AD NUMBER: AD0160345
LIMITATION CHANGES
TO: Approved for public release; distribution is unlimited.
FROM: Distribution authorized to U.S. Government agencies and their contractors; Administrative/Operational Use; 7 Apr 1958. Other requests shall be referred to Army Medical Research Laboratory, Fort Knox, KY.
AUTHORITY
USAMRL ltr dtd 27 Feb 1970

BEST

AVAILABLE

COPY

UNCLASSIFIED

AD 160345

Armed Services Technical Information Agency

ARLINGTON HALL STATION
ARLINGTON 12 VIRGINIA

FOR MICRO-CARD CONTROL ONLY

1 OF 1

UNCLASSIFIED

US ARMY MEDICAL RESEARCH LABORATORY

FORT KNOX, KENTUCKY

REPORT NO. 328
7 April 1958

PERCEPTION OF RELATIVE FREQUENCY AS A FUNCTION OF THE NUMBER OF STIMULUS AND RESPONSE CATEGORIES*

*Subtask under Psychophysiological Studies, USAMRL Project No. 6-95-20-001, Subtask, Basic Problems of Vision and Perception in Human Behavior.

RESEARCH AND DEVELOPMENT DIVISION
OFFICE OF THE SURGEON GENERAL
DEPARTMENT OF THE ARMY

REPORT NO. 328

PERCEPTION OF RELATIVE FREQUENCY AS A FUNCTION OF THE NUMBER OF STIMULUS AND RESPONSE CATEGORIES*

by

R. A. Gardner
with the technical assistance of
J. B. Forsythe, R. E. Kling
and
R. Kalla

from

Experimental Psychology Department
US ARMY MEDICAL RESEARCH LABORATORY
Fort Knox, Kentucky

*Subtask under Psychophysiological Studies, USAMRL Project No. 6-95-20-001, Subtask, Basic Problems of Vision and Perception in Human Behavior.

Report No. 328
Project No. 6-95-20-001
Subtask USAMRL S-2
MEDEA

ABSTRACT

PERCEPTION OF RELATIVE FREQUENCY AS A FUNCTION OF THE NUMBER OF STIMULUS AND RESPONSE CATEGORIES

OBJECT

To study multiple-choice behavior by means of a prediction response in the case where relative frequency is the only basis for prediction.

SUMMARY AND CONCLUSIONS

A prediction response was used to study a form of decision behavior. The number of stimulus-event categories that could occur was varied from two to eight. Each stimulus-event category was associated with one of the available response categories. The only property of the stimulus-event that could be used for prediction was relative frequency. The relative frequency of response to the most frequent stimulus-event category was a function of the number of categories. Asymptotic response levels were reached if at all, only very late in the series of 420 trials. For the last 140 trials all groups with more than three categories were responding at frequencies significantly different from those of the presented stimulus-events.

RECOMMENDATIONS

That this research be continued to investigate the separate effects of stimulus-event category number and available response category number and their interaction.

Submitted 30 October 1957 by:
R. A. Gardner, 1st Lt, MSC
with the technical assistance of
J. B. Forsythe, Pfc
R. E. Kling, Sgt
R. Kalla, Sp 3

APPROVED: ARTHUR J. RIOPELLE
Head, Experimental Psychology Department

APPROVED: FLOYD A. ODELL
Technical Director of Research

APPROVED: JOSEPH R. BLAIR
Lt Colonel
Commanding

PERCEPTION OF RELATIVE FREQUENCY AS A FUNCTION OF THE NUMBER OF STIMULUS AND RESPONSE CATEGORIES

I. INTRODUCTION

This experiment is concerned with behavior as a function of the relative frequency of stimulus-events in a series. Specifically, the response consists of successive predictions of stimulus-events when a relative frequency is the only basis for predictions. The condition where two categories of response are available has been extensively studied (2, 4, 5, and 7). The data of these studies indicate that the relative frequencies of response in such situations are a direct function of the relative frequencies of the stimulus-event categories. Furthermore, many authors have concluded that the relative frequencies of response are asymptotic at levels which tend to match the presented frequencies. This last phenomenon will be referred to here as matching behavior.

As yet, conditions in which more than two response categories are available to S have received little experimental attention. The present experiment was designed to study the average relative frequency of response to the most frequent stimulus-event category as a function of the number of choices; where a choice is defined as a discrete response category which consists of the prediction of a discrete stimulus-event category.

II. EXPERIMENTAL

A. Apparatus

The basic unit of apparatus consisted of an aluminum panel, 24 x 12 in, set in a three-sided wooden cabinet at an angle of 30° to the base of the cabinet. The base edge of the panel was 32 in from the floor. Set in the aluminum panel at 1.5-in intervals in an arc of radius 12-in were 0.25-in telephone jacks. The number of exposed jacks varied with the conditions of the experiment. All jack holes that were not used during a condition of the experiment were firmly sealed with brass cap nuts. Each exposed jack was labeled with 0.75-in high block letter of the alphabet. Anchored to the cabinet was a 0.25-in telephone plug. Twelve such cabinets were used in each experimental session. They were arranged in two rows of six cabinets with 9 ft between rows and 2 ft between adjacent cabinets. The stimulus categories consisted of 3 x 5 ft block letters corresponding to those labeling the exposed jacks. These were projected individually on a screen placed 6 ft in front of the center of the first row of cabinets.

