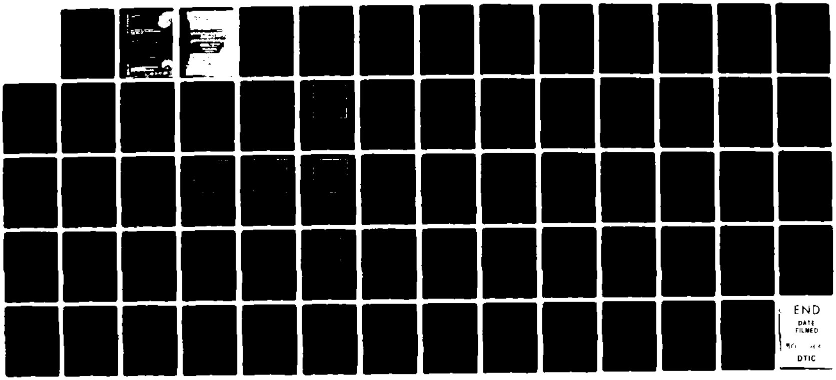
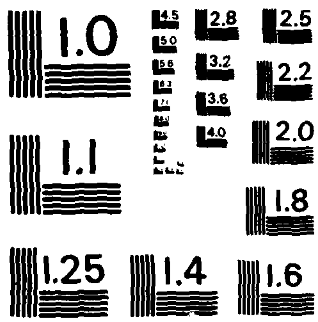
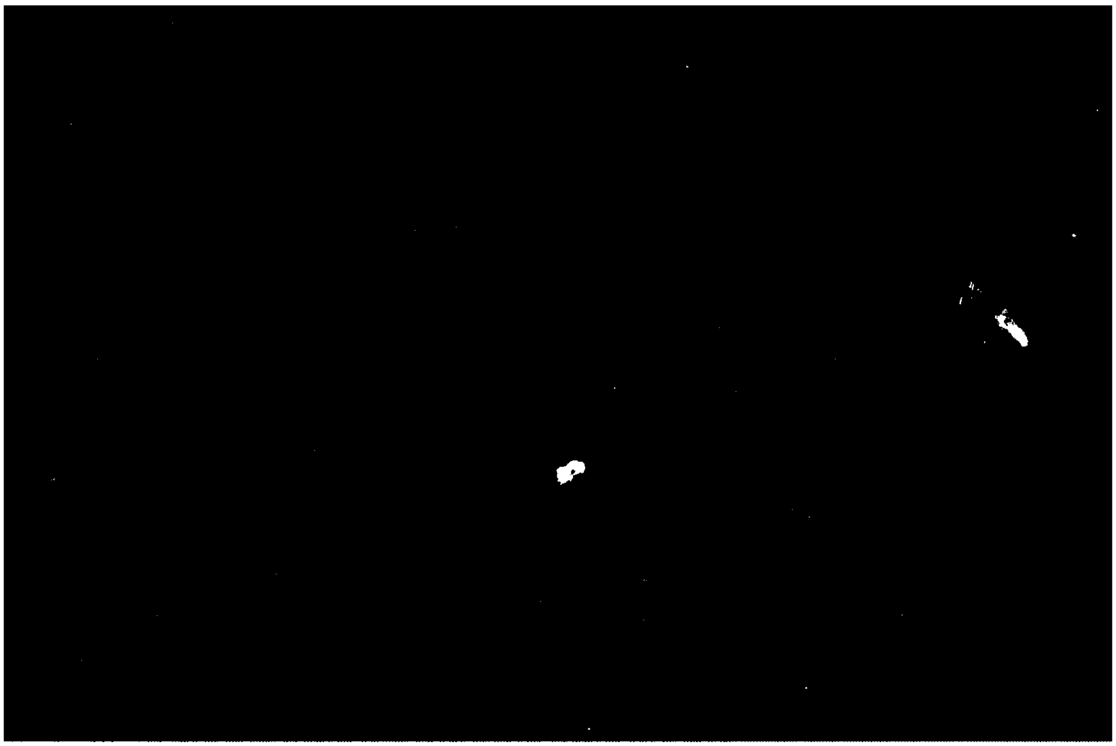
AD-A133 394 MAZE-SOLVING AS A PERFORMANCE MEASUREMENT TOOL FOR HUMAN OPERATIONS UNDER..(U) AIR FORCE AEROSPACE MEDICAL RESEARCH LAB WRIGHT-PATTERSON AFB.. S L WARD ET AL. 1/1

UNCLASSIFIED AUG 83 AFAMRL-TR-83-052 F/G 5/10 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS - 1963 - A



SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFAMRL-TR-83-052	2. GOVT ACCESSION NO. AD-A133394
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) MAZE-SOLVING AS A PERFORMANCE MEASUREMENT TOOL FOR HUMAN OPERATORS UNDER TIME-STRESS	5. TYPE OF REPORT & PERIOD COVERED Technical Report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Sharon L. Ward Richard J. Poturalski, Capt, USAF	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS AFAMRL, Human Engineering Division, AMD, AFSC, Wright-Patterson AFB OH 45433	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62202F, 7184-27-02
11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE August 1983
	13. NUMBER OF PAGES 67
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

AFAMRL Primary Investigator: Capt Richard J. Poturalski, AFAMRL/HEC,
Tele: (513) 255-3438

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Performance Measurement
Maze-solving
Time-stress
Planning Task
Cognitive Task

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report documents a series of experiments to investigate the usefulness of maze-solving as a performance measurement tool for humans under time-stressed conditions. Three experiments were carried out: the first to establish an appropriate difficulty level for the mazes, the second to test the effect of repeated exposures to the same set of mazes on performance, and the third to measure the effect of time-stress on maze-solving performance. Subjects were asked to solve a maze by guiding a small dot from the beginning of the maze to the goal. Subjects had control over the dot direction but not

Cont'd

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) over its speed. Stress was induced by increasing the dot speed. Results from the first experiment indicated that the largest and most complex maze configuration considered showed the most reliable difference between dot speeds, and this configuration was used in both of the subsequent experiments. The second experiment demonstrated that the same set of mazes could be presented to subjects four or five times with no appreciable learning effects. The third and last experiment quantified the differences in performance due to dot speed. There were significant and reliable differences in score among dot speeds. Errors made in solving mazes were examined qualitatively as well, and the errors found were symptomatic of a shortened planning horizon: incorrect turns at early critical decision points, and attempts to return to the centerline of the maze prematurely.

SUMMARY

This report documents a series of experiments to investigate the usefulness of maze-solving as a performance measurement tool for humans under time-stressed conditions. Maze-solving was considered as a candidate task because it seemed particularly suitable for assessing the effect of stress on planning abilities.

Three experiments were carried out: the first to establish an appropriate difficulty level for the mazes, the second to test the effect of repeated exposures to the same set of mazes on performance, and the third to measure the effect of time-stress on maze-solving performance.

The mazes that were used were all of the same basic structure: square mazes whose rows and columns were defined by a lattice of small squares. A pre-determined number of barriers connected these small squares in a pseudorandom fashion. Mazes were shown one at a time on a CRT and subjects were asked to solve them by guiding a small dot from the origin, at the left center intersection, to the goal at the right center intersection. Subjects had control over the dot direction but not over its speed. Stress was induced by increasing the dot speed. At the end of each trial, the subject was shown his/her score, which was defined as the ratio of the shortest possible time for solution to the actual elapsed time.

Results from the first experiment indicated that the largest and most complex maze configuration considered showed the most reliable difference between dot speeds, and this configuration was used in both of the subsequent experiments. The second experiment demonstrated that the same set of mazes could be presented to subjects four or five times with no appreciable learning effects. The third and last experiment quantified the differences in performance due to dot speed. There were significant and reliable differences in score among dot speeds. Errors made in solving mazes were examined qualitatively as well, and the errors found were symptomatic of a shortened planning horizon: incorrect turns at early critical decision points and attempts to return to the centerline of the maze prematurely.

Item analysis was performed for all of the mazes for which that was possible, and recommendations for selection of mazes in future test batteries were made. Individual mazes for which item analysis results are available are shown in Appendices A and B.

PREFACE

This report documents a series of in-house experiments conducted under the C^3 Operator Performance Engineering (COPE) program at the Human Engineering Division of the Air Force Aerospace Medical Research Laboratory (AFAMRL).

The authors wish to extend their appreciation to TSgt Danny Bridges of AFAMRL for software support and facility coordination, to TSgts Ed Skuya and Dan Hughes of AFAMRL for software development, and to Ms. Suzanne Kelly of AFAMRL for data analysis.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

SECTION 1

INTRODUCTION

This study examined maze-solving as a candidate performance measurement tool for assessing the effect of stress. Maze-solving was chosen because we believed it would be sensitive to changes in cognitive abilities brought on by stress. The particular form of stress investigated in this study was time stress, induced by increasing the tempo of the task.

There are many instances in which a reliable and sensitive measure of human problem solving performance is useful. A task which is sensitive to various kinds of stress on the human subject and which captures some aspects of cognition and planning enables a more accurate and full assessment of the impact of environmental stress on human performance.

Tracking tasks have often been used to assess performance. While tracking tasks are useful, they chiefly measure motor skills rather than cognitive and planning abilities. An alternate performance task which provides a sensitive measure of these latter abilities would be a real asset. The maze-solving task appears to have some advantages over tracking tasks in measuring cognitive and planning performance. It is more complex, it avoids using the human as a nullmeter, and it incorporates aspects of problem-solving.

A search of the literature showed a remarkable lack of research using mazes for human subject experiments. The one notable exception to this (Porteus, 1965) used mazes to test intelligence and personality traits. The techniques were sufficiently different from our methods that the mazes and method of scoring did not transfer to this application. One result from this previous work which was of some interest, however, was that the maze test showed a difference between pre- and post-lobotomy patients, whereas other standard intelligence tests did not. There was, therefore, some previous indication that the maze solving concept investigated here would be sensitive to some forms of physiological and psychological stress.

Some results from this study have been published elsewhere (Poturalski and Ward, 1982; Ward and Poturalski, 1982). Those papers emphasize, respectively, the reliability and sensitivity to time stress obtained with this task, and the nature of the errors made under time stress. This report includes details from those two papers as well as an analysis of performance measurement characteristics for individual mazes.

SECTION 2

GENERAL METHODS

This study was carried out in a series of three experiments. In each experiment, the subjects were presented with a battery of mazes to solve. Mazes were presented one at a time on a CRT display, and subjects were required to solve the maze by guiding a small dot from the beginning of the maze to the goal. For all presentations, subjects were not given any time to preinspect the mazes. Subjects had control over the dot direction, but not over its speed. Stress was introduced by increasing dot speed.

The first experiment determined the maze size and number of barriers which would maximize the sensitivity and reliability of the task. The second quantified any learning effect of repeated exposures to the same maze. The third and last part investigated the effect of time-stress on performance.

The mazes used in this study all had the same general structure: square mazes with a certain number of rows and columns and a certain number of barriers in the maze. A typical maze with ten rows and columns and 95 barriers is shown in Figure 1.

Mazes were generated by a computer algorithm in which a predetermined number of barriers were placed at random throughout the maze. Positions of barriers were then altered until there was at least one path from each intersection of the maze to any other intersection. The alterations guaranteed that (1) a solution path did exist, and (2) no portion of the maze was completely blocked off from the remainder.

After a battery of mazes had been generated, it was further screened to insure that there were no duplicates and to eliminate any mazes for which the optimum solution path was unusually long or unusually short. For the second and third parts of the study, additional maze characteristics were used to select mazes for presentation. These additional characteristics will be described later.

In some preliminary investigations, the maze generation procedure was slightly different from that described above. Initially, the number of barriers *per row and column* was fixed rather than the total number of barriers *per maze* being fixed. We found, however, that the latter method of generation produced mazes that were equally or slightly more difficult than our initial method. Since the less constrained method of fixing only the total number of barriers had the additional advantage of producing a richer variety of mazes, we used this generation method throughout.

We restricted our attention to mazes that had either eight or ten rows and columns. We found that smaller mazes could almost always be solved at a glance, and felt that larger mazes might be too complex for use in stressful environments. We considered only mazes with an even number of rows and columns so that the starting point and goal could be placed symmetrically on each side. The number of barriers used for the 8 x 8 mazes were 54 or 60 (38% or 42% of available barrier positions), and for 10 x 10 mazes 88 or 95 (40% or 43% of available positions). We found that it was extremely difficult to obtain viable mazes if the number of barriers approached 50% of the available positions and that mazes with barriers in fewer than 35% of the available positions were very easy to solve. Thus, the range of maze complexity investigated in these experiments included all mazes of this sort that were both feasible and complex enough to be of interest.

Maze presentation and data collection were done using a PDP 11/34 with a VT-11 CRT display. Subject control of dot direction was via a joystick with a four-way trim switch. The trim switch positions corresponded to dot direction changes: up, down, left, and right. Other controls on the joystick were not activated. Subjects were required to make their input during the time the dot was within an intersection. Trim switch movements initiated before the dot reached an intersection were ignored by the computer. The data recorded for each trial included all subject inputs, dot direction changes, and the times of their occurrence as well as summary measures. Times were accurate to within one sixtieth of a second.

The actual displayed size of the maze on the CRT screen was 6.4 inches square for the 8 x 8 mazes and 7.9 inches square for the 10 x 10 mazes. The length of each small hallway, from the end of one intersection to the beginning of the next, was twice the hallway's width. For the first experiment, subjects were instructed to seat themselves at whatever viewing distance they preferred. Most subjects used a viewing distance of 2.5 to 3 feet. A few subjects, however, placed themselves very close to the screen and traced a solution path with their fingers on the screen while simultaneously guiding the dot. To control this source of variability among subjects, the viewing distance was fixed at approximately 3 feet for the second and third experiments.

Stress was introduced by varying the speed of the dot the subject was controlling. There were two qualitatively different dot speed conditions: a "stopping dot" and a "moving dot" condition. For the stopping dot condition, the dot moved along a path at a fixed velocity until it arrived at a decision point. It then stopped until

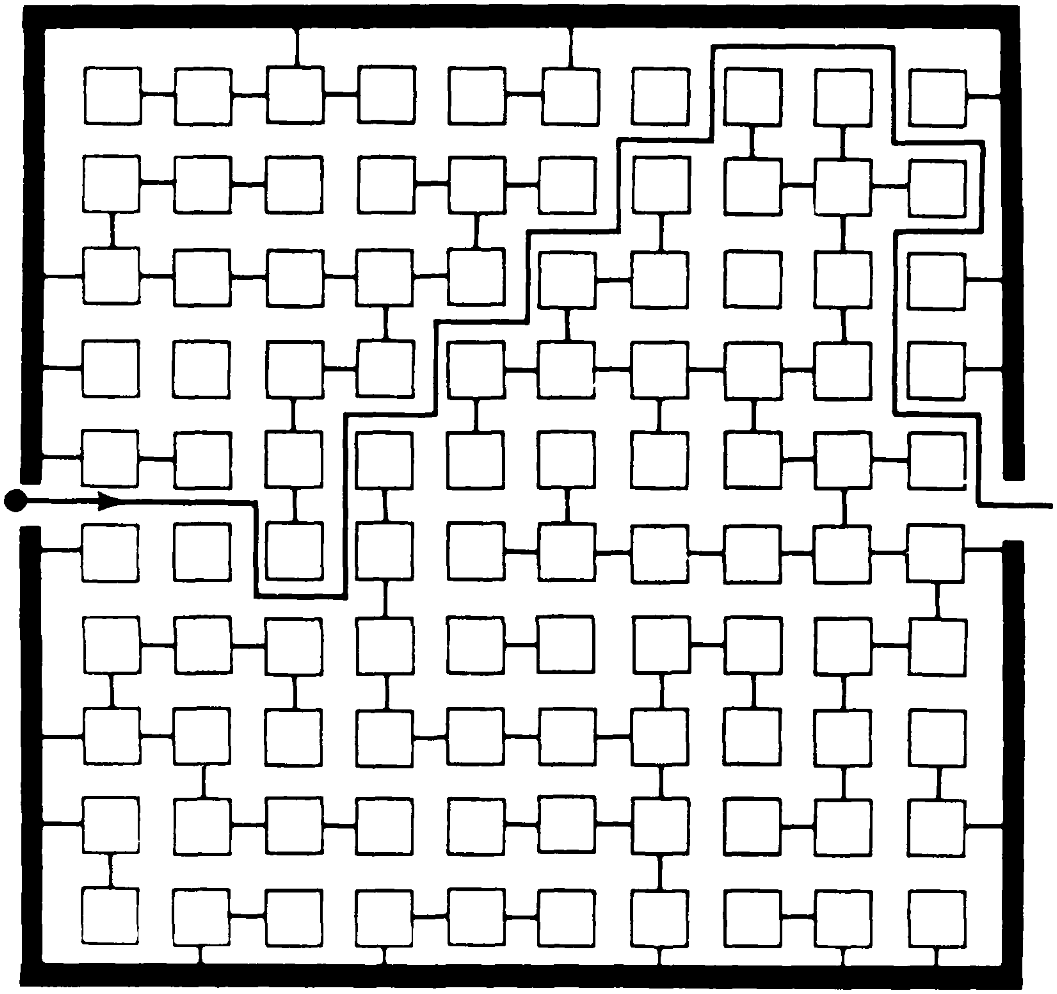


Figure 1. SAMPLE 10x10 MAZE WITH 95 BARRIERS. OPTIMUM PATH SHOWN.

the subject indicated the chosen direction of movement. For the moving dot conditions, the dot continued to move in a straight path until the subject indicated a change in direction or it passed out of the intersection. If the dot hit a barrier, it reversed direction. For the moving dot condition, the range of speeds investigated was 1.0 to 5.0 intersections per second. For the stopping dot condition, dot speed was 1.33 intersections per second. The dot always moved at a constant speed during any given trial.

Subjects were shown a score at the end of each trial which was proportional to the ratio of the shortest possible time to complete the maze to the actual time. Possible scores ranged from 0 to 100, and 100 was a perfect score. Average scores for each session were posted near the experimental area to provide motivation and promote competition among subjects.

Each part of the study was preceded by one or more training sessions. The training sessions were identical to the experimental sessions, except that different mazes were used. The training sessions were used to allow subjects to gain proficiency in maze-solving in general, without allowing any practice on the particular mazes that would be used in the experiment itself. Subjects were trained until their average scores reached asymptotic values. The number of required training sessions ranged from three, for subjects who had never been exposed to the task before, to one, for experienced and proficient subjects.

SECTION 3
EXPERIMENTAL RESULTS AND DISCUSSION

A. Part 1

The first experiment sought to isolate a particular maze configuration which would show reliable and sensitive differences among a range of dot speeds. Maze configuration was varied by altering the size of the maze (8 rows by 8 columns or 10 rows by 10 columns) and the number of barriers in the maze (54 or 60 for the 8 x 8 mazes; 88 or 95 for the 10 x 10 mazes).

Three dot speed conditions were investigated; stop, slow (1.0 intersection per second) and fast (4.0 intersections per second). The slow dot speed was chosen empirically to make the task fairly easy and the fast speed to make it fairly difficult. The stopping dot was regarded as a baseline nonstressed condition, in which subjects could control the pace at which they made choices.

In this first experiment, we wished to avoid the possibility of subjects learning specific mazes. Accordingly, the 12 subjects who participated were assigned at random to three groups of 4 subjects each. Every maze was shown in combination with each of the three dot speeds, one for each subject group. Dot speed by maze configuration combinations were assigned to each group so that each group was exposed to all three dot speeds in a balanced way and to all of the mazes in the battery. The battery consisted of 18 mazes of each configuration.

An analysis of variance was performed on the scores obtained, using maze configuration, specific mazes within configuration, dot speed, and subject groups as independent variables. A split-plot design (subjects assigned to one of the three groups) with specific maze nested within configuration was used. Results are shown in Table 1. There were significant ($p \leq .01$) main effects for configuration, specific maze within configuration, and dot speed. None of the other effects were significant.

A plot of average scores is shown in Figure 2. There was no significant difference between the stopping dot and slow moving dot conditions. The fast moving dot condition yielded significantly ($p \leq .01$) lower scores than the other two conditions. The most complex maze configuration yielded the lowest scores and the largest difference between fast and slow dot speed conditions. Relatively low average scores were desirable because we wanted to avoid any possible ceiling effects. Therefore, the 10 x 10 maze with 95 barriers was used for the other two parts of the study.

Figure 3 shows one of the 10 x 10 mazes with 95 barriers used in this experiment. The figure shows the shortest solution path and deviations from it as a function of dot speed condition. Deviations from the shortest path shown on this figure include all deviations made by any of the subjects. The fast moving dot speed induced more errors than either the slow moving dot or the stopping dot conditions. Errors tended to reflect an inadequate planning horizon rather than motor coordination problems. For the moving dot conditions, errors made because the dot moved straight ahead rather than turning would indicate either motor coordination difficulties or an inadequate planning horizon. Errors made because the dot turned in the wrong direction, on the other hand, could only indicate difficulty with the planning horizon. There were more of these latter kinds of errors in the sample maze shown, and these kinds of errors increased with dot speed. Furthermore, errors made in the fast dot speed condition were not as readily corrected; attempts to correct errors often seemed to reflect an intent to return to the centerline of the maze too soon. These tendencies all are symptomatic of an inadequate planning horizon.

The stopping dot condition, included as a baseline nonstressed condition, did not seem to fulfill that role. Subjects' comments and examination of their strategies indicated that this was a qualitatively different task. The fact that subject input required at each decision point rather than only at turns induced some subjects to continually flick the trim switch while scanning, simply to keep the dot in motion. This type of control activity interfered with the planning process and made this condition ineffective as a baseline. Accordingly, the stopping dot condition was not used in either of the second two parts.

B. Maze Screening

The mazes generated for use in the first experiment were screened only to insure that there were no duplicates and no mazes with unusually short or long optimum path lengths. We wished to improve on this, so that the mazes in a battery would be more nearly homogeneous in difficulty. We used the performance data obtained in the first experiment to refine our maze selection procedure.

TABLE 1

ANALYSIS OF VARIANCE RESULTS FOR PART 1, EFFECT OF MAZE CONFIGURATION ON SCORE

SOURCE	M.S.	F-STAT	df	PROB	R^2
Configuration	4367.8	17.05	3,213	< .01	.03
Maze within Config	950.1	3.71	68,213	< .01	.14
Speed	91766.2	353.35	2,6	< .01	.40
Subject Group	1672.7	–			.01
Config x Speed	165.1	.69	6,426	NS	.002
Config x Group	256.2	–			.12
Speed x Maze within Config	245.7	1.03	136,426	NS	.07
Speed x Group	259.7	–			.003
Residual	238.1	–			.22

Median scores obtained for each of the dot speed conditions for each maze were correlated with the following maze characteristics:

a. Shortest **path length,** also used to screen mazes in the first experiment;

b. **Total number of decision points,** stopping dot;

c. **Total number of turns;**

d. **Number of left turns** required, "backing up" moves;

e. **Maximum excursion up** from the centerline required, in intersections;

f. **Maximum excursion down** from the centerline required;

g. **Number of centerline crossings;**

h. Total **height** of excursions, the sum of items (e) and (f);

i. Total **area** bounded by the optimum path and the centerline, in blocks;

j. **Area/height** ratio; and,

k. **Area x height** product.

