A NOTE ON THE CORRELATIONS BETWEEN TWO MAZES

FRANK C. DAVIS AND EDWARD C. TOLMAN

University of California

This note is to be read as an appendix to the larger study, "The inheritance of maze-learning ability in rats" (1). As Hunter and his students have recently indicated, there is at present considerable question (2) as to the possibility in the case of rats of measuring reliable individual differences at all. All the intercorrelations found by these authors (as yet reported) are low—both as between different parts of the same maze or problem box and as between maze and problem box.

In the larger study referred to, our only evidence for the reliability of our maze was obtained by correlating one-half of the same maze learning against a second half. The purpose of the present minor investigation was to see if we could in reality get a correlation between this maze and a second one.

Twenty-one white rats[1] were run first for ten consecutive days one trial per day, in maze A (fig. 1) and then after an interval of six days for ten more consecutive days, one trial per day, in maze B (fig. 2). Records were obtained for errors[1] and time in each day's trial and for the total number of perfect runs achieved in the ten days.

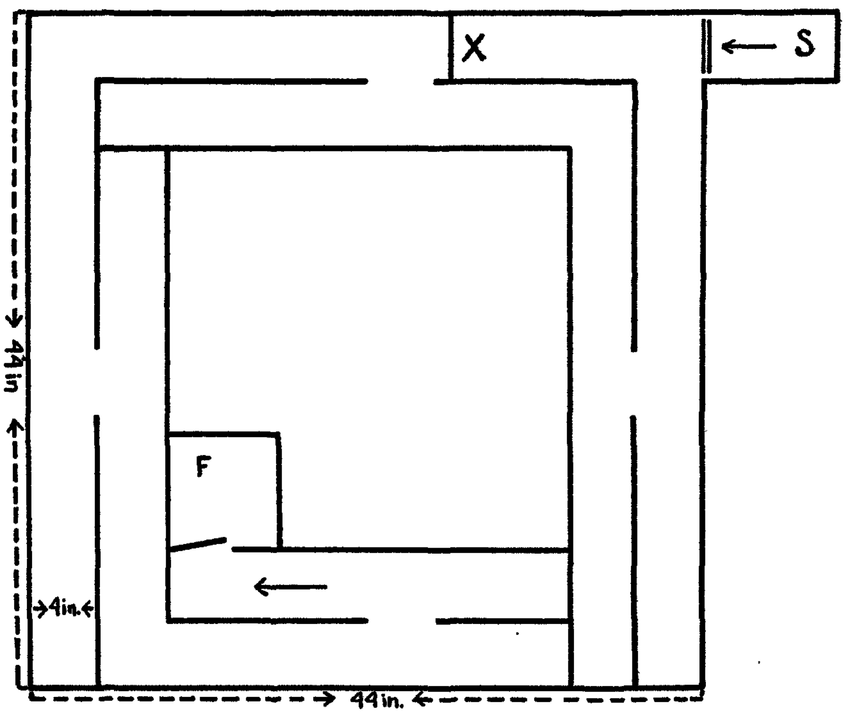
The records of 2 of the 21 rats had to be discarded because on one or more days they failed to complete the run and reach the food box. Of the remaining 19 rats, 7 males and 12

[1] Though white, the rats contained a small amount of known wild blood.

An entrance of a full body's length or more into a blind *while going forward*, or a retracing of more than one true path section—irrespective of how much longer it might be or how many blinds it might involve—were each counted as one error. This was, of course, quite arbitary. But it was used in the larger study and there proved quite satisfactory.

females, 13 were of approximately the same age, i.e., sixty days, at the beginning of the experiment. The remaining 6, all females, were older—approximately ninety days at beginning of the experiment. This fact of the greater age of the latter 6, which is known to affect time (3), and also the fact that due to changes being made in the building, the living conditions for

Figure 1.

