THE RELIABILITY OF THE PROBLEM BOX AS A METHOD FOR MEASURING THE LEARNING ABILITY OF THE RAT

By

WILLIAM T. HERON

Submitted to the Department of Philosophy and Psychology and the Faculty of the Graduate School of the University of Kansas in partial fulfillment of the requirements of the degree of Master of Arts.

Approved by

W. S. Hunter

Department of Philosophy and Psychology

Date June 2nd 1921.

TABLE OF CONTENTS

THE RELIABILITY OF THE PROBLEM BOX AS A METHOD FOR MEASURING THE LEARNING ABILITY OF THE RAT

INTRODUCTION

The data presented in this thesis were obtained in the psychological laboratory of the University of Kansas during the school year 1920-21. The work was suggested by and done under the direction of Professor Walter S. Hunter to whom the writer is grateful for advice and criticism.

The problem which forms the basis for the following work is one which has received much attention from those who are dealing with tests applied to humans. In view of this fact, it is rather curious that of all those who have applied tests of various kinds to animals, none have seen fit to inquire into the reliability of the tests used. Certainly this question of reliability is one of extreme importance if the results obtained in testing are to be considered valid sources from which to draw conclusions.

The purpose of the present experiment, then, is to determine the degree of reliability of the problem box as a test of the learning ability of the white rat. A comparison will also be made of maze and problem box learning, and other problems essential to the study of reliability will be discussed.

The meaning of reliability may be illustrated by an

example given by Spearman[1]. He says, "suppose three balls to be rolled along a well-kept lawn; then the various distances they go will be almost perfectly correlated to the various forces with which they are impelled. But let these balls be cast with the same inequalities of force down a rough mountain side; then the respective distances eventually attained will have but faint correspondence to the respective original momenta." In the first case the distance the balls traveled would be a reliable test of the amount of force with which they were impelled because the chance factors influencing their progress along the lawn would not be sufficient to destroy the correspondence between the impelling force and the distance traveled. In the second case, however, these factors are so great that they destroy the correspondence and thus the distance traveled is rendered unreliable as a measure of the forces involved.

The theoretical application of this illustration to the learning process in the problem box is plainly seen. The hypothetical capacity, called the learning ability of the rat, would correspond to the impelling force. The problem box is represented by the lawn or mountain-side whichever the case may be. The measure which is used to indicate the animals' progress in learning to operate the box is analogous to the distance traveled by the balls. In order to test the reliability of the problem box, therefore, the following assumptions must be made:

first, that this hypothetical capacity, learning ability, exists; second, that it is a constant factor with no essential variations within the experimental period; third, that the solution of the problem box involves this constant factor to a degree not essentially obscured by other factors involved in problem box learning.

That these three are reasonable assumptions to make is indicated by the fact that they have not, up to the present time, been questioned by anyone who has used the apparatus and animals in question. The present experiment is particularly concerned with investigating the validity of the third assumption. The first two are assumed to be true although the effect which they would have, if untrue, upon the investigation will be discussed.

HISTORICAL SUMMARY

The problem box, in one form or another, has been in use for years. Thus Watson[2] says, "The problem box method in animal behavior is as old as civilization. It remained for Thorndike, in 1899, to perfect the method and make it serviceable for bringing out certain facts connected especially with what one roughly calls 'learning in animals.'" The form of the box necessarily varies greatly because of the different animals used. It is obviously impossible for the rat to operate the same kind of a box as would be suitable for the monkey,

consequently, each experimenter has varied the boxes which he has used to suit his animal's capabilities.

The problem box has also been used in human testing. Healy's[3] puzzle box is probably the one in most common use. The lid of this box is fastened by a combination of strings and catches , and it can only be opened by following a definite sequence of five or six steps. The box is so constructed that the whole procedure to be learned can be studied by direct observation. Healy states the purpose of this test as follows: "It may bring out abilities or defects in manipulative powers, in the ability to analyse a slightly complicated physical situation, in powers of attention and continuity of effort."

The inclined plane box which is the one used in this experiment has been used in its simpler form, by Watson[4], Richardson[5], and Ulrich[6]. In its improved form it was used by Basset[7] in his work on the effect of brain weight upon learning ability. Inasmuch as it is the purpose of this experiment to investigate the apparatus rather than the problems of the above investigators, it is not necessary that their papers be reviewed.

The reliability of the maze used in this experiment is the subject that was investigated by Miss Maupin. Her work was done in the same laboratory as the present experiment, and she has made practically the same approaches to the problem as are made in this study.

The methods which have been used by Spearman[8], Burt[9], Wyatt[10], Krueger[11], Simpson[12], and others in their investiga-

gations of the reliability of tests applied to humans, are outlined by Whipple[13] as follows:

"Let A1-the 1st. series of observations of the trait A.

A2-the 2nd " " " " " " A.

Now if the outcome of the test is not disturbed by chance or constant errors, then the ranks of the several S's , should be the same in both. The constant errors, of course, are not due to the test and must be controlled otherwise. "If, however, chance errors are too obstrusively present, this fact will be revealed by a low correlation between A1 and A2. In practice a test whose coefficient of reliability is less than .60 or .70 is in need of rectification". ---"It should be understood that A1 and A 2 need not be independent series of tests given by different E's at different sittings, but may be made up from data obtained at a single sitting, though, as a rule, two sets of data are secured and the correlation is calculated between the first half of the first performance added to the last half of the second performance and the last half of the first performance added to the first half of the second performance"

"The point is that, in figuring the coefficient of reliability.....it is essential that the discrepancies between the two series of measurements of the same trait should really be of a 'chance' character. Suppose, for example,

that A represents memory for nonsense syllables and has been tested by requiring the subjects to memorize a dozen sets of syllables. It would be wrong, then to constitute A1 of the first six sets and A2 of the second six sets, because the latter half-dozen would be affected by a constant factor--that of practice--to an extent different from the first half-dozen. It would be better to constitute A1 from the odd and A2 from the even-numbered tests."

The correlations obtained by either of these two methods are known as the coefficients of reliability of the tests to which the correlations apply. The first of the two methods, i.e. the correlation of the results of the test obtained at two separate sittings, is the one most generally used. There are difficulties, however, in applying it to tests given to humans. This is indicated by Burt[14], in speaking of the calculation of the coefficient of reliability for the 'dotting test' which he gave to English school boys. He says, "It was not possible to obtain the Preparatory boys at the Laboratory (where this particular test was carried out) on more than one occasion, but that occasion was the same for every boy, viz. after morning school. Hence no reliability coefficient was obtainable for the Preparatory School, though the reliability was probably high. Each of the Elementary children was able to come twice. The reliability coefficient for the two series thus obtained for

them was .86. Part of the manipulation of the apparatus while the subject is actually at work devolves,---upon the superintendent of the experiment; so that the personal equation is in this test likely to be appreciable, though perhaps small; accordingly, as the series were all superintended by the same operator, the reliability coefficient is probably slightly too high."

For the sake of comparison the reliability coefficients for various tests applied to humans are enumerated in the following table:

TABLE I

COEFFICIENTS OF RELIABILITY FOR HUMAN TESTS

Name of test.	Investigator.	Coefficient of Reliability
1. Ebbinghaus Completion test	Simpson	.92
	Spearman and Krueger	.76
2. Hard Opposites	Simpson	.97
3. Memory of words	"	.73
	Spearman and Krueger	.92
4. Easy Opposites	Simpson	.93
5. Cancellation of A'S	Simpson	.72
6. Memory of Passages	"	.90
	Winch	.65 .68

Name of Test	Investigator	Coefficient of Reliability
7. Adding	Simpson	.91
	Spearman and Krueger	.76
	Burt	.50
	Brown:	
	Speed	.68
	Accuracy	.30
8. Geometrical forms	Simpson	.90
9. Learning pairs	"	.93
10. Scroll	"	.76
11. Recognizing forms	"	.40
12. Completing words	"	.92
	Brown	.70
	Burt	.68
	Burt and Moore	.58
	Wyatt	.89
13. Drawing lengths	Simpson	.72
14. Estimating lengths	"	.48
15. Mirror drawing	Burt and Moore	.52
16. Immediate memory	Burt	.70-.93
	Wyatt	.75-.76
	Abelson	.74-.81
17. Invention of stories	Whipple	.50
18. Word building	Wyatt	.88
19. Analogies	Burt	.71

All the tests in the above table may be considered reliable, according to Whipple's standard, with the exception of Recognizing Forms, Estimating Lengths, Mirror Drawing, and the Invention of Stories where the coefficients lie between .40 and .52. It may not be amiss to point out that all of these tests, with the exception of the last, are based on muscular co-ordination. It is, therefore, possible that the reliability of a test will be found to be inversely proportional to the amount of muscular co-ordination involved in it.

APPARATUS AND GENERAL METHOD

The problem box which was used in this experiment is similar to that employed by Basset and described by him in his paper on"Habit Formation in a Strain of White Rats with less than Normal Brain Weight". The measurements of the box were exactly the same as those which he used in constructing his apparatus. It was found necessary, however, to make some slight modifications in the incline plane and the electro-magnetic attachment. The plane which was first constructed according to Basset's directions did not appear to be delicate enough in action. It would sometimes operate and sometimes not, even though the rat appeared to place the same amount of weight upon the same point on the plane. Some trouble was also encountered in keeping the contact-points clean in order that sufficient contact would be made to operate the magnet. To obviate these difficulties the plane was reconstructed according to the diagram below. (Figure I.) A shows the support for the plane. It is simply a piece of brass bent at right angle at both ends, and screwed to the top of the table which the box is placed upon. B shows the plane which fits in between the upright pieces of A and is attached to them at the pivot C. This plane

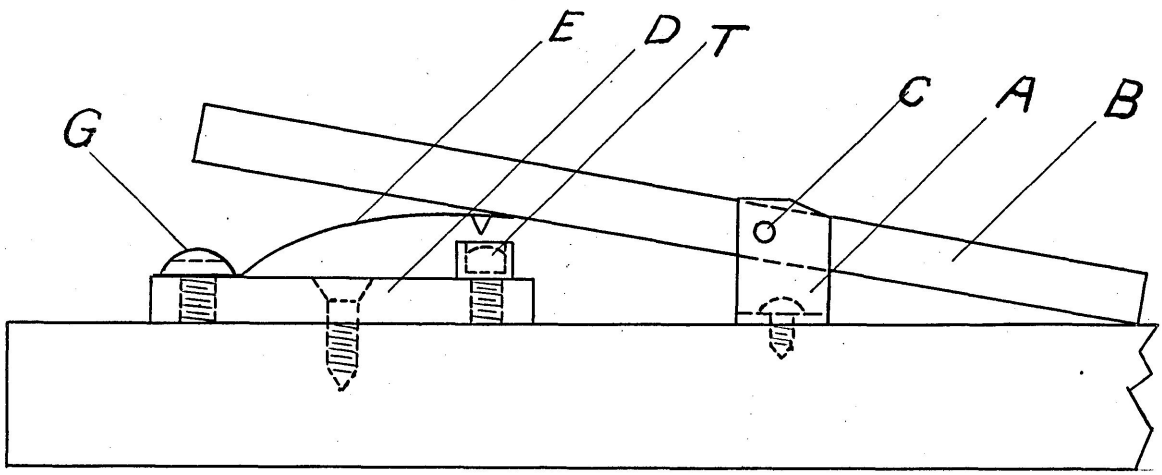


Fig I

Inclined Plane Apparatus.

is made of wood with 2 inches of it on the lower end and 3 inches on the upper. Thus the plane has a constant tendency to tip down and operate the contact apparatus. D is a piece of hard rubber 2 inches long, 3/16 inches thick, and 3/4 inches wide. E is a piece of hard spring copper of the same width and slightly longer than D. F is a screw with a large head into which a depression is bored for the purpose of holding a small amount of mercury. When E is pressed down, it makes contact with the mercury by means of the small point at its upper end. G is another screw which holds E to D and to which one wire is attached. The other wire is attached to screw F. Since B is not balanced a certain amount of the resilience of E is required to hold up the end of the plane resting upon it, but still it leaves plenty of clearance between the mercury and the contact point of E. This makes the plane fool-proof so that the proper amount of weight applied above the pivot C will invariably make a good contact and thus operate the magnet.