Responses were recorded electrically on Gorrel and Gorrel Electrex charting paper. The projector and the recording unit were situated 3 ft behind the second row of cabinets. The experimental room was air conditioned, partially sound-proofed, and dimly lighted by two, 100-watt lamps reflected off the walls of the room. The lamps were located halfway between the two rows at each end of the aisle between them.

B. Procedure

The block letters were presented in series obtained from a table of random numbers. The series was 140 trials long. The relative frequency of each letter was fixed for each 140 trials by the design of the experiment for that condition. The relative frequencies for groups of trials shorter than 140 were permitted to vary randomly. The sequence used for the first 140 trials was reversed for the second 140 trials and then presented forward again for the third 140 trials, making 420 trials in all.

The Ss were seated before the cabinets facing the projection screen in groups of 12. No S served in more than one condition. Each condition was administered at two sessions so that N for each condition was 24. Half of those receiving a given condition received it before noon and half after noon. They were instructed to predict on each trial the letter which would appear on the screen next. Instructions urged them to make as many correct predictions as possible, and to make their choices as rapidly as possible. The darkening of the screen between presentations indicated the beginning of each trial. Each S then indicated his choice by inserting his telephone plug into the jack labeled with the appropriate letter. Each S held his response plug in his right hand. The telephone plug and jack arrangement prevented S from making more than one choice per trial. With the appearance of the next letter on the screen, each S withdrew his plug and prepared for the beginning of the next trial. A bank of 12 lamps located near the recording apparatus indicated to E when each S had made a choice and when each S had withdrawn his plug.

An attempt was made to maintain a uniform pace of 4 sec per trial, and such a pace was well within the capacity of the individual S. In practice this was not possible for groups of 12, as it was necessary for all 12 plugs to be inserted together before each presentation, and for all to be withdrawn before a new trial could begin. The bulk of the time losses occurred during the very early trials during which E was setting the pace. Most groups fell into the 4-sec rhythm very soon and approximated it very closely during the greater part of each experimental session.

C. Design

The basic design consisted of presenting one stimulus-event category, A, for a fixed proportion of the trials while varying the number of alternative categories. For each of these, one response category was available to S and suitably labeled on the response panel. Two A proportions were employed, .70 and .60. In Experiment I the non-A trials were divided equally among the other stimulus-event categories. There were four .70 conditions and four .60 conditions. For the .70 groups these were two, three, four, or seven choices. For the .60 groups these were two, three, five, or eight choices.

In Experiment II there were two groups under three-choice conditions and these had unequal proportions for all three stimulus-event categories. For one of these groups, the presented proportions were .70, .20, and .10. For the other group they were .60, .30, and .10. For convenience these groups will be referred to as the 70-20-10 and 60-30-10 groups, respectively.

D. Subjects

The Ss were 240 young men engaged in basic infantry training at the US Army Training Center, Armor, Fort Knox, Kentucky. Their ages ranged from 17 to 26 years, with 58% reporting themselves as just 22, and 10% as older than 22. The reported education ranged from five years of elementary school to four years of college, with 46% reporting themselves as completing four years of high school only, and 10% as completing at least one year of college. Group to group variation in age or education was slight.

E. Counterbalancing

The letters which served as stimulus-event categories were partially counterbalanced in the following manner. In Experiment I the letter L served as category A and the letter S as one of the non-A categories for half of the .70 groups, while L and S were interchanged for the other half. The letters T and H were alternated in the same manner for the .60 groups. In Experiment II, L and S were alternated as the .70 and .20 categories for the 70-20-10 groups, while the letter N served as the .10 category for both halves of the group. The letters

T and H were alternated as the .60 and .30 categories for the 60-30-10 group and the letter D served as the .10 category for both halves of this group.

The spatial arrangement of the jacks on the response panels was partially counterbalanced in the following manner. Imagine that the positions of nine jacks are numbered from left to right; 8, 6, 4, 2, 1, 3, 5, 7, and 9 with position 1 at the center of the arc. For the .70 groups, L, S, and N occupied the three center positions; 1, 2, and 3, and each of their six possible spatial arrangements appeared on an equal number of response panels. For the .60 groups, T, H, and D replaced L, S, and N, respectively. For three-choice groups, only these three center positions were employed. For two-choice groups, only the L and S or T and H jacks were employed. For half of the four-choice .70 group, D appeared in position 4, and for the other half in position 5. For half of the seven-choice .70 group, D, H, T, and B appeared in positions 4, 5, 6, and 7, respectively, and for the other half in positions 5, 4, 7, and 6, respectively. For the five-choice .60 group, N and L were alternated in positions 4 and 5. For half of the eight-choice .60 group, N, L, S, B, and F appeared in positions 4, 5, 6, 7, and 8, respectively, and for the other half in positions 5, 4, 7, 6, and 9, respectively.