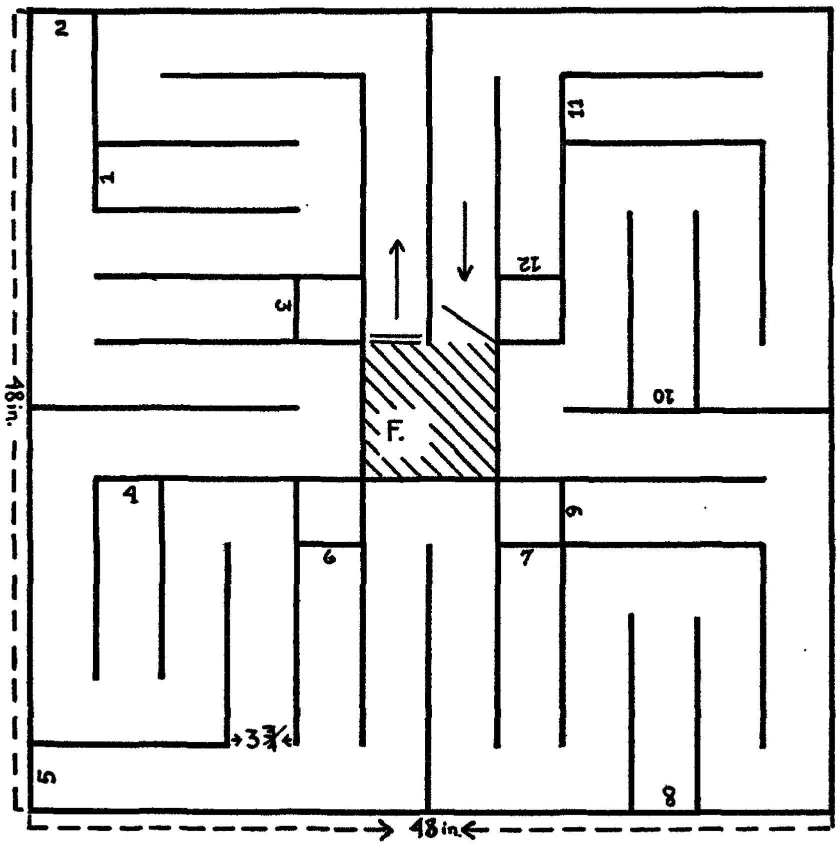


Maze A.

these latter 6 were decidedly disturbed during the running of the *second* maze, led us to be somewhat doubtful of the significance of including their results with those of the other 13. For these reasons we have in each case computed two independent sets of correlations—one using the results for the 13 younger rats only, and the other using those for all 19.

In the previous study (in which a maze practically identical[2] with maze A was used) an indication of the reliability of the error criterion was sought by correlating scores in odd runs

Figure 2.



Maze B.

against scores in even runs with the following results (82 cases). (All correlations in this paper have been computed according to the Pearson Product Moments formula.)

[2] The maze used there differed in one slight particular. The partition here marked *X* was shifted slightly nearer the starting box, thereby making one more, very short, blind not contained in the present maze.

Errors:

All runs included	r = 0.276
Runs 3 to 10 only	r = 0.379

The fact that 3 to 10 gave the higher correlations seemed to indicate that the error scores in runs 1 and 2 were largely a matter of chance. Such being the case we have used runs 3 to 10 only in all subsequent computations. Applying Brown's formula (4) to the 3 to 10 correlation, we got:

Errors: r = 0.550

or the probable correlation of the total error scores for all 8 runs with another similar total maze score.

This figure, while indicating enough reliability to make further work seem worth while, was not, of course, high. Examination of individual error records suggested, however, that if a rat's record on any given trial were to be scored not in absolute terms, but merely as greater than, less than, or equal to the average for all 82 rats, higher consistencies would have been obtained. In other words, it looked as if when a rat made an unusually great number of errors in any given trial (due perhaps to excitement), or an unusually small number (due perhaps to some lucky chance), that those extremes were not true measures of his real ability; that, in short, as long as a rat made more errors than the average for the given trial it had little significance whether he made *4 more or 40 more*. And, similarly, as long as he did better than the average, it had but little significance whether he did 1 less or 5 less errors than the average. The rat that made 40 more errors in one trial was no more likely to make a similarly excessive number in the next trial than the rat that made only *4* more errors than the average in the first of the two trials. And, similarly, the rat who by a lucky fluke made 4 errors less than the average in, say, the third trial was not any more likely to make a perfect run on the fourth trial than the rat who made, say, only 2 errors less than the average in the third trial.