These variables were chosen because they seemed to have some potential relationship to maze difficulty.

Correlation analysis and stepwise regression were performed separately for each maze configuration. The description of the results will be limited to the 10x10x95 maze configuration, since that configuration was used throughout the rest of the study. For the regression analysis, five different dependent measures were considered:

a. Median score, stopping dot condition;

b. Median score, slow moving dot condition;

c. Median score, fast moving dot condition;

d. Difference in median score between the stopping dot and slow moving dot; and

e. Difference in median score between the fast and slow moving dot conditions.

Both training and experimental data were included in the analysis. Each part included 18 mazes of the 10x10x95 configuration, for a total of 36 mazes.

Most of the linear correlation coefficients between the dependent and predictor variables were not statistically significant. Table 2 summarizes those correlations that were significant at $p \leq .05$. The same table shows the results of the regression analysis. Clearly, maze characteristics had a larger impact on performance as dot speed increased.

All of the maze characteristics found to have a significant relationship with performance were used to screen mazes used in the second and third experiments. In response to subjects' comments, we included three additional maze characteristics in our screening procedure. The first two were the number of intersections from the starting point to the first decision point and to the first turn in the optimum path, respectively. If either of these values was comparatively large, the effective size of the maze was reduced. The third criterion was that for a given maze battery, the proportion of mazes in which the optimum solution path made its first turn in the "up" direction should be approximately equal to the proportion in which the first turn was "down."

Acceptable ranges for each of these maze characteristics were established, and are shown in Table 3. Any maze generated for the second or third experiment whose characteristics fell outside any one of these ranges was not used in the battery. This procedure resulted in a rejection rate of approximately 30%.

C. Part 2

The second experiment assessed the effect of repeated presentation on performance. A battery of 24 mazes was presented on a daily basis for eight days, with only the order changed from one day to the next. The eight days were the five work days in one week plus Monday through Wednesday of the following week, for all subjects. All 24 mazes were 10x10 with 95 barriers, and a speed of 2.0 intersections per second was used throughout. Four subjects participated in this part, all of whom had also been included in the first experiment.

Table 4 presents the results from the analysis of variance performed on the scores obtained in this part. A full-factorial design with exposure as a fixed factor and maze and subject as random factors was used. Significant effects ($p \leq .01$) were (1) the main effects of exposure, maze, and subject and (2) the interaction effect of maze and exposure.

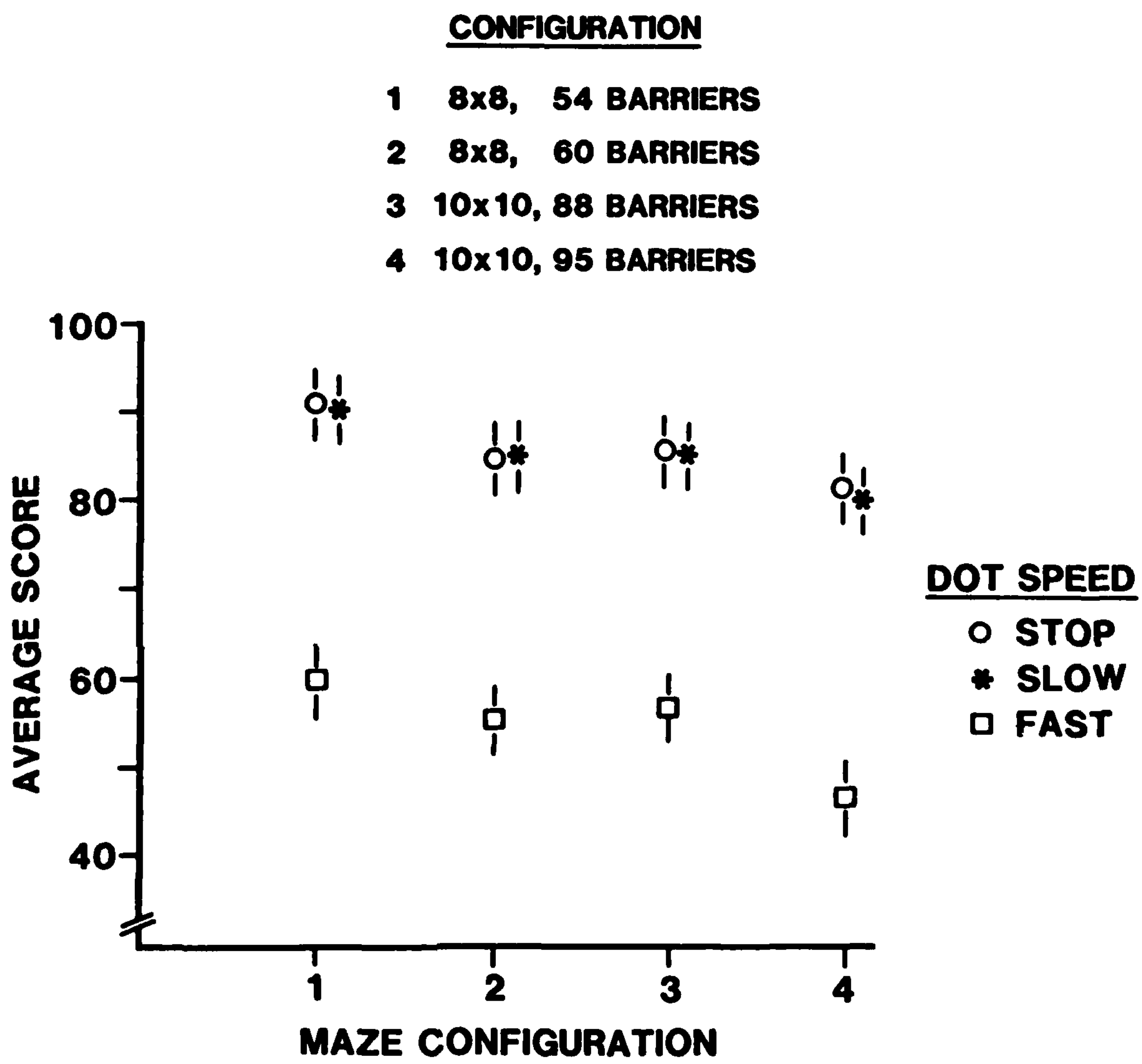
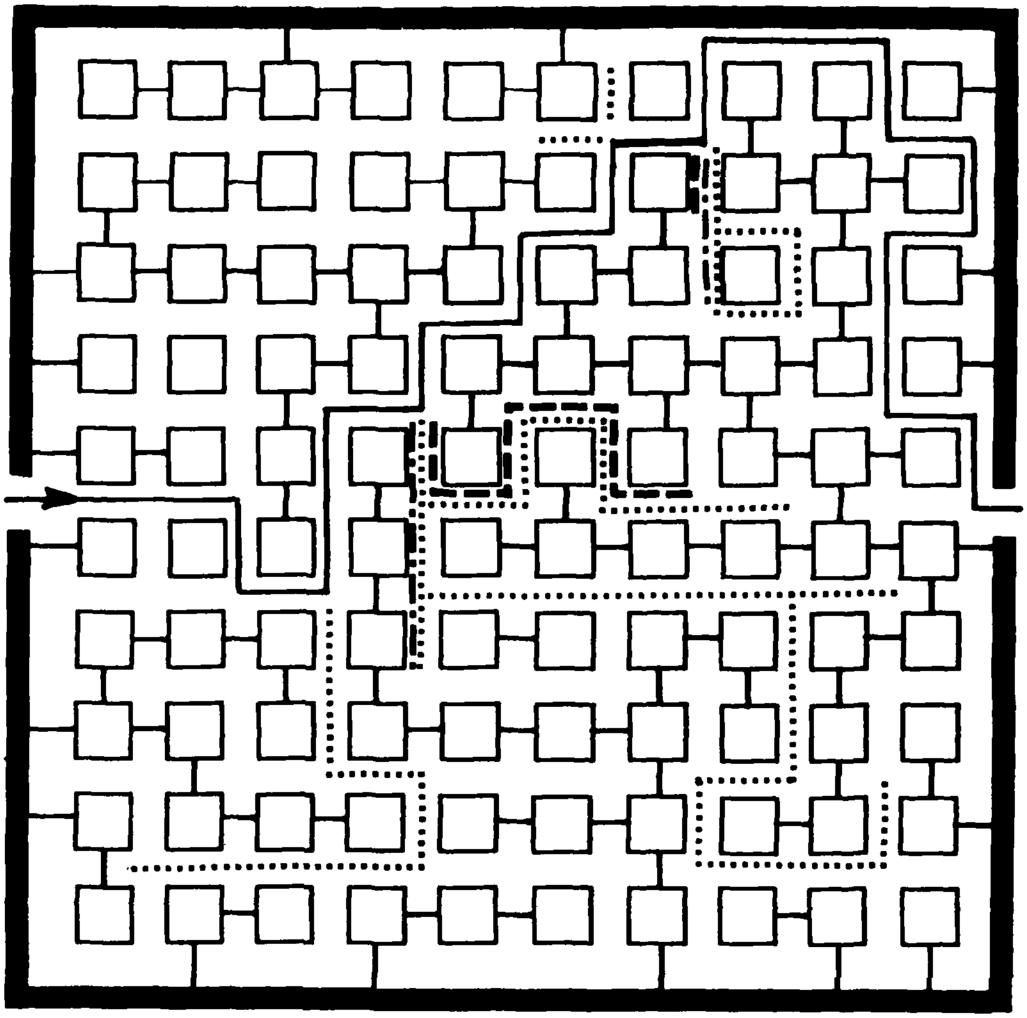


Figure 2. SCORES AS A FUNCTION OF DOT SPEED AND MAZE CONFIGURATION; MEANS AND 95% CONFIDENCE INTERVALS.



	AVERAGE SCORE
STOPPING DOT	83
SLOW MOVING DOT	86
FAST MOVING DOT	36
SHORTEST PATH	

Figure 3. DEVIATIONS FROM THE SHORTEST PATH, MAZE 7, PART 1

Figure 4 shows average score as a function of exposure. Scores increase gradually (and nonmonotonically) from the first day through the eighth. Tukey comparisons indicated significant differences ($p < .05$) for days 1, 2, or 3 versus 8 and for day 2 versus 6. On the basis of these results, we felt that a given maze could be presented four or five times without any appreciable learning effects.

Figure 5 illustrates the interaction between individual mazes and exposure. The figure shows the average for the first three days for each maze, and the average for the last three days. Mazes are arranged in ascending order of average score over all eight days. On the whole, repeated exposure did increase scores. There were, however, a few notable exceptions for which repeated exposure left scores unchanged or lowered them. Maze 5 was such an exception; performance on this maze will be examined in detail below and compared with performance on Maze 1, a more typical maze.

Mazes 1 and 5 are shown in Figures 6 and 7, respectively. The shortest solution path is indicated, along with deviations from it as a function of exposure. For Maze 1, there do not appear to be any qualitative differences among exposures in the types of errors made. This maze was the most difficult maze, as measured by average score, and remained difficult throughout the course of the experiment.

For Maze 5, performance appeared to degrade with repeated exposure. For the first two days, there were very few deviations from the shortest path, and the average score of 93 was comparatively high. On subsequent days, scores ranged from a high of 85 on day 4 to a low of 52 on day 6. The scores for this maze were bimodally distributed: it was either solved fairly easily or with some difficulty, and the tendency to solve it easily was highest during the first two days.

D. Part 3

The final experiment in this series investigated the effect of dot speed on maze performance using a within-subject design and the maze configuration chosen in the first experiment. Each of the 24 mazes in the battery was shown in combination with each of 3 dot speed conditions: slow, medium, and fast (1.67, 2.50, and 5.00 intersections per sec respectively). All of the mazes in this battery were screened using the same procedure as in Part 2; none of the mazes were repetitions from that part. The subjects were the same four as participated in Part 2. One of the 24 mazes was mistakenly not presented at the fast dot speed, so the data from this maze were not included in any analysis.

Analysis of variance was performed on the scores obtained, using a full-factorial design with dot speed as a fixed factor and maze and subject both as random factors. Table 5 summarizes the results from this analysis.

This experiment confirmed our earlier finding that increased dot speed produced reliable decrements in performance. Average scores for the slow, medium, and fast dot speeds were 87.2, 78.7, and 57.8 respectively, and these differences were statistically significant. The main effects of subject and maze were also significant.

The interaction of maze with dot speed was significant, and average scores for each maze and dot speed are shown in Figure 8. Mazes are arranged in ascending order of average score across dot speeds. The general tendency was for the more difficult mazes to show a larger difference among dot speeds than the easier ones.

The maze-solving strategies used in this task were further examined by tracking the actual paths used for each of three representative mazes, and these traces are shown in Figures 9, 10, and 11. Each figure shows a shortest path, and deviations from it as a function of dot speed. Figure 9 shows paths used in solving Maze 1, the maze with the lowest overall score. This maze had an early critical decision point. Subjects tended to choose the correct direction when the dot was moving slowly, but made more errors at this decision point with both the medium and fast dot speeds. The difference between the medium and fast dot speed conditions was that subjects tended to guide the dot back to the center line of the maze too quickly with the fast dot speed condition. These differences in chosen paths reflected a decreased planning horizon with increased speed.

Maze 15, shown in Figure 10, was of medium difficulty as measured by overall score. The average scores for slow and medium dot speed were both high in comparison to the average for the fast condition. Paths chosen in the fast dot speed condition showed a difficulty in making the correct choice at an early decision point. This difficulty was not apparent for the other two conditions.

Maze 13, shown in Figure 11, demonstrated that sometimes you can get lucky. For this maze there was, again, an early critical decision point. The correct choice was the default: if the subject took no action at this decision point, the dot would move straight ahead, on the correct path. In the fast dot speed condition, subjects chose this path, presumably because of reduced time to examine alternatives. When increased planning time was available, subjects sometimes chose the wrong path. The position of the critical decision point very early in the maze may have contributed to the increased error rate for this maze as compared with others (for example, Maze 1). Maze 13 is atypical of the general pattern of performance on this task, as can be seen in Figure 8. Results for this maze did, however, substantiate the effect of increased dot speed on reduced planning horizon found with other mazes.

In addition to overall score, several other dependent measures were examined to assess the effect of dot speed on maze-solving strategy. These additional variables were the following:

a. **Total Path Length;**

b. Subject's **optimum path length,** defined to be the observed path length exclusive of retraced portions or any doubling back;

c. Total **overshoot distance,** summed over all decisions made after the dot had crossed the center of the intersection, defined to be the path length between the center and the decision point and back again;

d. **Bounce distance,** defined as the distance the dot traveled when the subject allowed it to bounce between the barriers.

e. **Retrace distance,** the portion of the total path length spent retracing a path to get back to an earlier decision point; and

f. Total **undershoot distance,** summed over all decisions made before the dot had crossed the center of intersection, defined to be the path length between the decision point and center.

Analysis of variance results for total path length and overshoot distance confirmed results obtained with the score: dot speed significantly degraded performance, and the more difficult mazes showed larger differences among dot speeds than easier ones. Bounce and retrace distances were bimodally distributed: either zero or some relatively large value. The frequency of zero values for these two distances decreased with increased time stress, again supporting the results obtained using the score as the dependent variable.

Subjects optimum path length was almost always as short as the computer-generated solution. The frequency of cases in which the subject's optimum path was longer than the shortest possible path did increase with increased dot speed: 11% for the fast dot speed, 4% for medium, and 1% for slow.

Undershoot distance revealed some differences in subject strategies, as shown in Figure 12. The two better subjects, as judged by overall average score, showed an increase in undershoot distance from the medium to slow dot speed condition. The two poorer subjects, on the other hand, showed a decrease in undershoot distance with decreased dot speed. It appeared that the better subjects were ready to make their inputs sooner than the poorer subjects. This could have happened either due to increased planning horizons or better motor skills in anticipation of the next required step.

E. Comparison of Parts 1, 2, and 3

Figure 13 shows a summary of average scores obtained during the three parts of this study as a function of dot speed. Only data from the 10x10 mazes with 95 barriers are included here.

Results from Parts 2 and 3 show scores to be a nearly linear function of dot speed. Results from Part 1 show average scores below this line, indicating poorer performance.

The difference between Part 1 results and the remainder was not attributable to differences among subjects. The four subjects who participated in Parts 2 and 3 were also subjects in Part 1. The average scores of these four subjects in Part 1 were 84 and 42 for the slow and fast dot speed conditions, respectively. These averages were close to the averages over all subjects of 80 and 47. Furthermore, comparatively lower scores in Part 1 contradicted subjects' comments that the mazes in Parts 2 and 3 seemed more difficult.

The comparatively lower scores obtained in Part 1 may have been due to a difference in "set" between this part and Parts 2 and 3. In the second two parts, the only maze configuration used was 10x10 with 95 barriers. In these two parts, subjects had more opportunity to become proficient at this maze configuration than they did in Part 1. In the first part, this configuration was mixed with three other easier configurations. This finding emphasizes the need to maintain a comparable overall level of difficulty between maze batteries if the results will be used for any between-battery comparisons.

TABLE 2

CORRELATION AND REGRESSION ANALYSIS BETWEEN SCORES AND MAZE CHARACTERISTICS, FROM PART 1

	SIGNIFICANT CORRELATIONS*		REGRESSION ANALYSES		
DEPENDENT MEASURE	MAZE CHARACTERISTIC	r	MAZE CHARACTERISTICS INCLUDED	R^2	P
Stopping Dot Score	Area x Height	-.34	Area x Height	.12	<.05
Slow Dot Score	None		Area x Height	.08	NS
Fast Dot Score	Path Length	-.38	Number of Decision Points	.34	<.01
	Number of Decision Points	-.51**	Area x Height		
	Area	-.42			
	Height	-.43**			
	Area x Height	-.46**			
Difference, Stopping vs Slow Dot	Number of Left Turns	-.36	Total Number of Turns Number of Left Turns	.21	<.05
Difference, Fast vs Slow Dot	None		Total Number of Turns Number of Left Turns	.19	<.05

*All correlations listed are significant at $p \leq .05$
**Significant at $p \leq .01$

TABLE 3

ACCEPTABLE RANGES FOR MAZE CHARACTERISTICS

CHARACTERISTIC	RANGE
Path Length	15-33 Intersections
Number of Decision Points	7-19
Total Number of Turns	8-20
Number of Left Turns	0-2
Height	3-7 Intersections
Area	4-38 Blocks
Area x Height	0-200
First Turn	0-3 Intersections
First Decision Point	0-3 Intersections

TABLE 4

ANALYSIS OF VARIANCE RESULTS FROM PART 2, EFFECT OF REPEATED PRESENTATION ON SCORE

SOURCE	M.S.	F-STAT	df	PROB	R^2
Exposure	1177.1	4.04	7,161	<.01	.04
Maze	1747.2	6.53	23,69	<.01	.18
Subject	4560.0	17.04	3,69	<.01	.06
Exposure x Maze	291.5	1.58	161,483	<.01	.21
Exposure x Subj	271.9	1.47	21,483	NS	.03
Maze x Subject	267.6	–			.08
Exposure x Maze x Subject	184.7	–			.40

