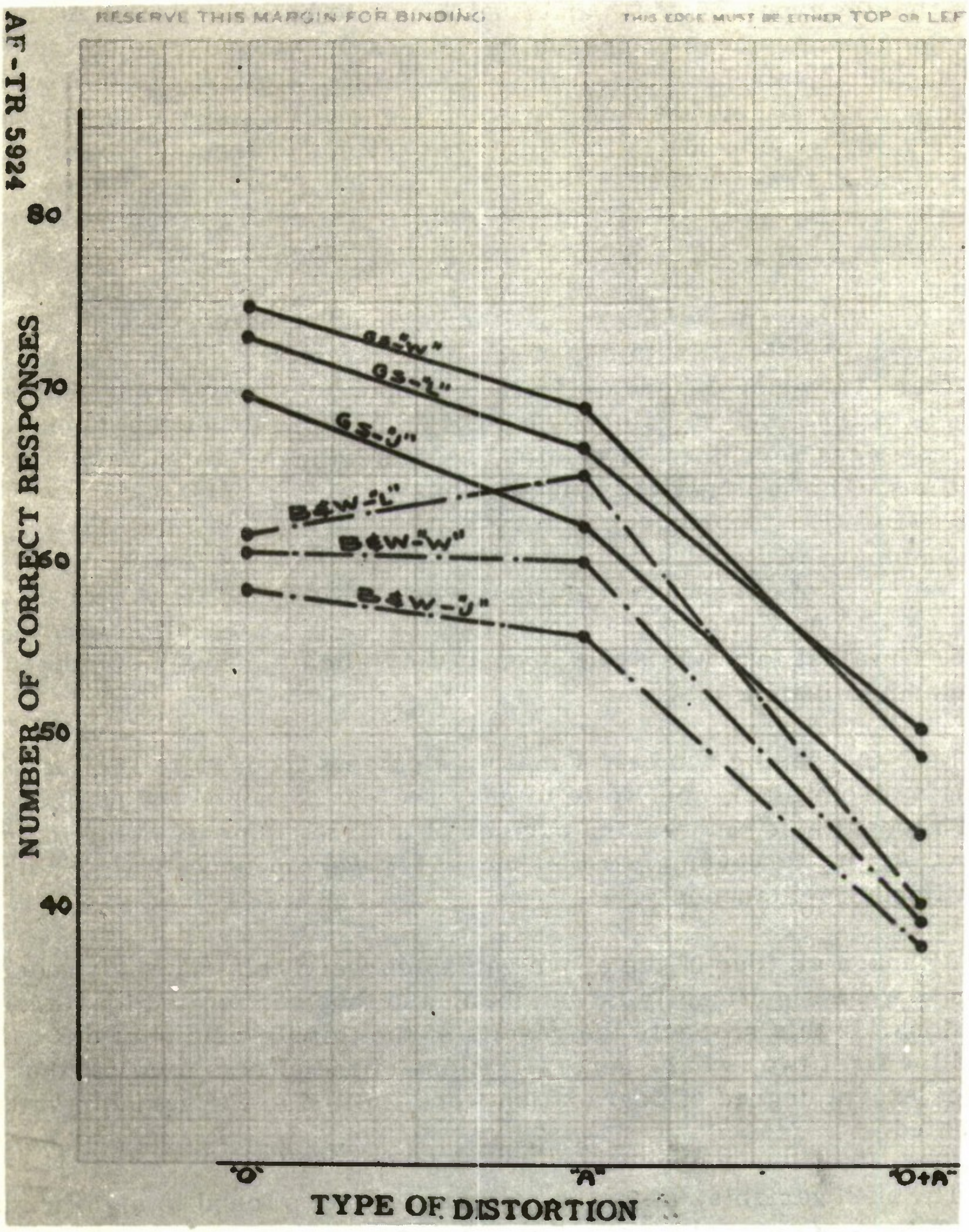

Figure 7. Graph showing the interaction between printed brightness, stimulus type, and type of degradation.

