A perceptual maze test sensitive to brain damage

A. L. BENTON, A. ELITHORN[1], M. L. FOGEL, AND M. KERR[1]

From the Departments of Neurology and Psychology, University of Iowa, U.S.A., and the Royal Free Hospital and National Hospitals for Nervous Diseases, London

Following a report (Elithorn, 1955) that performance on a perceptual maze was sensitive to changes due to age and to small focal cerebral lesions, one of us (A.L.B.) decided to include this test in an extensive battery of psychological tests which were given to a group of brain-damaged patients and a comparable group of control subjects. The present paper reports the findings with this test and records briefly the relationships of these findings to the results obtained with the other tests. Some of the latter results have been reported in previous papers (*e.g.*, Bechtoldt, Benton, and Fogel, 1962; Benton, 1962; Benton and Fogel, 1962).

DESCRIPTION OF THE TEST

The test material (Fig. 1) consists of 20 patterns each of which has a lattice background of paths or tracks. At a variable number of intersections dots are superimposed. In each pattern the subject is required to find a pathway from bottom to top, which passes through the greatest possible number of dots. There are two restrictions. First, the subject must keep to the paths or tracks and must not cut across from one path to another. Secondly, at any junction point the pathway chosen by the subject must continue forwards, *i.e.*, he may fork left or right but must not double back. Each pattern is so designed that there is only one correct series of dots through which the pathway can pass, although there may be several

[1]Members of the external scientific staff of the Medical Research Council, London.

alternative routes through the correct series. In order that the subject may know when he has found a correct pathway, the maximum number of dots which can be obtained in each pattern is printed underneath the pattern. The patterns are not identical with those used in the preliminary investigation. They were printed from movable type which gives a clearer and more uniform reproduction. Further, in producing the movable type the opportunity was taken to redesign the lattice so that the angles at the apices are 60°. This change makes the distances between the interstices equal both horizontally and along the lines of the lattice. The items were arranged in approximate order of difficulty and were presented in this order, each test being preceded by the administration of a demonstration pattern. Throughout this investigation the standard form of the test was used, that is to say, the items were presented with the identifying letter in the lower right-hand corner from the patient's point of view. The presentation of these patterns inverted provides an alternative version of essentially equivalent difficulty (Elithorn, Kerr, and Mott, 1960).

SUBJECTS

The subjects were 100 patients from the neurological and neurosurgical services of the University Hospitals of Iowa City and 100 control patients of comparable age and education from these services or from the general medical wards. The criteria of selection were that there should be adequate evidence of disease or injury involving the cerebral hemispheres in the first group and that there should be no suspicion of such damage in the control group. In the brain-damaged group it was possible to
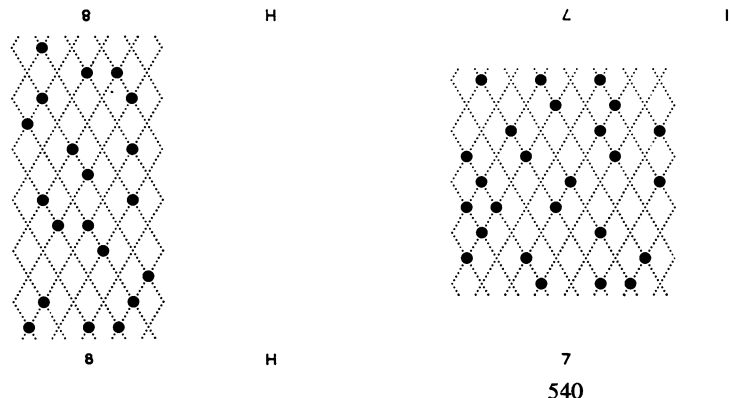

FIG. 1. *Sample patterns.*
Item H was solved in one minute by 66% of the control patients and 23% of the brain-damaged patients.
Item I was solved by 87% of the controls and 55% of the brain-damaged patients.

localize the lesion to one hemisphere in 83 patients and in 55 patients to restrict the localization to the anterior or posterior half (pre- or post-rolandic) of one or both hemispheres. The average age of the brain-damaged group was 42.00 ± 1.27 years, that of the control group 42.02 ± 1.16 years. The mean educational levels in terms of years of schooling were 10.23 ± 0.30 and 10.13 ± 0.24 respectively. All patients were within the range of 15 to 66 years, were capable of understanding directions and cooperating in the tests, and were able to undergo one hour of psychological testing without discomfort. None had sensorimotor deficits which affected test performance. Any patient who was acutely ill, whose behaviour raised the question of psychosis, who had a history of admission to hospital for a psychiatric disorder, or whose history suggested mental deficiency dating back to childhood was excluded from consideration. Many patients in the brain-damaged group were out-patients who were earning a living or successfully running a household and were seen when they returned for a check-up visit. No patient with a history of seizures, head injury followed by unconsciousness, or any history or findings suggestive of cerebral disease or injury was included in the control group.