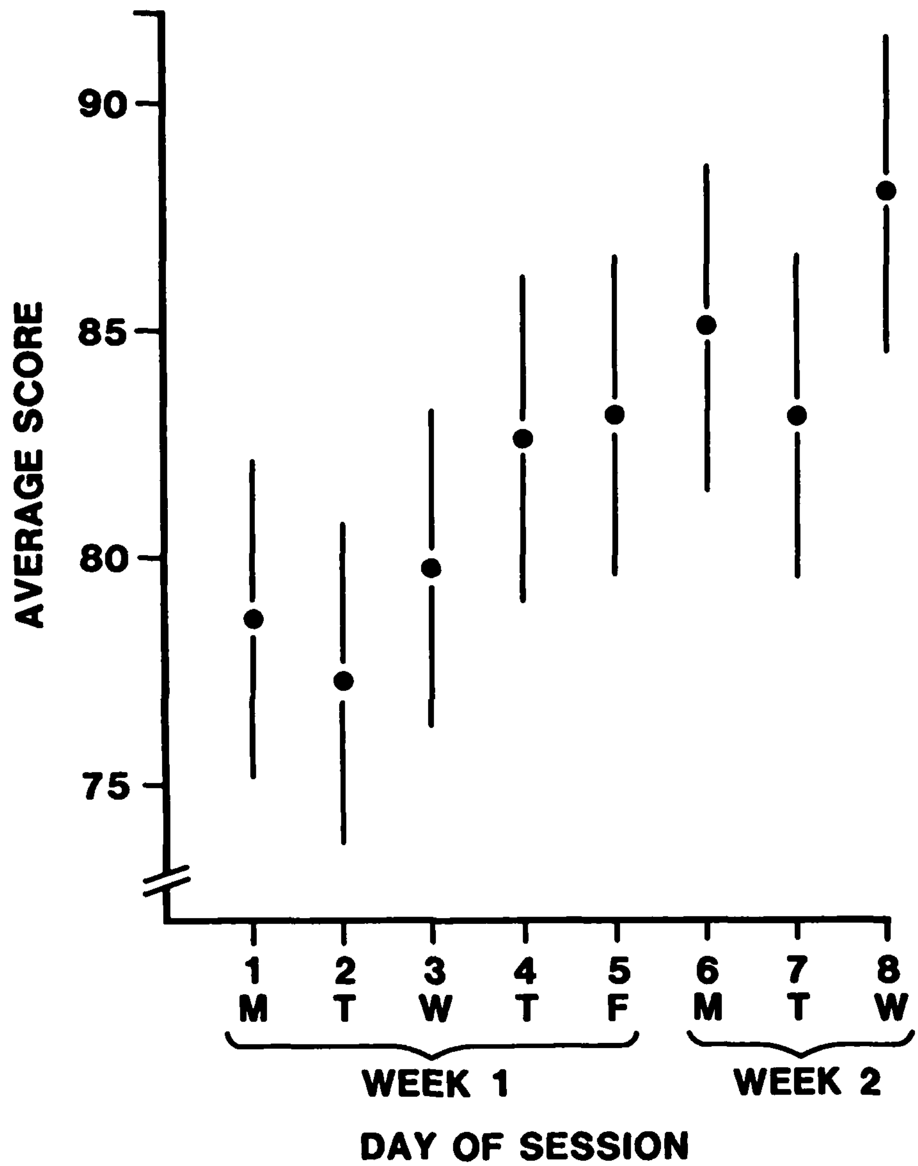


Figure 4. SCORE AS A FUNCTION OF EXPOSURE; MEANS AND 95% CONFIDENCE INTERVALS

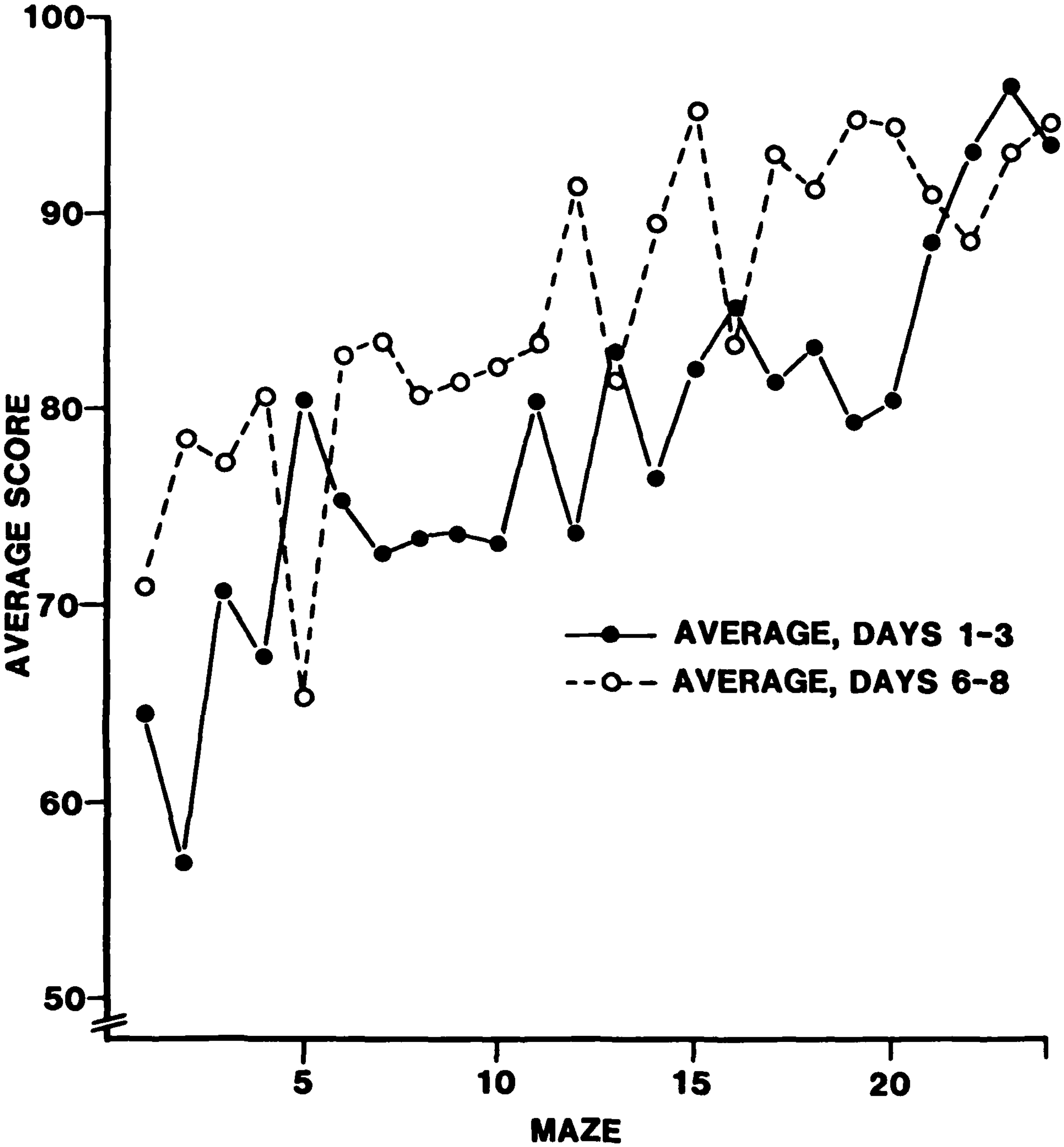
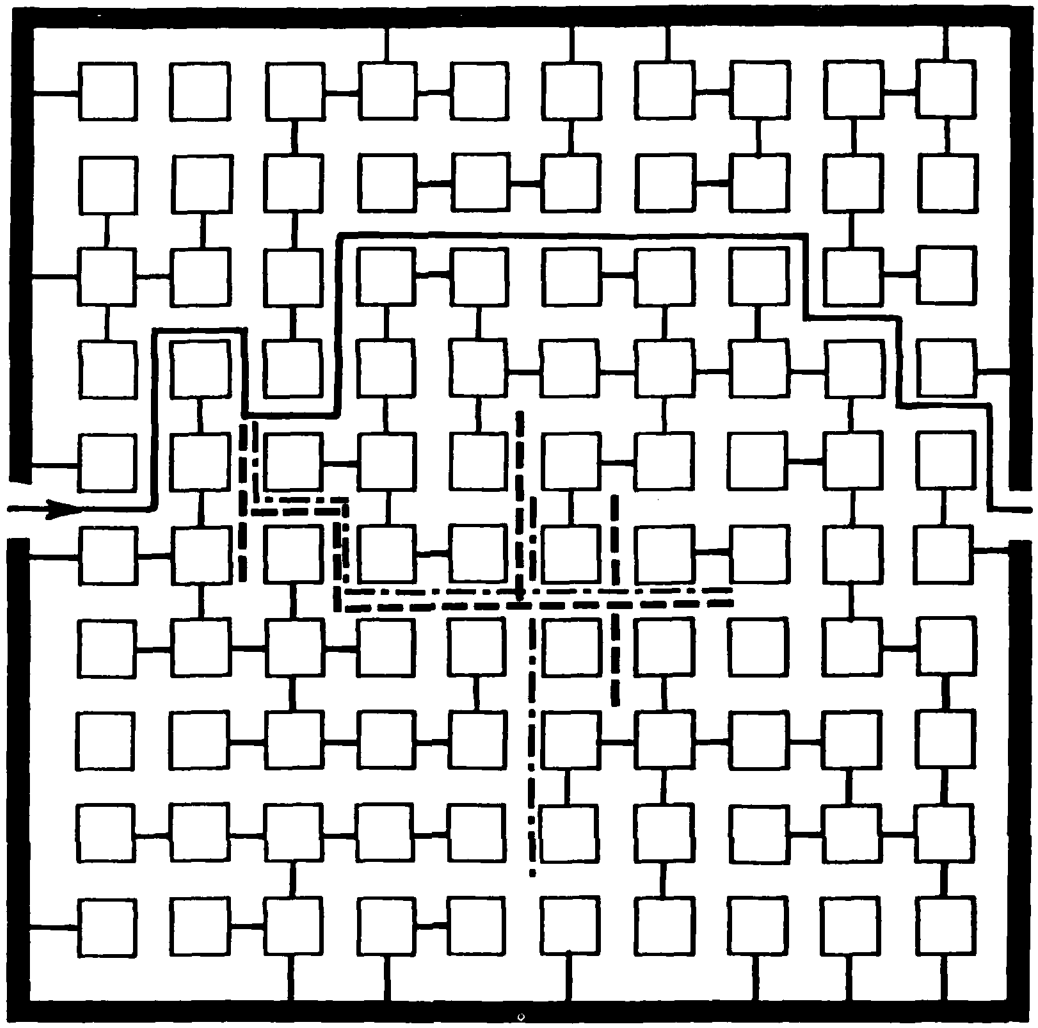
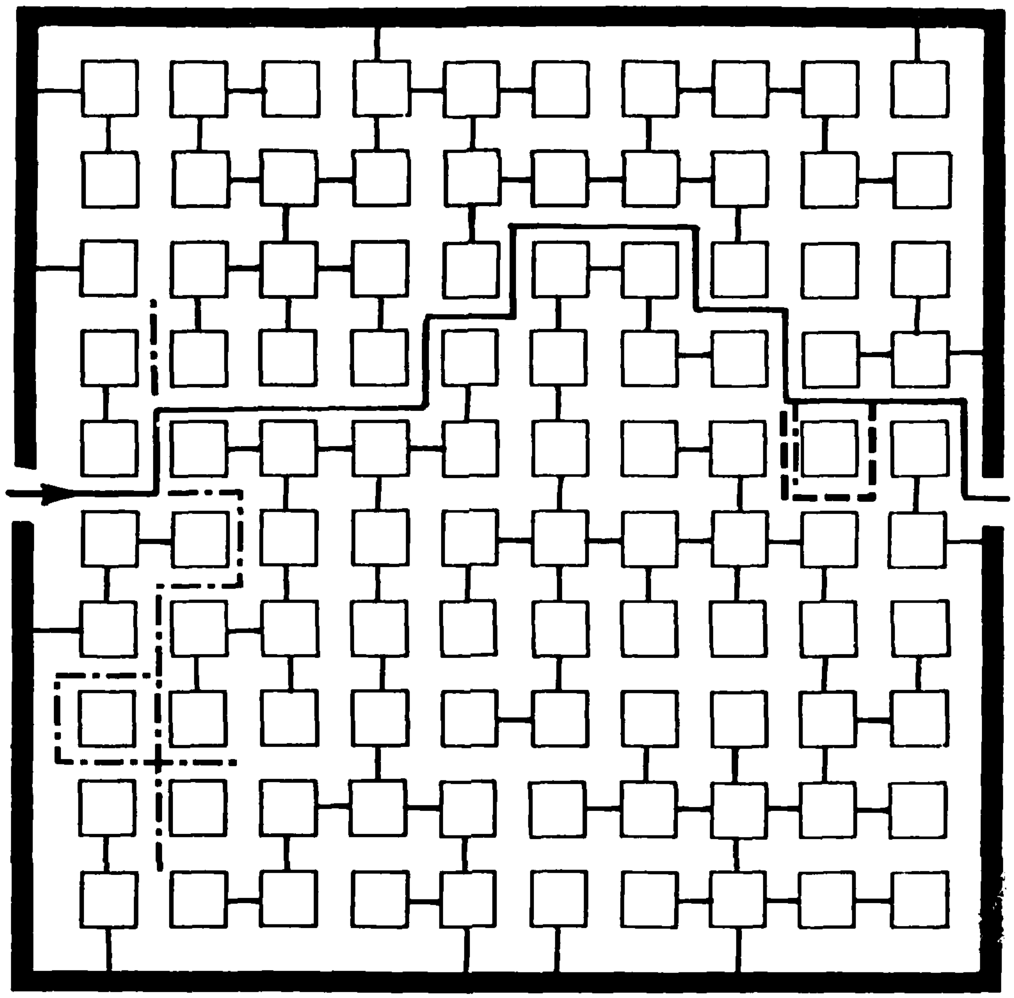


Figure 5. AVERAGE SCORES FROM PART 2 FOR DAYS 1-3 AND DAYS 6-8, SHOWING THE INTERACTION BETWEEN MAZE AND EXPOSURE. MAZES ARE ARRANGED IN ASCENDING ORDER OF AVERAGE SCORE.



		AVERAGE SCORE
----	DAYS 1 AND 2	59
-.-.-	DAYS 7 AND 8	71
——►	SHORTEST PATH	

Figure 6. DEVIATIONS FROM THE SHORTEST PATH, MAZE 1, PART 2.



		AVERAGE SCORE
– – – –	DAYS 1 AND 2	93
—·—·—	DAYS 7 AND 8	73
——➤——	SHORTEST PATH	

Figure 7. DEVIATIONS FROM THE SHORTEST PATH, MAZE 5, PART 2.

TABLE 5

ANALYSIS OF VARIANCE RESULTS FROM PART 3, EFFECT OF DOT SPEED ON SCORE

SOURCE	M.S.	F-STAT	df	PROB	R^2
Speed	21105.0	35.41	2,44	<.01	.26
Maze	1728.3	5.53	22,66	<.01	.23
Subject	1023.8	3.28	3,66	<.05	.02
Speed x Maze	596.0	2.31	44,132	<.01	.16
Speed x Subj	192.4	.75	6,132	NS	.01
Maze x Subj	312.5	--			.12
Speed x Maze x Subject	257.9	--			.21

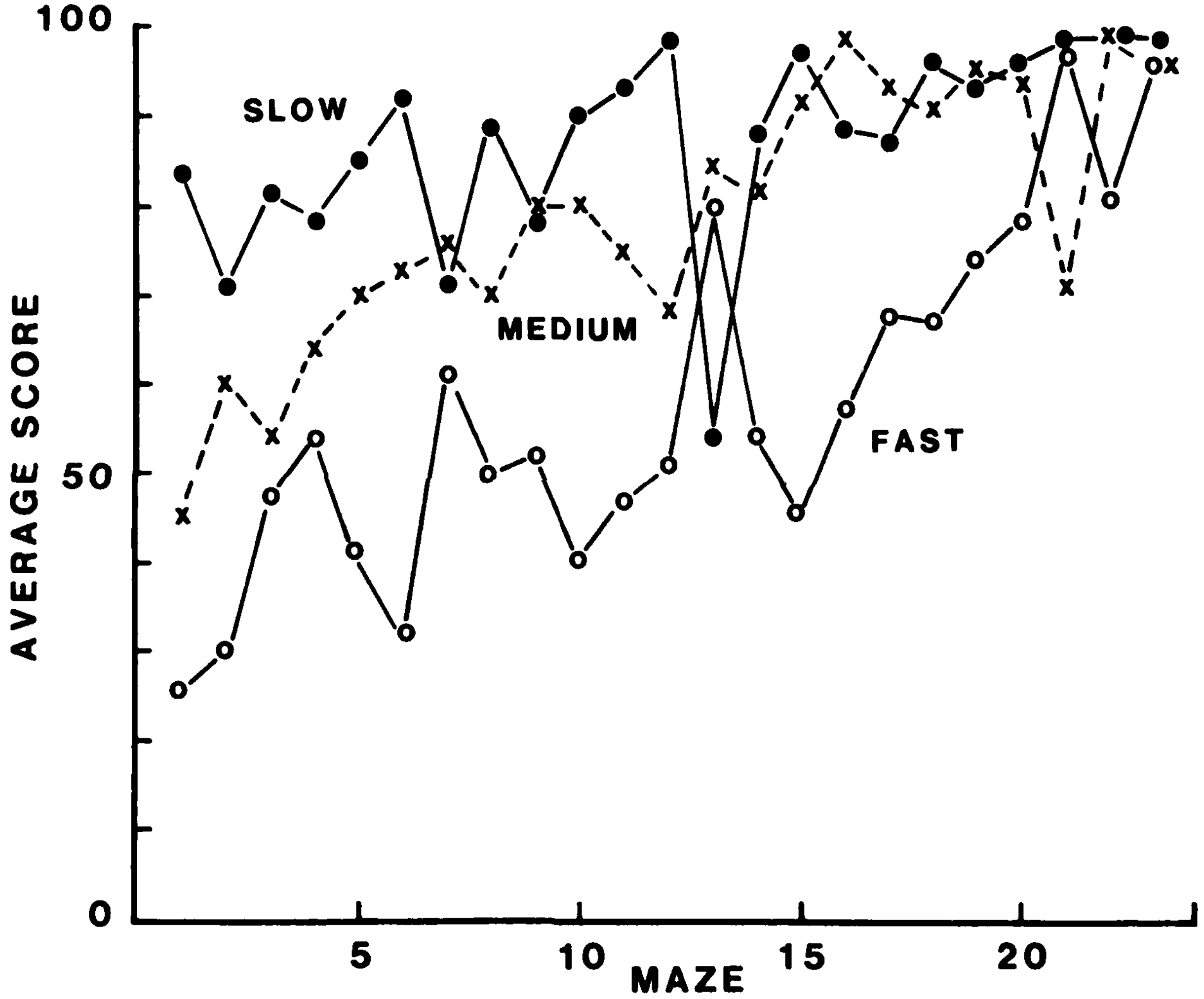
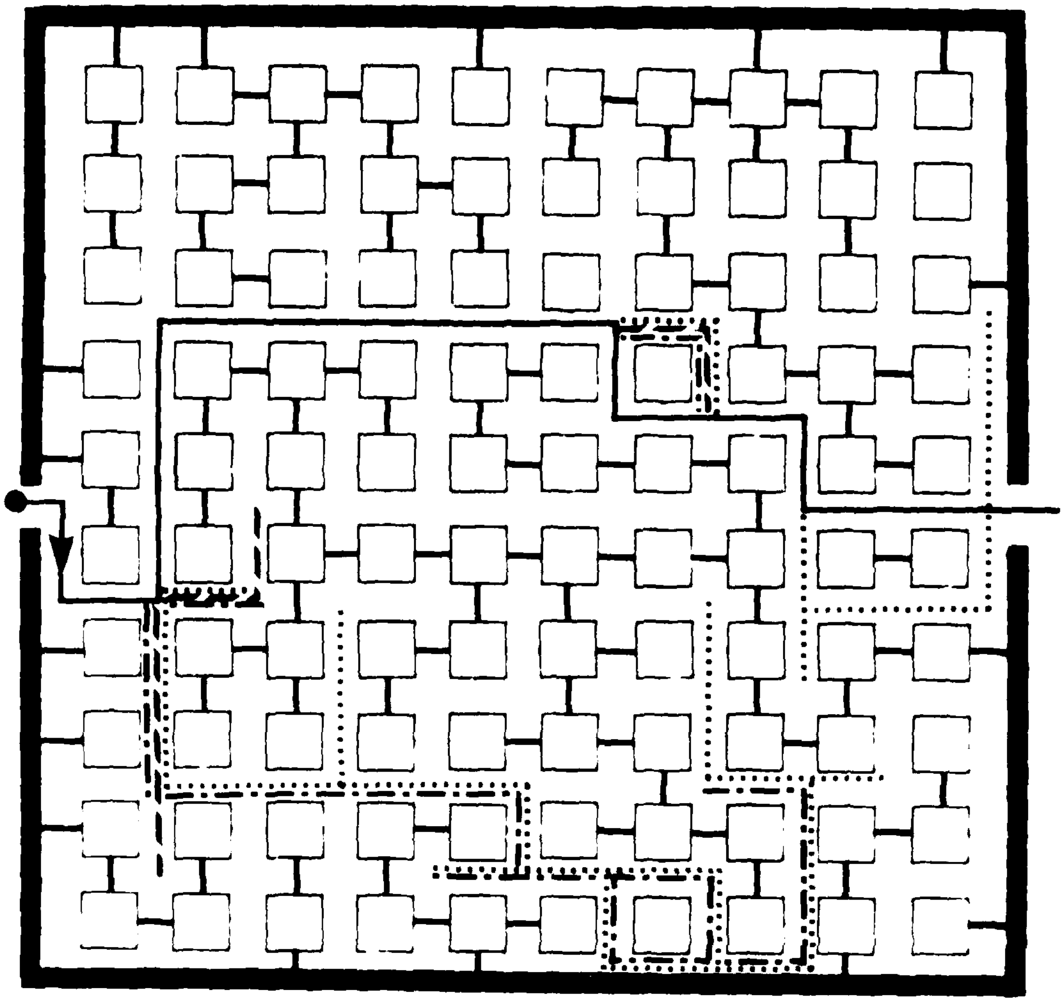
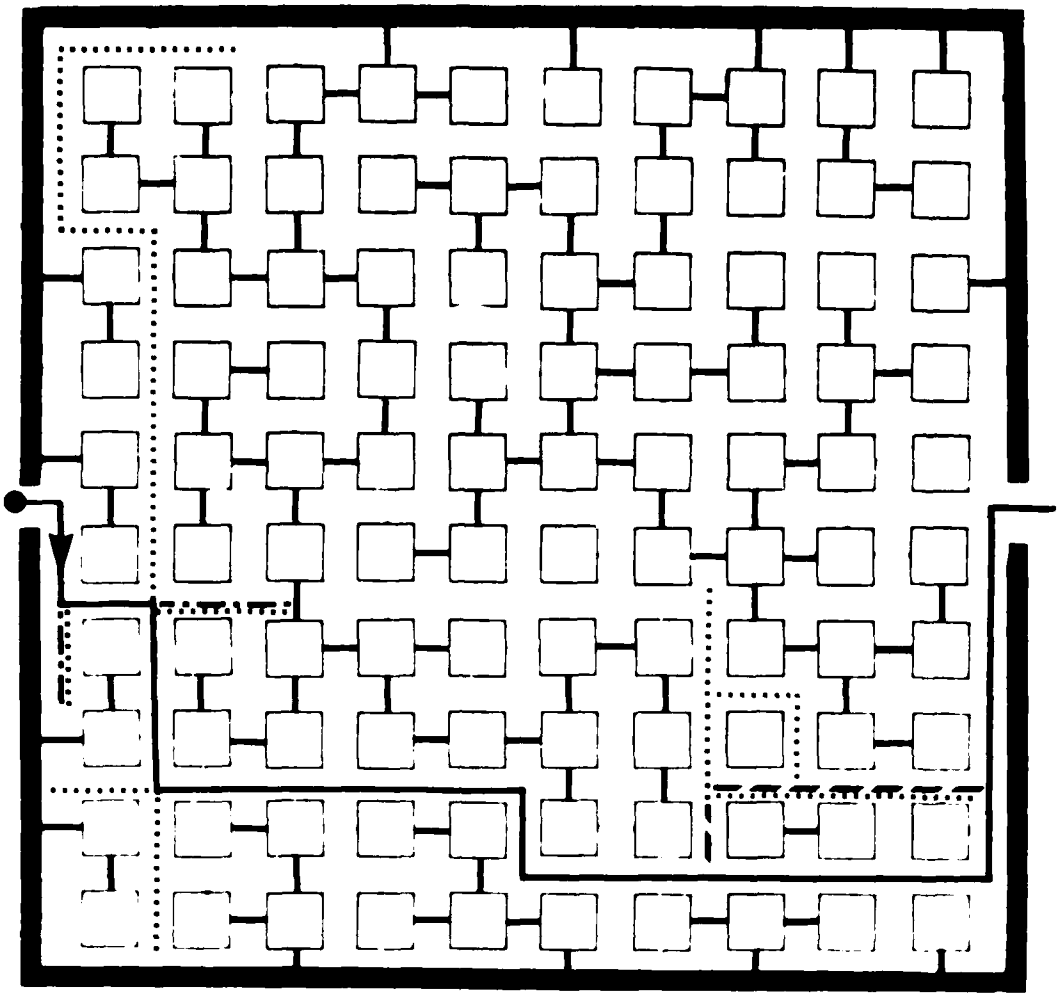


Figure 8. AVERAGE SCORE AS A FUNCTION OF DOT SPEED AND MAZE DIFFICULTY.