It was found in the electro-magnetic apparatus that the sharp metallic click of the armature against the core of the magnet when contact was made would almost invariably frighten the animal. This fright was often so great that an association would be quickly set up between

the plane and the noise so that the animal would avoid stepping on the plane. The character of the results as given below of the preliminary group of animals will demonstrate the seriousness of this factor. In order to remedy this objection, it was found necessary to cover the armature with a piece of soft rubber. This did not entirely eliminate the noise, but it deadened the metalic click so that the remaining noise did not seem to frighten the animal.

The control cage placed around the box was 24" x 28¼ x 10". It was always kept in a constant position with regard to the box. The animals were allowed to enter the enclosure of this cage from a release box in the same manner as usual in a maze. The ground-plan (Figure II) will show the relationship of the cage, release-box, incline plane, and problem box. X indicates the position of the food within the box.

Total time only was recorded for the problem box. It would be almost impossible to record accurately the distance traversed in this apparatus with a camera lucida attachment unless the animals were in some way prevented from climbing up the sides of the box. Static time could have been recorded but it was not thought practical to do so

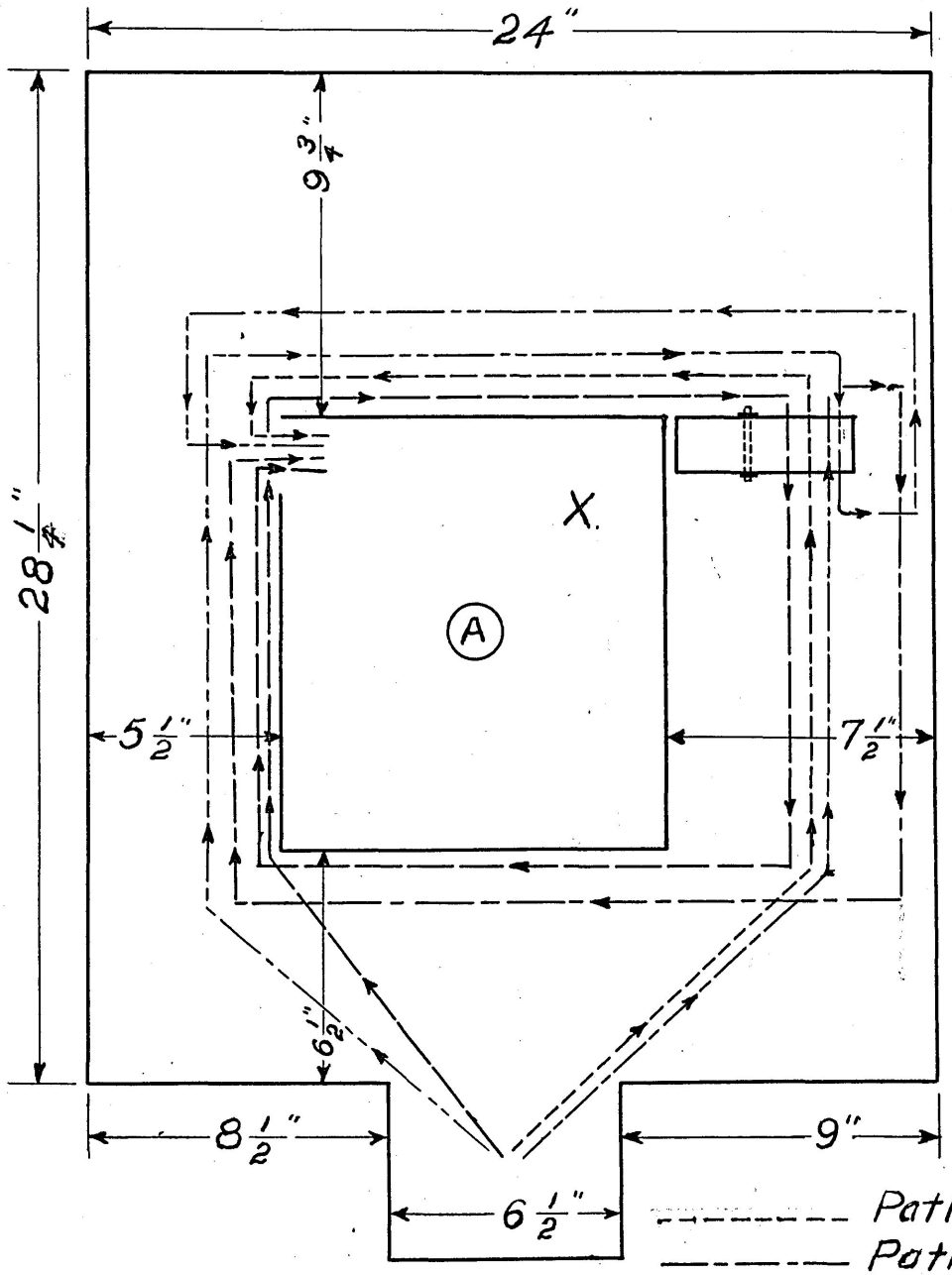


Fig. II

Ground plan of Apparatus.

because the animal seldom came to rest in the problem box (as he does in the maze).

The maze used in this problem was the Watson circular maze with the camera lucida attachments. The arrangement of the stops and openings are shown in the accompaning diagram (Figure III) of the ground plan of the maze. The true path is indicated by the broken line.

The method of using the maze did not differ from that generally used by other investigators. Each animal is released from the box R and the stop watch is started as soon as he enters with his whole body into the first alley. The watch is stopped as soon as he has his entire body within the food-box F. Static time is kept with a cumulative stop-watch. This watch is started each time that the animal stops in such a way as to be no longer traversing horizontal distance. Rearing up on the hind feet and examining the sides of the maze while keeping the hind feet on one spot was considered to be static time. The watch was stopped again as soon as the animal started to move in the horizontal plane. The static time was thus accumulated and by subtracting it from the total time recorded by the first watch, the sum of the net time was calculated.

The path which the animal traverses is recorded by

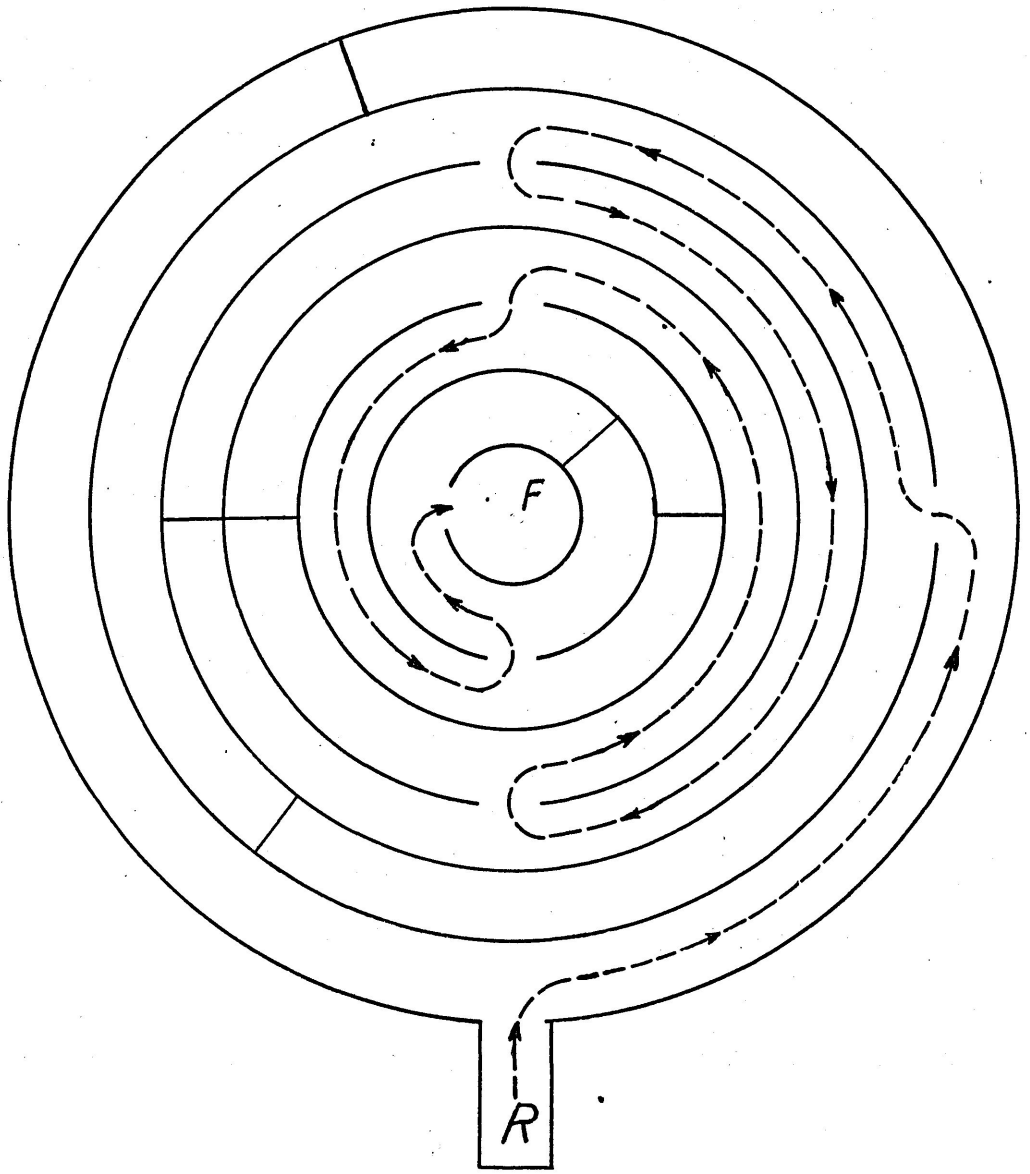


Fig III

Ground plan of Maze showing true path

R is release box.

F is feed box.

means of the camera lucida attachment. These tracings are later measured by means of a chartometer. The amount of surplus distance traversed can be computed by subtracting the distance required for a perfect run from the total amount of distance traversed.

SUBJECTS

The subjects used in this experiment were white rats of laboratory stock. They were divided into four different sets according to the training given them. These sets and the training given to each are as follows:

Set I. This set of ten rats will be known as the preliminary set because they were used by the experimenter for the purpose of practicing and perfecting his method. They were given completed learning by being run one trial per day until they had reached the criterion of perfection as described below. (See data and results.)

Set II. There were 28 animals in this set. They will be known as the total learning rats. They were not started in the experiment until the apparatus and method were perfected. They were given completed learning in the same manner as were Set I.