III. RESULTS

A. Experiment I

For each S the number of choices of the most frequently presented stimulus-event category, A, was tabulated in blocks of 20 trials. The means for each condition were converted to fractions representing the average relative frequency of response to category A. These values will be referred to as P(A). Figure 1 represents P(A) for the .70 groups in blocks of 20 trials and Figure 2 represents the same information for the .60 groups. The relative frequencies demanded by the design were only maintained precisely for blocks of 140 trials. For this reason the data were pooled in three periods of 140 trials each for further analysis. Data for the .70 groups for the three periods are presented in Figure 3 and Table 1 and for the .60 groups in Figure 4 and Table 2. The first block of 20 trials has been omitted from the means for the first period because S tends to distribute his initial responses equally among the available choices, regardless of the conditions. This may be seen for block 1 in Figures 1 and 2. Thus, the trend of these first trials is opposite to that of all subsequent trials.

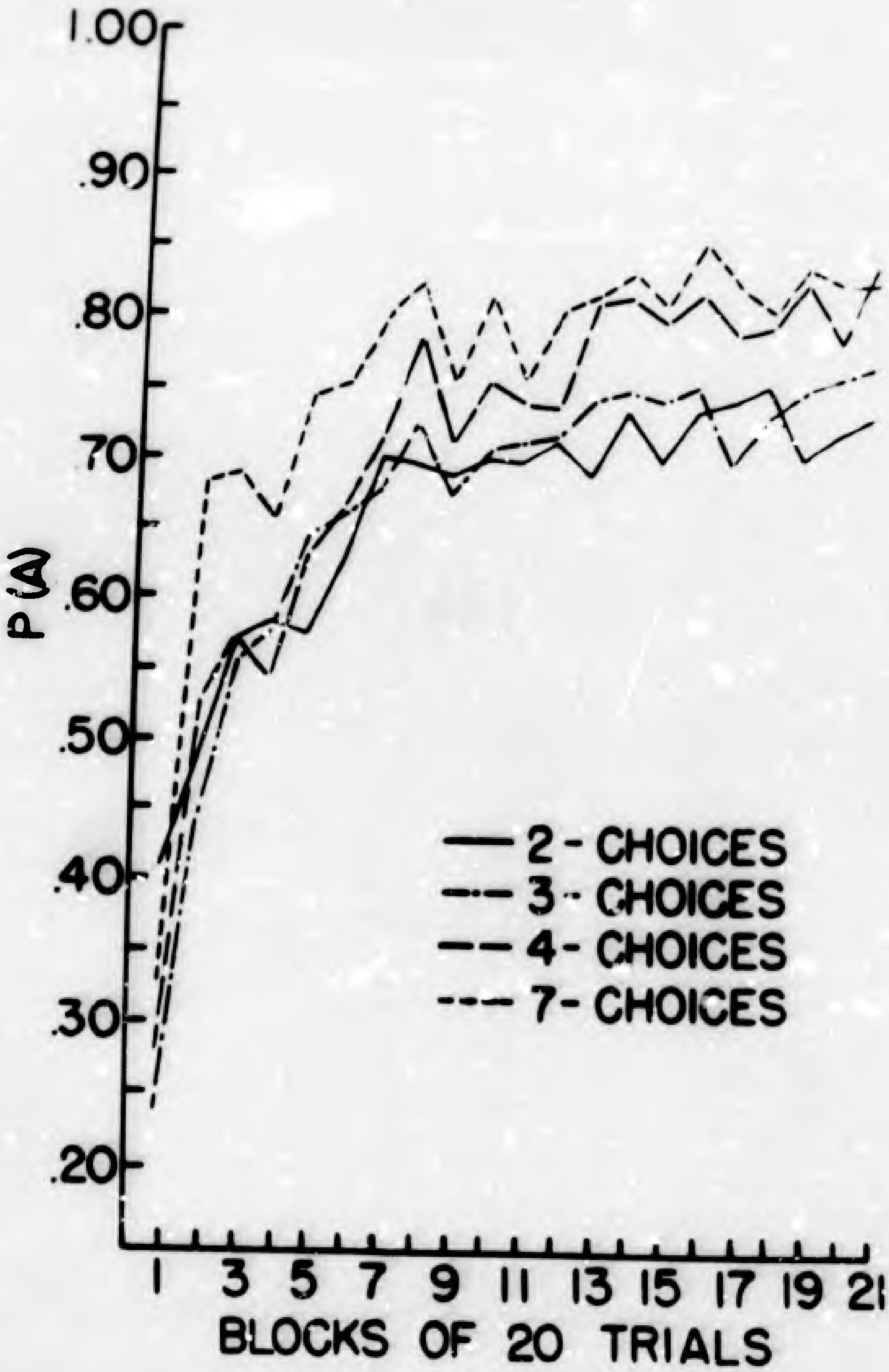

Fig. 1. P(A) scores of , 70 groups for successive blocks of 20 trials.