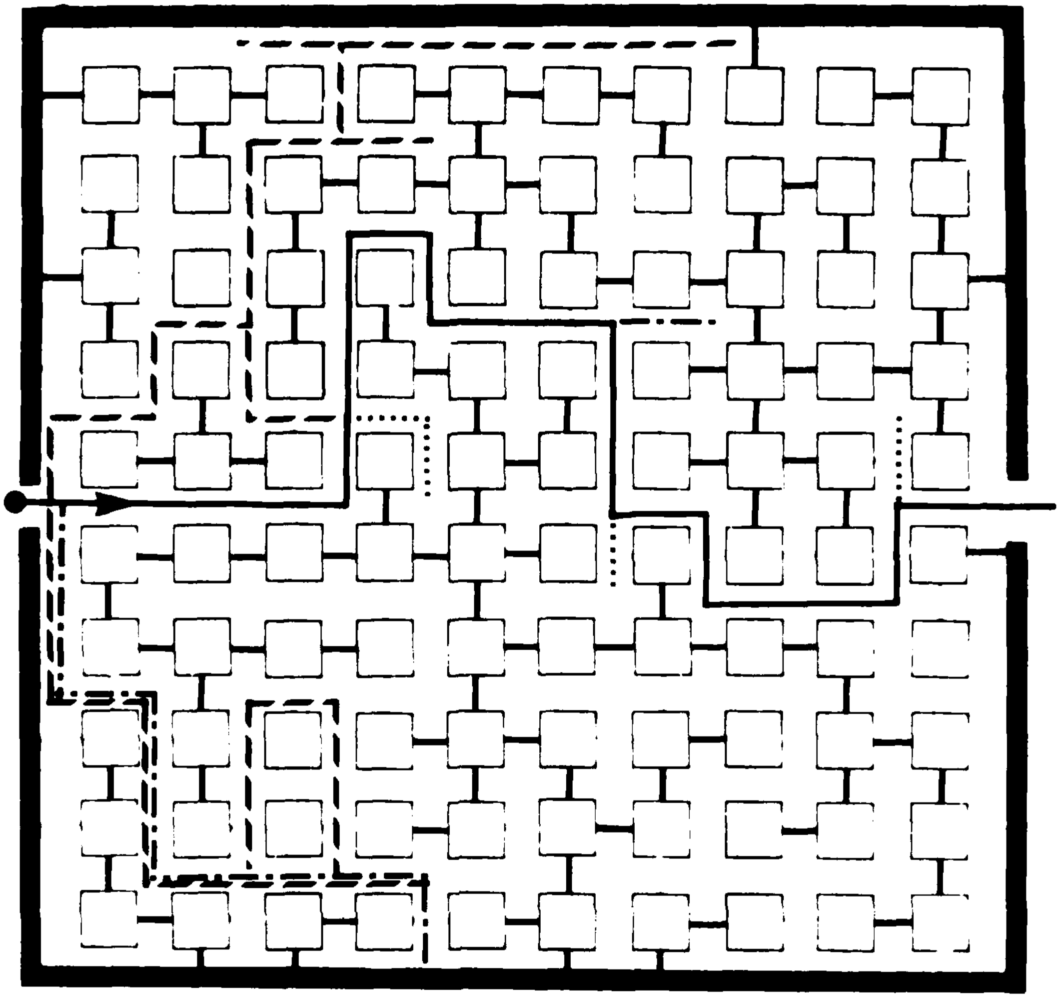
		AVERAGE SCORE
- - - - -	SLOW SPEED	83
. — . — . —	MEDIUM SPEED	45
. . . .	FAST SPEED	26
——→——	SHORTEST PATH	

Figure 9. DEVIATION FROM THE SHORTEST PATH, MAZE 1, PART 3.



	AVERAGE SCORE
SLOW SPEED	97
MEDIUM SPEED	92
FAST SPEED	46
SHORTEST PATH	

Figure 10. DEVIATION FROM THE SHORTEST PATH, MAZE 15, PART 3.



		AVERAGE SCORE
- - - - -	SLOW SPEED	54
.—.—.—	MEDIUM SPEED	84
....	FAST SPEED	80
——►——	SHORTEST PATH	

Figure 11. DEVIATION FROM THE SHORTEST PATH, MAZE 13, PART 3.

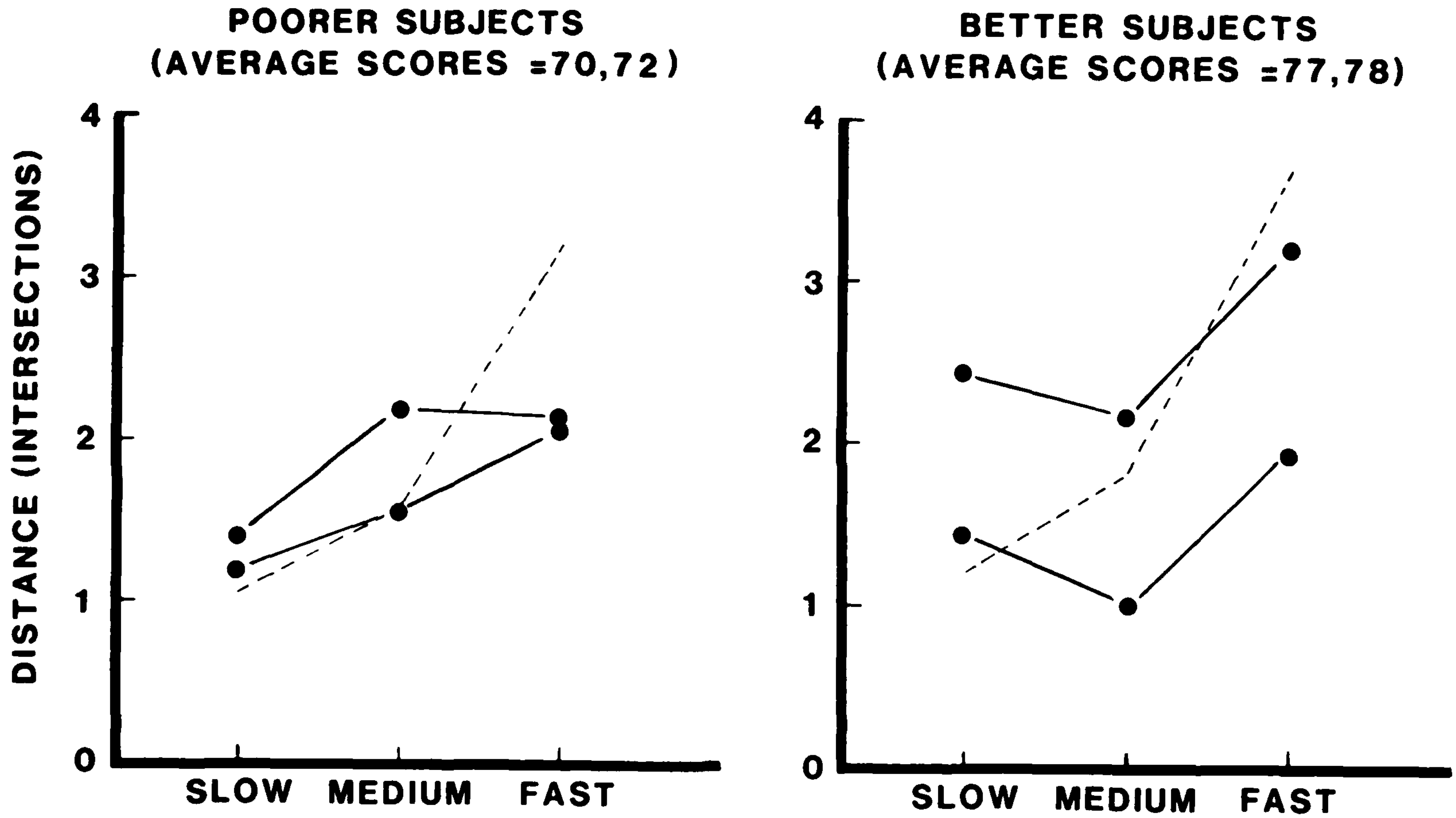


Figure 12. MEAN UNDERSHOOT DISTANCE BY SUBJECT AND DOT SPEED

—●— OBSERVED MEAN ---- ASSUMING CONSTANT UNDERSHOOT TIME

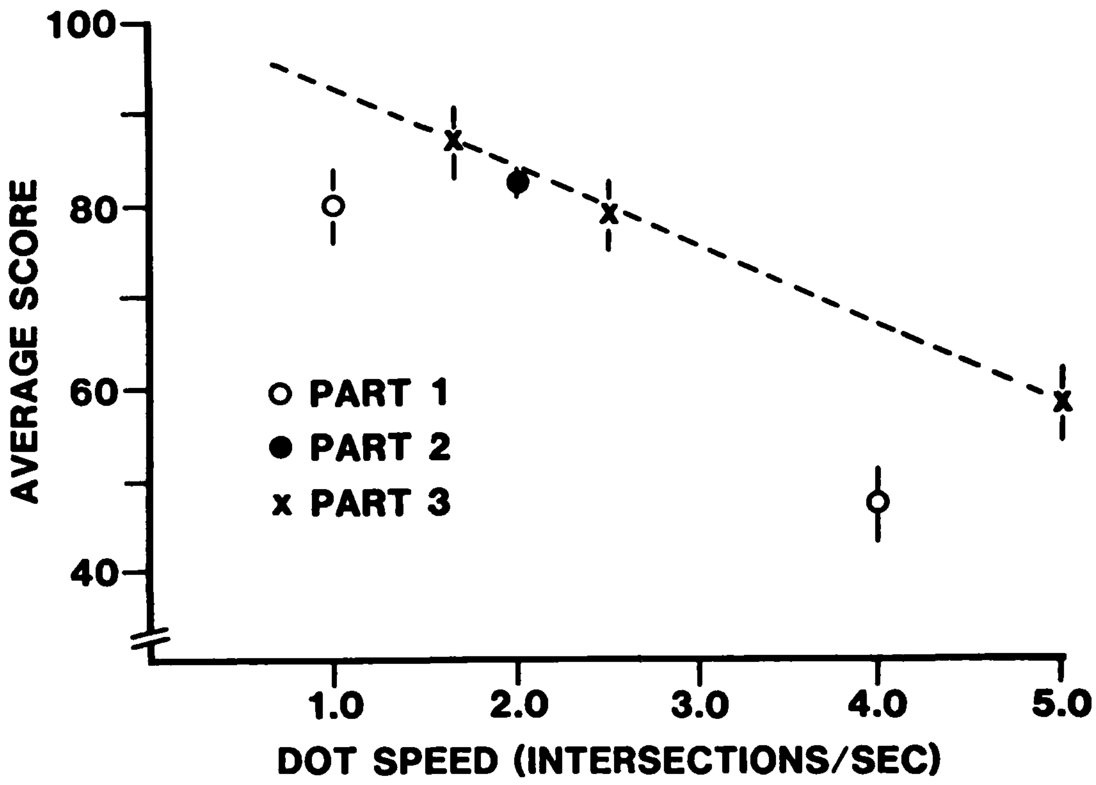


Figure 13. COMPARISON OF SCORES OBTAINED IN THE THREE PARTS; MEANS AND 95% CONFIDENCE INTERVALS.

SECTION 4

TESTS AND MEASUREMENTS PROPERTIES

This section discusses three topics that need to be considered when designing a test battery of mazes for use in future studies. The first is the reliability of each individual maze in showing differences among levels of the independent variables, as assessed by item analysis. The second is the practical consideration of the duration of each session. The last is the determination of the number of mazes and number of subjects required for a given level of precision.

A. Item Analysis

The purpose of item analysis is to quantify the reliability of each individual component of a test battery. The test battery can then be refined and improved by careful selection of these individual items.

This section presents results of item analysis for all of the 10x10 mazes with 95 barriers used in Parts 1 and 3. Mazes from Part 2 were not included in this analysis because those mazes were only presented at one dot speed and item analysis would require at least two.

Item analysis was performed by computing a reliability index for each maze in the battery. The reliability index was defined to be the correlation between scores obtained for a particular trial and the scores obtained in the corresponding dot speed conditions, averaged across all mazes and replications. Tables 6 and 7 show the results for Parts 1 and 3, respectively.

For Part 3, reliability indices fall into roughly three groups: .60 and above, near .40, and below .00. Clearly the two mazes with negative reliability indices should not be considered for use in future test batteries. One of these mazes, number 13, has been discussed in detail in the previous section, and the reason for the difference in performance between this maze and the majority of mazes in the battery has been examined.

The four mazes with reliability near .40 were examined further to see if they were suitable for inclusion in future test batteries. Three of the four were found to be unsatisfactory in one way or another. Mazes 23, 9, and 7 showed very small differences in average score for at least two out of the possible three pairs of dot speeds. Mazes 23 and 7 had highly variable scores for the medium dot speed condition. On the basis of these results, mazes with reliability near .40 or lower should probably not be included in future test batteries.

The reliability indices from Part 1 did not fall into such easily identified groups. The three mazes with reliability near .40 should probably not be used in future test batteries, and the two mazes with reliability between .50 and .60 would be of questionable utility. A reasonable guideline appears to be to choose those mazes with reliability of .60 or higher.

B. Duration of Sessions

The average duration of a single trial in Part 3 for dots speeds of 1.67, 2.50, and 5.00 intersections per second were 45.1, 35.5, and 27.7 seconds, respectively. The total session duration was approximately 20 to 30 minutes, including intertrial intervals (whose durations were under subject control) and task performance time of approximately 15 minutes. Sessions of this duration seemed comfortable to the subjects, by their comments. Based on pilot studies, sessions longer than 40 to 45 minutes may induce some fatigue effects and should be avoided if possible.

C. Sample Size Requirements

The number of mazes included in the battery and the number of subjects used in the experiment will play an important role in determining the precision of the results. This section will outline methods for calculating the approximate precision expected for a given number of mazes and subjects. For purposes of this discussion, precision is the minimum detectable difference in score between the levels of the independent variable. A detectable difference is one which is expected to be found statistically significant, if a true difference of what magnitude actually exists.

Sample size requirements for two different experimental designs will be discussed. The first is a full factorial design, such as the one used in Part 3. In such a design, each maze is shown to each subject at each level of the independent variable A (e.g., dot speed). The second is a partially nested design with two independent variables A and B (e.g., dot speed and G-stress). In such a design, a different set of mazes is used for each level of factor B. Each maze is shown to each subject at each level of factor A, within one level of factor B. Such a design would be necessary if the total number of exposures to any given maze would otherwise exceed four or five trials per subject.

TABLE 6

ITEM ANALYSIS, PART 1

IN ORDER OF AVERAGE SCORE		IN ORDER OF RELIABILITY	
MAZE	RELIABILITY	MAZE	RELIABILITY
1	.76	7	.96
2	.53	9	.95
3	.88	8	.93
4	.42	15	.88
5	.67	3	.88
6	.36	18	.87
7	.96	13	.85
8	.93	12	.84
9	.95	17	.79
10	.67	1	.76
11	.38	10	.67
12	.84	5	.67
13	.85	14	.65
14	.65	16	.56
15	.88	2	.53
16	.56	4	.42
17	.79	11	.38
18	.87	6	.36

Sample size for each reliability index = 8 (4 subj x 2 dot speeds)

TABLE 7

ITEM ANALYSIS, PART 3

IN ORDER OF AVERAGE SCORE		IN ORDER OF RELIABILITY	
MAZE	RELIABILITY	MAZE	RELIABILITY
1	.86	6	.91
2	.76	15	.87
3	.60	1	.86
4	.43	10	.78
5	.72	11	.77
6	.91	8	.76
7	.39	2	.76
8	.76	5	.72
9	.43	20	.71
10	.78	22	.70
11	.77	18	.66
12	.65	16	.66
13	-.34	12	.65
14	.62	14	.62
15	.87	19	.60
16	.66	17	60
17	.60	3	.60
18	.66	4	.43
19	.60	9	.43
20	.71	7	.39
21	-.21	23	.33
22	.70	21	-.21
23	.33	13	-.34

Sample size for each reliability index = 12 (4 subj x 3 dot speeds)

For both of these designs, "maze" and "subject" are considered to be random factors, and independent variable A and B are fixed factors.

Differences among levels of an independent variable can follow any one of a number of patterns, and this pattern influences the significance level. The pattern which is least likely to show a significant difference is the one in which all but the two extremes fall half-way between them. This was the pattern assumed here.

1. **Case 1: Full Factorial Design**

In order to estimate precision, one must first obtain, or guess, an estimate for within-condition variability. The mean square (MS) error term used in the analysis of variance for testing factor A is an estimate for this variability. If the MS error term is available from some previous study, it can be used to determine the sample size required. For this design, the correct error term to use is determined somewhat by the results, as follows:

IF:	THEN THE MS ERROR TERM IS:
a. A x Subjects not significant	MS_{AxM}
b. A x Maze not significant	MS_{AxS}
a. A x Maze and A x Subject both significant.	$MS_{AxS} + MS_{AxS} - MS_{AxMxS}$

In Part 3, where A corresponded to dot speed, A x Subject was not significant and the correct error term was MS_{AxM}. For the maze-solving task, this may be a fairly common outcome. For all of the following calculations in this section, we have assumed that the A x Subject error term would not be significant.

As an aside, note the importance of keeping the A x Maze interaction as small as possible, since this interaction will be the error term for testing the main effect of A. The A x Maze interaction will tend to become large when mazes are not of roughly the same difficulty: difficult mazes will tend to show a larger difference among levels of A than easy ones. So, one way of increasing precision is to choose the test battery carefully.

With the above assumption, the minimum detectable difference between two levels of factor A is

$$\Delta_A = \sqrt{\frac{2a}{ms}}\,\sigma\phi \tag{1}$$

where a = the number of levels of factor A;

m = the number of mazes;

s = the number of subjects;

σ = population within-condition standard deviation, estimated by $\sqrt{MS_{AxM}}$;

and ϕ = value obtained from the Pearson and Hartley charts for power of the F-test (Scheffe, 1959).

There are four parameters used to determine the value of , as follows:

υ_1 = numerator degrees of freedom for testing factor A = a-1;

υ_2 = denominator degrees of freedom for testing factor A = (a-1) (m-1);

α = significance level of the F-test;

β = power of the F-test (probability that if a difference of this magnitude truly exists, the F-test will find it significant);

Using formula (1), one can determine the expected precision of test results. The number of mazes and subjects should be chosen so the expected precision is acceptable for the application at hand.

As an example, values of Δ_A were computed for a range of values for a, s and m, and are shown in Table 8. For this example, α = .05 and β = .90 were used. The value of σ = 25 was taken from Part 3.

TABLE 8

SAMPLE VALUES FOR MINIMUM DETECTABLE DIFFERENCE, FULL-FACTORIAL DESIGN A x MAZE x SUBJECT

Number of Levels of A	Number of Subjects	M= Number of Mazes 5	10	20	30
a = 2	2		28.8	19.1	15.3
	4		20.4	13.5	10.8
	8		14.4	9.5	7.6
a = 3	2	48.8	30.9	20.7	16.6
	4	34.5	21.9	14.7	11.7
	8	24.4	15.5	10.4	8.3
a = 4	2	49.6	32.1	21.7	17.3
	4	35.1	22.7	15.3	12.3
	8	24.8	16.0	10.8	8.7

$\alpha = .05$ $\beta = .90$ $\mathcal{J} = 25$

2. Case 2: Partially Nested Design

A similar procedure can be followed to estimate precision for this second type of design. Once again, some estimate of within-condition variability is needed. One can guess at this, or use the results from Part 3 as a rough guide. If maze-solving has been used in a partially nested design before, results from that test may be helpful in estimating the variability. As in Case 1, the correct error term to use is determined somewhat by the results. For purposes of this discussion, we will assume that A x Subject, B x Subject, and A x B x Subject interactions are all nonsignificant. In light of results from Part 3, these assumptions appear quite reasonable.

Using the above assumptions, the minimum detectable difference between two levels of factor A is

$$\Delta_A = \sqrt{\frac{2a}{bms}}\,\sigma\phi \tag{2}$$

where a, m, and s are as above;

b = the number of levels of factor B;

σ = population within-condition standard deviation, estimated by $\sqrt{MS_{AxM\,(B)}}$;

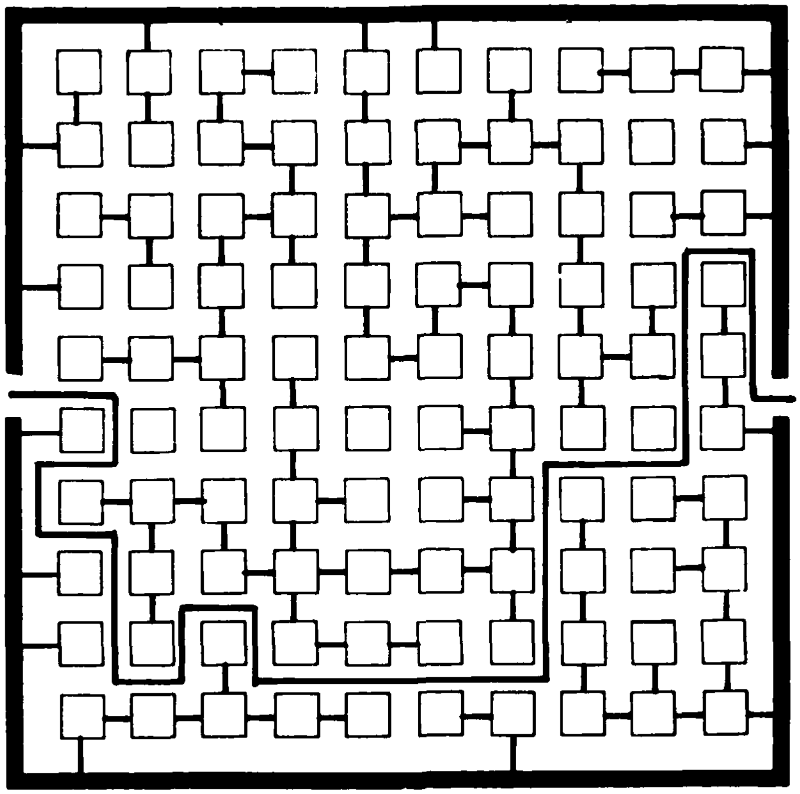
ϕ = value from Pearson and Hartley charts, using degrees of freedom υ_1 = a-1 and υ_2 = (a-1)b(m-1).

Similar formulas can be obtained for the minimum detectable difference among levels of factor B, and for the interaction term of A x B. Appendix C provides these formulas and outlines the sample size calculations developed in this section.

SECTION 5
CONCLUSIONS

This feasibility study has shown the maze-solving task to be a reliable and sensitive measure of stress induced by time contraints. This study showed that the number of incorrect paths chosen increased as the dot speed increased, and errors that did occur were not as readily corrected. Furthermore, the errors that occured with increased dot speed were typically errors of commission, rather than simple failures of motor coordination. These tendencies indicated that the decrement in performance scores with increased dot speed was due to shortened planning horizon, rather than the limits of motor skill.