Set III. This set, composed of 22 animals, was given

six trials in the problem box, one trial per day. They were then allowed to rest for sixty days, at the end of which time they were run one trial per day for six days in the maze.

Set IV. The 20 rats in this set were run for six trials in the problem box and then given the same period of rest as Set III. At the close of this period they were again returned to the problem box for six more trials.

The age of the rats when they were started in the experiment varied between 50 and 70 days. Previous to being started in training all the animals were thoroughly tamed by being handled and fed by the experimenter. The animals were considered sufficiently tamed when they would allow the experimenter to pick them up without making an effort to escape. The animals, previous to their use in the problem box, were allowed to eat on the table within the control cage for several days. The problem box was removed at such times, but the incline plane was allowed to remain.

During the whole period of experimentation including the 60-day periods of rest the animals were cared for by the experimenter. Every effort was made to keep the animals in perfect physical health. The temperature of the room in which they lived was kept as constant as possi-

ble. The customary food of bread and milk was given the animals with an occasional variation of lettuce leaves and sunflower seeds. The cages were kept thoroughly cleaned and the drinking water was renewed daily.

The experimental work was done at night beginning at about seven o'clock. In cases where there was an exception to this rule, the variation was not more than three hours. No animals of any particular set were allowed to feed before the others. When an animal had finished his trial, he was placed in a box and kept there without food until the others of his set were all through and then all were fed together. During experimentation all animals were fed on a chair and no food was placed in their cages. They were allowed to eat to satiety while on the chair.

With the conditions thus controlled, it is believed that the following results are as representative as it would be possible to make them.

DATA AND RESULTS

TOTAL LEARNING

The degree of perfection to which the rat must attain before he is said to have arrived at completed learning is our first consideration. Perfection may be estimated in terms either of speed or of accuracy. Investigators lay emphasis upon one or the other by requiring either that the animal complete the trial in a certain amount of time or that he perform the operation without making any errors. In the present experiment it was decided to use the latter criterion of perfection for the reasons which are stated below. The rat was trained, therefore, until he could perform the operation of opening the box and entering it without making any false movements. Upon repeating such a performance three days in succession it was considered that he had completely learned the problem.

There were four possible paths which the animal might take in an errorless performance. These four are shown in Figure II. A record was kept of the path which the animal took in the performance of each errorless tiral. Although it was not necessary that the animal take the same path in each of the three perfect trials which con-

stituted the criterion of completed learning, yet in practice it was found that the animal invariably took the same route each time. It was also found that in the great majority of cases the animals took the route numbered three in the sketch (Figure II). This is probably due to the fixation of an association between food and the position of the door.

The criterion of accuracy, instead of speed, was chosen; first, because when the rat had attained the degree of accuracy required, he was performing the problem in a miminum of time. Second, this criterion is more comparable to that which is generally used in the maze. This reason is not important as far as the present work is concerned because the comparison which is here made between the maze and problem box did not involve completed learning records in either case. The third and last reason for using the criterion of accuracy rather than speed is, I believe, the most important of all. The latter criterion, as used by Basset, required that the animal complete the whole operation of opening the box and entering it within a period of four seconds, and, furthermore, the rat could not spend more than two seconds in going either from the entrance of the control cage to

the plane or from the plane to the entrance of the problem box. In other words, the rat might perform the operation in the allowed time of four seconds, but if he spent 2.5 seconds in going from the entrance of the control cage to the plane and 1.5 seconds in going from the plane to the entrance of the problem box, he would not make a perfect trial. Such a criterion seems to depend too much upon the physiological capacity of the rat to cover a given amount of distance in a given time. One animal may be making perfectly accurate movements yet because his rate of locomotion is slow, he is not considered to be making perfect trials. At the same time, another animal may be making less accurate movements but because his speed of moving is rapid, he is said to be making perfect trials. From this point on, the learning of the first animal is not concerned with solving the problem of accuracy but rather in increasing his rate of travel.

In the criterion used in the present experiment, the above physiological factor is ruled out. When the animal has learned to solve the problem, he is done and it is not necessary to give him training to increase his speed of locomotion.

Some difficulties, however, were found in consistently

carrying out the criterion of learning outlined above. Five of the 28 rats (Set II) formed a position habit of going around the box two times before stepping at the proper place on the plane and entering the box. These rats would almost invariably run too close to the box on the first time around. They would not, however, try to enter the door upon making this circle but would continue at full speed past it. All of these rats in Set II were given from 20 to 58 trials more than the average number of trials necessary for the rest of the rats to complete their learning according to the criterion. The average number of trials given to the rats with the position habit is 37 more than the average number of trials given to the others. Despite this these rats never broke themselves of the habit. They did, however, reduce their time until it was on a par with the rats who were making perfect trials, although they were, of course, traveling twice as far.

The question to be faced is what shall we do with the records of these rats in making our calculations. Because of the fact above mentioned that the rats reduced their time of travel to an equality with that of the other rats, it was decided that the criterion, upon the basis of which these records could be used in computations , would be three successive trials the average

time of which did not exceed the average time of the three perfect trials of the other rats. Using this as the criterion of completed learning with these rats, it is found that the average number of trials necessary for them to complete the learning is 56 which is 21 more than the number necessary for the rats using the other criterion. This gives some indication of the seriousness of the criticism made of an arbitrary time limit as a criterion of learning if the learning is to consist merely in learning the problem and not also in an increase in speed of muscular contraction.

The following tables give a summary of the records made by the various animals in learning the problem. Table I gives the records for the rats in the preliminary set. Table II gives the results for the rats which have been used in making the calculations which will follow. It was thought perfectly justifiable not to use the records of the preliminary set, because of the change of method and apparatus which was made during their learning period.

TABLE I

Summary of data of the preliminary Set

Rat No.	Total trials for completed learning.	Total time for completed learning.	Average time per trial.	Maximum time for one trial.	Minimum time for one trial
1	40	7201	180.0	4847	3
2	39	7205	184.7	5640	3
3	43	5382	125.1	490	5
4	28	8151	290.8	2107	3
6	37	1552	41.9	325	4
7	24	5090	212.0	3000	7
8	32	2190	68.4	515	3
15	47	14008	298.0	4020	5
16	21	4655	221.1	1738	4
17*	58	6587	113.5	1734	4
Totals*	312	55434	1622.2	22682	37
Averages*	34.6	6159.3	180.2	2520.2	4.1
Medians*	37	5382	184.7	2197	4
Totals	370	62021	1735.7	24416	41
Averages	37	6202.1	173.5	2441.6	4.1
Medians	38	5984.5	182.3	1922.5	4

*Rat 17 developed a position habit. Computations marked (*) do not include his records. The following computations include his data.

TABLE II

Summary of data of Total Learning Rats

Rat No.	Total trials for completed learning.	Total time for completed learning.	Average time per trial.	Maximum time for one trial.	Minimum time for one trial.
18	54	1737	32.1	200	3
20	43	771	17.9	135	3
24	66	2160	32.7	550	3
25*	57	2737	41.6	859	3
27	40	1384	34.6	199	3
29*	61	2751	45.1	519	3
31	70	1779	25.4	420	2
32	17	4963	291.9	2346	5
34	12	2764	230.3	1155	4
35	47	1950	41.5	501	2
36	42	4096	97.7	1404	2
37	36	2808	78.0	523	4
38	10	1010	10.1	526	6
39	13	2609	200.7	791	8
40*	42	1161	27.6	259	3
42	28	1222	43.6	226	2
43	28	1331	47.5	646	3
44	22	1620	73.6	353	3
45	45	1093	25.4	391	2
46*	67	3229	48.3	593	2

TABLE II (cont.)

Rat No.	Total trials for completed learning.	Total time for completed learning.	Average time per trial.	Maximum time for trial.	Minimum time for one trial.
47	28	2285	81.6	826	4
48	18	1048	58.2	380	4
49*	53	2061	38.8	542	3
51	64	1122	17.5	130	2
52	39	1401	35.8	284	4
53	21	1527	72.7	460	4
54	26	1287	49.5	310	5
55	39	2113	54.2	1270	3
Totals*	808	44080	1694.1	14885	84
Averages*	35.1	1916.8	73.6	647.1	3.6
Medians*	36	1620	47.5	460	3

*Rats marked (*) developed a position habit. Computations marked (*) do not include their records. The following computations include their data.

Totals	1088	56069	1853.9	16798	95
Averages	38.8	2003.1	66.2	599.9	3.4
Medians	39.5	1758	44.3	510	3

The difference, which the changes in apparatus and method made, can be readily noticed upon reference to Tables I and II. The total number of trials required and the median of these trials do not show much variation between the two sets; in the time required there is great difference. Thus, for the preliminary set the median for the total amount of time required for learning is 5984.5 seconds while in the other set it is less than one-third as much, being 1758 seconds. The median of the average times per trials in the first set is 182.3 while for the second group it is 44.3 which is only about one-fourth as much. Again the median of the maximum time required for one trial in the first group is 1922.5 seconds while for the second group it is only 510 seconds. Since these rats were all from the laboratory stock there is no reason to believe that these differences are owing to any other factor than the change in method and apparatus.

Table III shows the results of dividing the time for the rats, not including the preliminary set, into tenths according to the method whic h will be described later. This method of division, as suggested by Miss Vincent, is as follows: If the rat has required 27 trials to complete the learning, then the first tenth will be the time for the first two trials and .7 of the time for the third; the second tenth will be the remaining .3 of the third trial plus trials four and five and .4 of trial six. This process is continued until all of the 27 trials are used. If the rat requires only ten trials to complete the learning, then the time for every trial would be one-tenth. If he requires seven trials, then one tenth would be constituted from .7 of the time of a trial.