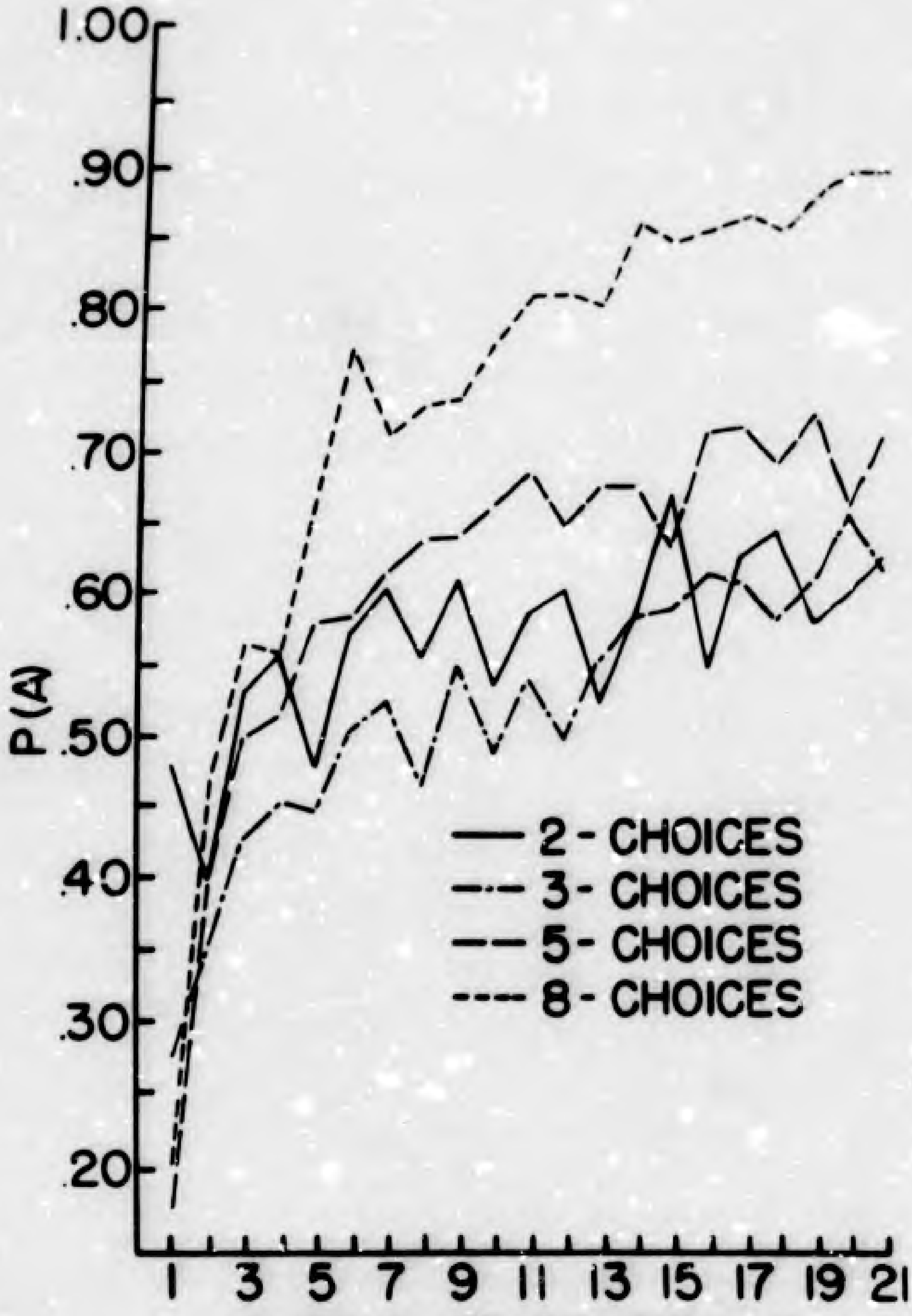

Fig. 2. P(A) scores of .60 groups for successive blocks of 20 trials.