APPENDIX A
MAZES USED IN PART 1



MAZE 1

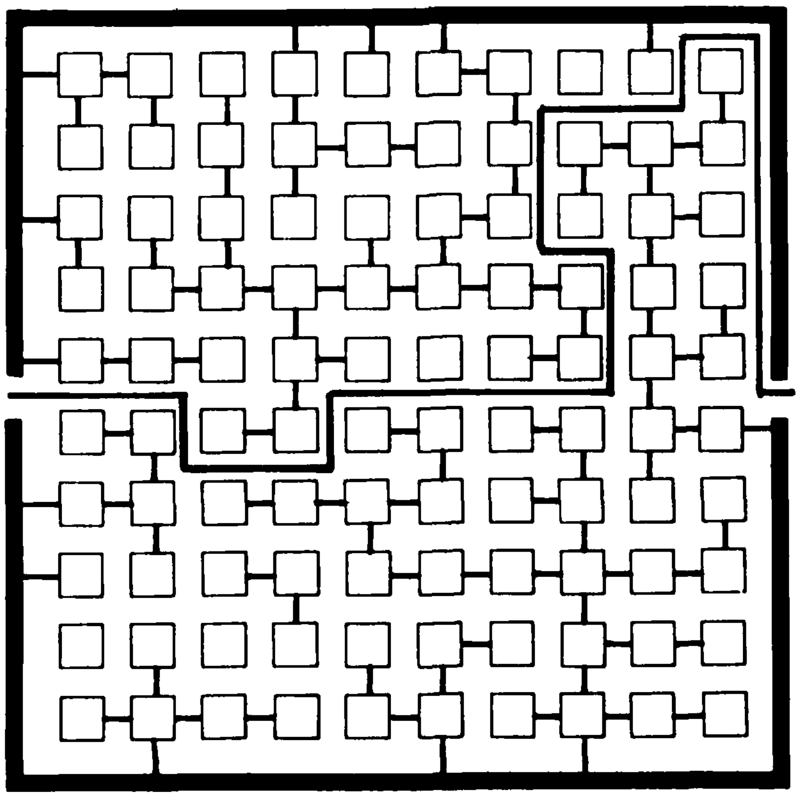
CODE HO

RELIABILITY .76

AVERAGE SCORE

SLOW 59

FAST 32



MAZE 2

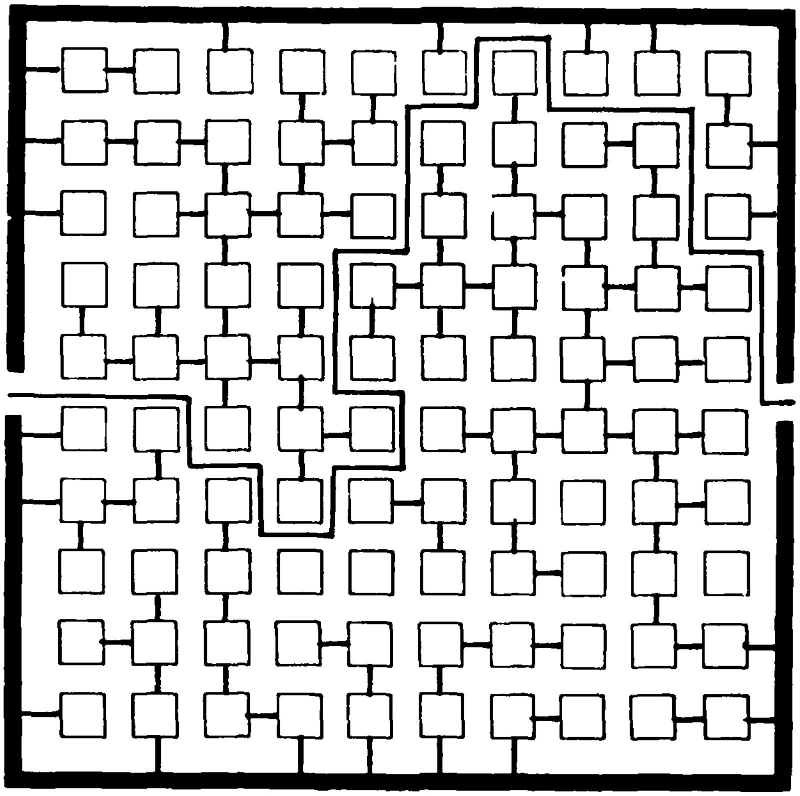
CODE HI

RELIABILITY .53

AVERAGE SCORE

SLOW 57

FAST 33



MAZE 3

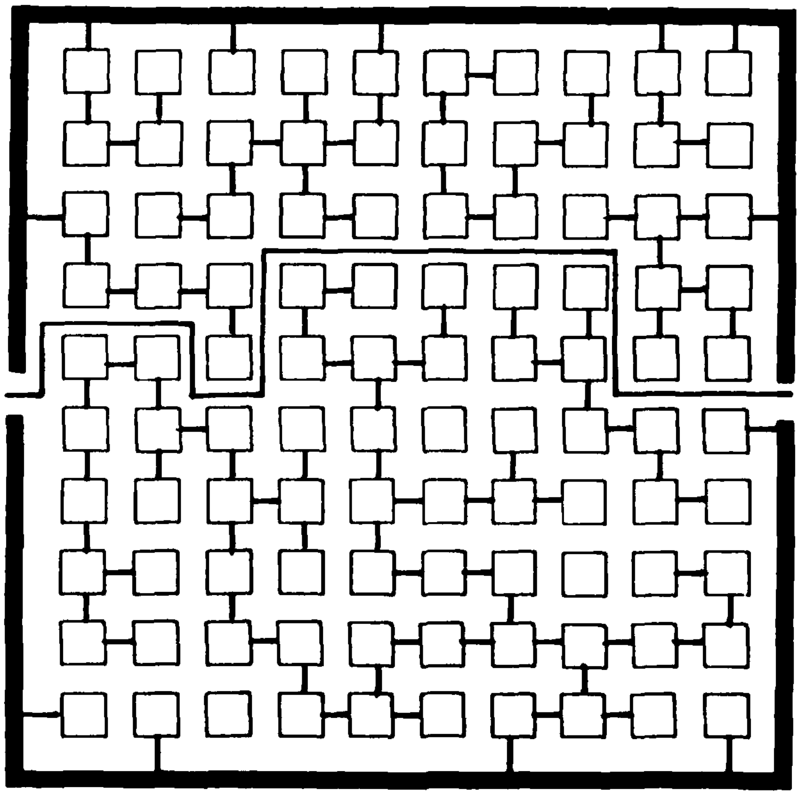
CODE HQ

RELIABILITY .88

AVERAGE SCORE

SLOW 75

FAST 35



MAZE 4

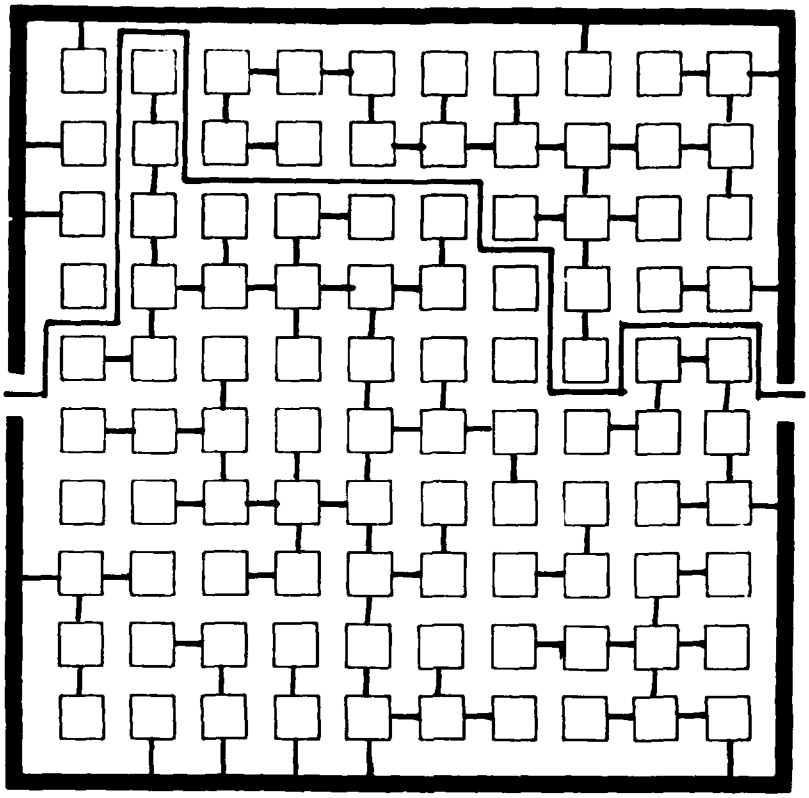
CODE HC

RELIABILITY .42

AVERAGE SCORE

SLOW 62

FAST 43



MAZE 5

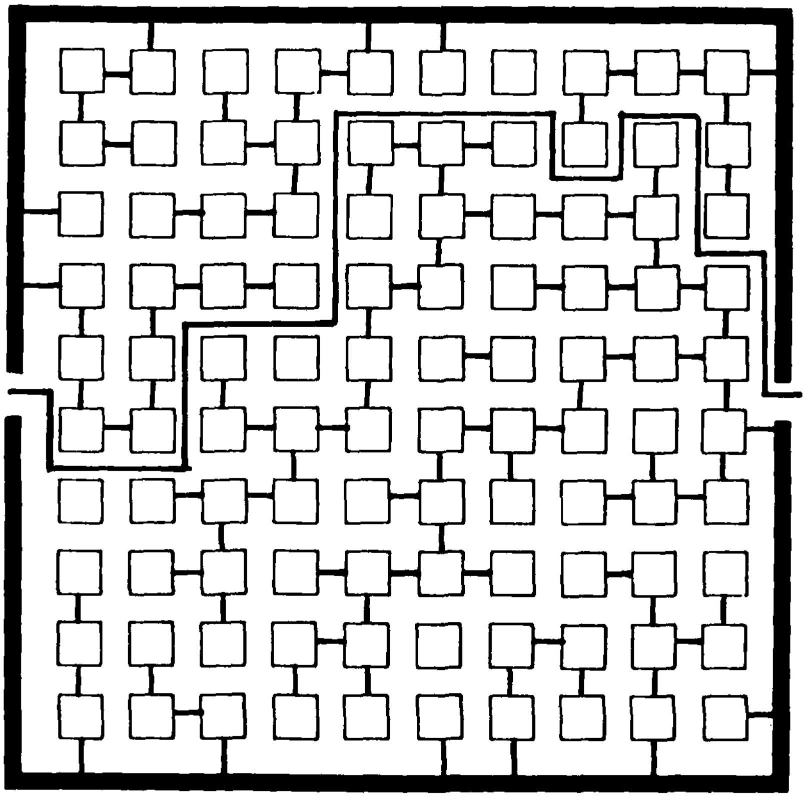
CODE HF

RELIABILITY .67

AVERAGE SCORE

SLOW 68

FAST 42



MAZE 6

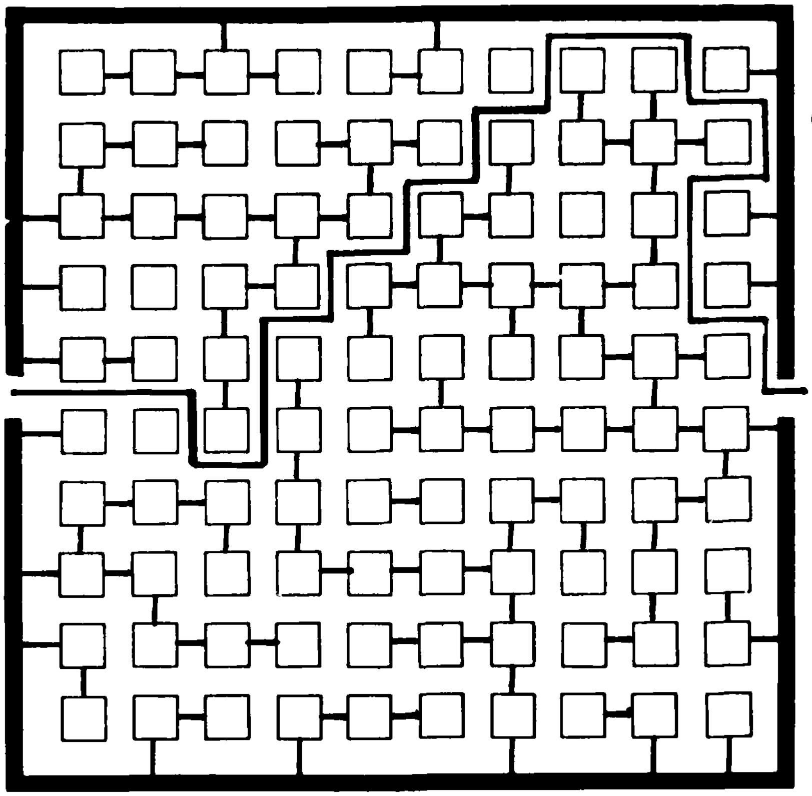
CODE HG

RELIABILITY .36

AVERAGE SCORE

SLOW 76

FAST 56



MAZE 7

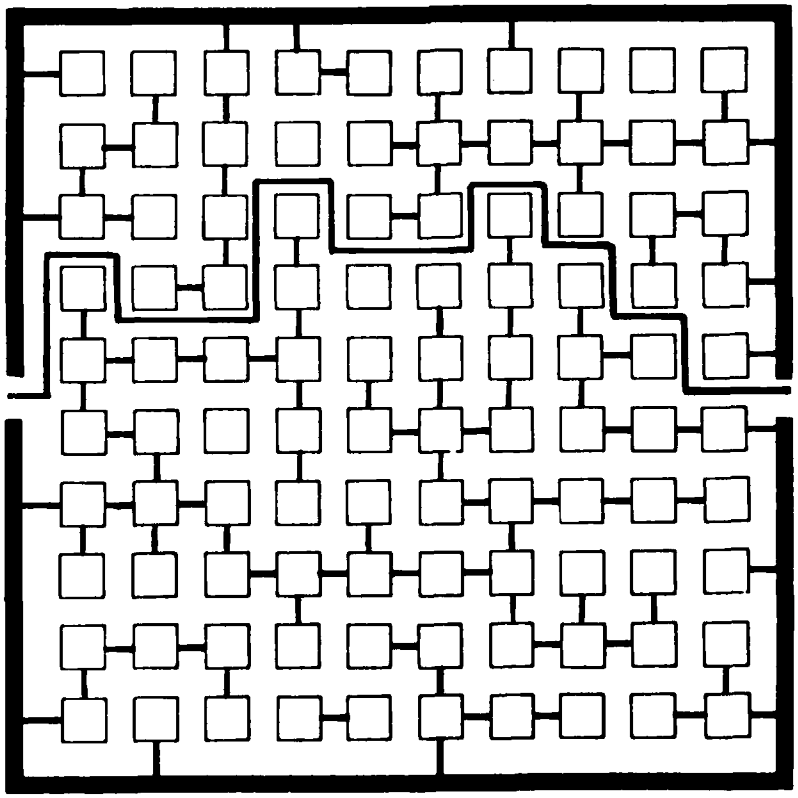
CODE HP

RELIABILITY .96

AVERAGE SCORE

SLOW 86

FAST 36



MAZE 8

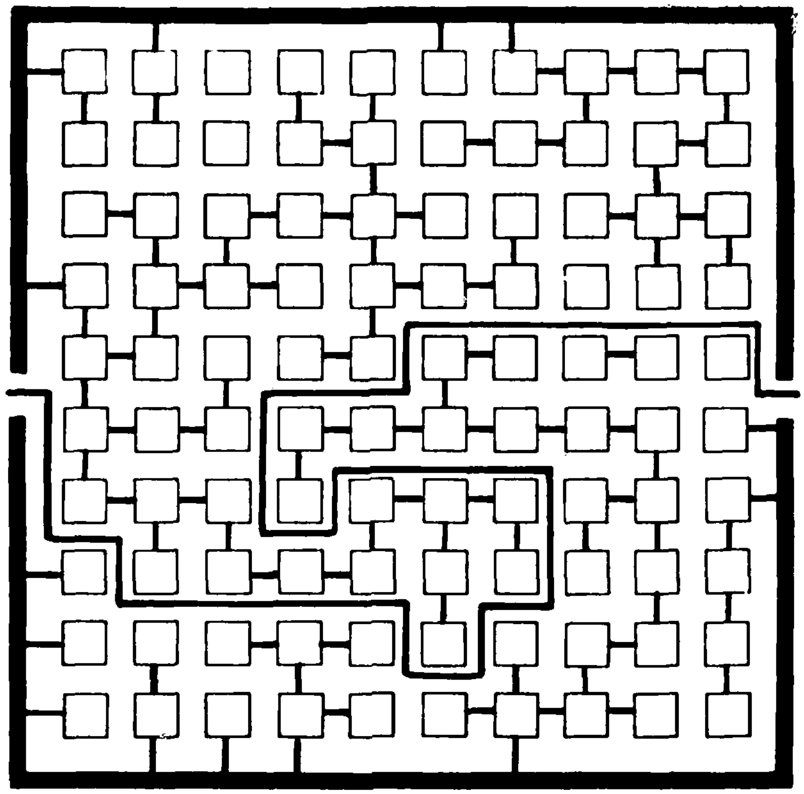
CODE HA

RELIABILITY .93

AVERAGE SCORE

SLOW 85

FAST 40



MAZE 9

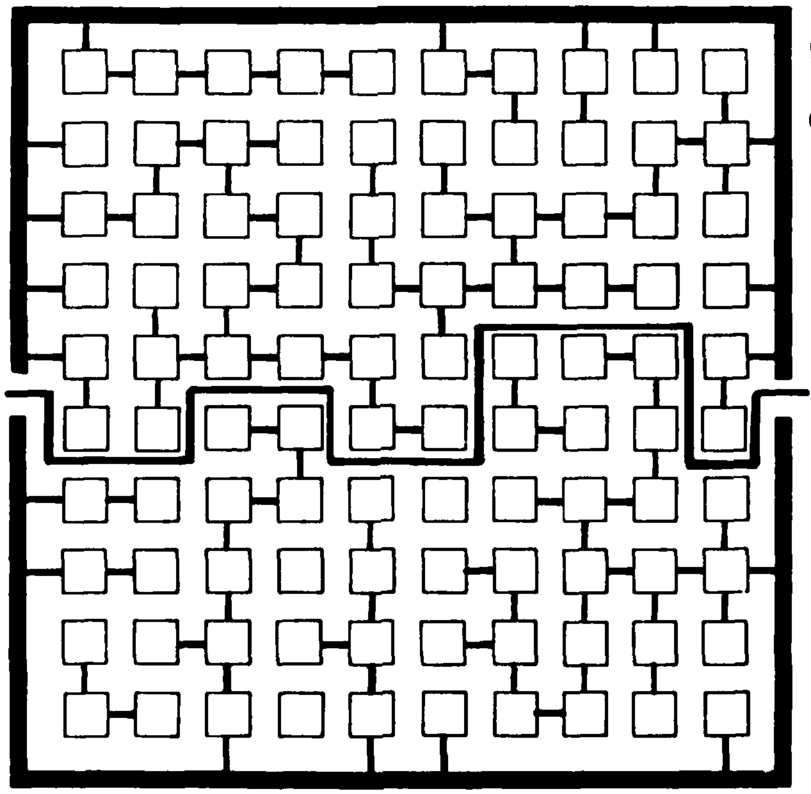
CODE HK

RELIABILITY .95

AVERAGE SCORE

SLOW 86

FAST 32



MAZE 10

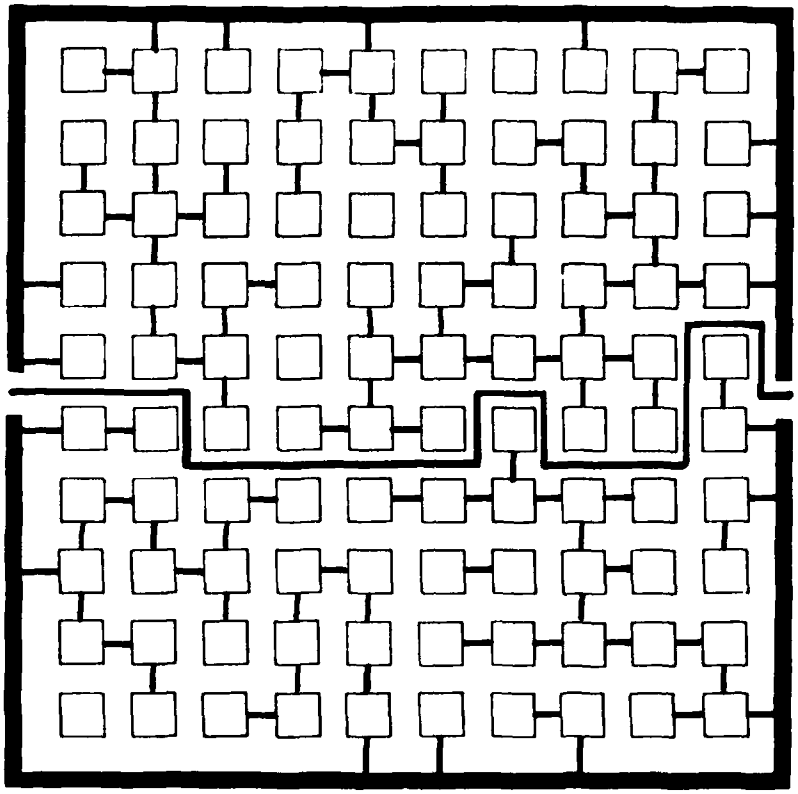
CODE HB

RELIABILITY .67

AVERAGE SCORE

SLOW 79

FAST 50



MAZE 11

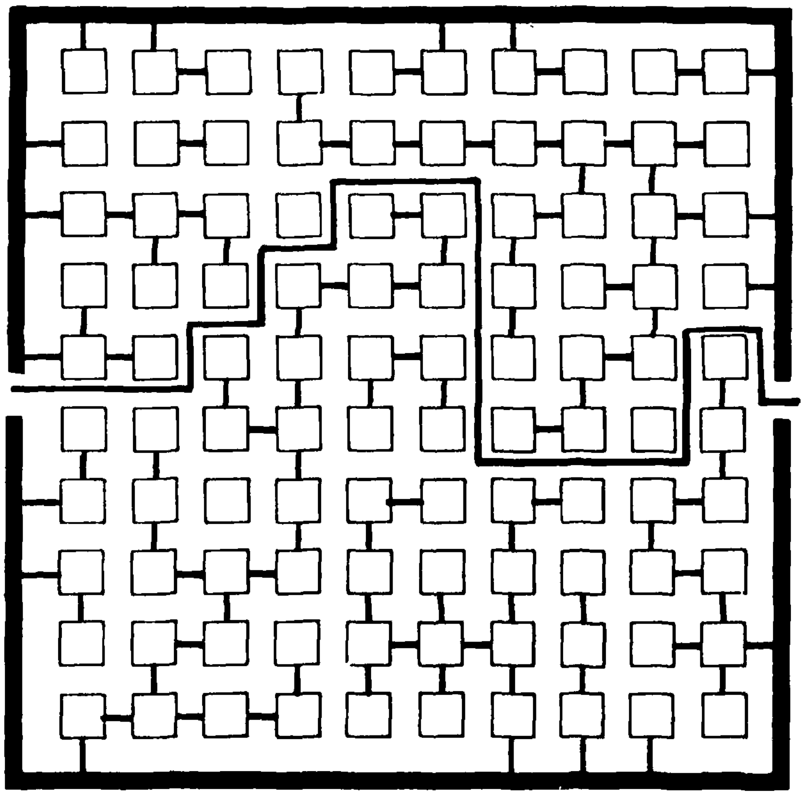
CODE HL

RELIABILITY .38

AVERAGE SCORE

SLOW 77

FAST 59



MAZE 12

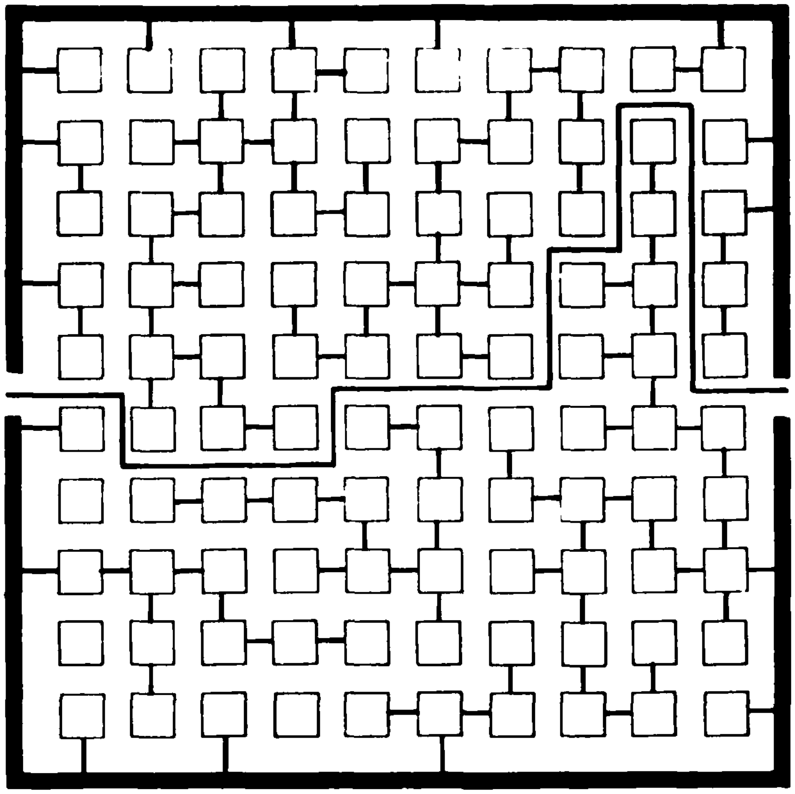
CODE HR

RELIABILITY .84

AVERAGE SCORE

SLOW 88

FAST 54



MAZE 13

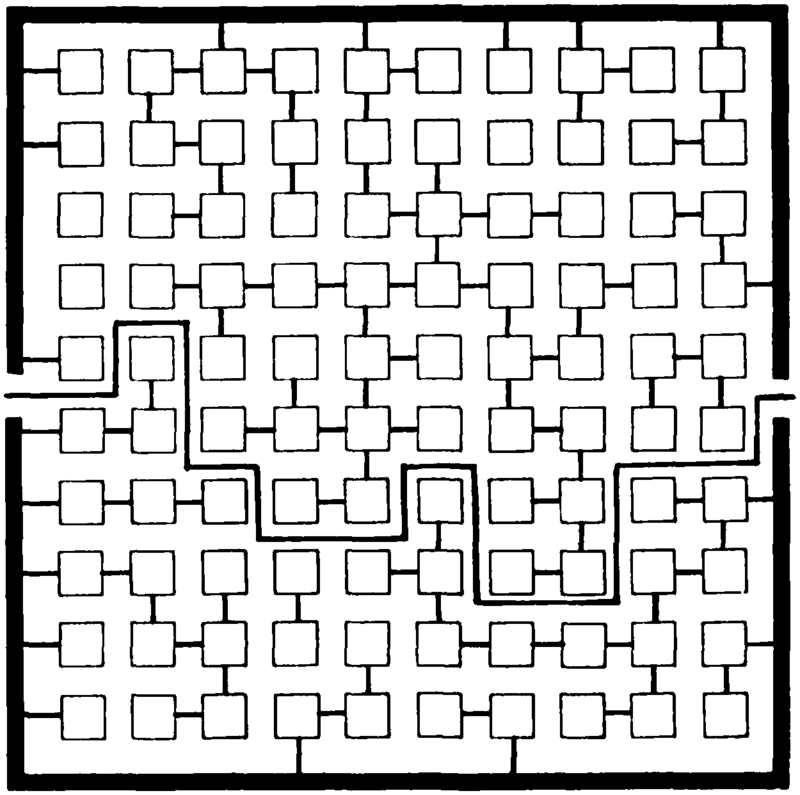
CODE HD

RELIABILITY .84

AVERAGE SCORE

SLOW 89

FAST 42



MAZE 14

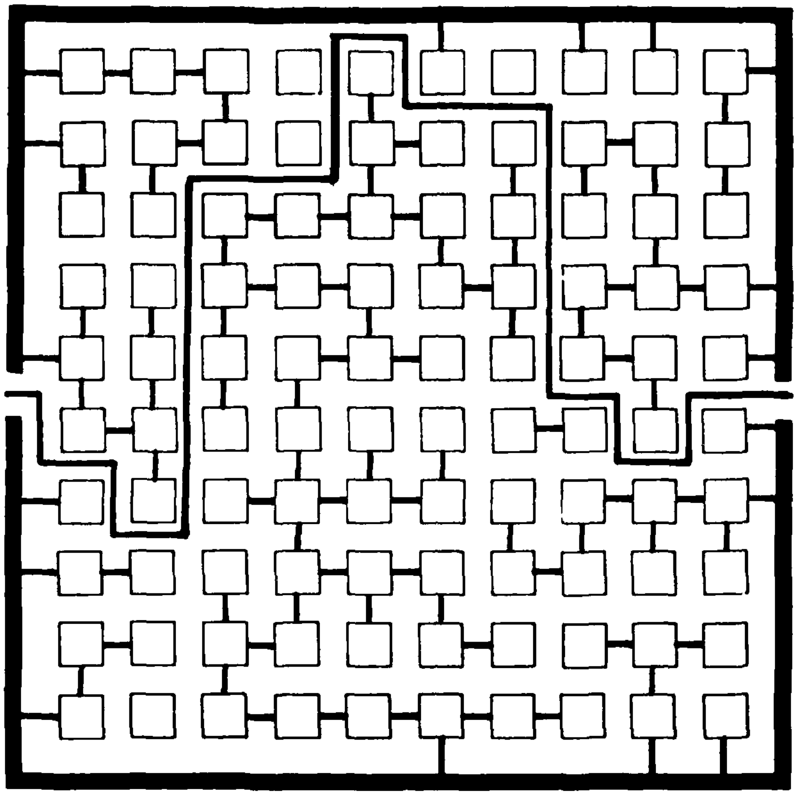
CODE HN

RELIABILITY .65

AVERAGE SCORE

SLOW 84

FAST 60



MAZE 15

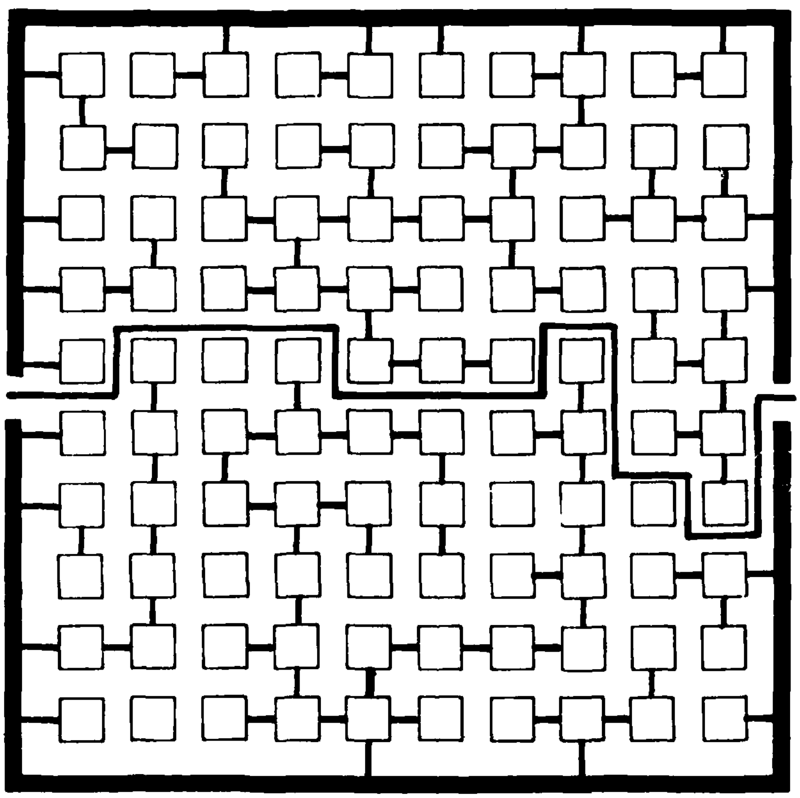
CODE HJ

RELIABILITY .88

AVERAGE SCORE

SLOW 88

FAST 51



MAZE 16

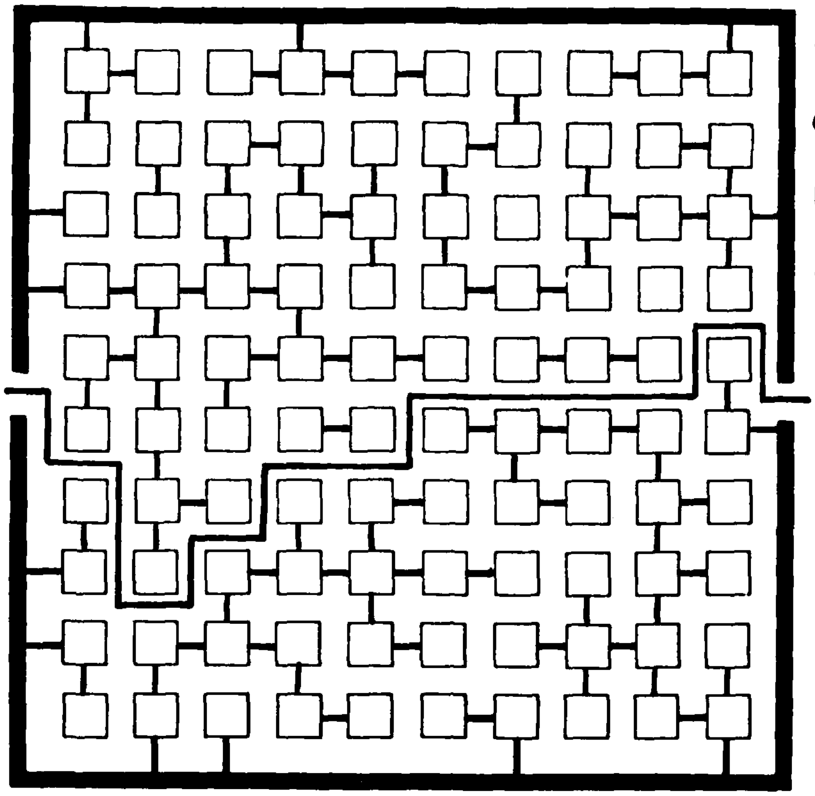
CODE HH

RELIABILITY .56

AVERAGE SCORE

SLOW 86

FAST 62



MAZE 17

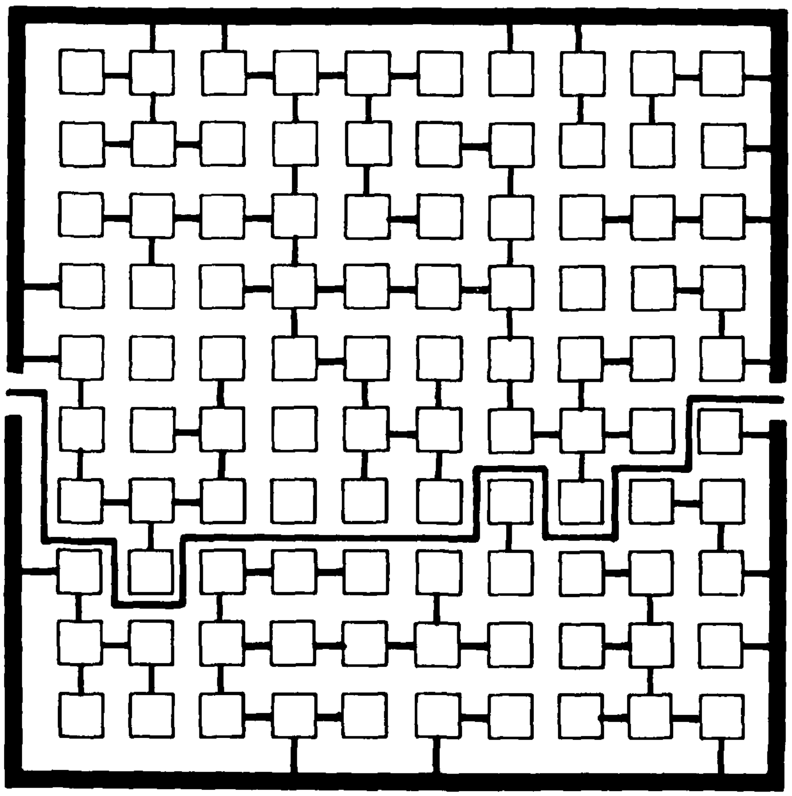
CODE HE

RELIABILITY .79