The learning curve for these rats made by plotting minutes against trials does not deviate greatly from the ordinary type of learning curve. The only variation noticeable is the great rise at the second trial. Such a marked rise does not generally occur in learning curves, although there may be a slight increase. I believe that this rise may be attributed to the fact that the food is always, in the case of the problem box, in such a position that it may be sensed by the animal. This being the case, the first time the animal is placed in the box

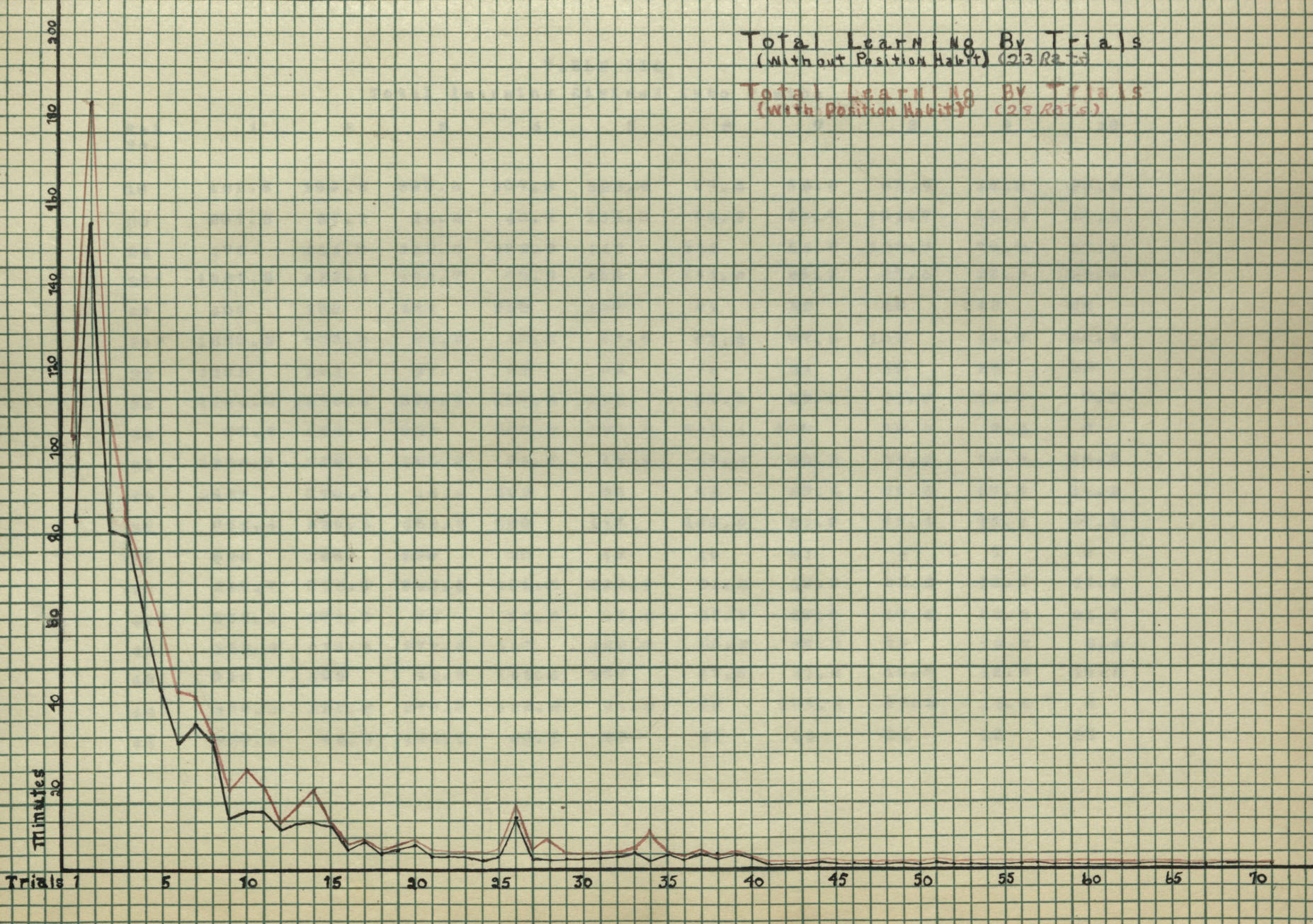

Total Learning By Trials
(Without Position Habit) (23 Rats)
Total Learning By Trials
(With Position Habit) (25 Rats)
Minutes
200
180
160
140
120
100
80
60
40
20
Trials 1
5
10
15
20
25
30
35
40
45
50
55
60
65
70

TABLE III

Total learning divided into Tenths

Rat No.	1	2	3	4	5	6	7	8	9	10
18	422.8	199.2	397.6	107.8	323.6	87.4	86.6	35.5	45.8	30.6
20	207.5	37.1	14.4	46.2	135.8	72.2	56.7	101.7	80.2	19.2
24	476	414.8	113.6	482.6	145	216	56.6	138.8	79.2	37.4
25*	1367.5	432.3	194.7	70.5	459	111.6	68.1	27.5	29.3	26.5
27	409	180	111	220	160	61	32	45	142	24
29*	1690.3	608.9	53.3	40.1	47.9	51.8	77.2	102	44.6	35.4
31	1183	201	37	83	68	79	34	40	26	28
32	1704.2	1556.2	998.1	297.5	211	59.2	35.1	66.9	24.9	9.9
34	87.6	356.4	822	600.4	200.6	146.4	520	19.2	6.4	5
35	1136.8	212.6	300.9	72.7	43.5	37.3	49.2	31.2	29.3	36.5
36	2367	870.8	46.2	102	68	53	28.4	25.6	20.8	14.2
37	954.4	339.4	353.2	599	117	199.2	51.8	88.4	75.4	30.2
38	526	180	200	28	12	10	31	6	8	9
39	249.7	823.9	945.5	156.5	290.9	33.3	19.9	63.5	13.9	11.9
40*	556.8	271.6	73	91.2	35.4	26.2	28.8	30.6	23.2	24.2
42	332.2	213.4	277.6	147	159.8	28.2	17.2	22	13.2	6.4
43	813	103	48.6	141.4	100	47.8	31.6	19.8	15.6	10.2
44	465.2	395.2	270.8	136.2	41.6	76	126.2	82.6	18.2	8
45	630.5	94.5	102.5	60.5	26	41	34	21	45	38

TABLE III (cont.)

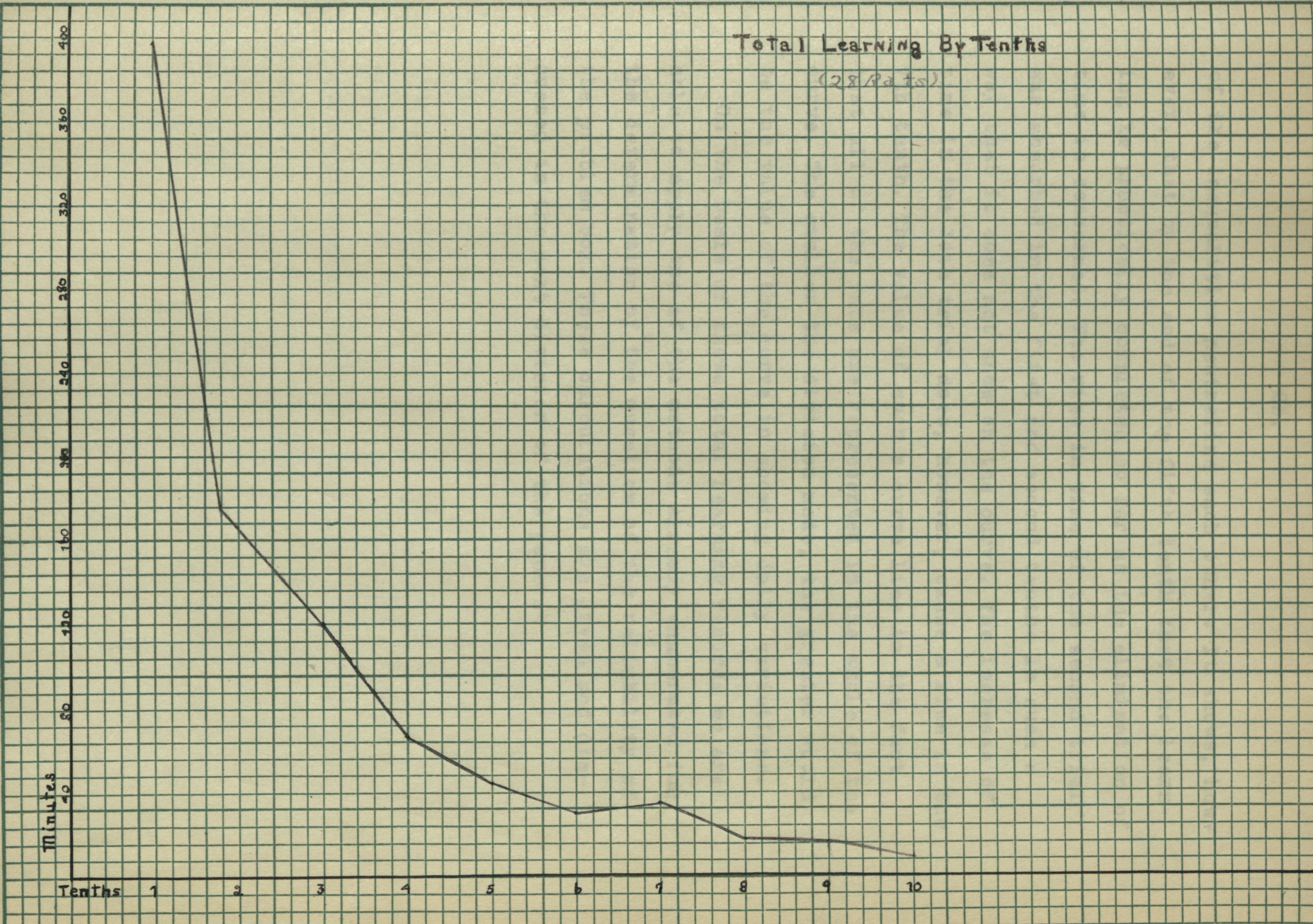
Rat No.	1	2	3	4	5	6	7	8	9	10
46*	1951	613	108.7	156.9	93.4	97.2	75.9	47.5	54.4	31
47	1133.8	264.6	499.8	178.4	38.4	70.2	22.4	31.2	37	9.2
48	419	279.2	91.6	64.4	51.8	47.6	39.4	28.2	17.8	8
49*	830.6	149.8	310.8	165.8	80	45.8	378.7	28.1	42.6	28.8
51	398.6	355.8	120	57.4	40.2	27.8	29.8	36	31.4	25
52	450.6	302.4	261.9	160.1	43.5	55.9	43	39.4	22.2	22
53	501.1	759.9	118.5	34.9	22.1	15.3	35.6	19.2	11	9.4
54	402.8	448	230.4	40.8	33	60.2	33.2	14.6	21	13
55	1504	126.4	61.1	154.9	104.6	57	30.4	32.4	22.1	20.1

Rats marked (*) developed a position habit.

his timidity keeps him form going on any exploratory expeditions to the top of the box or around the sides of the control cage, and since he can sense the food within the box, he endeavors to find an opening around the sides of the box through which he might arrive at the satisfaction of his hunger. He thus encircles the box and is practically certain to step on the plane at the right place within a relatively short period of time. On the second trial, however, the situation is less fearsome and the curiosity of the rat prompts him to make a thorough exploration which may lead entirely away from the plane for a relatively long period of time. On the third trial, this curiosity having been more or less appeased, the main object of the rat is to get the food and consequently he again stays close to the sides of the box.

There is practically no difference between the curve excluding the rats with the position habit and that including them. The only deviation of any note is that made in the 34th trial where there is a rise in the curve including the rats with the position habit due to the great time which it took rat No. 49 to work the problem that day.

The curve plotted on the basis of tenths is naturally smoother than the trial curve due to the fact that many


Total Learning By Tenths
(28 Pats)
Minutes
400
360
320
280
240
200
160
120
80
40
Tenths
1
2
3
4
5
6
7
8
9
10

of the inequalities neutralize each other. It shows, however, that by the end of the fifth tenth that the learning is practically completed so far as the elimination of time is concerned. The rise in time of the second trial was absorbed in the first and second tenths in such a way that it does not show on the curve. The increase of time in the 34th trial does, however, cause a rise in the 7th tenth. The value of such a curve lies in the fact that all of the rats are represented in the whole of the curve. The smaller the division made, the greater would be the value of the curve in showing daily variations.

On the basis of these total learning records we are able to apply one of the methods already enumerated for the determination of the question of the reliability of the problem box. These calculations are given and discussed in the next division of the paper.

RELIABILITY:

TOTAL LEARNING

The calculation of the coefficient of reliability from the total learning records must be made by the second method described by Whipple, i.e., the correlation of the sums of the odd and even trials. It is an assumed fact that if the practice effects are distributed normally for each rat over the entire learning period, then if chance errors are not operating to a serious degree to vitiate the results, there should be a high correlation between the odd and even trials.