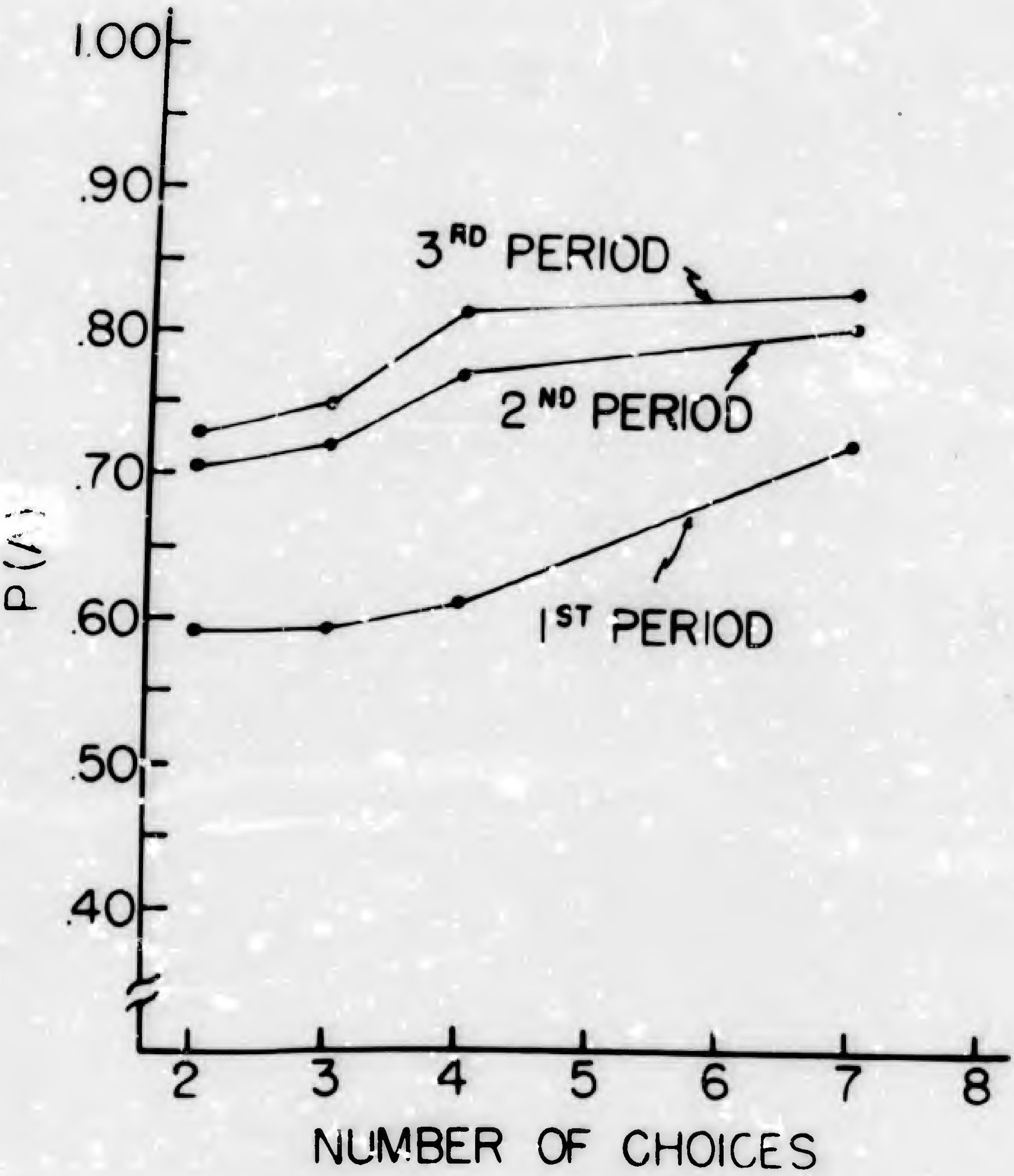

Fig. 3. P(A) scores of .70 groups for the three periods of Experiment I.