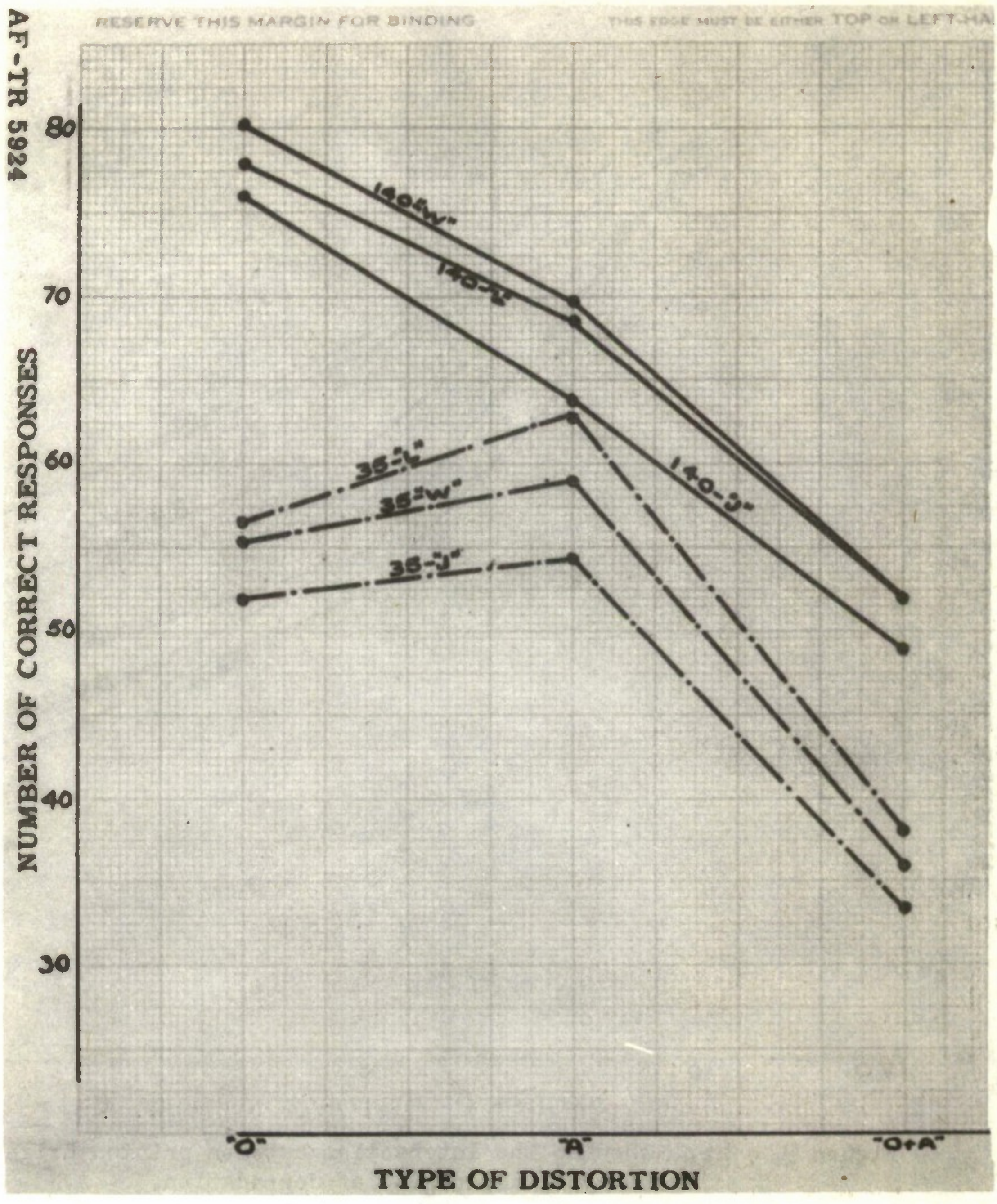

Figure 8. Graph showing the interaction between matrix size, stimulus type, and type of degradation.