The correlation calculated by this method, using the Spearman Foot-rule formula of $R = 1 - \frac{6\Sigma d}{n^2-1}$ $P.E. = \frac{.13}{\sqrt{n}}$ was found to be only .17±.081 including the records of the rats with the position habit, and .09±.089 excluding the records of these rats. These coefficients are insignificant, thus indicating that there is some factor or factors which are seriously affecting the relative standings of the various rats with regard to odd versus even trials.

In the problem box, as in all other apparatus in

which the animal is used, the first few trials are occupied with exploratory movements and the operation of the apparatus comes purely as an accident during the course of these movements. It is very possible then that the chances operating in these trials would be sufficient to destroy the correlation between the odd and even trials. If such is the case, it could scarcely be said, on the basis of our low correlation, that the problem box is unreliable, for this correlation would be due to the novelty of the situation which soon ceases to operate.

In order to test the effect of these first trials upon the correlation, the time of the first trial was deducted from the total time of the odd trials and the time of the second trial was deducted from the total time of the even trials and the correlation was again made. Then the time for the first and third trials was subtracted from the total for the odds and the time of the second and fourth from the evens and the correlation made.

The following table gives these correlations and also the ones already mentioned.

Data correlated.	Correlation. R.	P.E.
Total time odd trials vs. total time even trials (excluding rats with position habit).	.09±	.089
Total time odd trials vs. total time even trials (including rats with position habit).	.17±	.081
Total time odd trials minus 1st. trial vs. total time even trials minus 2nd. trial (including rats with position habit).	.50±	.081
Total time odd trials minus 1st. and 3rd. trials vs. total time even trials minus 2nd. and 4th. trials (including rats with position habit).	.35±	.081

From the correlations given above it would appear that the chances for error in the first and second trials operate to destroy the correspondence between the odd and even trials. When the first and second trials are ruled out, the correlation is at once raised to .50 which is six times the probable error. If, however, we eliminate the third and fourth trials the correlation also drops to .35 which is only a little over four times the probable error. This would indicate that beginning with the 3rd. and 4th. trials, the animals cease to be influenced so greatly by chance, and that upon the efficiency of the

factor which they introduce into the situation will depend the rapidity with which the problem is learned.

There may be still another factor which would operate to destroy the correlation between the odd and even trials. This factor is the presence of periodic streaks of decreased efficiency on the part of the rat. For example, rat 27 has a periodic increase of time roughly every fourth trial. It took him 40 trials and 1384 seconds to complete the problem. If the time for every fourth trial is added, beginning with the first trial which is the longest (199 seconds), it is found that 733 seconds or more than one half of the total time occurs in these ten trials. In other words, the average number of seconds for all the other trials combined is only 21.7 . This is an extreme case selected from this group of rats, but that such a phenomenon may occur to a greater or less degree in other animals will not, I believe, be denied, especially in view of the fact that periods of inefficiency, have been commonly observed in human as well as in animal behavior. It seems hardly necessary to point out the effect which such a phenomenon would have upon a correlation if it occurred frequently enough. This effect would be especially great if the periodicity should fall always in the odd or even set of trials for any particular rat.

In order to obviate the most of the effect which might result from such periodic occurrences the odd and even tenths(method of division given above) of the total learning time were correlated. These correlations are as follows:

No.	Data corrected.	Correlation.
1.	Total time odd tenths vs. total time even tenths (including rats with position habit)	.23± .081
2.	Total time odd tenths minus first tenth vs. total time even tenths minus second tenth (including rats with position habit)	.35± .081

Thus by making this division in the learning time we get only a slightly better correlation; i.e..23 instead of .17. However, by eliminating the first two tenths, we get the same correlation as was obtained by eliminating the first four trials. In all cases, of course, the first two tenths will include the first two trials and in the majority of cases they will also include the third and fourth. In dividing the time in this manner we have, to a larger extent than in odd versus even trials, the factor of unevenly distributed practice effects entering to disturb the correlation. Thus, by eliminating the detrimental effects of periodic decreases in efficiency, we

introduce at the same time this other disturbing factor of practice effects. It appears, therefore, that these two factors neutralize each other so far as their effect upon the correlation is concerned.

The conclusion from the data and calculations presented here on the question of reliability is that the problem box does not possess a very great degree of reliability, as measured by the correlation of odd versus even trials. Some of the factors which may operate to produce this low correlation have been mentioned and an attempt made to obviate them. And if by eliminating them a high correlation had resulted, future investigators would be enabled to forsee such difficulties and allow for their influence in drawing conclusions on the basis of data obtained in the problem box.

The attempted elimination of these factors, however, did not, in this case, result in an increase of the coefficient of reliability to the extent that the problem box could be considered reliable according to Whipple's standard. In order, therefore, to determine whether or not our attempted elimination was successful, we should apply other methods for the calculation of the coefficient of reliability. This is done in the next section.

RELIABILITY:
CORRELATIONS OF PARTIAL LEARNING IN THE PROBLEM BOX

The purpose of this section of the experiment is to determine whether or not the rats retained their same relative standing in the total time required to make six trials in the problem box after an interval of sixty days had elapsed between the last trial of the first six trials and the first trial of the second six. Records taken in this manner may be used in the calculation of the coefficient of reliability by the first method mentioned above, and quoted from Whipple. The assumption is that if the apparatus is reliable, the animal should tend to retain his relative rank when tested under the same conditions but at two different times. If there is no such correspondence, then it is not possible to say which of the two measures should be used in placing the rat. Consequently, it would be impossible to get reliable data in experiments upon the effects of extraneous factors upon the learning process.

The data here presented are based on the records of 20 rats. Their records are correlated trials against

trials and totals against totals. The following table gives the results of these calculations.

TABLE IV

Correlations 6 days P.B.-60 days-6 days P.B.

P.E.±.098 for all coefficients

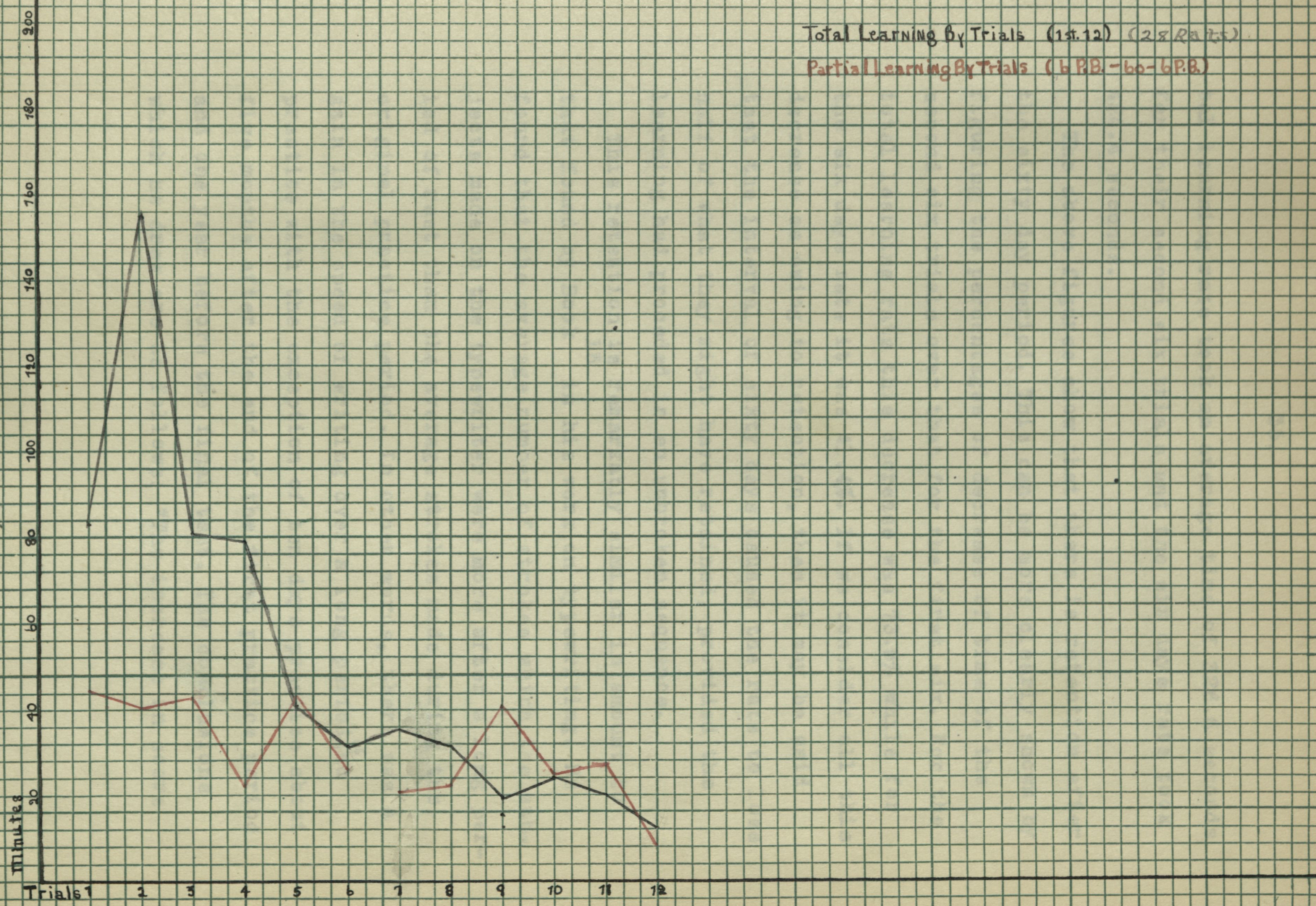
P.B. 1st. 6 days. Trials	1	2	3	4	5	6
P.B. 2nd. 6 days. Trials						
1	.006	-.09	-.26	.08	.03	-.22
2	.08	-.007	.22	-.06	.14	-.2
3	.37	-.11	-.007	.17	.09	.03
4	-.16	-.09	-.19	.03	.03	-.03
5	-.26	.08	-.16	-.17	-.11	-.007
6	.006	.19	.007	.12	-.11	-.007

Data

No.	Correlated	Correlation
1.	Total time for first six trials versus total time for second six trials	-.098
2.	Total time minus time for trials one and two for first six trials versus total time for second six trials minus time for trials one and two	-.22

None of the above correlations are significant. The highest coefficient is .37 which is not quite four times the probable error. The correlation of the total time for the two periods of learning is only -.098 which equals the probable error. This coefficient is raised by eliminating the time for the first two trials in each period. Even after this raise, however, it still remains insignificant.

The curve of learning shown on the following page for the twelve trials which these rats were given is plotted with trials on the X-axis and minutes on the Y-axis. It is much more irregular than the curve for the rats which were given total learning. There is, however, a gradual drop and the first trial of the second period starts on a lower level than the last trial of the preliminary six trials. There is a slight rise, however, in the eighth trial which brings this point on the curve a little above the corresponding point of total learning. The curve shows a very pronounced rise in the ninth trial. This was owing to the erratic behavior of two rats, Nos. 98 and 111. Rat 98 took 626 seconds in this trial as against 144 in the preceding trial, while rat 111 required 632 seconds in trial 9 as against 99 seconds in the trial before. If the records of these two rats are eliminated, the total time required for trial 9 is 1465 seconds as opposed to 1490 seconds for trial 8. Thus, trial 9 would be on the same level as trial 8 if the curve were plotted without the records of these rats. The total learning rats, however, required much more time for the first 6 trials than did the six day rats. The average time for the former for the first six trials was 1238.8 seconds, while for the latter it was only 725.7 seconds . The to-


Total Learning By Trials (1st. 12) (28 Parts)
Partial Learning By Trials (6 P.B. - 6o - 6 P.B.)
200
180
160
140
120
100
80
60
40
20
Minutes
Trials 1 2 3 4 5 6 7 8 9 10 11 12

tal learning rats took an average time of 326.6 seconds for their second six trials and the six days rats took 424.4 seconds.