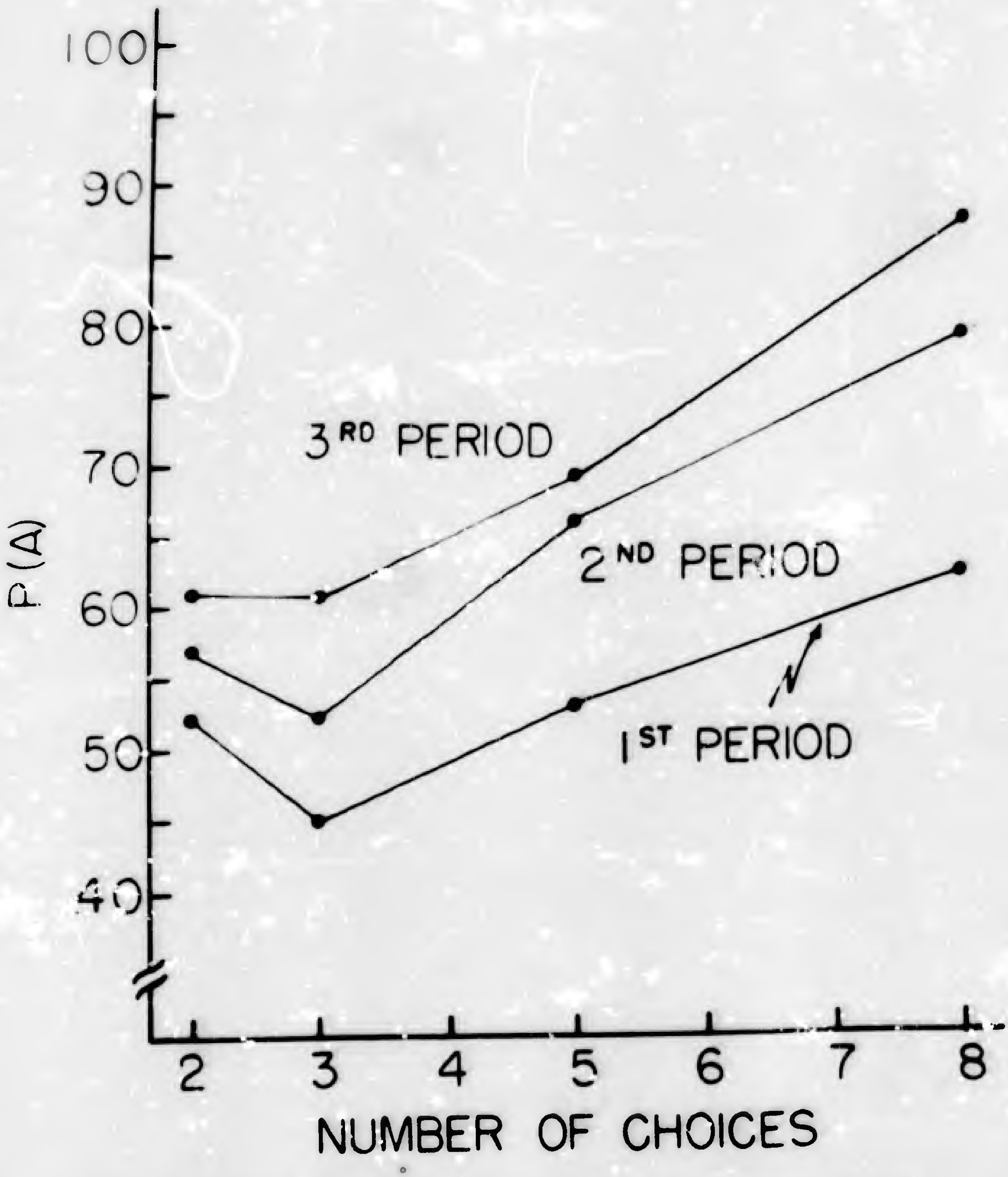

Fig. 4. P(A) scores of .60 groups for the three periods of Experiment I.

TABLE 1

P(A) FOR THE .70 GROUPS OF EXPERIMENT I

		Number of Choices				
Period	Trials	2	3	4	7	F
1	21 - 140	.592	.593	.609	.721	4.79**
2	141 - 280	.705	.719	.767	.803	2.35
3	281 - 420	.728	.746	.811	.829	3.04*
t (Per. 2 vs Per. 3)		2.09*	2.20*	2.74*	2.15*	

* $p < .05$

** $p < .005$

TABLE 2

P(A) FOR THE .60 GROUPS OF EXPERIMENT I

		Number of Choices				
Period	Trials	2	3	5	8	F
1	21 - 140	.523	.452	.532	.621	5.20**
2	141 - 280	.570	.524	.660	.789	15.46**
3	281 - 420	.611	.608	.692	.871	20.70**
t (Per. 2 vs Per. 3)		4.53**	3.96**	2.37*	4.25**	

* $p < .025$

** $p < .005$

A separate analysis of variance was computed for each period for each of the two presented proportions. Each of the six analyses which resulted was based on a 4 x 2 factorial analysis of variance with the number of choices and the letters employed as category A (L vs S for the .70 groups and T vs H for the .60 groups) as variables. The effects of the different letters were evaluated as follows. For the .70 groups, all of the Fs calculated for L vs S or their interaction with the number of choices variable were less than 1.0. For the .60 groups, all Fs for T vs H were less than 1.0 and all of the interactions were less than 2.0. Thus, the effects of this variable appear to have been negligible and linearly related to the number of choices.

The Fs computed for number of choices appear in Tables 1 and 2. The null hypothesis of Bartlett's test for heterogeneity of variance was untenable with $P < .05$ for the first period of the .60 groups and the second period of the .70 groups. However, F for the first of these was high enough and for the second, low enough, not to require an alteration in decision as to the reliability of the differences among means (6, p. 78-88).

The rise between the second period, trials 141 to 280, and the third period, trials 281 to 420, was evaluated as follows. For each S a difference score was obtained by subtracting his P(A) for the second period from his P(A) for the third period. Separate ts were computed for each group and these appear in Tables 1 and 2.

All eight groups reached or exceeded matching behavior by the third period. For the two- and three-choice .70 groups, P(A) slightly exceeded .70 by the last period, but neither t was significant. For the four- and seven-choice .70 groups, P(A) was significantly higher than .70 for both the second and third periods with $P < .02$. The two- and three-choice .60 groups barely reached P(A) scores of .60 by the last period with $t < 1.0$ in both cases. For the five-choice .60 group, P(A) was significantly greater than .60 by the last period with $P < .02$. For the eight-choice .60 group, P(A) was significantly greater than .60 with $P < .001$ for both the second and third periods.