METHOD OF ADMINISTRATION AND SCORING PROCEDURE

Each patient was seated comfortably at a table and the material was presented with the following explanation. The explanation was not made verbatim but was, if necessary, elaborated in order to make sure that the subject understood the nature of the task.

'In this test you have to find the best of several possible paths. I will show you a series of designs like this (demonstration pattern). In each one the paths are represented by dotted lines. You see that the paths keep crossing each other. At some of the meeting points there are large black circles. You are to find the path from bottom to top which goes through the largest possible number of circles. There are two main rules: First, stay on the paths and do not cut across the white spaces. Second, keep going forward and do not double back. At each intersection you can go left or right, but not backwards (left, right, and backwards demonstrated). Down here there is a number. Each pattern has a number like this which tells how many circles you have to go through. In this pattern (demonstration) you have to find a path which goes through

three circles. You may start anywhere along the bottom line. Here is a path which does not go through any circles at all (demonstrated); if you started here you could find a path which would go through two circles (demonstrated). In the centre of this pattern there are three circles but it is not possible to take a path which goes through all three without going back or across a white space, and neither one of these is allowed. But here there are three circles through which you can draw a path without breaking any of the rules (demonstrated). For each pattern there is only one series of circles which will give you the correct answer, but it may be possible to go through these in slightly different ways. In this pattern you could get the correct solution from several starting points (demonstrated; then demonstration pattern reversed). Now you try it.'

If the patient failed, the points were demonstrated again and he was then given a fresh demonstration pattern.

The patient was then told that he would have a minute to do each pattern and the 20 test patterns were then presented in alphabetical order. No further corrections or demonstrations after failure were given. The maximal time for scoring performances was one minute for each pattern. However, in order to avoid frustration and discouragement, if at the end of one minute the patient seemed to be on his way toward completing his response or was engrossed in the task, he was permitted up to 30 seconds to do so. However, such an 'over-time' performance was scored as a failure. For each pattern successfully solved the time to the nearest second was recorded.

RESULTS

The distribution of scores for both controls and brain-damaged subjects is shown in Table I. The rows b and c are the scores obtained within the 60-second limit. Using a 'below distribution' criterion, 16 (16%) of the brain-damaged group scored less than 6, the lowest score in the control group. Forty-three (43%) of the brain-damaged subjects scored less than 9, while only five subjects (5%) in the control group scored less than this. Employing the criterion of the optimal cutting score (here 10 or less) the perceptual maze test correctly identified 71.5% of the combined groups.

TABLE I

DISTRIBUTION OF MAZE SCORES

Score	a	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
n 60 seconds																																				
-ols	b							1	1	3	4	6	10	10	12	13	13	14	9	4																
-damaged	c	2	3	1	1	2	7	7	10	10	6	9	9	6	9	6	4	6	2																	
n 30 seconds																																					
-ols	d					3	3	7	7	6	10	9	11	11	16	6	6	3	2																	
-damaged	e	2	5	2	6	12	16	6	9	7	6	12	6	8	2	0	0	1																		
hted score																																					
-ols	f											1	0	1	0	2	2	5	5	4	5	2	2	7	11	1	9	4	8	6	9	4	5	1	3	2	1
-damaged	g	2	0	3	1	0	1	2	1	0	7	6	5	5	6	3	5	6	2	6	4	5	1	4	5	4	4	2	3	3	2	1	0	1			

Since time scores were kept, it is possible to determine whether better discrimination (or at least no impairment of discrimination) could be achieved by reducing the time allowed for each pattern. The rows d and e give the distribution of scores obtained if a credit is given only if the subject solves the pattern in 30 seconds or less. It is clear that this greatly reduced the discriminative power of the test. The lowest score recorded by the controls was 3 and only 9% of the brain-damaged group scored less than this. If a 5% cut-off point is used, then only 15% brain-damaged patients were identified.

It was suggested in a previous paper that the time scores be used to derive a score weighted with an additional credit for each item solved in less than 30 seconds. While this procedure gives a wider range and better distribution of scores with normal subjects (Elithorn et al., 1960), the present results suggest that it does not increase the value of the test as an index of brain damage and may even reduce it (Table I, rows f and g). Using such weighted scores, only 10 (10%) of the brain-damaged group scored less than the lowest score made by a member of the control group and only 39 (39%) scored less than the 5% level for the control group. However, the optimal cutting score correctly identified 70·5% of the combined groups.