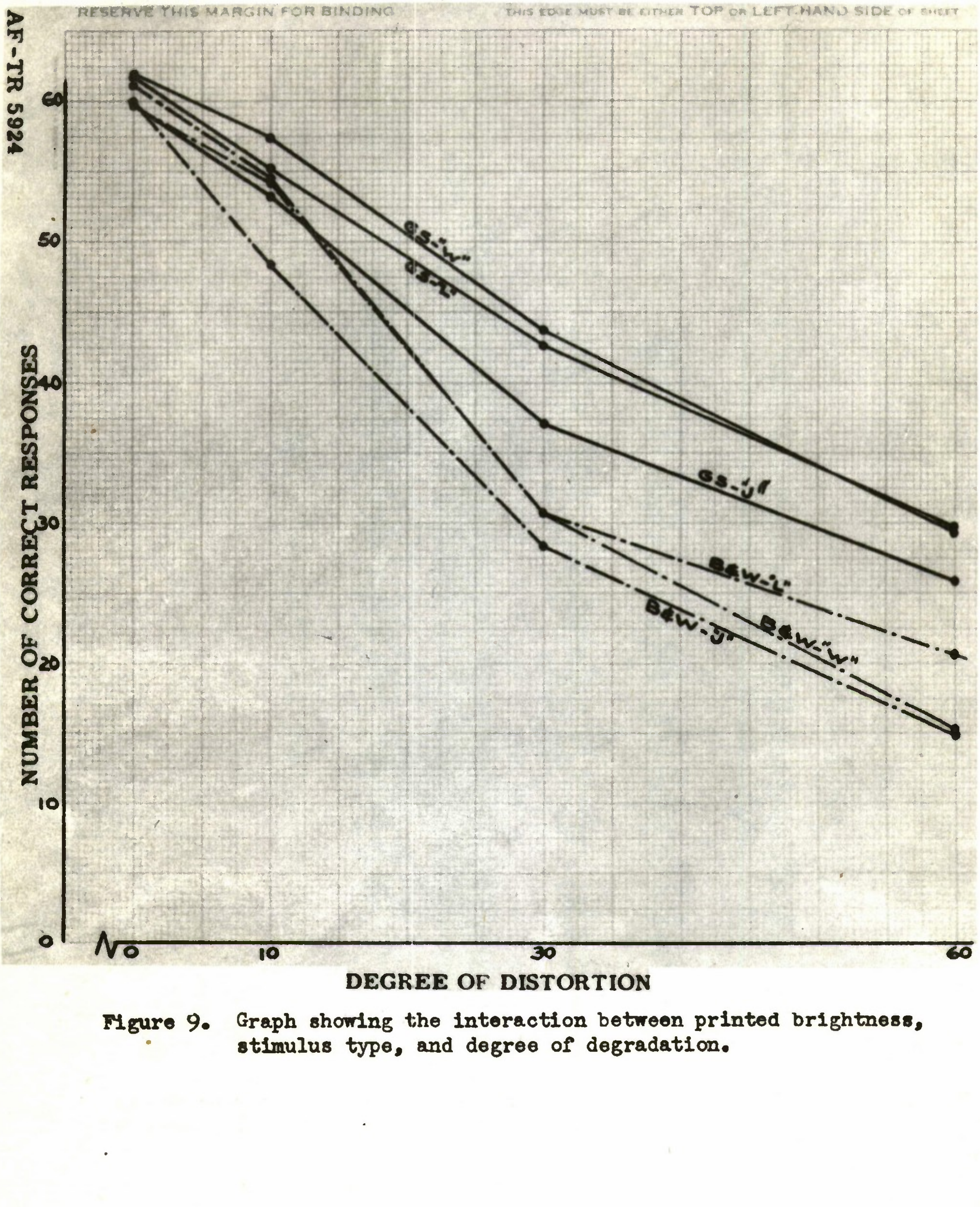

Figure 9. Graph showing the interaction between printed brightness, stimulus type, and degree of degradation.