AVERAGE SCORE

SLOW 98

FAST 54



MAZE 18

CODE HM

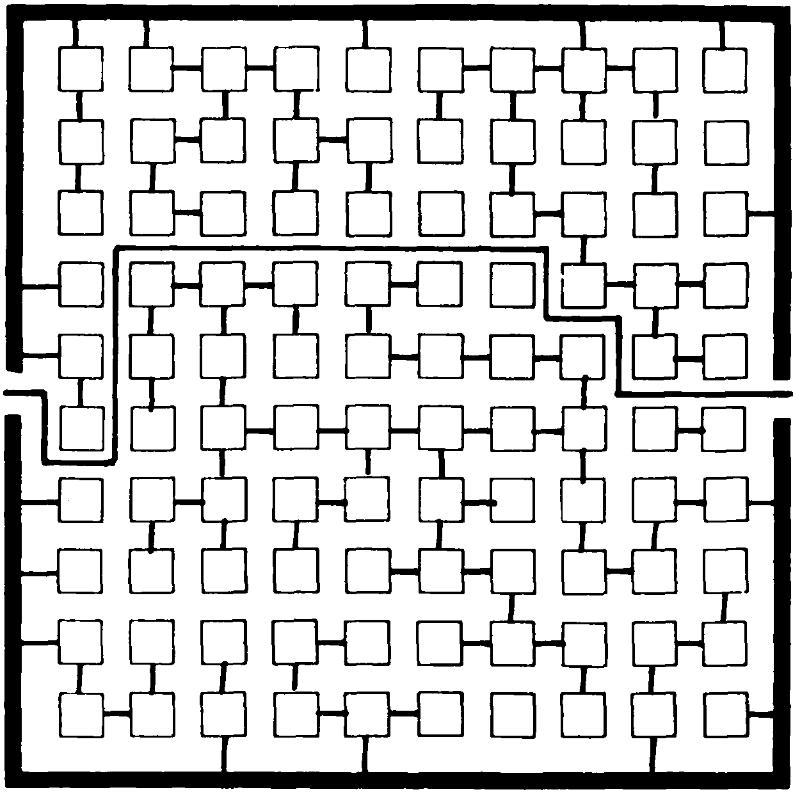
RELIABILITY .87

AVERAGE SCORE

SLOW 98

FAST 62

APPENDIX B
MAZES USED IN PART 3



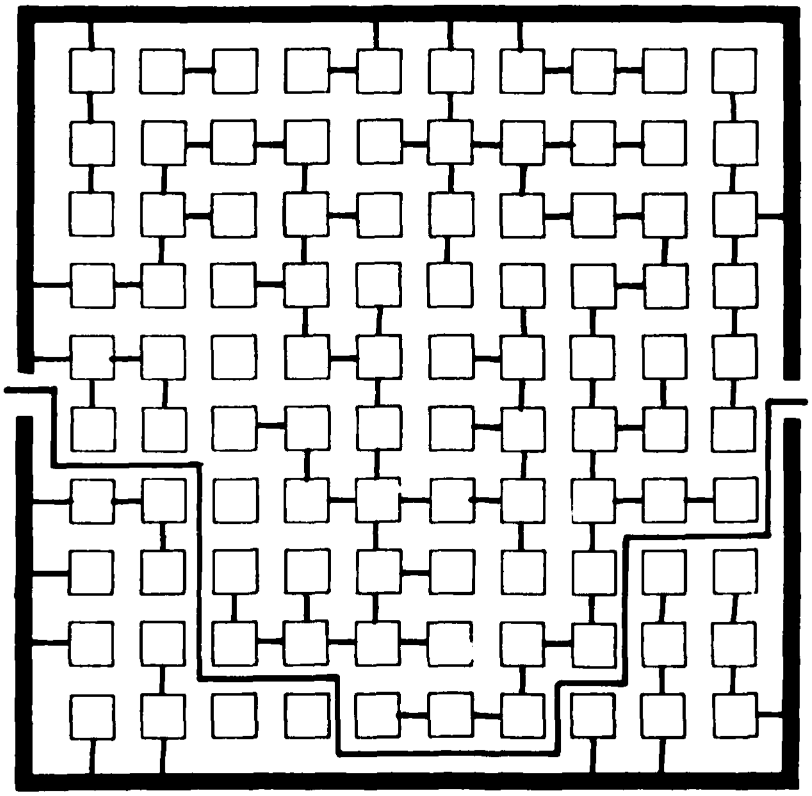
MAZE 1

CODE XQ

RELIABILITY .86

AVERAGE SCORE

SLOW 83
MEDIUM 45
FAST 26



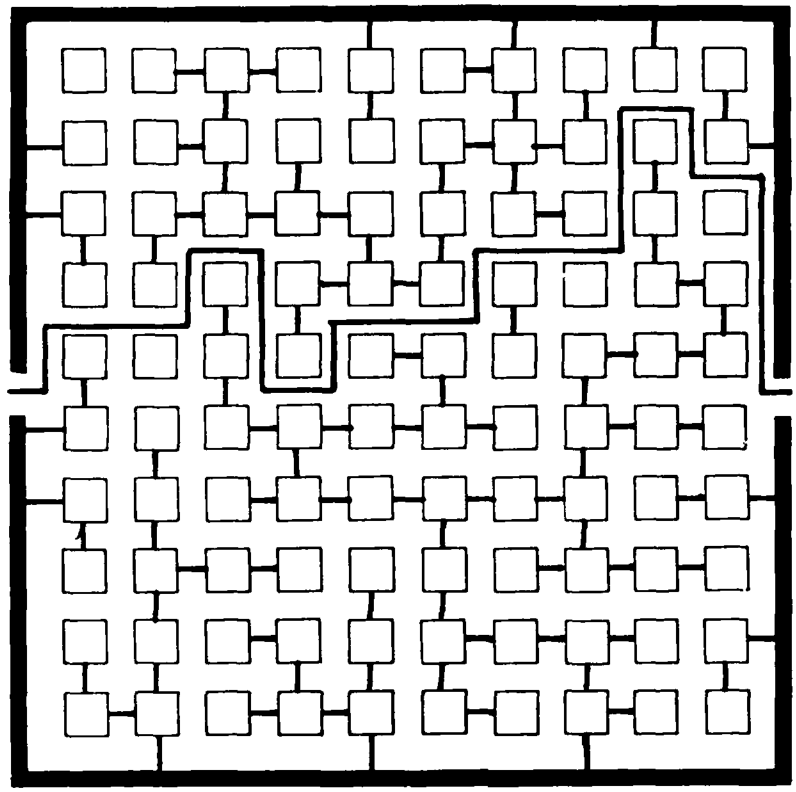
MAZE 2

CODE XF

RELIABILITY .76

AVERAGE SCORE

SLOW 72
MEDIUM 61
FAST 30



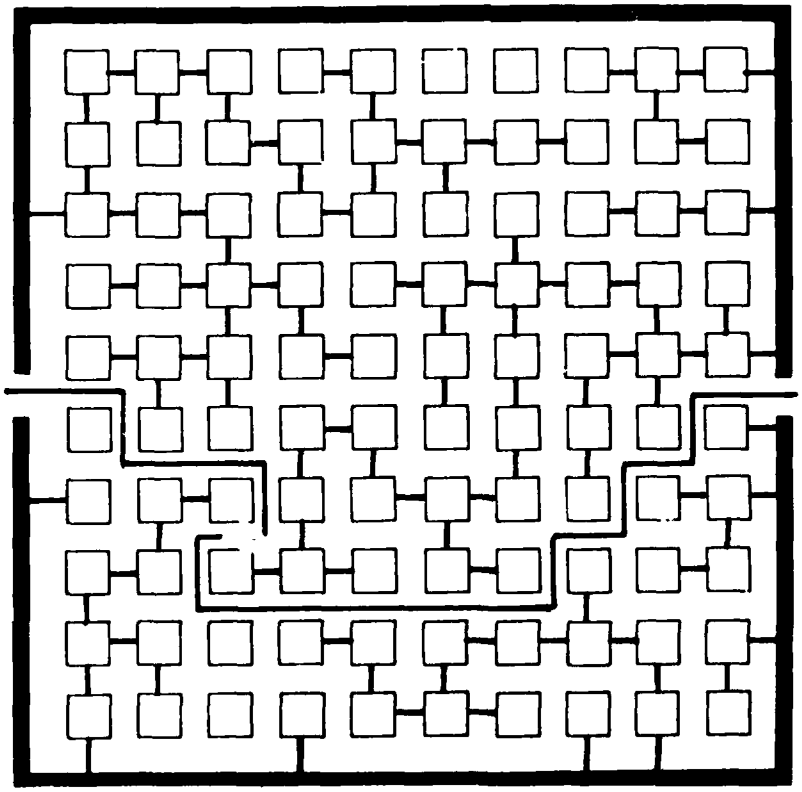
MAZE 3

CODE XU

RELIABILITY .60

AVERAGE SCORE

SLOW 81
MEDIUM 54
FAST 47



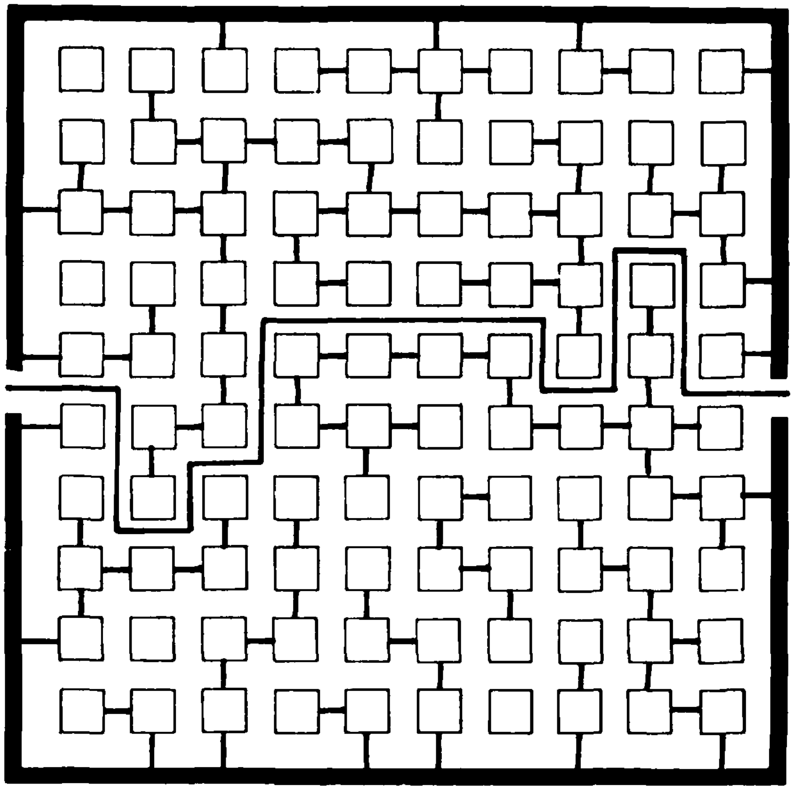
MAZE 4

CODE XG

RELIABILITY .43

AVERAGE SCORE

SLOW 78
MEDIUM 64
FAST 53



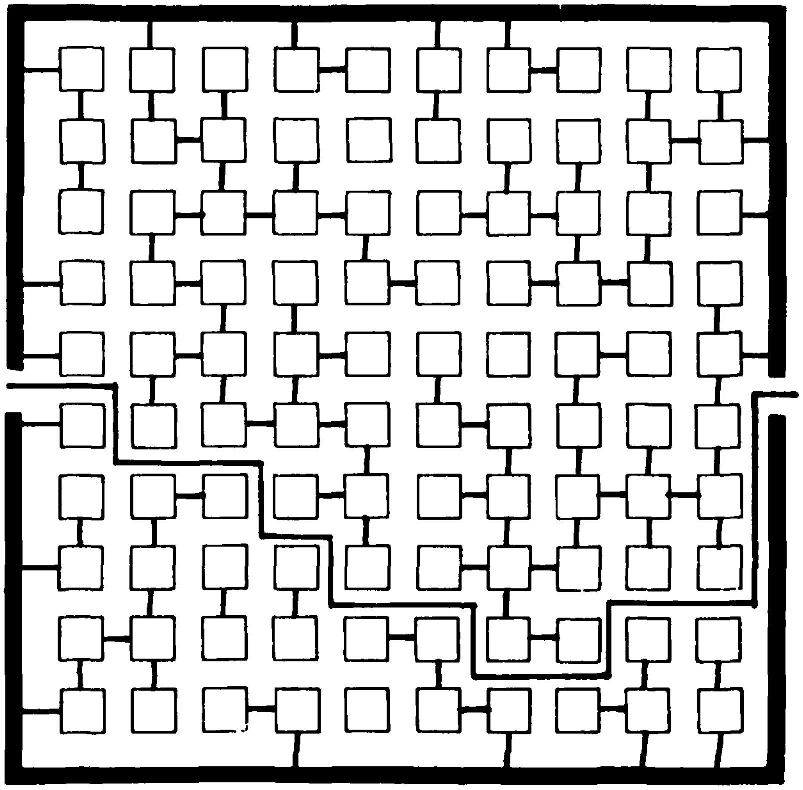
MAZE 5

CODE XS

RELIABILITY .72

AVERAGE SCORE

SLOW 85
MEDIUM 70
FAST 41



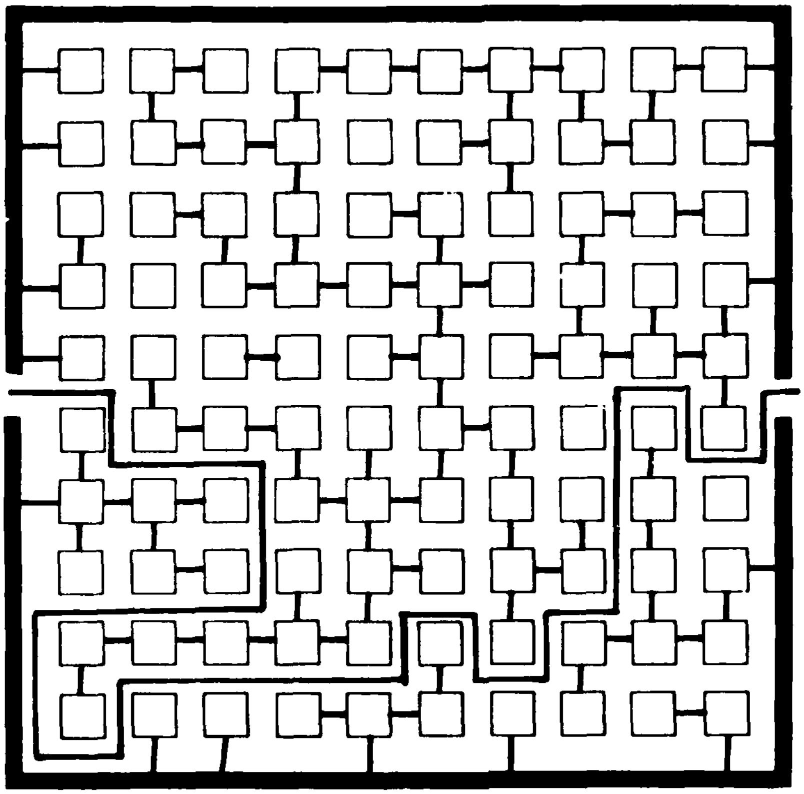
MAZE 6

CODE XE

RELIABILITY .91

AVERAGE SCORE

SLOW 92
MEDIUM 72
FAST 33



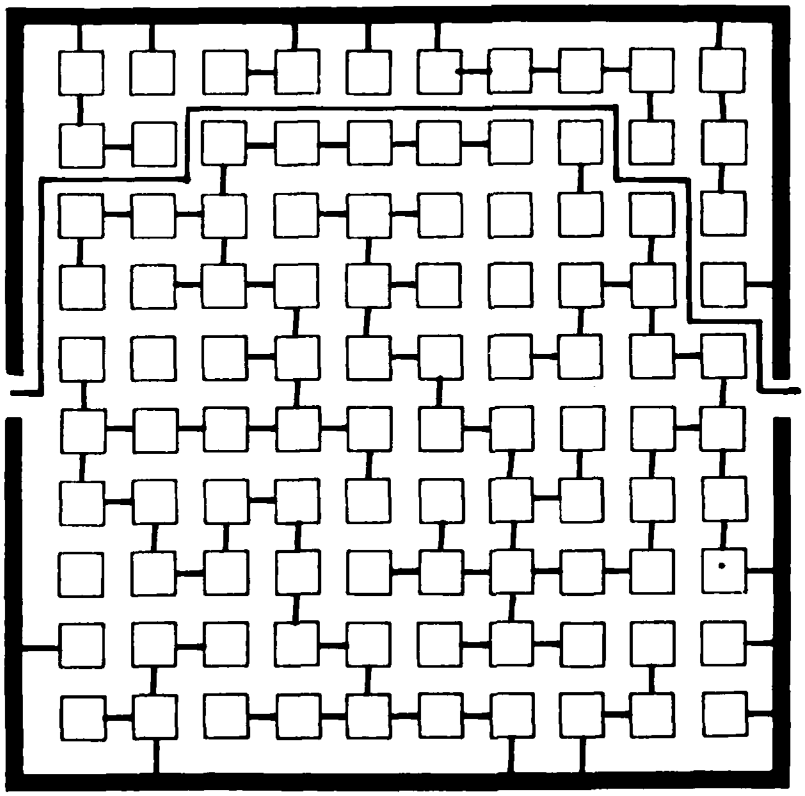
MAZE 7

CODE XW

RELIABILITY .39

AVERAGE SCORE

SLOW 71
MEDIUM 75
FAST 61



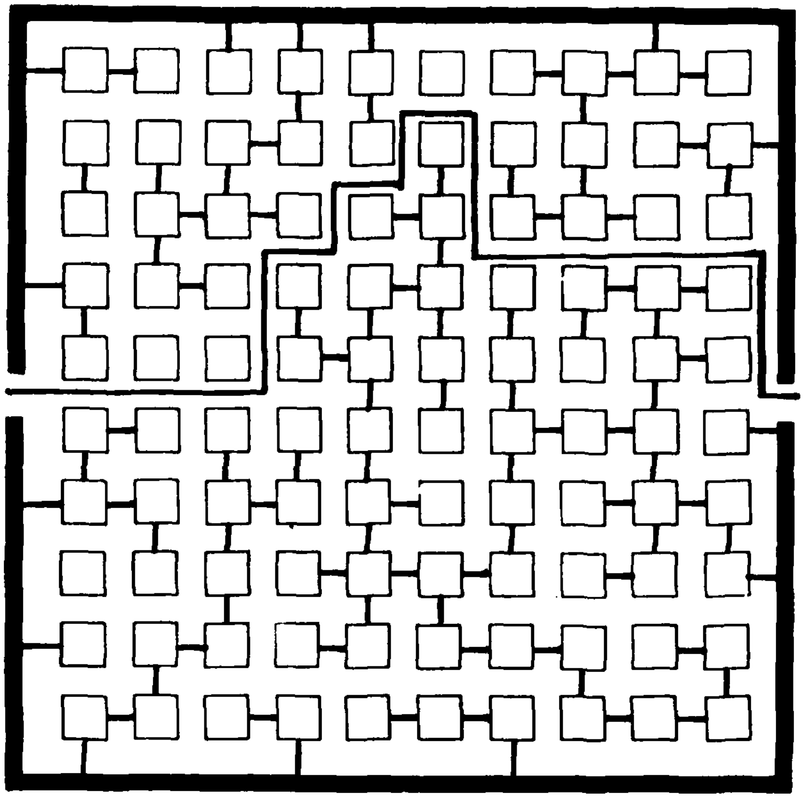
MAZE 8

CODE XD

RELIABILITY .76

AVERAGE SCORE

SLOW 88
MEDIUM 70
FAST 50



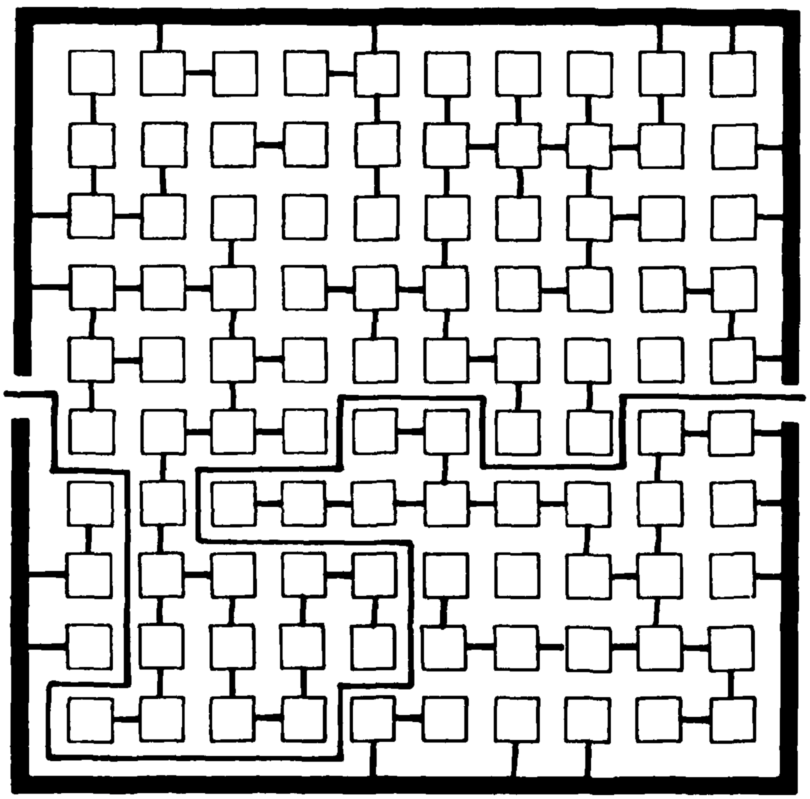
MAZE 9

CODE XO

RELIABILITY .43

AVERAGE SCORE

SLOW	79
MEDIUM	80
FAST	52



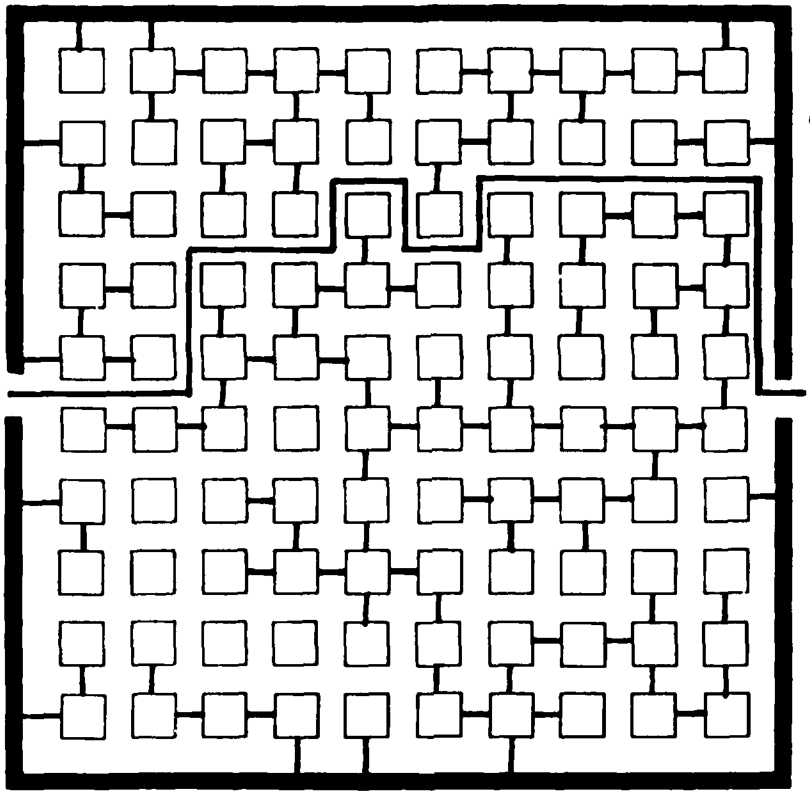
MAZE 10

CODE XX

RELIABILITY .78

AVERAGE SCORE

SLOW	90
MEDIUM	80
FAST	40



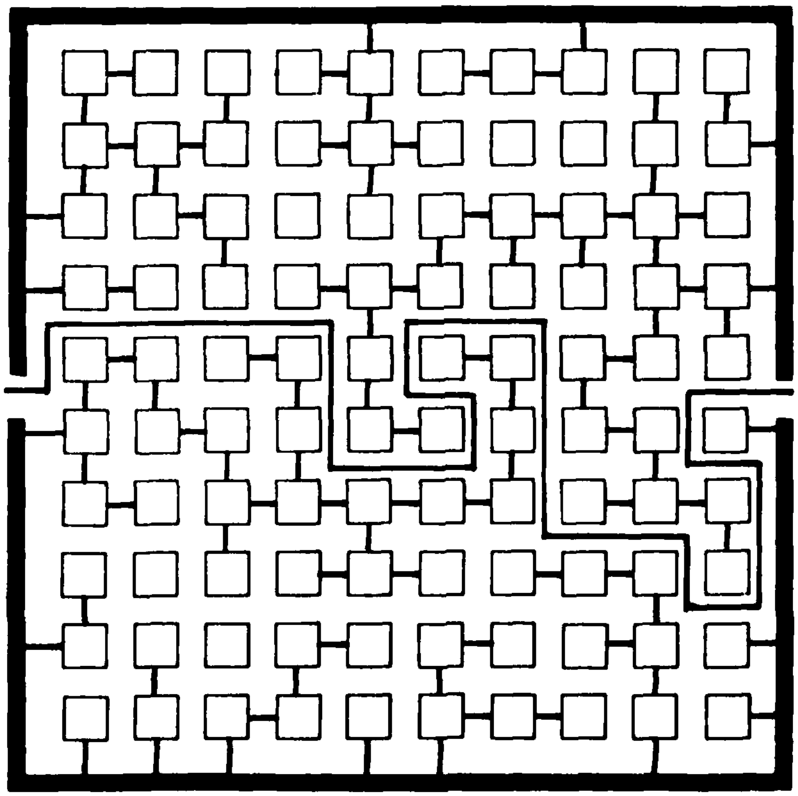
MAZE 11

CODE XP

RELIABILITY .77

AVERAGE SCORE

SLOW 93
MEDIUM 75
FAST 47



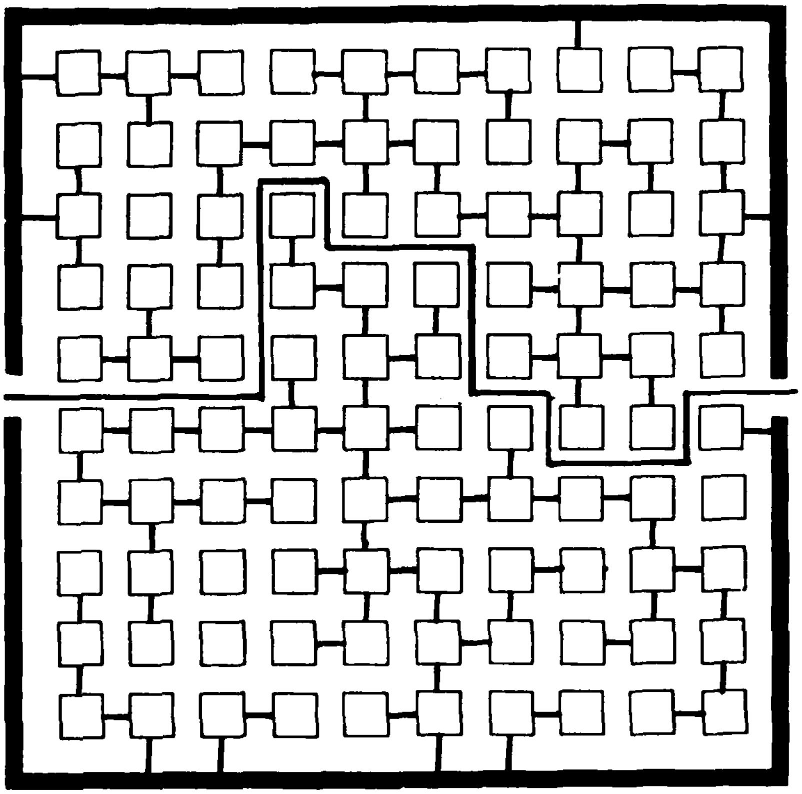
MAZE 12

CODE XR

RELIABILITY .65

AVERAGE SCORE

SLOW 99
MEDIUM 68
FAST 51



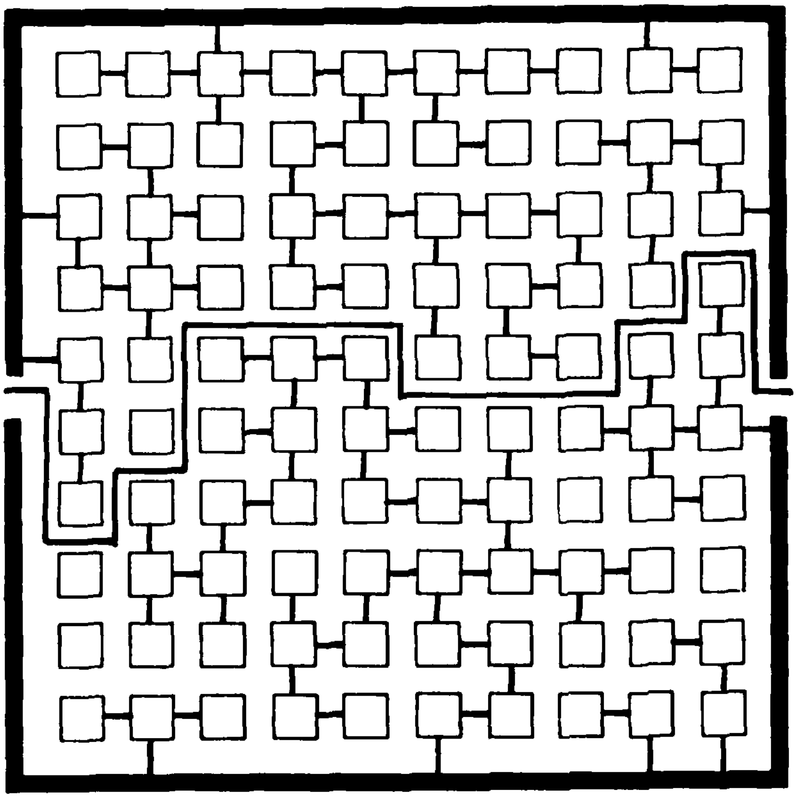
MAZE 13

CODE XC

RELIABILITY -.34

AVERAGE SCORE

SLOW 54
MEDIUM 84
FAST 80



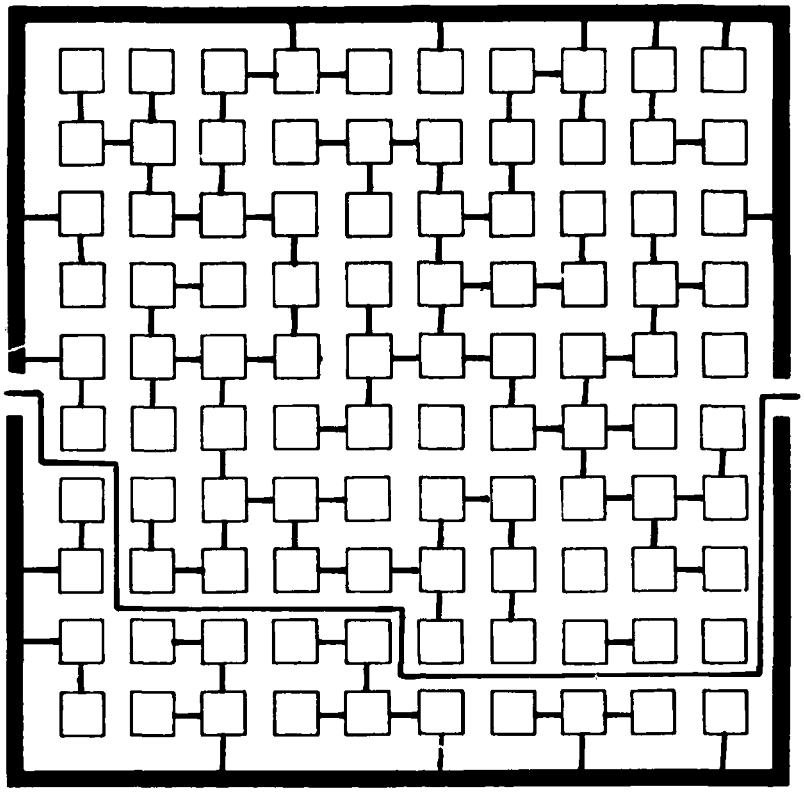
MAZE 14

CODE XK

RELIABILITY .62

AVERAGE SCORE

SLOW 87
MEDIUM 82
FAST 54



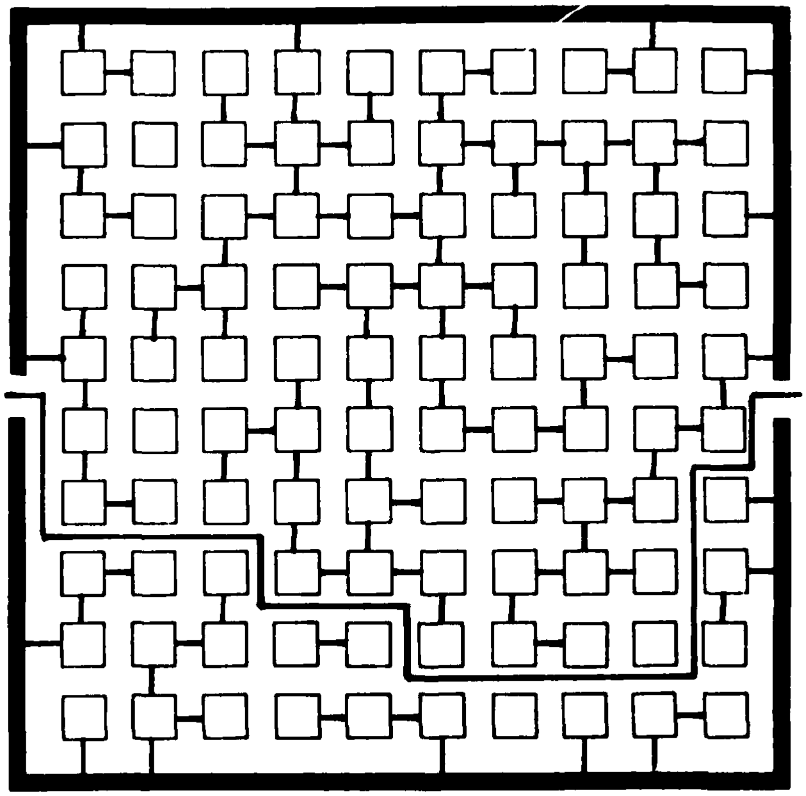