The above figures show that there is retention after the sixty day period. This can be more clearly shown by figuring the percentages of decrease of time of the second six trials over that for the first six. For the total learning rats this decrease was 73.5% while for the six days rats it was 41.4%. If we consider all this decrease as owing to retention, then it may be said that the interval of sixty days caused the rats to lose half of what they might have retained, provided the learning had proceded in an unbroken sequence.

This retention is remarkably similar in amount to that found by Book[15] in his work on typewriting. He found that the average number of strokes which the subjects made on the typewriter were more after a long period of rest than the average strokes made during the original practice period. In other words, there was a gain in the speed of writing over a long period of no practice with the exception of ten days practice about five months after the end of the first practice period and one year before the final test. The conclusions which he drew are as follows: "The increase in score

shown by our second memory series was owing, so far as we could make out, rather to the disappearance, with the lapse of time, of numerous psycho-physical difficulties, interfering associations, bad habits of attention, incidentally acquired in the course of learning, interfering habits and tendencies, which, as they fade, left the more firmly established typewriting associations free to act."[16]

Watson criticizes Book's conclusions as follows: "The curve of the original learning is not reproduced here, but the fact is that he stopped practicing on a spurt. Had he taken the records for ten more days, his average would probably have been around 1700 strokes per ten minutes instead of 1503. Furthermore, the practice gained from the ten days in June, 1906, was not without marked effect upon the succeeding and final memory tests. Hence, it seems that we need only to conclude that there is a tremendously great and surprising permanence in the functioning of the typewriting habit, but nothing supernatural in character."

However applicable Watson's criticism of stopping on a spurt may be to the work done by Book, it certainly is not a valid explanation of the fact that after sixty days the rats in this experiment started on a lower level than that which they had reached in their preliminary

training. The problem of retention and the increase in skill during a period of no practice is one which can be easily dealt with, using animals as subjects, the only hindrance being that no introspective analysis can be made. Disregarding that fact, however, the problem still possesses enough significance to make it exceedingly interesting.

This factor of retention is a troublesome one in estimating the validity of the coefficient of reliability obtained by correlating the results obtained at different times. Thus, Burt[17] says, in speaking of the mirror drawing test, "The correlation at the Elementary School between the total results of the first series of sittings and those of the second was .52. From its very nature the Mirror drawing experiment cannot strictly be held to test quite the same capacity on repetition. The observed figures indicate considerable retention by the subjects of the effects of the first sitting; and doubtless the degree of retention is not the same for all. Indeed, the second series was undertaken partly with the hope that it might furnish a test of the retentiveness of improvability rather than of improvability itself. No method, however, presented itself of isolating this second capacity, and of differentiating the subjects

accordingly. In any case the difference of retentiveness does not seem to have been great. And of the various substitutes for a reliability coefficient that one might devise, the figure cited above probably gives the best approximation for these particular experiments."

Inasmuch as the coefficients of reliability for the problem box, as given above, do not approximate even .50 the conclusion must be forthcoming that, so far as our results can show, the problem box is an unreliable test when used with the white rat, unless other factors can be found which will account for this low coefficient, or unless it is shown that this factor of retention is sufficient to produce this result.

There is left, however, one more approach to the question of reliability, and since we have found factors in both other approaches which may or may not affect our coefficients of reliability, as determined by those methods, the problem box should be given one more attempt to prove itself reliable. This approach is made in the following section.

RELIABILITY:

CORRELATION OF LEARNING IN THE MAZE WITH THE LEARNING IN THE PROBLEM BOX

The method which has been frequently employed to determine the reliability of intelligence tests given to human beings is to correlate the standing in the intelligence test with the standing in school grades or with the estimation of the subject's ability by a competent judge. In animal behavior, however, it is, of course, impossible to give the animal an intelligence test or to estimate its ability by obeervation, but an approximation to the above method may be made by correlating the results of two different tests to which the animals have been subjected.

In order to do this the 22 animals used in this section of the experiment were first given six trials in the problem box, then allowed to rest 60 days after which they were run for six trials in the maze. Six trials were used in order that the problem should not be so thoroughly learned that retention would greatly affect the results, and also in order that the age factor might be more easily controlled. The fact that the rats were 60 days older

when they were put in the maze than when they were in the problem box should not affect the learning in the maze, since Miss Hubert[18] found no age effects in such a short period of time.

Correlations were made between the various measures taken in the maze, distance, net time, and gross time, and time in the problem box. These correlations are given in the following tables. The probable error for all coefficients is .093.

TABLE IV
Time in Problem box versus Net time in Maze.

Maze Trials	1	2	3	4	5	6
Problem box Trials						
1.	-.06	.043	-.13	.036	-.12	-.019
2.	-.13	.019	.006	.12	.17	.07
3.	-.055	-.097	-.097	-.11	.006	-.097
4.	.14	.24	.17	-.13	.093	-.019
5.	.192	.08	-.13	-.055	.036	-.097
6.	.069	.14	-.055	.036	.19	.19

TABLE V
Time in Problem Box versus Gross time in Maze

Maze Trials	1	2	3	4	5	6
Problem box Trials						
1.	-.15	.16	-.05	.18	.105	.25
2.	-.006	.043	-.08	.23	.08	.16
3.	-.105	.03	-.09	.15	-.105	-.05
4.	.043	.09	-.33	-.105	-.11	.043
5.	-.105	.14	-.006	.09	.006	.11
6.	-.03	.07	-.105	.006	-.12	.14

TABLE VI

Time in Problem Box versus Distance in Maze

Maze Trials	1	2	3	4	5	6
Problem box Trials						
No.						
1	-.097	-.106	-.097	.12	.14	.135
2	.006	.12	-.048	.265	.13	.07
3	-.19	-.06	-.019	.17	-.13	-.03
4	.12	.01	-.19	-.19	-.081	.036
5	-.11	.24	.07	-.043	.15	.12
6	.006	.12	-.081	-.03	-.097	.08

TABLE VII

Correlation of Totals

No.	Things Correlated	Correlation
1	Total time P.B. versus total Net time Maze	.019
2	" " " " " Grose " "	.006
3	" " " " " Distance "	.093
4	" Net" maze time. " " " "	.70
5	" Gross " " " " "	.67

A glance at the correlations presented here shows that the experimenter has been unable to find a relationship between the time required for six trials in the problem box and the various measures of the learning for

six trials in the maze. The highest correlation found was -.33 which is between the time for the fourth trial in the problem box and gross time for the third trial in the maze. The coefficient, however, is only three times the probable error and therefore cannot be considered significant.

It is not possible to interpret any of the experimental data here presented in such a way as to show why there is not a relationship between the learning process in the maze and that in the problem box. From a priori grounds, however, two possible reasons may be cited. The first is that in the problem box the food is always so placed that it may stimulate the olfactory and possibly the visual sense organs of the rat. The possible influence of this factor has already been discussed on page 24. The second is that in the problem box, the rat is not compelled to make so many choices between the right and the wrong way to go. That is, in the maze the animal has to make a choice at every doorway between the true path and the cul-de-sac, while in the problem box he only must make a choice between which end of the plane he will step on. Whether these two differences are sufficient to account for the absence of relationship between

the animals' learning in the two pieces of apparatus or whether they have anything whatever to do with the situation we have no way of knowing. It should be mentioned here that the preceding section of the paper showed that on the average there is a large amount of retention after an interval of sixty days. This retention may have caused inhibition or facilitation in the learning of the maze so that we are not altogether justified in drawing a dogmatic conclusion that there is no relationship between these two learning processes. If it be granted however that this retention had little or no effect, it is hard to see how it should have such an effect as to render the coefficients insignificant, then there must be something wrong with either the maze or the problem box as tests of the rat's learning ability. One or the other of them must be an unreliable method of testing the rat's ability. The principle involved here is the same as that which is involved in making correlations between the results of the application of the same test at two different times. If the two tests are reliable and are testing the same capacity, then the results obtained in each should be the same under the same conditions.

Our data fails to reveal any correspondence between

results in the maze and those in the problem box, consequently, these two tests are measuring two different things or one or the other of them is not reliable, unless other factors in the method of application can be found which did not operate with an equal influence in both tests.

Inasmuch as the correlations in this and in the preceding section were made with six trials only, it is necessary to determine what the relationship is between six trials and total learning. This involves the problem of an adequate criterion of learning which requires less time to arrive at than the criterion of completed learning. This problem and its relationship to the above correlations is treated in the following section.

AN ADEQUATE CRITERION OF LEARNING

The purpose of the calculations made in this section is to determine whether or not there is some part of total learning which will place the rat equally well as to his ability in the solution of the problem as do the results of total learning. If such a part could be found and if it is a small enough part of total learning, then a great saving of time to the experimenter would result. It is also necessary to treat this problem in the present study in order to justify our use of only six trials in the correlations of resubts of the problem box obtained at two different applications, and in the correlation of the maze and the problem box learning.

Lashley[19] has done some work on the criterion of learning in the maze. He used, however, only one standard of comparison, i.e. number of trials preceding the first perfect run, as a possibility of possessing the same degree of reliability as the generally used criterion, i.e. three perfect runs in succession. He found a correlation coefficient of .62 with a probable error of .063. And he concludes that if there is found to be a difference in the average number of trials required to

arrive at one perfect run between several groups of animals then these groups of animals will show the same relative difference in the average number of trials required for completed learning.

There is a question involved in making such a calculation which has not been mentioned by Lashley. That is, is it right to use trials rather than time or distance as a measure of the learning progress of the rat? Trials, it is true, have been more generally used for this purpose but in the case of my own data, the correlation between the total trials and the total time is .0103 with a probable error of .081. This coefficient is only one-eighth as large as the probable error. Such a non-correspondence of trials and time is also shown if we examine some of the individual records. The most striking example of this is shown in the case of rats 20 and 38. Thus it took rat 20 forty-three trials to complete the problem with a total time of only 771 seconds, while rat 38 used only 10 trials but a total time of 1010 seconds. It is difficult, on the basis of the present data, to decide definitely in favor of either trials or time. It would seem, however, that time is a more-delicate measure of the rat's progress than are trials. Correlations have been made, however, using both measurem and

these, together with the various parts of learning which have been correlated, are listed in Table VIII.