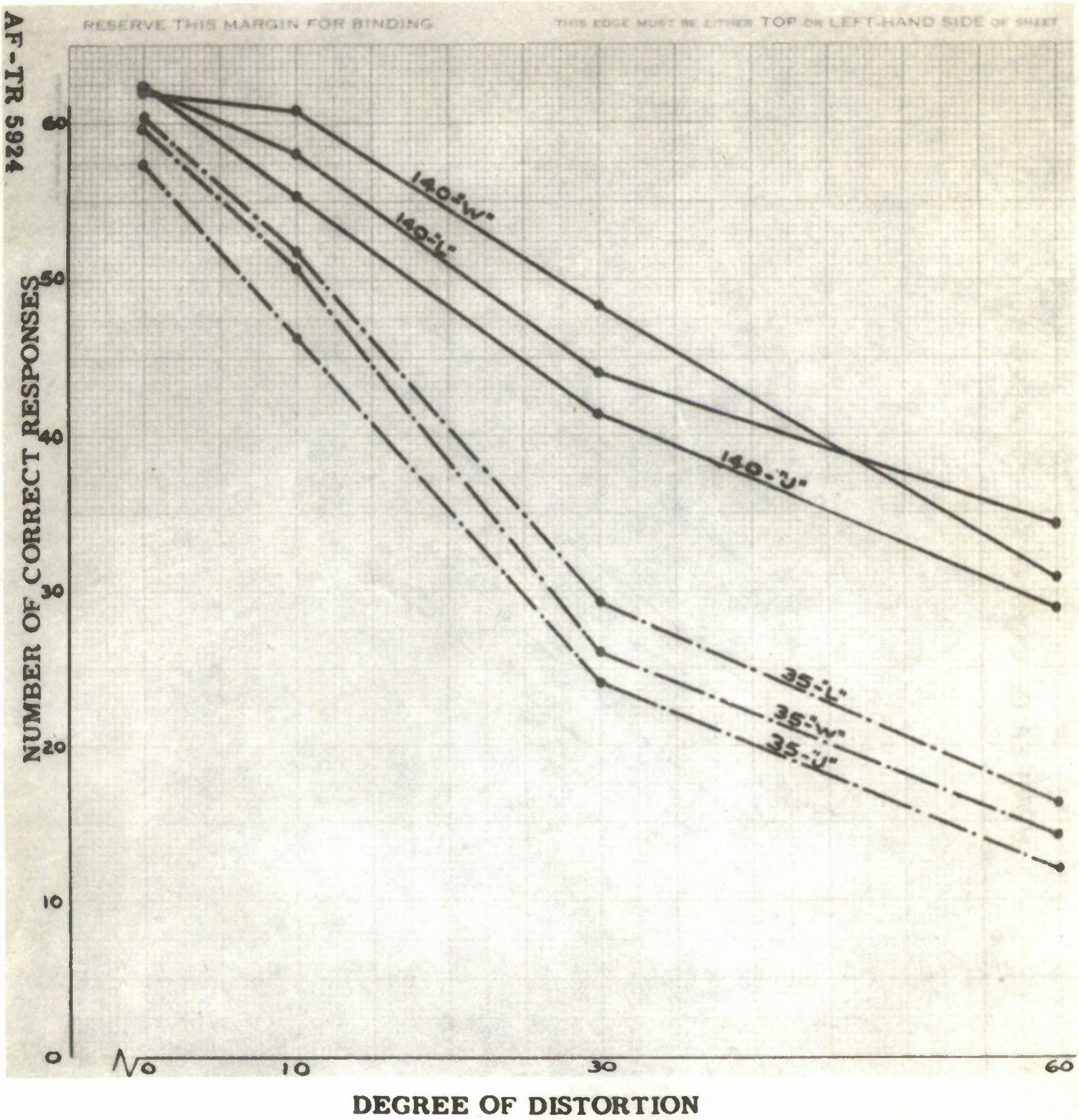

Figure 10. Graph showing the interaction between matrix size, stimulus type, and degree of degradation.