In a previous article (3) the hypothesis was advanced that the matching behavior of the two-choice groups represents a special case so that P(A) for all multiple-choice groups should exceed that for the two-choice groups. When the P(A) scores of the two-choice groups were compared with those of the corresponding three-choice groups, t values were less than 1.0 even during the last period. However, the P(A) score for each group having more than three choices was greater than that of the corresponding two-choice group during both the second and third periods with $P < .05$ with one exception. This was the five-choice .60 group during the second period (t for this comparison was 1.92). In the cases of the seven-choice .70 group and the five-choice .60 group the assumption that variances were equal to those of corresponding two-choice groups was not tenable at the .05 level of confidence. The P values for these comparisons were arrived at as suggested by Lindquist (7, p. 98) with 23 degrees of freedom rather than 46. It should be noted also that each P value quoted here refers to a two-tailed test.

B. Experiment II

For the 70-20-10 group of Experiment II, P(A) for periods 1, 2, and 3 was .612, .744, and .758. While these values are consistently higher than those shown in Table 1 for the three-choice .70 group of

Experiment I, all three t values were less than 1.0. When compared with a hypothetical $\bar{P}(\underline{A})$ of .70, for the second period $t = 1.95$ ($P > .05$) and for the third period $t = 2.47$ ($P < .025$). When compared with $\bar{P}(\underline{A})$ for the two-choice .70 group of Experiment I, t did not approach significance for any period. The rise from the second to the third period was not significant. For the 60-30-10 group of Experiment II, $\bar{P}(\underline{A})$ for periods 1, 2, and 3 were .509, .639, and .665. While these values are consistently higher than those shown in Table 2 for the three-choice .60 group of Experiment I, only that for the second period is significant with $t = 2.90$, $P < .01$. When these values were compared with a hypothetical $\bar{P}(\underline{A})$ of .60 for the last two periods or with $\bar{P}(\underline{A})$ for the two-choice .60 group of Experiment I, no t was significant at the .05 level of confidence. For the 60-30-10 group also, the rise from the second to the third period was not significant.

IV. DISCUSSION

These data seem to indicate that response frequencies in this situation are a function of the number of choices. After a sufficient number of trials, all groups under conditions of more than three choices responded to the most frequent stimulus-event category with a relative frequency higher than the presented frequency. Whether or not the terminal values found in this experiment do in fact represent asymptotic performance cannot be determined in a precise manner. When $\bar{P}(\underline{A})$ for the second block of 140 trials is compared with $\bar{P}(\underline{A})$ for the third block of 140 trials, all ten groups show increments and these increments are significant beyond the .05 level for eight of the ten groups. Comparisons between adjacent blocks of fewer trials were not made because frequencies were not fixed for blocks of less than 140 trials. Thus, fluctuations between adjacent blocks of fewer trials were apt to reflect random fluctuations in the presented frequencies. Even if this were not the case, it would be difficult to determine asymptotes because the increments in $\bar{P}(\underline{A})$ were very small. The mean rise from the second to the third block of 140 trials was only .03 for the .70 groups of Experiment I and .06 for the .60 groups of Experiment I. The error variance for 140-trial blocks was small enough to demonstrate the reliability of this rise, but this might not be the case with the error variance for blocks of fewer trials. The present experiment was not designed to yield a direct answer to this question. Such evidence as it provides indicates that $\bar{P}(\underline{A})$ continued to rise for a considerable number of trials, possibly for the whole series of 420 trials.

The matching behavior commonly found with two choices has attracted a certain amount of theoretical interest. For the most part, theoretical treatments have either assumed that the terminal $\bar{P}(\underline{A})$ will be independent of the number of alternative responses, or have implied this by omission of that variable when defining the problem. Grant, Hake, and Hornseth (4) have shown that in the two-choice situation the proportions of negative feedback associated with the two alternatives will be equal to each other in matching behavior. In a previous article (3) the present author has shown that, if we assume that S will attempt to maintain a similar equilibrium among the different sources of negative feedback, than multiple-choice situations should not yield matching behavior. Instead, if the relative frequency of category A is fixed, P(A) should rise as a function of the number of choices. The derivation of an algebraic model from this is given in detail elsewhere (3). This model gives the average probability of an A -response, P(A), from the following formula:

$$P(A) = 1 - \frac{n(1 - A)}{A(n^2 - 1) + 1}$$

where A is the relative frequency of the most frequent category and n is the number of non-A categories in the case where all of the non-A category events occur with equal frequency. It should be noted that the number of choices as defined by the operations of this experiment is equal to n + 1.

With A = .70, and two, three, four, and seven choices, this formula predicts that P(A) will be .70, .81, .86, and .93, respectively. With A = .60, and two, three, five, and eight choices, it predicts P(A) scores of .60, .71, .84, and .91, respectively. Except for the two-choice groups, these predictions overestimate the data considerably when the unit of comparison is a 140-trial block. In addition, this negative feedback model differentiates between combinations of stimulus-event frequencies when the non-A frequencies are unequal. Specifically, for 70-20-10 conditions and 60-30-10 conditions, it predicts that the average frequency of response to the .10 categories will approach zero. The groups of Experiment II, test this prediction. They responded to the .10 category with relative frequencies slightly higher than .10 even during the last 140 trials. ($t < 1.0$ for both cases). Thus, the data of this experiment do not seem to confirm the predictions of this model. Nevertheless, the notion that the relative distribution of responses in this situation is governed by relative amounts of negative feedback might still furnish a solution, albeit with some more fortunate mathematical expression.