To see if these assumptions were correct, we tried rescoring for trials 3, 5, 7, 9, and 4, 6, 8, 10, respectively, simply in terms of

1, 2, 3, for each trial accordingly as they were below, equal to, or above the average for that trial; and the resultant much higher intercorrelation was obtained.

Errors 3, 5, 7, 9, vs. 4, 6, 8, 10
$$r = 0.509$$

Applying Brown's formula to this, we got as the probable reliability coefficient of the score obtained from all 8 trials added together:

$$r = 0.675$$

a much more respectable figure.

TABLE 1

n = 13

4 RUNS VS. ANOTHER 4 RUNS	8 RUNS VS. ANOTHER 8 (BY BROWN'S FORMULA)	MAZE A VS. MAZE B
Errors, crude score:		
Maze A 0.268	0.422	0.658
Maze B 0.019	0.037	
Errors corrected score:		
Maze A 0.452	0.623	0.656
Maze B 0.339	0.506	
Time crude score:		
Maze A 0.119	0.213	0.160
Maze B 0.186	0.317	
Time corrected score:		
Maze A 0.627	0.771	0.456
Maze B 0.575	0.730	
Perfect Runs		0.537

Turning now to the present experiment, both procedures were again tried:[3] (*a*) of correlating *crude scores* for *errors* in trials 3, 5, 7, 9 vs. those in trials 4, 6, 8, 10 and (*b*) of correlating *corrected scores* (i.e., scores computed simply in terms of below, equal to, and above the average in each trial.[3] (See tables 1 and 2, column 1, rows 1, 2 and 3, 4.)

[3] In the case of maze A the averages used for computing the corrected score were those obtained from the 82 rats of the larger experiment. In the case of maze B they were averages from the 19 rats of the present experiment.

As expected, the corrected score gives very decidedly higher correlations.

Applying Brown's formula to these results, we got the expected correlation of maze A with another similar maze A (assuming that the learning of the first maze would not affect the result from the second) and of maze B with another similar maze B. (See tables 1 and 2, column 2, rows, 1, 2, 3 and 4.)

With these *calculated* probable correlations in mind between each maze and its own double, we may examine the *actual* correlations obtained between the two mazes (tables 1 and 2, column 3, rows 1, 2, 3 and 4).

TABLE 2

n = 19

4 RUNS VS. ANOTHER 4 RUNS	8 RUNS VS. ANOTHER 8 (BY BROWN'S FORMULA)	MAZE A VS. MAZE B
Errors crude score:		
Maze A 0.141	0.247	0.373
Maze B 0.334	0.501	
Errors corrected score:		
Maze A 0.339	0.506	0.389
Maze B 0.603	0.752	
Time crude score:		
Maze A 0.052	0.099	−0.096
Maze B 0.322	0.487	
Time corrected score:		
Maze A 0.553	0.712	0.313
Maze B 0.664	0.798	
Perfect runs		0.344

Assuming that going from maze A to maze B is about the same as going from either maze to its own double, then these figures are (at least for 13) certainly as high as, if not higher than, we would have expected from the calculated figures. A second point to be noted is that the correlations are practically identical whether crude scores or corrected scores are used, in spite of the fact that the calculated figures led us to expect higher correlations from the corrected scores. This was surprising. The only explanation we can think of is that whereas an occasional divergent result may very much upset the total score for only 4 runs, it will have a much less distorting effect upon the score for all 8 runs of the maze-learning. Or, in other words, the tendency to make one or more divergent results may after all be really characteristic of individual rats. And, although the chances are small that such divergent results will occur in equal amounts in alternate halves of one and the same maze-learning, they may well tend to occur to about equal amounts in each of two separate maze-learnings.