MAZE 15

CODE XH

RELIABILITY .87

AVERAGE SCORE

SLOW	97
MEDIUM	92
FAST	46



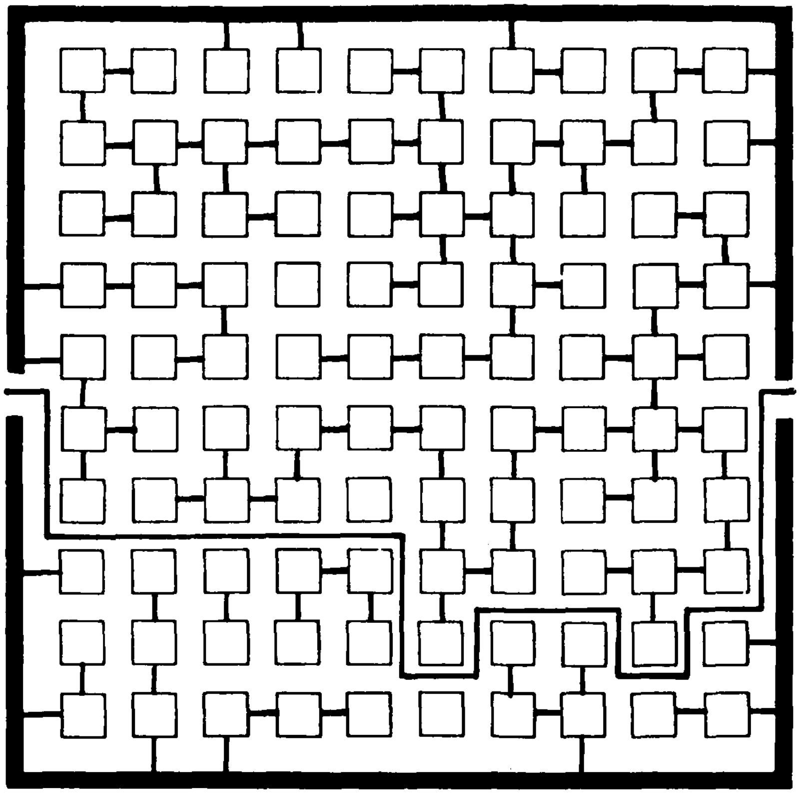
MAZE 16

CODE XT

RELIABILITY .66

AVERAGE SCORE

SLOW	89
MEDIUM	99
FAST	57



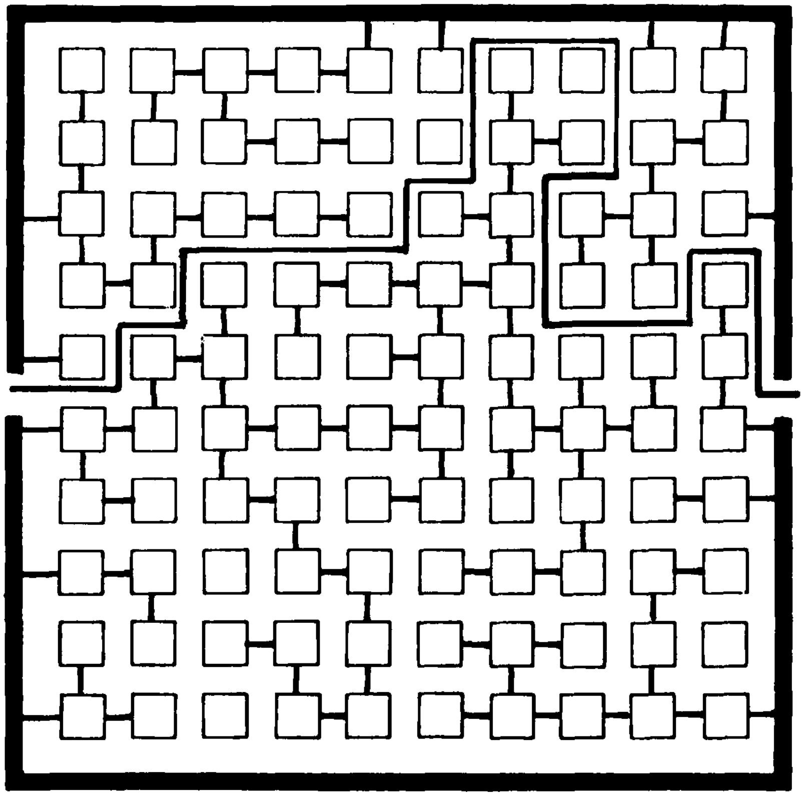
MAZE 17

CODE XM

RELIABILITY .60

AVERAGE SCORE

SLOW	87
MEDIUM	93
FAST	68



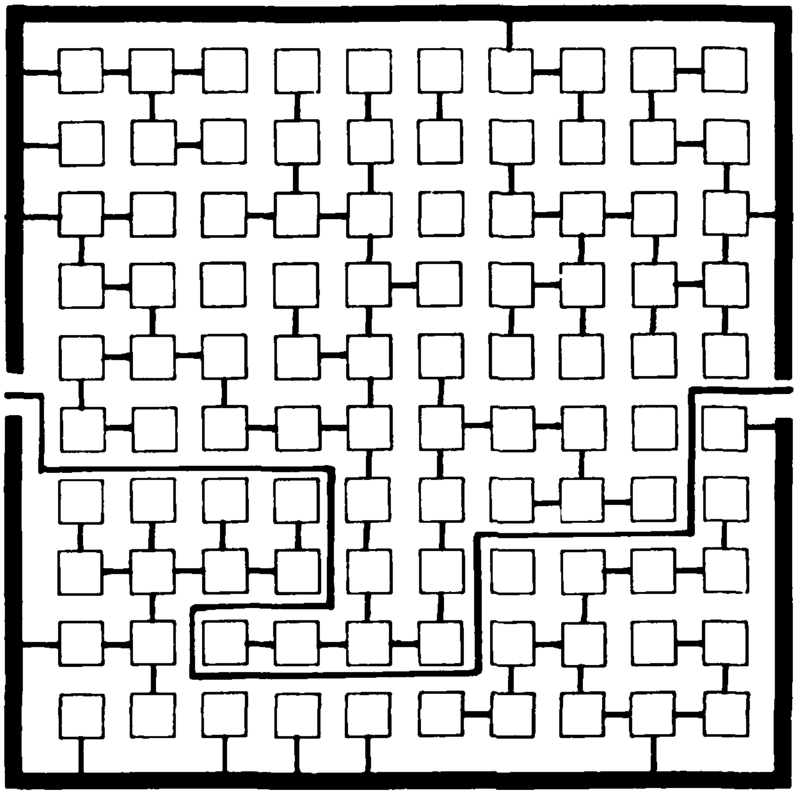
MAZE 18

CODE XV

RELIABILITY .66

AVERAGE SCORE

SLOW	96
MEDIUM	91
FAST	67



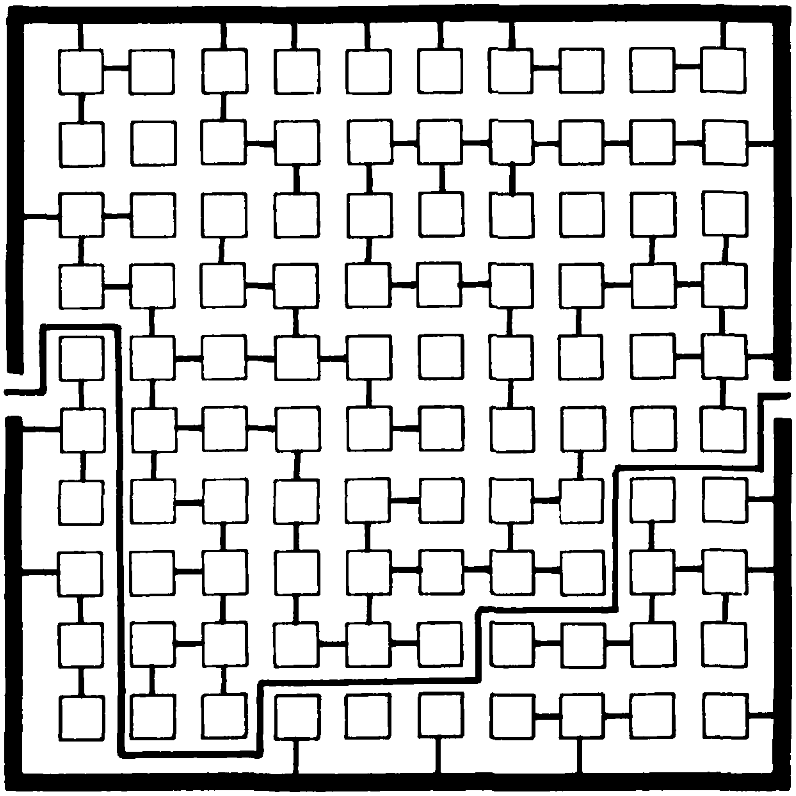
MAZE 19

CODE XB

RELIABILITY .60

AVERAGE SCORE

SLOW	94
MEDIUM	95
FAST	74



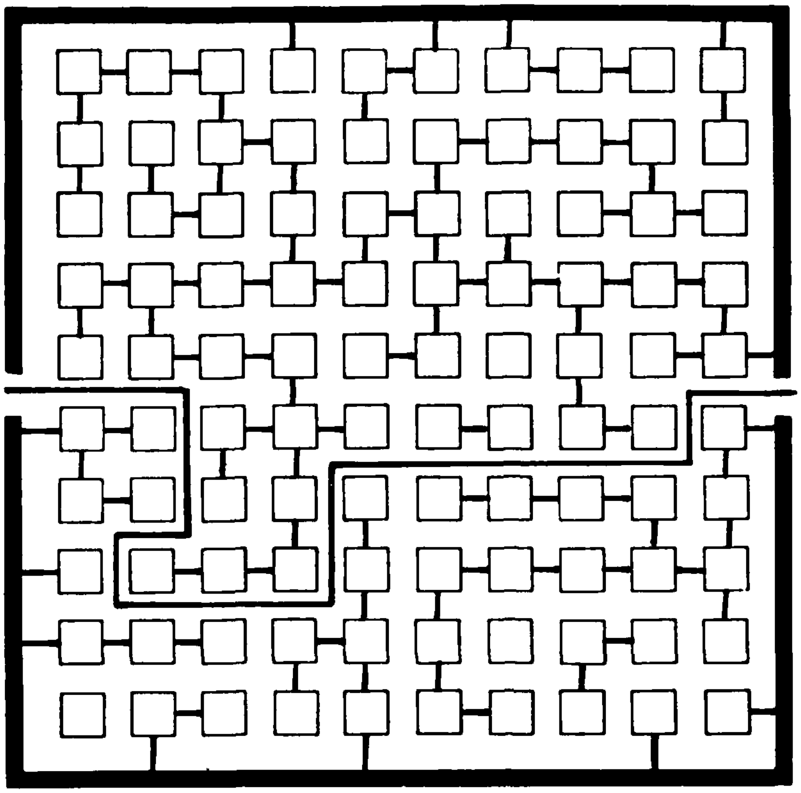
MAZE 20

CODE XN

RELIABILITY .71

AVERAGE SCORE

SLOW	96
MEDIUM	94
FAST	78



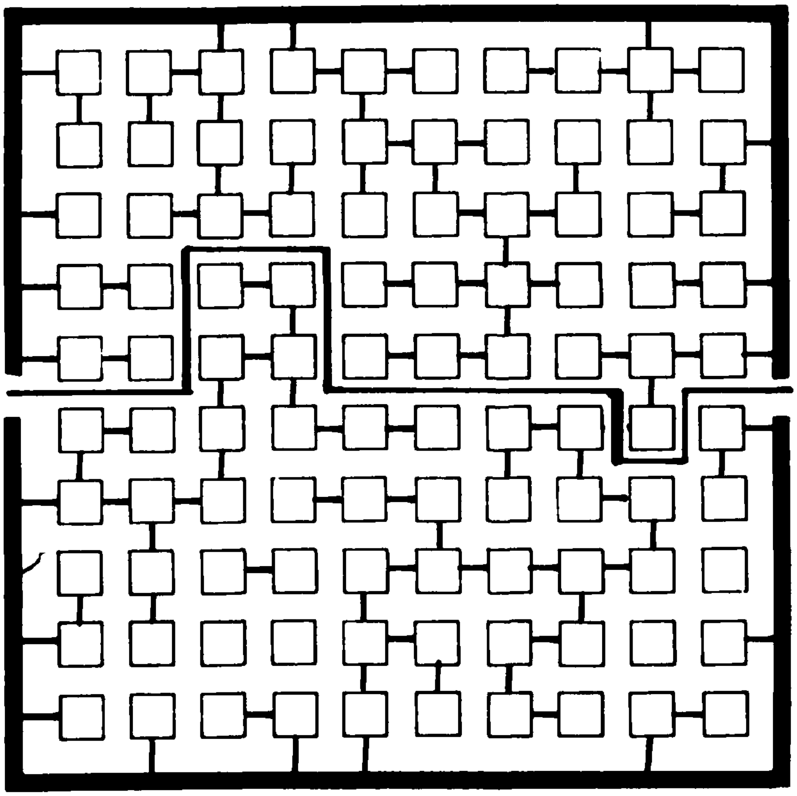
MAZE 21

CODE XI

RELIABILITY -.21

AVERAGE SCORE

SLOW 98
MEDIUM 71
FAST 98



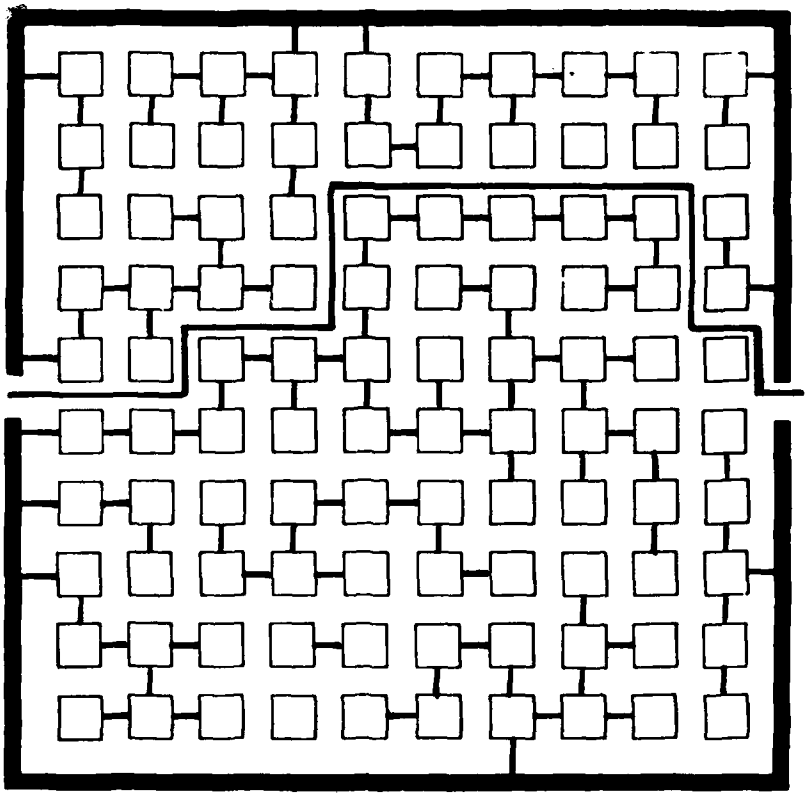
MAZE 22

CODE XA

RELIABILITY .70

AVERAGE SCORE

SLOW 99
MEDIUM 99
FAST 81



MAZE 23

CODE XL

RELIABILITY .33

AVERAGE SCORE

SLOW 99
MEDIUM 96
FAST 95

APPENDIX C

Sample Size Calculations

Case 1: Full Factorial Design

Design: A x Maze x Subject

Number of Levels: a x m x s

Factor A fixed, factors Maze and Subject random.

Significance test for factor A, assuming that the A x Subject interaction is not significant:

$$F = MS_A \div MS_{A \times Maze}$$

$$df = (a-1), (a-1)(m-1) = \upsilon_1, \ \upsilon_2$$

Then the minimum detectable difference between the two levels of factor A is

$$\Delta = \sqrt{\frac{2a}{ms}} \, \sigma \phi(\upsilon_1, \upsilon_2)$$

where

$$\sigma \approx \sqrt{MS_{AxMaze}}$$

and

$\phi(\upsilon_1, \upsilon_2)$ is obtained from the Pearson and Hartley charts for power of the F-test, using degrees of freedom $\upsilon_1 = a-1$ and $\upsilon_2 = (a-1)(m-1)$

Case 2: Partially Nested Design

Design: A x B x Maze x Subject,
maze nested within Factor B (a different set of mazes used for each level of Factor B)