An alternative to the method of reducing testing time by lessening the time allowed per pattern is to reduce the number of items. Examination of the results obtained with each pattern showed that some items discriminated very poorly. When items B, E, F, L, and T were eliminated, then on the remaining 15 items, 29 (29%) patients from the brain-damaged group scored less than 4, the minimum for the control subjects. If the 5% criterion is used then 40% of the brain-damaged patients fell below the cutting point compared with 43% using a similar criterion with the standard 20 patterns. The optimal cutting score was 70·5%, almost equal to that of the 20 pattern version. It seems likely therefore that further studies will show that the shorter version may be useful where testing time is limited.

The problem as to what ability or abilities this test measures will be largely deferred, to be discussed in detail in another publication devoted to an analysis of the correlation matrices derived from all the tests used in the Iowa battery. However, it is germane to ask whether, as a test which appears to demand spatial skills, there is any relationship between the degree of impairment shown and the site of the lesions. Further, it has been suggested that maze tests are particularly sensitive to frontal lobe damage (Porteus and Kepner, 1944). With the present series of patients it has been possible to localize 19 lesions as frontal and 36 as post-rolandic, while 45 were presumed to involve both frontal and post-rolandic tissue. Forty-five lesions could be localized to the left hemisphere and 38 to the right; 17 involved both hemispheres. Table II shows the relationship between the locus of the lesion and the frequency of scores less than 9 (5% criterion score for the 20 patterns).

It is clear that a grossly impaired performance on the maze is no more common with frontal lesions than with posterior lesions but it does appear to be associated with lesions which are in the right hemisphere rather than the left. This difference in relative frequency of defective performance in the patients with right and left hemisphere lesions reaches the 0·05 level of significance ($\chi^2 = 3·99$, d.f.1, $0·02 < p < 0·05$). However, even if further testing confirms that the contingency is significant statistically, it may be that this association represents merely the fact that the lesions localized to the right tended to be larger than those localized to the left because of the absence of aphasia as an early symptom leading to earlier admission to hospital. That this interpretation would not account for this contingency is suggested by the fact that patients with right hemisphere lesions not only perform worse than those with left hemisphere lesions, but also perform worse than those with bilateral cerebral damage.

The principal interest of the present test material lies in the possibility that it may test a fairly discrete ability which contributes to an appreciable degree to intelligent behaviour (Elithorn, 1955). Its immediate utility in the clinical field depends on its power to discriminate brain-damaged patients compared with

TABLE II

PATIENTS SCORING BELOW CRITERION CLASSIFIED BY LOCUS OF LESION[1]

	Left		Right		Bilateral		Total	
	No. Below Criterion	Total in Group	No. Below Criterion	Total in Group	No. Below Criterion	Total in Group	No. Below Criterion	Total in Group
Frontal	0	6	5	10	2	3	7	19 (37%)
Post-rolandic	5	16	8	18	0	2	13	36 (36%)
Unlocalized	9	23	8	10	6	12	23	45 (51%)
Total	14	45 (31%)	21	38 (55%)	8	17 (47%)	43	100 (43%)

[1]The 5% cut-off point corresponding to a maze score of 8 or less.

the discriminative powers of other clinical tests. The Iowa battery contained 27 other tests, which are listed in the Appendix. If the 'below distribution' criterion for cut-off scores is used, then the maze ranked fifth in efficiency in detecting brain-damaged subjects. However, the below-distribution criterion suffers from the severe disadvantage that a catastrophic performance by a single normal subject leads to complete rejection of the test concerned. On four of the 28 tests in the Iowa battery, at least one control subject achieved no better score than the lowest achieved by any patient in the brain-damaged group. A better criterion is to take as the cutting score the highest score which is surpassed by at least 95 % of the normal group. Using this criterion, the maze correctly identified 43 % of the brain-damaged group and ranked second among the 28 tests. If the 1 % criterion is adopted, taking as a cutting score the highest score which is surpassed by at least 99 % of the normal group, the maze test ranked fifth.

The clinical psychologist seeks in the first instance a battery of tests which will give him a good chance both of detecting the presence of organic damage and of being able to localize its focal components. The value of any given test in a battery will depend not only on its discriminative power *per se*, but also on its ability to contribute information which is additional to and supplements that supplied by the other tests in the battery. This component will clearly be dependent on the composition of the battery. Hence it is of some interest to consider, as a first approximation, the additional contribution to discrimination which can be achieved by pairing the maze with other tests in the Iowa battery. When paired with each of the 27 other tests in turn, the perceptual maze gives an average identification of 50 % of the brain-damaged patients. This is not bettered by any other tests, 'blocks and sticks', the next most sensitive by this criterion, giving an average figure of 49 %. The highest discrimination shown by the maze in combination is that found when it is paired with oral picture description (57 %). The only other pairing which gives as high a discrimination is oral picture description combined with blocks and sticks. It must be remembered, however, that each test misidentifies five normal patients which may lead to misidentification of 10 % normal subjects for the combined tests. In the two pairings quoted above, the misidentification of normal subjects is 10 for both pairs.