We may turn now to the second of our criteria—time. For it, also, we may correlate scores in odd runs vs. those in even runs—and we may use both crude scores and scores corrected in analogous fashion to that in which we used them for errors (tables 1 and 2, column 1). Here the uncorrected scores are nil, whereas the corrected scores are very respectable. Applying Brown's formula, these correlations are raised (tables 1 and 2, column 2).

Finally we turn to the actual correlation between the two mazes (column 3).

The noticeable feature about these figures for time is that in all cases they are much lower than the figures of columns 1 and 2 would have led us to expect. And this is particularly true for the *corrected* scores where the preceding figures were sufficiently high to be thought to have considerable significance. This low intercorrelation for time is interesting and suggests that time, in spite of its apparently high reliability within a single maze, as shown at least by the corrected scores, is nevertheless a somewhat ambiguous measure. Evidence that such is the case will be presented below.

Let us turn first, however, to the results for our third measure, perfect runs. In the case of this measure, no significant internal correlations within each maze by itself seemed possible. Therefore we can present merely the actual correlation obtained between the two mazes. This was: 0.537 ($n = 13$) and 0.344 ($n = 19$), (tables 1 and 2).

It will be observed that these figures are higher than those for time. They are much more nearly like those for errors.

Finally, we computed the intercorrelation between errors, time, and perfect runs for each maze by itself. They are shown in table 3.

They come out as expected: the greater the number of errors, the greater the time; the greater the number of errors, the smaller the number of perfect runs; and the greater the time, the smaller the number of perfect runs.

TABLE 3

(n = 13)	(n = 19)
Maze A	
$r_{ET} = 0.930$	$r_{ET} = 0.844$
$r_{EP} = -0.926$	$r_{EP} = -0.886$
$r_{TP} = -0.850$	$r_{TP} = -0.636$
Maze B	
$r_{ET} = 0.580$	$r_{ET} = 0.666$
$r_{EP} = -0.717$	$r_{EP} = -0.784$
$r_{TP} = -0.023$	$r_{TP} = -0.271$

E = number of errors, trials 3 to 10 (crude score).
T = time, trials 3 to 10 (crude score).
P = number of perfect runs in the ten trials.

TABLE 4

(n = 13)	(n = 19)
Maze A	
$r_{ET.P} = 0.718$	$r_{ET.P} = 0.785$
$r_{EP.T} = -0.710$	$r_{EP.T} = -0.843$
$r_{TP.E} = 0.080$	$r_{TP.E} = 0.450$
Maze B	
$r_{ET.P} = 0.812$	$r_{ET.P} = 0.759$
$r_{EP.T} = -0.870$	$r_{EP.T} = 0.841$
$r_{TP.E} = 0.693$	$r_{TP.E} = -0.542$

But let us examine the partials shown in table 4.

Comparing the partials (table 4) with the originals (table 3) the outstanding point is that the correlations between time and perfect runs were in every case changed from negative to positive.[4] In other words, if the effect of errors on time is eliminated, the true correlation between time as such and perfect runs is positive, not negative. In other words, the slow-running (cautious ?) rat (the rat with the bigger time), when the effect of errors upon time has been eliminated, is the more perfect; while the rapid running (slap-dash ?) rat is the less perfect. The latter tends to overrun; to continue to make a few final errors because of too much speed. He gets going so fast that he runs into a blind before he knows it. (Indeed, such behavior is familiar enough to those who have watched rats perform.) If, now, this interpretation is correct, it appears that time must be an ambiguous measure. In part, it is a mere measure of errors and, in part, of an independent factor which we may perhaps call cautiousness or deliberateness. And this second factor may have no correlation with errors, and indeed, tends in the last stages of learning to actually reduce the number of errors.