Number of Levels: a x b x m x s

Factors A and B fixed, Factors Maze and Subject random.

Significance tests, assuming that A x Subject, B x Subject, and A x B x Subject interactions not significant.

a. Test for A:

$F = MS_A \div MS_{A \times M(B)}$
$df = (a-1), (a-1)b(m-1)$

b. Test for B:

$F = MS_B \div MS_{M(B)}$
$df = (b-1), b(m-1)$

c. Test for A x B:

$F = MS_{AxB} \div MS_{AxM(B)}$
$df = (a-1)(b-1), (a-1)b(m-1)$

Then the minimum detectable difference between two levels of a given factor are:

a. Factor A:

$$\Delta_A = \sqrt{\frac{2a}{bms}} \quad \sigma_{AxM(B)} \quad \phi(\upsilon_1, \upsilon_2)$$

where

$$\sigma_{AxM(B)} \approx \sqrt{MS_{AxM(B)}}$$

and

$\phi(\upsilon_1, \upsilon_2)$ as above, in Case 1, with $\upsilon_1 = a-1$ and $\upsilon_2 = (a-1)b(m-1)$

b. Factor B:

$$\Delta_B = \sqrt{\frac{2b}{ams}} \quad \sigma_{M(B)}\, o(\upsilon_1, \upsilon_2)$$

where

$$\sigma_{M(B)} \approx \sqrt{MS_{M(S)}}$$

and

$\phi(\upsilon_1, \upsilon_2)$ as above with $\upsilon_1 = b-1$ and $\upsilon_2 = b(m-1)$

c. Interaction A x B:

$$\Delta_{AB} = \sqrt{\frac{2ab}{ms}} \quad \sigma_{AxM(B)} \quad \phi(\upsilon_1, \upsilon_2)$$

where

$$\sigma_{AxM(B)} \approx \sqrt{MS_{AxM(B)}}$$

and $\phi(\upsilon_1, \upsilon_2)$ as above with $\upsilon_1 = (a-1)(b-1)$ and $\upsilon_2 = (a-1)b(m-1)$

REFERENCES

Porteus, S.D., **Porteus Maze Test: Fifty Years' Applications,** Pacific Books, Palo Alto, California, 1965.

Poturalski, R.J. and S.L. Ward, **Maze-Solving as a Tool for Measuring Performance Under Stress,** Annual Meeting of the Aerospace Medical Association, May 1982.

Scheffe, H., **The Analysis of Variance,** John Wiley & Sons, New York, 1959.

Ward, S.L. and R.J. Poturalski, **Changes in Maze-Solving Errors Due to Stress,** Annual Meeting of the Human Factors Society, October 1982.

☆U.S. GPO: 1983 659-005/212