In assessing the value of the perceptual maze test as a clinical aid in the detection of brain damage, we have so far considered the raw scores. Since the performance of brain-damaged subjects is determined in part by their premorbid level of ability it is often useful to relate their observed performance

to an estimate of the score they might be expected to achieve in the absence of any cerebral pathology. Premorbid performance on any test in which the score is in part a function of intelligence will correlate positively with an estimate of the subject's acquired knowledge and with the number of years devoted to his education and negatively with his chronological age. Since the data for age and number of years of education for each subject were available, simple and multiple regression equations for the maze scores were calculated for the control group using these two variables together with the score on the Wechsler comprehension subtest. From the multiple regression equation, a predicted score and hence a discrepancy score for each brain-damaged patient was obtained.

The relevant equations, where x_1, x_2, and x_3 denote comprehension score, age, and education respectively, and y the regression of maze scores on these variables, are:

SIMPLE REGRESSION
$$y = 7.71 + 0.52\,x_1$$
$$y = 14.25 - 0.02\,x_2$$
$$y = 6.76 + 0.66\,x_3$$

MULTIPLE REGRESSION $\quad y = 6.55 + 0.387x_1 - 0.006\,x_2 + 0.284\,x_3$

From the last equation the 95 % confidence limits for the predicted maze score for each individual have been calculated as $\pm$ 4.27. Using this as the cut-off point, 37 % brain-damaged patients are correctly identified, but as the confidence limits apply to both tails of the distribution, only two or three normal subjects could be wrongly identified using this method. Taking the cut-off point at a level which misidentifies five (5 %) normal subjects, the number of brain-damaged patients correctly identified increases to 44 %. The increase in sensitivity gained by taking account of age, educational level, or the Wechsler comprehension subtest score is not great, and would hardly seem to justify the additional calculations. Moreover, the Wechsler comprehension subtest is itself susceptible to brain damage and the criterion of educational level is applicable only to the United States. For the present control group the correlation between maze score and age is not high ($r = -0.330$), but a previous study gave a higher correlation ($r = -0.500$), and it is possible that a correction for age based on further studies would increase the discriminative power of the test.

SUMMARY

A perceptual maze test was given to 100 patients with cerebral disease and 100 control patients. Interest

was focused on the following questions: 1 The degree to which test performance differentiated between the brain-damaged and control patients; 2 the discriminating power of the test as compared to that of each of 27 other tests given to the same patients; 3 the association between test performance and locus of lesion in patients with focal cerebral lesions; 4 the additional contribution made to discrimination when the maze test was paired with other tests in the battery.

As an index of brain damage, the maze test proved to be one of the most sensitive of the 28 tests in the battery. There was a somewhat higher incidence of defective performance in patients with lesions of the right hemisphere than in those with lesions of the left hemisphere. However, the findings provided no evidence to support the thesis that performance on the test is impaired specifically by lesions localized in the frontal lobes. The combination of the perceptual maze test with a picture description test resulted in the correct identification of a proportion of brain-damaged patients which was not exceeded by any other pair of tests in the battery.

REFERENCES

Bechtoldt, H. P., Benton, A. L., and Fogel, M. L. (1962). *Psychol. Rec.*, **12**, 147.
Benton, A. L. (1962). *Confin. neurol. (Basel)*, **22**, 141.
——, and Fogel, M. L. (1962). *Arch. Neurol.*, n.s. 7, 347.
Elithorn, A. (1955). *J. Neurol. Neurosurg. Psychiat.*, **18**, 287.
——, Kerr, M., and Mott, J. (1960). *Brit. J. Psychol.*, **51**, 19.
Porteus, S. D., and Kepner, R. de M. (1944). *Genet. Psychol. Monogr.*, **29**, 3.

APPENDIX

The following tests were used in the investigation:—

Benton visual retention (form C), Benton visual retention (form G, multiple choice), blocks and sticks, colour cognition, colour naming, drawing, finger localization, finger touching, Gorham proverbs, minute estimation, oral arithmetic (calculation), oral picture description, paired associated verbal learning, paragraph reading, perceptual maze, Rey test, right-left discrimination, tactile form board, temporal orientation, W.A.I.S. (arithmetic, block design, comprehension, digit span, picture arrangement, and similarities), word fluency, writing, written arithmetic (calculation).