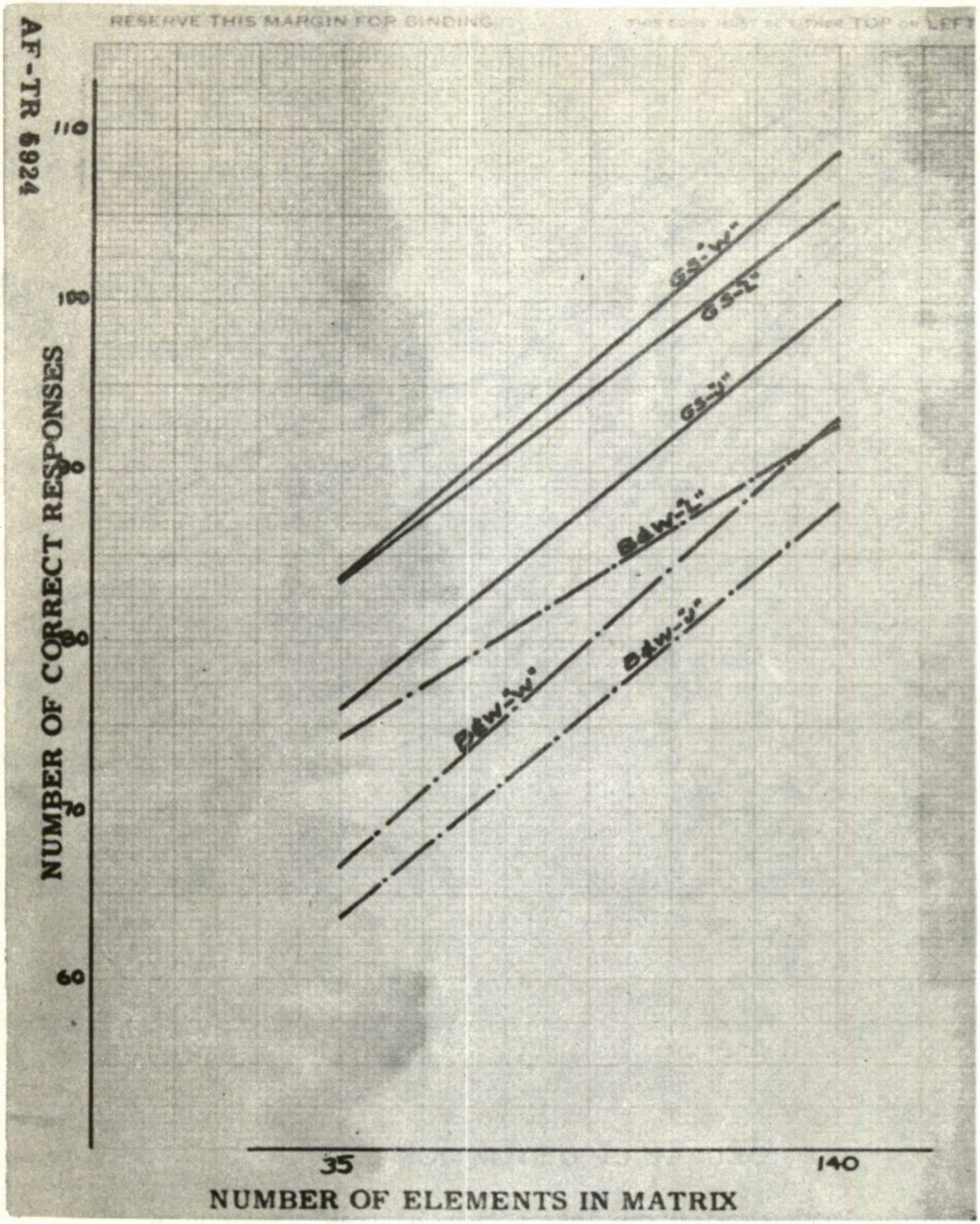

Figure 11. Graph showing the interaction between matrix size, printed brightness, and stimulus type.

patterns. It appears that proficiency level operates in such a way as to increase recognitions directly and independently and also indirectly by augmenting the increases produced by the various favorable stimulus changes, e. g., printed brightness.