TABLE VIII

CORRELATIONS MADE FOR THE DETERMINATION OF ADEQUATE CRITERION OF LEARNING

No.	Data Correlated	Correlation	P.E.
1	Total time vs. time including 1st. perfect trial	.44	.081
2	Total time vs. time including 2nd. perfect trial regardless of position	.61	.081
3	Total time vs. time including 2nd. perfect trial in succession	.72	.081
4	Total trial vs. trials including 1st. perfect trial	.08	.081
5	Total trials vs. trials including 2nd perfect trial regardless of position	.07	.081
6	Total trials vs. trials including 2nd. perfect trial in succession	.24	.081
7	Total time vs. time for 1st. six trials	.51	.081
8	Total time vs. time for 3rd. and 4th trials	.34	.081
9	Total time vs. time for 3rd., 4th., 5th., and 6th. trials	.46	.081

These correlations indicate that of the shorter criterions which are correlated with total time, the time for the first six trials is the best. It gives a correlation of .51 while the standard of comparison which Lashley used gives with my data a correlation of .44. The difference between these two standards as far as the correlation goes is not a significant one, but there would seem to be no reason for hesitancy in choosing which to use inasmuch as the six trials is by far the shorter of the two. The next shortest criterion to be chosen, if these two are rejected, is two perfect trials regardless of position. This criterion gives a correlation of .61. Two perfect trials in succession give a high correlation, .72, which is to be expected since in a great number of cases these two perfect trials are the first and second trials of the three perfect trials in succession which is the criterion used for completed learning.

As can be seen at a glance at the table, similar correlations made with trials did not give a single significant coefficient. The highest one, .24, which is for total trials versus trials including two perfect trials in succession is only three times the probable error, while the other two are less than the probable error. These

coefficients would tend to strengthen the conclusion that trials are not as good a measure of the learning process as is time.

The correlation made using the average time per trial which, of course, combines both time and trials are also lower than those made with time alone. They do, however, more nearly approach the coefficients made for time only, than do those made for trials.

When it was found that the correlation between the first six trials and total time was rather high, it was thought that by throwing out the first two trials a still higher correlation might be attained. Especially was this thought possible when the disregarding of the first two trials gave a much higher correlation between the odds and the evens. Similar results did not follow in this case, however, since the correlation fell to .46 when the first two trials only were omitted, and to .34 when the 3rd. and 4th. trials only were used.

We may conclude, then, that if we are to take anything less than total learning as indicative of the rat's ability in the problem box that the time for six trials is as good as any that can be found. The advantages of this is obvious in the saving of time, labor, etc. in calculating the effects upon the rat's ability in the problem

box of age, sex, drugs, etc. It should be mentioned, however, that in such experiments using this short learning period that negative results are not necessarily reliable. For instance, suppose that we are testing the effect of a certain drug upon the rat's ability to learn the maze. We will give the rats the drug and then run them six trials in the box and we find that the drug has had no effect on the length of time involved in the six trials. Are we justified in saying on that basis that the drug would have no effect upon the total learning time of the rats? We believe that the answer to that question would have to be in the negative for the following reasons: first, the effects might be so slight that they could not be detected by this rather coarse measurement of the rat's ability; second, the effects of the drug might not be noticeable in the coarser motor coodinations which are involved in the first six trials, but be very noticeable in the finer adjustments necessary for a perfect reaction. For these reasons, therefore, one should hesitate to draw conclusions from negative results in an experiment of this kind. On the other hand, any positive results which might be found would very probably be true, as far as they went. It is always possible, of course, that they would be much greater if the learning were carried further.

SUMMARY AND CONCLUSIONS

Three different methods have been used in this experiment to investigate the reliability of the problem box as a method of testing the learning ability of the rat. Although the results of these methods have not lead to as definite conclusions as it was hoped that they might, yet it is felt that the methods have been used to their fullest extent and value. The highest coefficient of reliability found was .50+. This was between the odd and even trials with the first two trials omitted. When these two trials were included, the coefficient was only .17+. If, therefore, the problem box is used as a method of testing learning, the results could be made more reliable by elimination of the first two trials.

There was found to be a large amount of retention after an interval of sixty days. This amount was about half as much as if the learning had not been interrupted.

The criterion of learning which would be the best to choose, if a shorter criterion than total learning is desired, would be the total time for the first six trials. This criterion gave a coefficient of .51 when

correlated with the total time required to arrive at the criterion of three perfect trials in succession. It is not valid, however, to draw conclusions from negative results when a criterion such as this is used.

There was found to be no correlation between the maze and the problem box. The principal reason for this would seem to be the difference in the position of the food in the two pieces of apparatus.

BIBLIOGRAPHY

1. Spearman, C. Correlations calculated from Faulty Data. Br. Jr. Psy. 3, 1910, pp. 271-95

2. Watson, J. B. Behavior, An Introduction to Comparative Psychology. New York 1914. p.93

3. Healy, The Individual Delinquent.

4. Watson, J. B. See reference number 2, page 95.

5. Richardson, F. A Study of Sensory Control in the Rat. Psych. Mon. Ser. No. 48

6. Ulrich, J. L. Distribution of Effort in Learning in the White Rat. Behav. Mon. 1915, vol. 2.

7. Basset, G. C. Habit Formation in a Strain of White Rats with less than Normal Brain Weight. Behav. Mon. Ser. No. 9.

8. Krueger, F. and Spearman, C. Die Korrelation zwischen verschiedenen geistigen Leistungsfahigkeiten. ZPs, 44=1907, pp.50-114.

9. Burt, C Experimental Tests of General Intelligence. Br. Jr. Psy. vol. 3, pp.-94-177

10. Wyatt, S. The Quantitative investigation of Higher Mental Processes. Br. Jr. Psy. vol.6,pp.109

11. Krueger, F. See reference number 8.

12. Simpson,B. R. Correlation of Mental Abilities. Columbia Educational Series. N.Y. 1912

13. Whipple, G. M. Manual of Mental and Physical Tests Vol. 1, p.52

14. Burt, C See reference number 9.

15. Book, W. F. The Psychology of Skill Missoula, University of Montana, 1908

16. Watson, J. B. Psychology from the Standpoint of a Behaviorist. Phil. 1919. p.305.

17. Burt, C. See reference number 9.

18. Lashley, K. S. The Criterion of Learning in Experiments with the Maze. Jr. An. Behav. 1917, vol.7, pp.66-7.

APPENDIX

EXPLANATORY NOTE TO TABLES

The time records in the following tables are all in terms of seconds.

The distance records are in terms of inches as measured by the chartometer.

The time records for individual trials in the total learning data which are marked (*) are the records of the rats with the position habit which were used as the criterion of completed learning.

TABLE I

TOTAL LEARNING RECORDS

RATS	18	20	24	25	27	29	31	33
TRIALS								
1	200	135	19	11	199	519	82	62
2	163	42	64	859	80	158	342	2346
3	33	14	40	412	28	87	140	740
4	16	15	9	26	102	243	43	281
5	P8	P5	5	35	50	229	75	812
6	P7	P5	9	35	11	415	81	175
7	42	13	550	15	102	403	420	175
8	21	P6	14	107	17	25	14	148
9	10	16	P5	14	11	165	110	56
10	P6	P4	P5	103	30	16	9	24
11	145	P5	16	156	22	17	30	36
12	51	10	P4	67	48	21	P3	3P7
13	87	10	150	44	51	P11	23	41
14	11	6	P4	27	25	10	13	42
15	142	P3	P3	40	25	11	6	3P6
16	65	15	7	29	119	5	5	3P7
17	63	20	6	13	56	7	5	3P5
18	12	6	40	15	38	10	P2	
19	13	11	52	4	40	5	4	
20	17	52	P3	14	26	5	10	
21	10	64	P3	8	19	6	5	
22	9	8	13	15	9	5	60	
23	14	7	10	20	19	14	4	
24	10	8	11	14	14	1P3	5	
25	34	50	9	225	3P4	9	4	
26	239	3P4	418	83	3P4	7	4	
27	33	8	45	20	7	5	4	
28	10	3P4	7	87	17	11	1P2	
29	11	18	12	52	3P4	10	5	
30	3P6	25	30	28	12	1P4	3P6	
31	8	9	31	10	24	11	19	

RATS	18	20	24	25	27	29	31	32
TRIALS								
32	3P40	19	24	7	3P5	7	25	
33	31	43	14	10	115	5	4	
34	10	12	23	28	19	7	6	
35	3P6	49	86	13	3P5	7	5	
36	7	6	23	15	3P3	6	6	
37	29	19	45	1P9	14	23	5	
38	20	3P3	8	9	3P4	4	1P2	
39	8	3P4	25	22	3P3	16	57	
40	3P4	7	10	1P3	3P3	22	1P2	
41	7	3P4	3P3	1P3		6	1P2	
42	7	3P4	3P3	8		10	5	
43	3P4	3P3	15	6		10	5	
44	8		8	1P3		51	1P2	
45	7		3P4	1P3		15	4	
46	3P4		19	7		13¢	15	
47	13		3P3	6		6	1P2	
48	10		3P3	6		10	4	
49	9		44	1P3		5	1P2	
50	3P3		40	1P4		6	13	
51	12		8	6		13	1P5	
52	3P4		39	5		8	5	
53	3P4		1P3	1P3		6	5	
54	3P4		20	7		7	4	
55			3P6	1P3*		4	4	
56			10	5*		8	4	
57			8	5*		8	6	
58			20	5		5	1P2	
59			11	1P2		5*	4	
60			9	5		5*	4	
61			16	5		4*	4	
62			1P3	1P3		5	4	
63			4	4		5	1P2	
64			3P3	5		9	3	
65			3P3	5		5	3	
66			3P3	5		5	4	

RATS	18	29	24	25	27	29	31	32
TRIALS								
67				1P2		5	10	
68				1P2		9	1P2	
69				4		8	1P4	
70				5		5	1P2	
71				5		12		
72				1P3		1P4		
73				5		5		
74				4		5		
75				12		5		
76				4				
77				12				
78				5				
79				5				
80				7				
81				5				
82				11				
83				7				
84				7				
85				4				
86				4				
87				1P2				
88				4				
89				6				
90				4				
91				10				
92				14				

RATS	34	35	36	37	38	39	40	42
TRIALS								
1	P20	501	1404	P34	526	100	259	77
2	338	167	1110	212	180	499	189	160
3	215	342	161	523	200	791	33	119
4	1155	61	47	309	28	699	41	141

RATS / TRIALS	34	35	36	37	38	39	40	42
5	173	94	725	P6	12	20	174	38
6	166	49	178	160	P10	338	76	26
7	19	30	71	12	31	49	39	28
8	637	26	33	189	1P6	11	13	226
9	26	15	22	40	3P8	3P11	11	33
10	1P6	161	17	150	3P9	67	37	68
11	1P5	118	9	15		3P8	12	58
12	1P4	19	4	156		3P13	9	2P6
13		40	5	63		3P8	14	100
14		26	16	359			30	55
15		13	6	45			30	16
16		21	14	10			16	9
17		17	80	32			12	1P4
18		19	7	48			3P5	10
19		1P5	28	23			9	1P4
20		8	6	159			11	1P4
21		8	11	1P13			8	8
22		10	10	1P7			5	10
23		15	20	29			6	6
24		1P3	16	1P9			5	1P2
25		12	6	1P4			9	7
26		9	5	35			2P6	1P3
27		1P3	7	22			8	1P2
28		9	1P3	32			5	1P2
29		14	12	8			9	
30		17	6	1P4			5	
31		1P4	5	1P13			8	
32		8	6	46			13	
33		10	8	27			3	
34		7	5	1P4			6	
35		1P4	8	1P5			4	
36		1P2	1P3	1P5			7	
37		12	1P3				5	
38		7	6				6	
39		7	5				9	