It would be seen, then, that in spite of the general practice to favor time as the one best single measure (5)—that in reality it is a decidedly ambiguous and uncertain one. And, indeed, this conclusion was, as we saw above, borne out by our actual correlations obtained between the two mazes. Those for time were the lowest of all, and this in spite of the fact that within the single maze the time reliabilities (at least for the corrected scores) were decidedly high. In other words, the time records for an individual rat in a given maze may be very consistent and yet, vary the conditions in some respect, i.e., change the maze, and they may become quite different. For, if we take the ability to make a small total of errors (i.e., a short distance if distance had been measured) as the one most certain criterion of maze ability, it appears that time tends to be correlated in two opposite directions with this ability. It tends, on the one hand, in a purely mechanical fashion to increase as errors increase; but, on the other hand, taken by itself and independent of this

[4] A similar positive partial correlation between time and perfect runs with error constant was found in the original study.

purely mechanical connection, it tends to have the opposite relation: i.e., the slower, the more cautious, the rat, the fewer the number of *final* errors. This would seem to mean that any change in conditions, such as changing from one maze to another, which would vary the equilibrium between these two factors, may quite change the relative time records. In the one maze the blinds may be short and the purely mechanical effect of errors upon time may be less important. In such a maze, the differences due to differences in degree of caution[5] may have relatively great weight in the total time score. The better rat will tend towards the larger time score. In a second maze, on the other hand, the effect of errors on time may bulk large—in which case the poorer rat will have the larger time score.

Before concluding, there is one more point of possible significance which comes out when we compare table 2 with table 1. For it will be observed that although the correlations in columns 1 and 2 in table 2 do not run very much lower than the corresponding ones in table 1, this is not true for column 3. There the correlations of table 2 are decidedly lower. And this perhaps means that although introducing variation in conditions (such as age) may not affect the reliability of each single maze per se they may tend to vary the order of merit as between two mazes.

We may sum up our conclusions as follows:

1. It must be admitted that in general our correlations run low (tables 1 and 2) and this would support the data presented by Hunter and his students.

2. It must also be admitted, however, that in the case of errors (for the 13 rats, table 1) the correlation between the two mazes ran rather higher than was to be expected from the correlations within the single mazes.

3. The "corrected scores" while raising the apparent correlations within each maze did not seem to raise the correlations as between mazes.

4. Time seems to be the most unreliable measure of all as *between* mazes. And this is to be explained, we believe, by the

[5] Using this concept of course in a purely behavioristic sense.

fact that time depends upon two independent factors—one of which tends to make it go down with maze-ability and the other of which tends to make it go up. And these two factors, while constant for any one maze, may tend to be different for different mazes.

5. The correlations obtained for perfect runs are similar to those obtained for errors.

6. Comparing table 2 with table 1 it appears that variations of age and environmental conditions lower the correlations *between* mazes more than they do the correlations within mazes.

REFERENCES

(1) Tolman, E. C.: The inheritance of maze-learning ability in rats. Jour. Comp. Psychol., 1924, iv, no. 1, 1–18.

(2) Hunter, Walter S.: Habit interference in the white rat and in human subjects. Jour. Comp. Psychol., 1922, ii, 29–59.

Hunter, Walter S.: Correlation studies with the maze in rats and humans. Comp. Psychol. Mon., 1922, i, no. 1, pt. 2.

Heron, William T.: The reliability of the inclined plane problem box as a measure of learning ability in the white rat. Comp. Psych. Mon., 1922, i, no. 1, pt. 1.

(3) Hubbert, H. B.: The effect of age on habit formation in the albino rat. Behav. Mon., 1915, ii, no. 6.

(4) Kelley, Truman L.: Statistical Method. MacMillan, 1923, p. 205–6.

(5) Hicks, Vinnie C.: The relative values of the different curves of learning. Jour. Animal Behav., 1911, i, 138–156.

Hubbert, H. B.: Time versus distance in learning. Jour. Animal Behav., 1914, iv, 60–69.