DISCUSSION

The results of Experiment III hold important implications for the subsequent design and operation of the previously discussed visual presentation system. For example, the results indicate that if coded messages are to be used, several alternative procedures may be successfully employed, but others should be avoided. This is indicated by the fact that a particular letter was on the average more frequently recognized when either presented alone or when in the context of a four-letter word than when included in a random four letter sequence. This further suggests that if coded text is to be used, it should involve the presentation of individual letters or sequences of letters which have "meaning" for the operator. Further, it would appear that this latter alternative would permit the successful use of apparently random, meaningless groups of letters, only if the operator had previously familiarized himself with these letter groups. This emphasizes the role of operator learning.

Other results of this experiment further point to the possible role of learning in the successful recognition of elemental-formed letters. A case in point is the finding that changing from black-and-white to gray scale was less effective in increasing numbers of recognitions than was increasing number of matrix elements from 35 to 140 -- a relationship in reverse of that obtained in Experiment II. A previously suggested hypothesis to account for this reveral is that subjects in Experiment II had more trials during which they learned to utilize the additional cues furnished by the use of the gray scale. This opportunity was not furnished in Experiment III. Because of the possible operational importance of this factor, it should be investigated in future research.

Unfortunately, at this point it cannot be definitely stated that the differences between the results of Experiments II and III are entirely attributable to learning. This has been previously pointed out, together with the suggestion that increases in matrix size might be a more potent determiner of correct recognition than the use of the gray scale for the 11 letters investigated in Experiment III. Thus, perhaps for some letters the use of gray scale produces better recognition, while for others increasing number of matrix elements is superior. This line of reasoning is based on the assumption that the individual letters of the alphabet depend for their correct recognition upon certain "critical portions". Further it might be assumed that some letters have more critical areas than others. If these "critical portions" are distorted, correct recognition may be seriously reduced or even completely eliminated. This hypothesis of selective letter sampling thus merits further research because of the importance of its engineering application. For example, if only one particular group of letters or characters is employed, differential printed brightness might be the important characteristic to include in the system because of the nature and extent of the critical letter areas involved. On the other hand a different group of letters involving different critical areas might be used with more success if the elemental printer were modified so as to include more elements.

Another important result which needs comment is that of the continued disproportionality of reductions in recognition produced by simultaneously combined omission and addition. If the data are analyzed in a posteriori fashion, it can be seen that in general 30 percent simultaneously combined omission and addition produces greater reductions than does either 60 percent addition or omission alone. It must be reiterated that this is based on mean values because further analyses show reversals. Interestingly enough these reversals occur only under two combinations of conditions: (1) with letters printed in black and white employing the 35-element matrix; and (2) with letters printed in the gray scale employing the 140-element matrix. It does not seem surprising that the disproportionality disappears when the gray scale and the 140-element matrix are simultaneously employed, since this has been demonstrated in earlier research. However, it is difficult to explain why the disproportionality should also disappear when the worst possible printing conditions - black and white with the 35-element matrix - are employed. It would appear that this again might be a result of a particular selection of subjects and letters.

Despite these possible reversals, conclusions based on mean values still suggest that the system should be either highly sensitive or highly insensitive so that either pure addition or pure omission is produced. In the event that simultaneously combined omission and addition is produced, it would appear that the best way to minimize this disproportionally greater reduction in recognition is to simultaneously employ the gray scale as well as the 140-element matrix. The data show that these recommendations hold equally well for individual letters as well as for meaningful and non-meaningful letter groups.

CONCLUSIONS AND IMPLICATIONS

Since the present study is the last of a series of three exploratory investigations on "Infomax" conducted at the University of Virginia, it may be well to summarize the suggestions and implications arising from the three studies both as to equipment design and for future laboratory research on the psychological (legibility) factors involved. It should be pointed out that in a number of important regards the results of the studies supplemented and reinforced one another. The tentative recommendations offered below are proposed under two headings: (1) for engineering design of the visual presentation equipment, and (2) for further psychological research.