Schipper (7) presents a possible alternative notion. He points out that, as the number of non-A alternatives increases, each non-A proportion must become smaller and smaller. If we assume that P(A) is governed by some comparison between the A-proportion and the average proportion of the alternatives, then we would expect it to increase with increasing numbers of non-A stimulus-event categories. Stated at this level of precision, Schipper's hypothesis does agree with the rank order of the groups of this experiment. But, at this level of precision, the negative feedback hypothesis outlined above also agrees with the rank order of this experiment. The procedure employed here does not permit us to differentiate between the two hypotheses. In his study (7), Schipper presents data obtained under a procedure which might be expected to separate the two. Unfortunately, he ran his groups for only 120 trials, which was, perhaps, too soon to expect statistically reliable differences to appear.

The paucity of theoretical attention to the major variable of this experiment and the lack of appreciation for the number of trials over which P(A) may be expected to rise, are reflected in the paucity of comparable data. In a previous study (3), data were presented for trials 286 to 450 for two- and three-choice conditions with .70 and .60 proportions. The procedure was comparable with certain exceptions which will be noted presently. The P(A) values for the two-choice groups were not significantly different from the presented proportions. With three choices and equal non-A frequencies, P(A) was .802 for the .70 group and .676 for the .60 group. Both were significantly higher than the presented proportions and the P(A) for corresponding two-choice groups. Cotton and Rechtschaffen (1) were able to replicate six conditions of that study and found no discrepancies greater than .026 and no significant discrepancies. Only the two-choice groups of the Cotton and Rechtshaffen study failed to reach a P(A) significantly higher than presented proportions. Their P(A) values for three-choice .70 and .60 groups with equal non-A frequencies were .805 and .660, respectively. The procedure of the replication study was nearly identical to that of the original. Noting the high degree of agreement between the two earlier studies and the disagreement of both with the three-choice data of the experiment, we might expect to discover the source of this disagreement in procedural differences. There were three major differences in technique: 1) Subject populations; the earlier studies used university undergraduates of both sexes, the present study used basic trainees in the U.S. Army; 2) Individual vs. group techniques; the earlier studies tested Ss individually, the present study tested them in groups of 12; 3) Stimulus differentiation; the stimulus-event categories of the earlier studies consisted of three identical signal lamps. The present study employed different letters of the alphabet.

There are some data relevant to the third point in Experiment III of the Gardner study. In that part of the study, stimulus differentiation (different colored vs. identical lamps) and intertrial interval (4-sec vs.

10-sec) were manipulated in a 2 x 2 factorial design for 70-20-10 proportions. A significant interaction effect was obtained when P(A) for a group with different colored lights and a 4-sec intertrial interval fell to .755. In the present experiment the 70-20-10 group with different letters of the alphabet and a 4-sec intertrial interval reached a P(A) of .758. The similarity is certainly striking. If it could be reproduced with the remaining procedural differences controlled, this would suggest that P(A) is sensitive to variations in stimulus differentiation and intertrial interval. A confirmation of this effect could open up a fruitful means of studying these two significant behavioral variables.

V. SUMMARY AND CONCLUSIONS

A prediction response was used to study a form of decision behavior. Categories of stimulus-events represented by block letters of the alphabet were presented in a series of 420 trials. The number of prediction responses which were available was the same as the number of stimulus-event categories and this number ranged from two to eight. The only property of the stimulus-event series that could be used for prediction was relative frequency. The relative frequency of the most frequent stimulus-event category was fixed at either .70 or .60. It was found that the relative frequency of response to the most frequently presented category was a function of the number of categories. With more than three choices the relative response frequencies were significantly different from the presented stimulus-event frequencies. The data indicated that asymptotic levels of response, if they do exist, occur after a much greater number of trials than that assumed in the design of most work in this area.

VI. RECOMMENDATIONS

That this research be continued to investigate the separate effects of stimulus-event category number and available response category number and their interactions.

VII. REFERENCES

1. Cotton J. W. and A. Rechtschaffen. Replication report on two and three choice verbal conditioning phenomena. J. Exper. Psychol. In press.

2. Estes, W. K. and J. H. Straughan. Analysis of a verbal conditioning situation in terms of statistical learning theory. J. Exper. Psychol., 47: 225-234, 1954.

3. Gardner, R. A. Probability-learning with two and three choices. Amer. J. Psychol., 70: 174-185, 1957.

4. Grant, D. A., H. W. Hake, and J. P. Hornseth. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. J. Exper. Psychol., 42: 1-5, 1951.

5. Hake, H. W., and R. Hyman. Perception of the statistical structure of a random series of binary symbols. J. Exper. Psychol., 45: 64-74, 1953.

6. Lindquist, E. F. Design and analysis of experiments in psychology and education. Boston: Houghton Mifflin, 1953.

7. Schipper, L. M. Prediction of critical events in context of different numbers of alternative events. J. Exper. Psychol, 52: 377-380, 1956.

UNCLASSIFIED