RATS	34	35	36	37	38	39	40	42
TRIALS								
40		7	1P S				5*	
41		8	1P 3				2P 5*	
42		1P 3	1P 2				4*	
43		1P 5					6	
44		13					4	
45		1P 2					6	
46		1P 8					7	
47		1P 10					8	
48							6	
49							14	
50							22	
51							6	
52							2P 6	
53							7	
54							3P 8	
55							10	
56							2P 5	
57							2P 4	
58							18	
59							9	
60							2P 3	
61							10	
62							9	
63							3P 4	
64							2P 3	
65							6	
66							4	
67							20	

RATS	43	44	45	46	47	48	49	51
TRIALS								
1	646	353	391	234	27	115	122	90
2	87	103	47	593	826	380	104	108
3	100	46	53	496	351	154	542	34
4	44	296	110	41	134	82	7	15

RATS	43	44	45	46	47	48	49	51
TRIALS								
5	30	156	59	346	55	42	49	20
6	15	107	7	143	1P9	42	22	130
7	22	117	52	140	119	33	34	3P4
8	11	55	3P3	123	326	31	14	3P3
9	24	43	3P3	22	128	27	48	60
10	68	15	32	310	81	34	30	111
11	51	18	42	144	20	17	14	70
12	40	42	17	40	1P3	24	45	35
13	51	9	8	24	15	20	25	93
14	17	125	7	20	21	14	120	12
15	14	25	3P3	12	55	18	99	45
16	25	3P3	3P3	11	8	3P6	18	22
17	11	76	7	18	9	3P5	38	3P6
18	18	8	44	20	7	3P4	68	11
19	6	11	3P3	16	7		25	3P3
20	9	3P5	17	119	11		11	12
21	9	3P4	3P3	7	13		17	3P4
22	6	3P3	9	19	7		25	12
23	3P3		8	19	17		1P3	3P4
24	7		5	81	18		24	6
25	6		3P3	6	8		18	20
26	3P4		11	12	1P4		13	3P3
27	3P3		18	17	1P3		4	3P3
28	3P4		12	5	1P3		22	7
29			14	1P3			5	14
30			3P3	8			5	3P4
31			5	22			7	3P3
32			6	36			1P6	8
33			5	5			7	4
34			8	22			353	3P4
35			3P2	4			3	3P3
36			3P3	41			8	6
37			30	16			6	6
38			4P4	1P3			5	3P3
39			3P3	15			4	302

RATS	43	44	45	46	47	48	49	51
TRIALS								
40			5	1P3			7	5
41			6	21			6	3P4
42			27	17			5	6
43			3P3	6			4	3P3
44			3P3	5			7	5
45			3P2	10			5	7
46				13			12	5
47				9			5	3P3
48				10			16	9
49				5			4	3P3
50				6			8	6
51				5			1P3*	8
52				5			4*	3P3
53				9			4*	1C
54				11			4	3P3
55				9				3P3
56				11				8
57				*6				3P2
58				1P5				5
59				5				3P4
60				11				3P3
61				10				6
62				3P2				3P3
63				1P2				3P3
64				11				3P4
65				1P2*				
66				5*				
67				1P2*				
68				5				

RATS	52	53	54	55
TRIALS				
1	1P14	19	47	30
2	120	450	207	1270
3	61	321	248	141

RATS	52	53	54	55
TRIALS				
4	284	460	310	70
5	106	55	13	33
6	117	70	129	51
7	43	15	104	21
8	3P10	18	29	18
9	65	16	22	12
10	19	9	3P5	20
11	157	7	20	3P8
12	27	7	16	25
13	11	8	3P5	9
14	29	10	21	30
15	109	32	32	103
16	3P5	3P4	12	9
17	12	7	9	61
18	13	6	8	11
19	12	3P4	7	25
20	9	3P4	5	8
21	12	3P5	3P4	21
22	18		9	25
23	15		9	3P5
24	16		3P5	5
25	9		3P5	11
26	18		3P5	7
27	3P4			7
28	8			8
29	7			12
30	10			7
31	16			7
32	3P4			3P4
33	3P4			8
34	7			5
35	7			3P5
36	10			9
37	3P4			3P5
38	3P4			3P4

TABLE II

6DAYS PROBLEM BOX-60 DAYS-6DAYS MAZE

	PROBLEM BOX	MAZE			
		Gross Time	Static Time	Net Time	Distance
Rat 57					
Trials					
1	78	1790	1261	529	255
2	62	1935	1291	644	478
3	94	105	22	83	74
4	114	99	14	85	92
5	311	147	30	117	139
6	435	40	7	33	44
Rat 61					
Trials					
1	682	184	61	123	53
2	143	149	21	128	78
3	40	400	89	311	190
4	40	267	68	199	163
5	134	154	10	144	118
6	70	172	65	107	85
Rat 62					
Trials					
1	33	172	46	126	64
2	81	170	70	100	60
3	110	51	2	49	65
4	248	23	3	20	36
5	15	25	2	23	42
6	166	11	0	11	26
Rat 64					
Trials					
1	118	282	89	193	91
2	229	124	33	91	81
3	215	36	3	33	42
4	50	57	10	47	74
5	22	87	26	61	107
6	66	36	6	30	58

PROBLEM BOX		MAZE			
		Gross Time	Static Time	Net Time	Distance
Rat 65					
Trials					
1	29	372	109	263	104
2	780	198	64	129	111
3	141	32	3	29	46
4	200	29	3	26	44
5	152	64	10	54	101
6	27	33	2	31	69
Rat 66					
Trials					
1	261	551	184	367	188
2	228	60	4	56	65
3	163	29	0	19	48
4	19	26	0	26	38
5	39	25	1	24	54
6	101	24	1	23	45
Rat 68					
Trials					
1	35	790	354	436	206
2	75	108	28	80	84
3	258	27	2	25	41
4	88	19	0	19	43
5	31	14	0	14	27
6	35	10	0	10	26
Rat 70					
Trials					
1	95	1155	510	645	220
2	24	249	87	162	88
3	152	166	26	140	116
4	130	130	30	100	121
5	18	97	13	84	129
6	94	69	5	64	79

PROBLEM BOX — MAZE

		Gross Time	Static Time	Net Time	Distance
Rat 71					
Trials					
	1	27	478	315	163	87
	2	113	335	177	158	101
	3	325	197	50	147	118
	4	24	44	3	41	54
	5	17	27	1	26	47
	6	45	22	1	21	47
Rat 72					
Trials					
	1	18	329	123	206	132
	2	492	432	143	299	238
	3	780	210	27	183	179
	4	116	32	3	29	36
	5	316	25	2	23	49
	6	8	16	16	0	40
Rat 73					
Trials					
	1	424	266	112	154	68
	2	225	358	82	276	175
	3	90	52	5	47	49
	4	25	57	7	50	77
	5	19	54	8	46	98
	6	141	27	2	25	67
Rat 75					
Trials					
	1	205	373	40	333	127
	2	555	204	53	151	109
	3	30	52	2	50	49
	4	20	451	35	416	173
	5	53	40	2	38	58
	6	18	60	3	57	68

	PROBLEM BOX	MAZE			
		Gross Time	Static Time	Net Time	Distance
Rat 76					
Trials					
1	920	1327	944	383	151
2	130	1848	1133	915	381
3	11	78	19	59	51
4	288	35	15	20	39
5	109	27	3	24	54
6	32	23	3	20	46
Rat 78					
Trials					
1	129	163	33	130	61
2	99	124	25	99	69
3	38	80	13	67	52
4	5	81	27	52	88
5	8	28	7	21	39
6	19	16	1	15	30
Rat 80					
Trials					
1	12	1520	1160	360	216
2	466	1620	1140	480	305
3	92	971	640	331	311
4	246	29	8	21	47
5	7	70	15	55	60
6	176	31	9	22	48
Rat 83					
Trials					
1	522	173	43	130	75
2	102	994	374	650	383
3	13	66	14	52	52
4	67	28	3	25	34
5	8	24	6	18	37
6	66	38	0	38	57

	PROBLEM BOX		MAZE		
	Gross Time	Static Time	Net Time	Distance	
Rat 84					
Trials					
1	390	213	18	195	118
2	56	51	6	45	54
3	12	40	5	35	64
4	15	18	0	18	40
5	44	18	2	16	38
6	7	12	0	12	30
Rat 85					
Trials					
1	13	233	78	155	77
2	67	462	152	310	187
3	144	66	7	59	65
4	980	25	3	22	34
5	441	29	13	16	36
6	34	22	5	17	36
Rat 86					
Trials					
1	178	220	20	100	102
2	242	465	125	340	258
3	411	85	20	65	99
4	79	28	4	25	51
5	83	35	9	26	65
6	1P3	10	0	10	29
Rat 87					
Trials					
1	212	1583	969	614	320
2	82	1323	818	505	129
3	46	51	10	41	50
4	42	41	18	23	48
5	133	13	1	12	26
6	190	10	0	10	27

PROBLEM BOX				MAZE	
		Gross Time	Static Time	Net Time	Distance
Rat 89					
Trials					
1	40	1225	850	375	183
2	11	321	125	196	120
3	10	520	180	340	133
4	326	162	52	110	97
5	9	60	22	38	43
6	53	125	57	68	94
Rat 90					
Trials					
1	220	421	186	235	143
2	186	368	114	255	151
3	13	322	66	256	195
4	31	51	7	44	62
5	25	15	2	13	29
6	25	9	0	9	28

TABLE III

6DAYS PROBLEM BOX-60DAYS-6DAYS PROBLEM BOX

RAT	91		92		93	
	1st. 6	2nd. 6	1st. 6	2nd. 6	1st. 6	2nd. 6
TRIALS						
1	38	26	84	21	335	20
2	45	6	135	17	143	7
3	240	47	48	182	148	13
4	179	18	42	14	89	236
5	56	29	25	36	180	3P3
6	64	18	47	36	517	10

RAT	94		95		96	
TRIALS						
1	303	423	14	14	20	20
2	56	64	6	83	65	56
3	248	5	17	8	388	77
4	141	60	44	3P5	120	50
5	93	10	77	7	390	58
6	63	7	104	30	58	100

RAT	98		100		101	
TRIALS						
1	105	41	102	133	29	125
2	195	144	239	18	27	63
3	50	626	252	28	293	90
4	23	45	13	13	153	67
5	274	53	9	13	31	8
6	48	14	134	7	175	15

RAT	104		102		103	
TRIALS						
1	414	30	83	38	25	114
2	39	13	33	354	47	49
3	110	71	25	56	84	75
4	31	18	14	224	25	79
5	240	7	34	5	10	228
6	51	7	22	9	1P9	58

RAT	105		108		111	
	1st. 6	2nd. 6	1st. 6	2nd. 6	1st. 6	2nd. 6
TRIALS						
1	195	6	SP15	24	33	127
2	180	81	337	25	813	99
3	16	8	15	215	193	632
4	143	13	24	10	86	48
5	455	35	151	3P6	57	139
6	17	4	197	26	102	3P5

RAT	112		113		114	
TRIALS						
1	127	50	385	24	479	46
2	21	11	50	2P7	84	226
3	200	9	27	12	257	68
4	200	100	19	25	20	73
5	88	95	251	14	54	565
6	151	17	63	5	69	116

RAT	117		119	
TRIALS				
1	51	19	126	36
2	19	57	41	110
3	18	239	107	62
4	14	104	56	47
5	34	33	331	27
6	115	34	49	19