1. Implications for Engineering Design of Presentation Devices.

a. The system should permit minimal degradation of the stimulus pattern. High degrees of degradation produce disproportionate losses in recognition.

b. If it is not technically feasible to control signal degradation in the equipment design, special effort should be made to avoid the simultaneous combination of addition and omission of elements. Two alternatives are to be considered: (1) a highly sensitive system, allowing all of the signal and some noise to appear, and (2) a highly insensitive system, precluding noise, but omitting some of the signal-elements. In either of the above it is assumed that the signal-to-noise ratio is greater than one.

c. Increasing the number of matrix elements will decrease loss of recognition maximally if only addition, or only omission, is present, but has less effect if both addition and omission are present simultaneously, in the stimulus pattern.

d. If signal and noise are printed in differential brightnesses according to their intensities ("gray scale"), recognition is improved. This improvement is best under conditions of only one type of degradation operating in the presenter at one time (i. e., addition or omission), but there is some improvement even when both types of degradation are present simultaneously.

e. Simultaneous use of "gray scale" and increased matrix size afford the greatest improvement in recognition, especially where both types of degradation are present simultaneously.

f. Although stimulus type is relatively independent of the other variables already discussed, it is suggested that if coded material is to be utilized in the presenter, the code consist of discrete letters or meaningful letter groups rather than meaningless groups of letters (jumbles).

g. The efficiency of a presentation device is a function of the training level of the operators who must use the equipment, in that a simpler system (i.e., one permitting simultaneous occurrence of addition and omission of elements) will still serve effectively for highly competent operators, accustomed to interpreting "garbled" signal patterns. A related psychological principle in visual perception involves the "set" or expectancy of the receiving operator. If he is thoroughly familiar with the population of messages which may be presented to him, his ability to interpret seriously deteriorated signal patterns will be much greater. Thus the problems of communication equipment design and human operator proficiency cannot be divorced.

2. Implications for Further Psychological Research.

The following are problems requiring "human engineering" research for adequate answers:

a. Factors determining legibility of numbers (presumably the same as those effective for letters).

b. The variable of type of print (e.g., square vs. circular elements, size of elements relative to stimulus pattern, Roman vs. pica type; upper vs. lower case type).

c. The variable of contrast ratio between signal and background.

d. The variable of color of stimulus pattern and background.

e. Verification of the influence of the variables already investigated in the light of equipment specifications from the engineers as to such factors as distribution of noise, limits of signal interference, sensitivity of the system, etc.

f. Investigation of the dependence of stimulus pattern intelligibility upon operator familiarity with the stimulus vocabulary and practice level of viewing.

BIBLIOGRAPHY

1. Laschever, N. L. Survey of Techniques Applicable to "Infomax". Internal Report, Advanced Development Branch, Communication and Navigation Laboratory, Engineering Division, USAF Air Materiel Command, April 1950.

2. Air Force Technical Report 5922, Factors Determining Legibility of Letters and Words Derived from Elemental Printers ("Infomax" Principle), I, USAF Air Materiel Command, April 1951.

3. Air Force Technical Report 5923, Factors Determining the Legibility of Letters and Words Derived from Elemental Printers ("Infomax" Principle), II, USAF Air Materiel Command, May 1951.

AD-A800 032

UNCLASSIFIED

ATI 184 797 * 366750

University of Virginia, Charlottesville

FACTORS DETERMINING THE LEGIBILITY OF LETTERS AND WORDS DERIVED FROM ELEMENTAL PRINTERS - THE THIRD OF A SERIES OF REPORTS ON THE "INFOMAX" PRINCIPLE, by Eugene R. Long, Sterling L. Reid, and Ralph Queal. Aug '51, 37 pp. incl. tables, graphs, charts.

Published by WADC, Research Div., Wright-Patterson Air Force Base, O. (AF Technical Report No. 5924)

USAF Contr. No. W33(038)-ac-21269

Psychology (63)
General (0)

Data - Presentation
Reading
Human engineering

(Copies obtainable from ASTIA-DSC.)

UNCLASSIFIED